Available online at www.sciencedirect.com

ScienceDirect

Underground Space 25 (2025) 327–349

www.keaipublishing.com/undsp

Research Paper

Intelligent lithology identification via spectral-image fusion

Zhenhao Xu *, Shan Li, Peng Lin, Qianji Li

State Key Laboratory for Tunnel Engineering, Jinan 250061, China
School of Qilu Transportation, Shandong University, Jinan 250061, China
Institute of Geotechnical and Underground Engineering, Shandong University, Jinan 250061, China

Received 20 October 2024; received in revised form 15 May 2025; accepted 31 May 2025
Available online 15 October 2025

Abstract

Lithology identification is of vital significance for fundamental geological research and engineering applications. Traditional methods rely on rock image features and often cause confusion among visually similar rocks. To enhance identification accuracy, spectral features are integrated to better represent rock composition. Nonetheless, spectral testing inevitably damages samples and is prone to challenges of the occurrence of similar spectra for different materials. This study explores the advantages of hyperspectral imaging technology, enabling the integration of spectral and image data without damage or contact. A novel spectral-image fusion method is proposed for lithology identification. It uses a dual-channel residual neural network that combines spectral and texture feature channels. Additionally, key features are effectively captured by spectral-spatial attention mechanisms. Finally, a customized transfer learning approach is utilized to fine-tune the pre-trained network on a small dataset for lithology identification at the tunnel site, facilitating rapid model adaptation. The research indicates that employing the ResNetX2-50 network for integrating spectral-image information yields optimal identification results, with a fusion accuracy of over 99% on the test set, which is 2 percentage points higher than the accuracy of a single spectral model and about 20 percentage points higher than the accuracy of a single texture model. Research findings provide robust technical support for non-destructive, non-contact, fast lithology identification in field investigations and construction projects.

Keywords: Spectral-image feature fusion; Spectral-spatial attention; Feature level fusion; Intelligent identification of lithology; Transfer learning

1 Introduction

Geological analysis is a core component of geotechnical engineering and fundamental to ensuring project safety, stability, and economic efficiency (Zhang et al., 2023; Li et al., 2023; Xu et al., 2023; Liu et al., 2025). Lithology identification plays a critical role in geology, geotechnical investigation, and engineering, supporting mineral/petroleum exploration and optimizing design, safety assessment, and risk evaluation for geotechnical projects (You et al., 2024; Rashid et al., 2023; Wang et al., 2024; Xu et al., 2024). Traditional methods for lithology identification primarily involve visual observation, thin section analysis, and techniques like X-ray fluorescence/ X-ray diffraction (XRF/XRD) (Gao & Jiao, 2022; Xu et al., 2022, 2025). However, these approaches have their inherent limitations. Visual observation is subjective and error-prone for similar-looking rocks; thin-section microscopy requires strict experimental conditions and sample processing, unsuitable for in-situ analysis; XRF/XRD demand destructive sample grinding (Xu et al., 2021; Muktadir et al., 2021; Barker et al., 2021). Hyperspectral technology offers distinct advantages for lithology identification with its non-destructive, in-situ, non-contact, and large-area detection capabilities, enabling rapid analysis of abundant spectral-spatial data (Lin et al., 2023; Zhang et al., 2024).

Hyperspectral remote sensing, prized for its high-resolution, non-destructive, and non-contact detection, is widely used in food safety, medicine, agriculture,

* Corresponding author at: State Key Laboratory for Tunnel Engineering, Jinan 250061, China.
E-mail address: zhenhao_xu@sdu.edu.cn (Z. Xu).
Peer review under the responsibility of Tongji University

https://doi.org/10.1016/j.undsp.2025.05.011

environmental monitoring, and resource exploration (Tan et al., 2021; Wang et al., 2022). It has become a key research area in machine vision, focusing on hyperspectral image acquisition, analysis, and processing. These images integrate spectral (wavelength-specific reflectance) and spatial (pixel layout) information. Hyperspectral image classification, a core task, assigns labels by analyzing object spectral reflection and absorption across bands (Audebert et al., 2019; S. Li et al., 2019; Yang et al., 2023). Classification methods are grouped into spectral-only, spatial-only, and feature-fusion approaches (Tejasree & Agilandeeswari, 2024). Early methods, like support vector machine (SVM) and *K*-nearest neighbor (KNN), relied on spectra but faced challenges due to hyperspectral images' complex spatial and spectral heterogeneity (Z. Li et al., 2024). Research shows that spatial-spectral feature fusion improves classification accuracy. Zhong et al. (2018) effectively integrated spectrum-space features using 3D convolutional neural network (3D CNN) and adopted the ResNet structure to alleviate the phenomenon of declining accuracy during back propagation of gradients. F. Li et al. (2019) proposed a novel hyperspectral image classification method based on a multi-feature fusion strategy that combined spectral space features from spectral-spatial feature learning (SSFL) with texture features from local binary pattern (LBP) images. Zhang et al. (2019) introduced a spatial-spectral feature fusion model based on predictive feature weighting, which utilized one-dimensional convolutional local spectral filters and two-dimensional spatial-spectral filters in an end-to-end manner for input into the fusion module. Wang and Zuo (2024) developed a two-branch 1D-2D CNN lithology mapping method based on Gaofen-5 (GF-5) images to characterize spectral features, local spatial features, and neighboring relationships. X. Li et al. (2024) designed Spe-Spa MN, a spectral and spatial feature fusion network used to predict soil carbon content through multi-scale feature fusion. Currently, a substantial body of relevant literature has demonstrated that the algorithm incorporating both spectral and spatial features confers significant advantages in hyperspectral image classification.

Using single hyperspectral features for lithological identification has limitations: spectral-only analysis may fail due to "same-spectrum foreign objects" (similar reflectance across different rocks) or weathering/metamorphism-altered spectral characteristics, while texture-only methods struggle with similar surface textures across rock types. Multi-source information integration has become a key research focus (Wang et al., 2021; Zhang et al., 2023; Lu et al., 2023). Hyperspectral imaging captures both spectral and texture features, and their fusion reveals complex rock properties unattainable by single-source data, providing comprehensive lithological analysis (He et al., 2018). This is particularly critical for weathered/metamorphosed rocks, where spectral data alone are compromised by physical/-chemical alterations; integrating inherent texture/structural information with spectra aids in identifying original and weathered rock traits. This approach expands feature dimensionality, enhances model robustness, and leverages spectral-texture complementarity to improve identification accuracy for heterogeneous, geologically complex samples, outperforming single-feature methods.

In summary, there are several issues with lithology identification based on hyperspectral images, including (1) mitigating the "curse of dimensionality" in high-dimensional spectral data and optimizing feature extraction (Ferrari et al., 2013); (2) achieving spectral-spatial correspondence with traditional feature extraction methods (Chen et al., 2021); (3) avoiding overfitting in fusion network designs (Hong et al., 2022); and (4) developing efficient models for new datasets with limited data (Ahmad et al., 2022).

To solve these issues, this study proposes an intelligent lithology identification method via deep fusion of spectral-image features with spectral-spatial attention guidance. It establishes a spectral-image fusion classification framework using transfer learning to address scarce on-site rock samples. First, principal component analysis (PCA) downsizes spectral-spatial dimensionality to extract representative features efficiently. A local co-occurrence matrix with adaptive sliding windows is introduced to refine texture feature extraction, addressing limitations of traditional global texture analysis. For hyperspectral data-based fine-grained lithology identification, the method ensures spectral-texture feature alignment through shared pixel-level class labels. A dual-channel residual neural network (based on ResNet) fuses spectral and texture features at the feature level, and spectral-spatial attention mechanisms are integrated to enhance critical regional features, thereby improving recognition accuracy. To overcome deep learning data/model limitations, transfer learning fine-tunes pre-trained networks on small geological datasets, boosting model convergence and generalization for tunnel site lithology identification.

2 Spectral-image fusion identification method

Imaging hyperspectral technology allows for the simultaneous acquisition of image and spectral information from rock samples, enabling the comprehensive capture of various characteristic features such as external morphology, internal structure, and chemical composition. By assigning unified labels to rock image and spectral features collected from the same location, and utilizing a lithology intelligent identification model that integrates spectral-image feature fusion and spectral-spatial attention guidance, the model continuously learns the characteristics of each lithology. Ultimately, the model can return a label for classifying the imported dual-channel data. At the same time, using transfer learning methods, pre-trained parameters based on laboratory data were loaded onto a deep transfer learning model based on on-site rock imaging spectral data, ultimately achieving intelligent identification of rock types in engineering sites. The framework of the

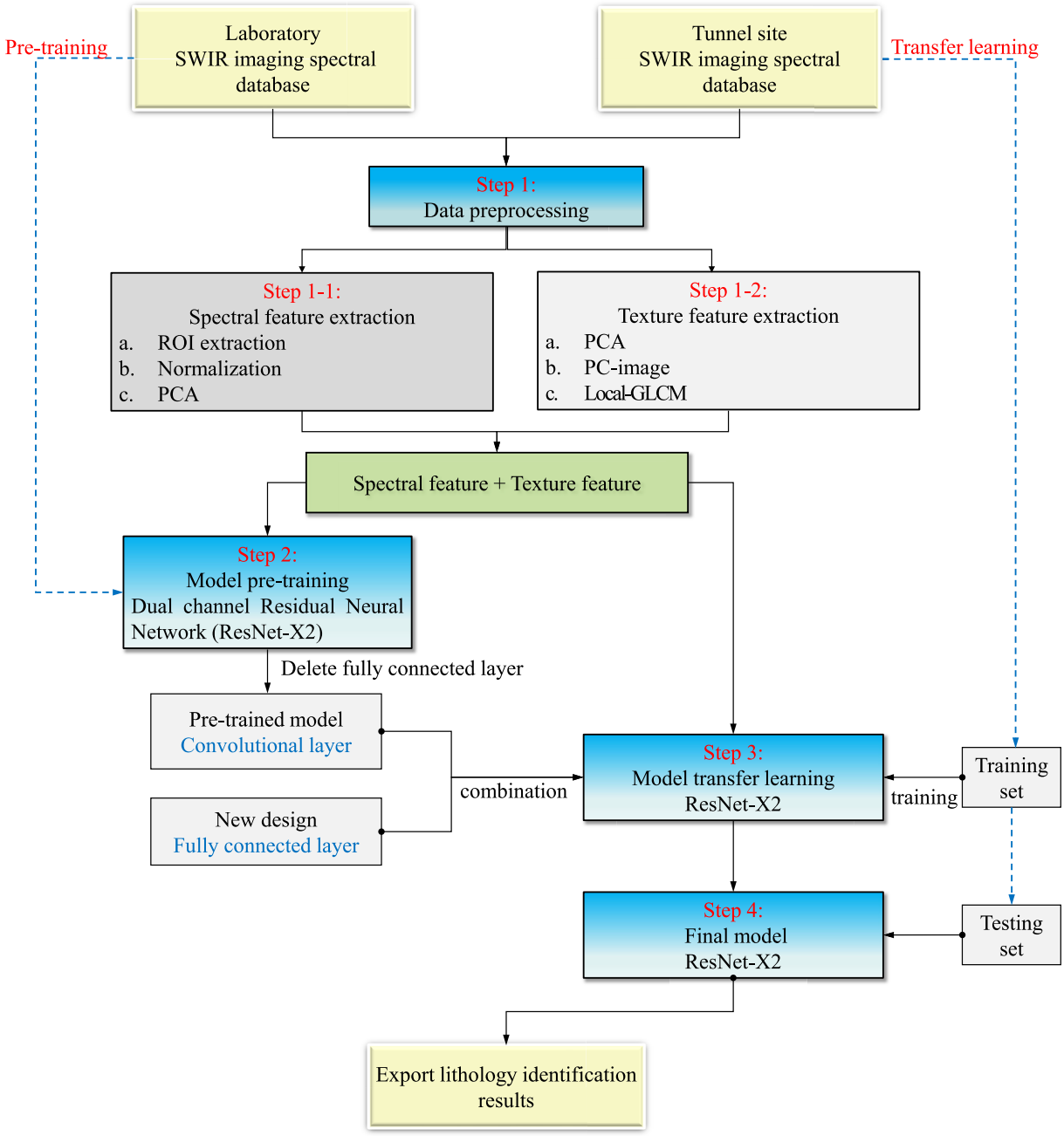


Fig. 1. Framework of fusion identification method.

fusion identification method using spectral and image of hyperspectral image (HSI) is shown in Fig. 1.

2.1 Feature extraction

Hyperspectral images encompass a wealth of spectral and image information. The spectral characteristics of rocks are primarily influenced by their mineral composition, internal crystal structure, and physicochemical properties. Image features, on the other hand, predominantly encompass the roughness, shape, and distribution patterns of the rock surface, offering crucial insights into the rock's formation process and its surrounding environment. In this study, normalization and PCA dimensionality reduction are employed to extract spectral features from region of interest (ROI), and the local gray-level co-occurrence matrix (Local-GLCM) is utilized to extract image features from PCA-reduced hyperspectral data.

2.1.1 Hyperspectral image preprocessing

Before utilizing imaging hyperspectral data for lithology identification, the initial step involves precisely extracting the ROI from the hyperspectral image. Subsequently, ROI is employed to acquire a high-quality dataset.

Identifying ROI not only considerably diminishes the complexity of data processing but also enhances the precision and efficiency of subsequent analyses. During the ROI determination process, it is imperative to consider the representativeness and diversity of the dataset, ensuring comprehensive coverage of various rock types and also encompassing regions of the same rock type under differing weathering and metamorphic conditions, thereby accommodating variations in rock physical properties and spectral features. This study randomly marked representative regions for each rock type. These regions encompass a vast amount of spectral information pertaining to rock imagery, furnishing high-quality input data for the training and testing of deep learning networks.

2.1.2 Spectral feature extraction

During rock spectral data collection, variations in instruments, lighting, sample backgrounds, and morphologies cause significant inter-pixel spectral differences across bands. Normalizing spectral data becomes a critical preprocessing step to enhance data stability, analyzability, and interpretability before deep learning model input. Once the ROI is defined, all pixel spectral values within it are standardized across bands to ensure inter-pixel comparability, emphasizing individual pixel spectral intensity changes and facilitating lithology differentiation.

Normalization is a widely used preprocessing method in signal processing and data analysis, aimed at achieving consistency in data values by adjusting the scale of the data. For each value x in the raw data, it is converted to a normalized value x' within a new scale based on the selected range [0, 1] and the minimum $x_{\min}$ and maximum $x_{\max}$ values of the raw data. The specific formula is as follows:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}. \tag{1}$$

Hyperspectral imaging's high dimensionality, while rich in material spectral details, increases computational costs and introduces redundancy, potentially degrading lithology identification efficiency. This study applies PCA to normalized spectral data for dimensionality reduction. By selecting optimal principal components, spectral dimensionality is reduced to improve model generalization. PCA, a widely used technique, mitigates issues like the "Hughes phenomenon" by linearly transforming data into a coordinate system maximizing variance. Key steps include forming the data matrix, computing the covariance matrix, solving for eigenvalues/eigenvectors, sorting them by eigenvalue magnitude, and projecting the original data onto the top eigenvectors to retain maximal information in lower dimensions.

2.1.3 Image feature extraction

In computer vision and image processing, texture represents the spatial distribution of pixel grayscale/color and their local variations, with regional/object textures characterized by pixel arrangement, grayscale changes, or color differences; extracting such features aids computational image understanding for tasks like classification and object recognition. This study uses Local-GLCM to extract texture features with maximal explainable variance from PCA-dimensionalized hyperspectral data, where GLCM quantifies texture via pixel intensity spatial relationships (directions/distances) to derive features (contrast, homogeneity, energy, correlation) reflecting rock surface properties. Unlike global-texture-focused traditional GLCM, Local-GLCM emphasizes local spectral-texture correspondences under unified labels for detailed lithological identification, constructing local co-occurrence matrices via a sliding window to calculate pixel-neighbor relationships and iteratively extract central-pixel features (contrast, energy, homogeneity, correlation) across the image. The study extracts four normalized (0–1) local texture features: contrast (C_1, measuring local intensity variation/edge strength), energy (E, indicating texture smoothness/uniformity), homogeneity (Hom, reflecting local intensity uniformity, inversely related to contrast), and correlation (C_2, assessing pixel-neighbor intensity linear dependence/trend consistency), enhancing model learning by encoding surface roughness, granularity, and distribution patterns through statistical spatial analysis. The formulas for these four texture features are as follows:

$$C_1 = \sum_{i,j} (i-j)^2 P(i,j), \tag{2}$$

$$E = \sum_{i,j} P(i,j)^2, \tag{3}$$

$$\mathrm{Hom} = \sum\nolimits_{i,j} \frac{P(i,j)}{1 + (i-j)^2}, \tag{4}$$

$$C_2 = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)P(i,j)}{\sigma_i \sigma_j}, \tag{5}$$

where $P(i,j)$ is the probability of pixel values i and j appearing together in a specified direction, μ_i and μ_j are the average values of rows and columns, and σ_i and σ_j are the standard deviations of rows and columns, respectively.

2.2 Fusion identification model with integrated spectral-spatial attention guidance

This study is based on ResNet and introduces a spectral-spatial attention mechanism to construct a dual-channel residual neural network that includes spectral feature channels and texture feature channels, achieving feature-level fusion of spectral data.

2.2.1 Enhancing key spatial-spectral features

In dual-channel residual networks that process multi-source data, not all features are equally important for task completion. If key features aren't effectively extracted before fusion, the performance may not surpass that of single-source models. So, spectral-spatial attention

mechanisms are added. These mechanisms prioritize important features, which improve processing efficiency, identification accuracy, and generalization by enhancing feature extraction.

A customized spatial transformer networks (STN) is added to the texture branch. Its goal is to focus on important spatial regions in 3D data. It has 3D convolutional and pooling layers for depth information. There is a specially designed fully connected layer for 3D affine transformation parameters. There is optimized initialization to find critical 3D regions. The improved STN (Fig. 2(a)) has two main parts. One is the localization network which uses convolutional, pooling, and fully connected layers to extract spatial features. Then it predicts transformation parameters like rotation, scaling, and shearing. This allows the network to focus on regions of interest by warping the image. The other part, the grid generator & sampler, turns the predicted parameters into transformation matrices. This dynamically adjusts the spatial layout and emphasizes features important for lithological identification. These changes create a trainable spatial transformation module. This module lets neural networks adaptively focus on local regions in 3D texture feature maps.

To address the importance of the spectral dimension in lithological tasks, the spectral attention for feature enhancement (SAFE) mechanism is proposed. It is tailored to the batch size × $H \times W \times D \times C$ ($C = 1$) input format, where H represents the height of the feature map, W represents the width of the feature map, D represents the depth of the feature map (spectral dimension), and C represents the channel of the feature map. As shown in Fig. 2(b), SAFE has several components. First, global average pooling generates a 1D vector of spectral channel global contexts. Then, a fully connected layer learns channel importance weights. Next, Softmax normalization converts these weights into a probability distribution. Finally,

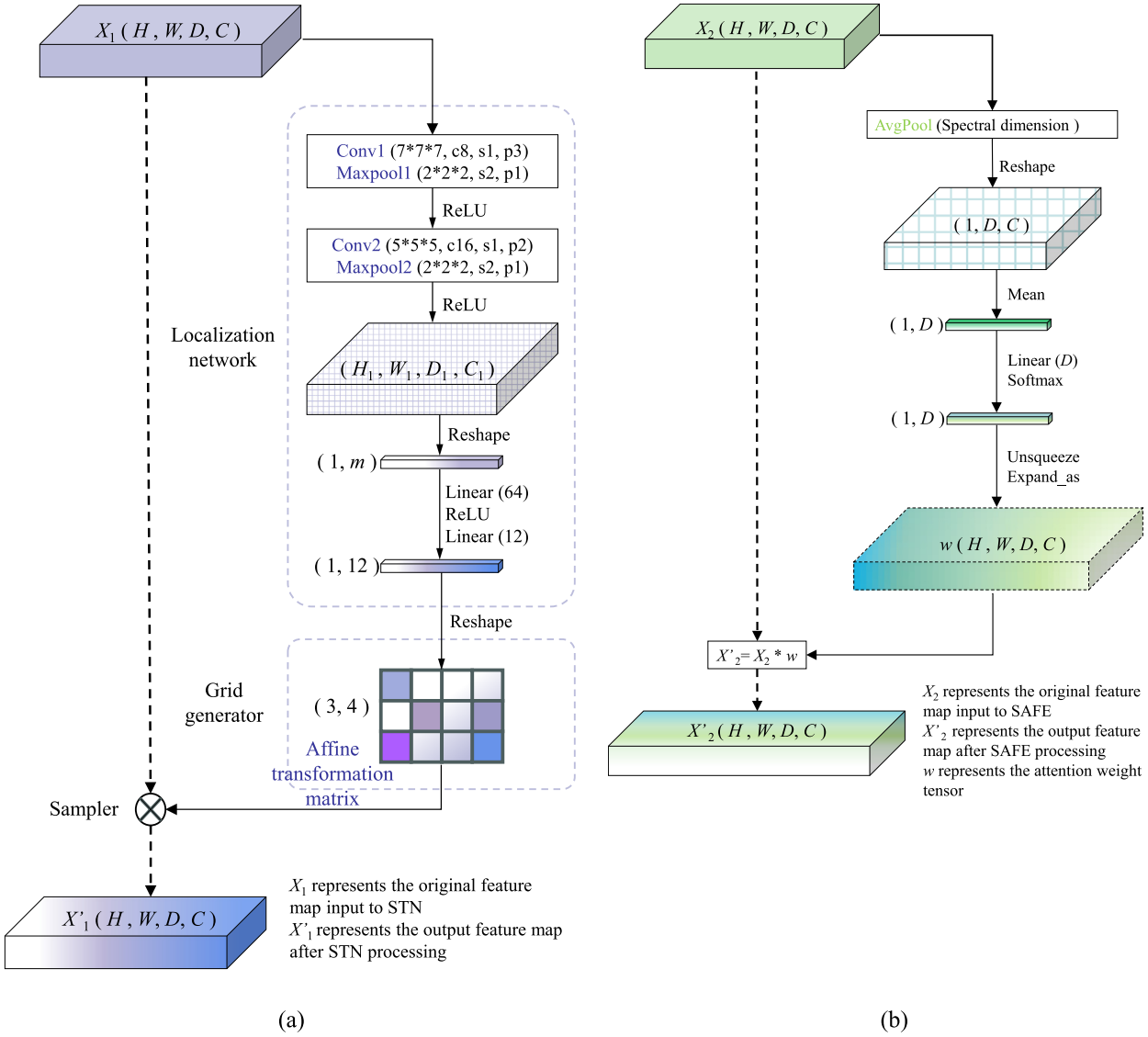


Fig. 2. Spatial-spectral attention mechanism. (a) STN spatial attention mechanism, and (b) SAFE spectral attention mechanism.

weight application emphasizes critical spectral channels and suppresses irrelevant ones. This dynamic adjustment during training enhances the model's ability to capture task-relevant spectral information and improves the representation of hyperspectral data.

2.2.2 Spectral-image feature fusion framework

In order to meet the requirements of feature-level fusion and enable the model to better integrate the features of multi-source data, this study modifies the traditional ResNet architecture, proposing a dual-channel ResNet, as illustrated in Fig. 3.

This network structure provides independent channels for each data source, allowing each channel to focus on extracting and transforming the features corresponding to its respective data source. This ensures that the diversity and richness of the data are fully utilized, while simultaneously enabling feature-level fusion of different sources within the model. Within the network structure, each data source passes through a residual neural network composed of convolutional layers, pooling layers, or other types of layers for feature extraction. Then, in the later stages of the model (when the convolutional layers transition into fully connected layers), features extracted from different source channels are integrated. Before reaching the feature-level fusion stage, each channel undergoes independent, in-depth processing and feature extraction, ensuring that richer and more complex feature representations are captured from each channel, thereby improving the overall model accuracy. Moreover, this model architecture enhances the flexibility and efficiency of transfer learning. By training with independent channels for different data sources, the model learns generalized spectral and texture features, which remain valuable for different tasks (even when new rock types not present in the original training set are introduced). In transfer learning applications, only the weights and biases of the fully connected layer after feature fusion need to be retrained. This method preserves the previously learned generalized features while improving the model's ability to adapt to new tasks, achieving a balance between training efficiency and model performance.

2.2.3 Transfer learning from laboratory to engineering applications

Transfer learning, a machine learning paradigm, enables deep learning models to utilize knowledge gained from a source task to expedite and optimize learning for a target task—particularly when labeled data is scarce. In this study, a tailored transfer learning strategy is employed for lithology identification, diverging from conventional approaches that rely on models pre-trained on large-scale generic datasets (e.g., ImageNet). The methodology proceeds in two stages: first, the model is pre-trained using a substantial dataset of lithology samples collected in controlled laboratory conditions which, while not matching the scale of publicly available benchmarks like ImageNet, is inherently more relevant to tunnel application scenarios and encapsulates detailed spectral and textural features of rocks specific to such environments; second, transfer learning is applied using in-situ hyperspectral data with the model leveraging shared parameter information between the source domain (laboratory hyperspectral data) and the target domain (field-collected hyperspectral data), where fine-tuned weights and parameters are integrated into a ResNet-based deep learning architecture to accelerate training, augment the hyperspectral rock imaging dataset, and enhance the model's generalization capability.

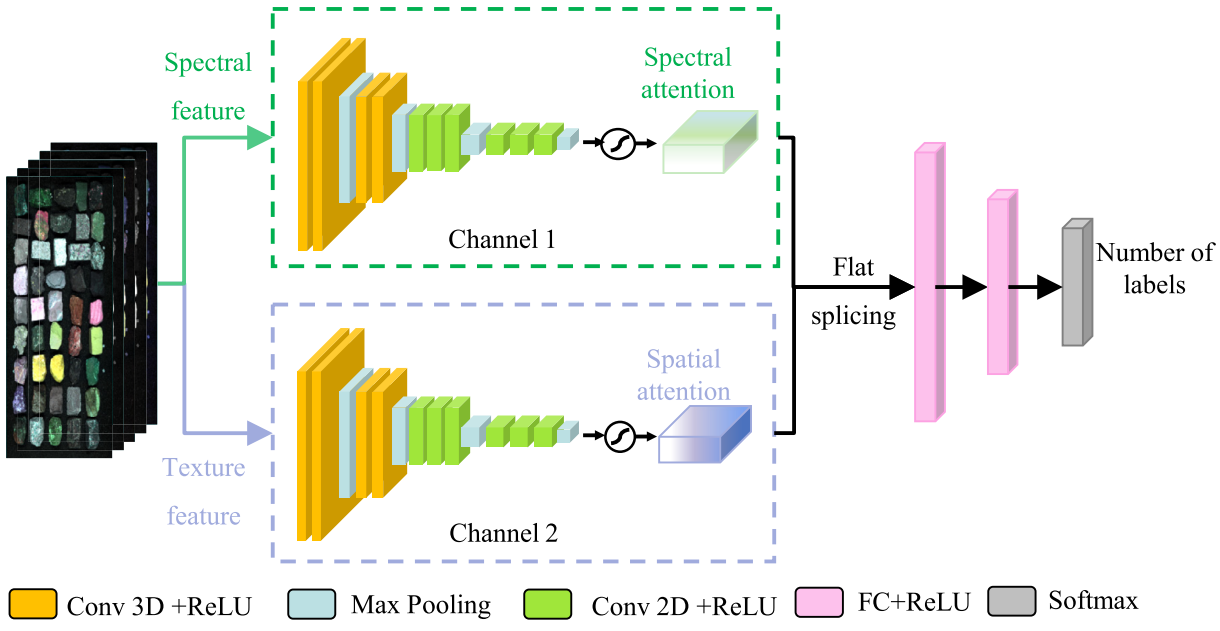


Fig. 3. Dual-channel residual neural network structure.

The obstacles encountered by the lithological identification transfer model predicated on graph information fusion in acclimating to tunnel environments include: (1) Training data for the model lacks variety and representativeness. This limits its ability to adapt to complex tunnel environments. Tunnel areas have different rock types compared to other regions, and rock data in tunnels is unevenly distributed. High data costs and limited samples cause overfitting. Solutions include dropout techniques and transfer learning to expand datasets, speed up training, and improve adaptability. (2) Using only one data type (like spectra or textures) faces challenges. Moisture, dust, and poor lighting in tunnels create noise and blurry features. Spectra may fail if rocks have similar chemical reflectance. Textures may mislead if surfaces look alike. The model fixes this by merging spectral (chemistry) and texture (physical structure) data for better accuracy. (3) Combining different data types causes issues like high dimensionality and mismatched scales. The model uses a dual-channel network: one channel processes spectra, another handles textures. Features are merged mid-network, and an attention mechanism highlights key areas. This improves rock recognition in complex tunnel settings where single data types struggle.

2.2.4 Reliability assessment

After constructing the lithology intelligent identification model, it is crucial to assess its identification performance. To objectively evaluate the model's effectiveness, selecting appropriate evaluation metrics is necessary to compare the performance of different models. In this study, commonly used metrics in machine learning and statistical classification problems are employed to assess the model's performance. The metrics include accuracy, precision (P), recall (R), and F_1-score to measure the overall accuracy of the model. A confusion matrix is also utilized to reflect the classification precision of each type of rock spectral data.

To evaluate both the overall accuracy and the accuracy for different rock types, the following metrics are used. Accuracy represents the proportion of correctly classified rock types over the total in the test set:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{6}$$

where TP (true positive) represents the number of correctly predicted rocks of a certain type, FP (false positive) represents the number of correctly predicted rocks of other types, FN (false negative) represents the number of incorrectly predicted rocks of a certain type, and TN (true negative) represents the number of correctly predicted rocks of other types.

The precision measures the model's ability to distinguish negative samples. The higher the precision, the more accurate the model's classification of rocks in that category and the stronger its ability to distinguish negative samples. For a certain rock category φ, P_φ represents the proportion of correctly predicted rock samples by the model to all predicted rocks of that type. The precision is weighted and averaged to obtain the overall accuracy P:

$$P_\varphi = \frac{\text{TP}_\varphi}{\text{TP}_\varphi + \text{FP}_\varphi}, \tag{7}$$

$$P = \frac{\sum_{\varphi=1}^{M} P_\varphi \times w_\varphi}{\sum w_\varphi}, \tag{8}$$

where φ represents the index of a certain type of rock, $\varphi \in [1, M]$, TP_φ represents the number of rock samples correctly predicted by the model as this type of rock, and FP_φ represents the number of other types of rock samples incorrectly predicted as this type of rock. M represents the total number of rock types, and w_φ represents the weight of the number of hyperspectral images of a certain type as a whole.

The recall measures the model's ability to discriminate between positive samples; the higher the recall, the better the model's ability to discriminate between positive samples. For a certain rock category φ, the recall R_φ indicates the proportion of the number of a certain rock sample correctly predicted by the model to the total number of rock samples of that category, and the recall is weighted and averaged to obtain the overall recall R:

$$R_\varphi = \frac{\text{TP}_\varphi}{\text{TP}_\varphi + \text{FN}_\varphi}, \tag{9}$$

$$R = \frac{\sum_{\varphi=1}^{M} R_\varphi \times w_\varphi}{\sum w_\varphi}, \tag{10}$$

where FN_φ denotes the number of rocks in a certain rock sample that are incorrectly predicted to be other types of rocks.

F_1-score is the weighted average of the precision, recall, and both precision and recall are considered. The range of values is [0, 1], with 1 representing the best output of the model and 0 representing the worst output result of the model.

$$F_1\text{-score} = \frac{2P \times R}{P + R}. \tag{11}$$

The confusion matrix reflects the accuracy of the classification of each type of rock spectral data and allows visualization of the model's prediction of the classification of different types of rock spectral data, as shown in Table 1. Each column of the confusion matrix represents the categories predicted by the model, while each row represents the actual categories. Therefore, by obtaining the following

Table 1
Confusion matrix and its basic indicators.

Confusion matrix		True label	
		Positive	Negative
Predict label	Positive	TP	FP
	Negative	TN	FN

four basic metrics, it is possible to clearly understand the categories in which the model is confusing.

3 Laboratory experiment and validation

3.1 Establishment of the hyperspectral image database

3.1.1 Sample sources

In this study, spectral data from 45 rock types were selected to construct a project-site rock imaging spectral database for validating transfer learning effectiveness. The dataset includes 15 igneous rocks (e.g., peridotite, basalt, granite), 15 sedimentary rocks (e.g., conglomerate, sandstone, limestone), and 15 metamorphic rocks (e.g., quartzite, marble, schist). Additionally, 50 distinct laboratory rock samples (ultrabasic, basic, neutral, acidic types, unrelated to the project-site dataset) were collected for pre-training.

Pre-training weight selection prioritizes target-task similarity: while transfer learning is widely used, substantial feature/data disparities between pre-trained and tunnel datasets can hinder parameter utilization or degrade model performance. Thus, we pre-trained using laboratory-collected hyperspectral lithology data, which, though smaller than ImageNet, better aligns with tunnel applications and contains specific spectral/textural features. Training utilized geologically similar data to the tunnel site, with consistent source/target domain formats, enhancing feature learning efficiency. These data's unique characteristics make them ideal for pre-training, enabling effective weight transfer to tunnel environments.

The dataset of the project site encompasses 45 distinct rock types, consisting of 390 912 data points, among which 200 412 are zeros. Its spatial dimensions are 1016 × 384 × 272 (Height × Width × Number of bands). The detailed distribution of sample sizes for transfer learning on the project-site dataset is illustrated in Fig. 4. The laboratory dataset, on the other hand, comprises 50 rock types. With a total of 476 544 data points, including 148 619 zeros, it is formatted as 1241 × 384 × 272. The tunnel dataset, collected from a tunnel in China, contains 17 types of granite and limestone. It is composed of 281 472 data points, with 190 908 of them being zeros, and has a size specification of 733 × 384 × 272.

3.1.2 Hyperspectral image acquisition and calibration

This study used a laboratory push scanning platform imaging system to collect imaging spectral data, as shown in Fig. 5. The imaging system works in push scan mode to collect hyperspectral data in the short-wave infrared band (1000–2500 nm) region.

Due to the differences in composition and structure of rocks in nature, their electromagnetic wave absorption and reflection characteristics vary. At the same time, the noise of measuring equipment itself, including instrument dark current and environmental electromagnetic interference, can also, to some extent, mask the electromagnetic

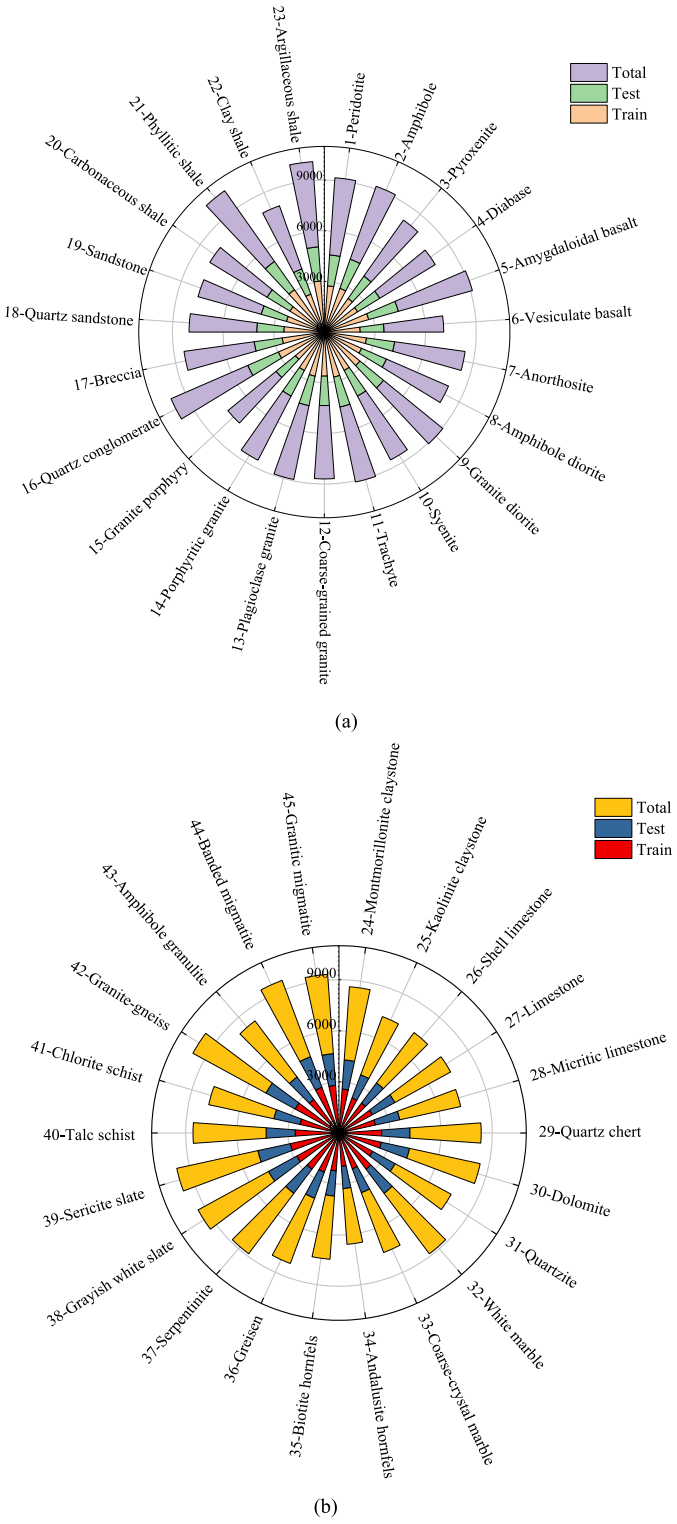


Fig. 4. Sample size for project site rock pre-training. (a) The first 23 rocks, and (b) the last 22 rocks.

characteristics of rocks themselves. Therefore, before collecting imaging spectral data, it is necessary to set the scanning parameters of the measurement system to minimize the influence of noise and improve the accuracy of lithology identification. A total of 45 raw imaging spectral data

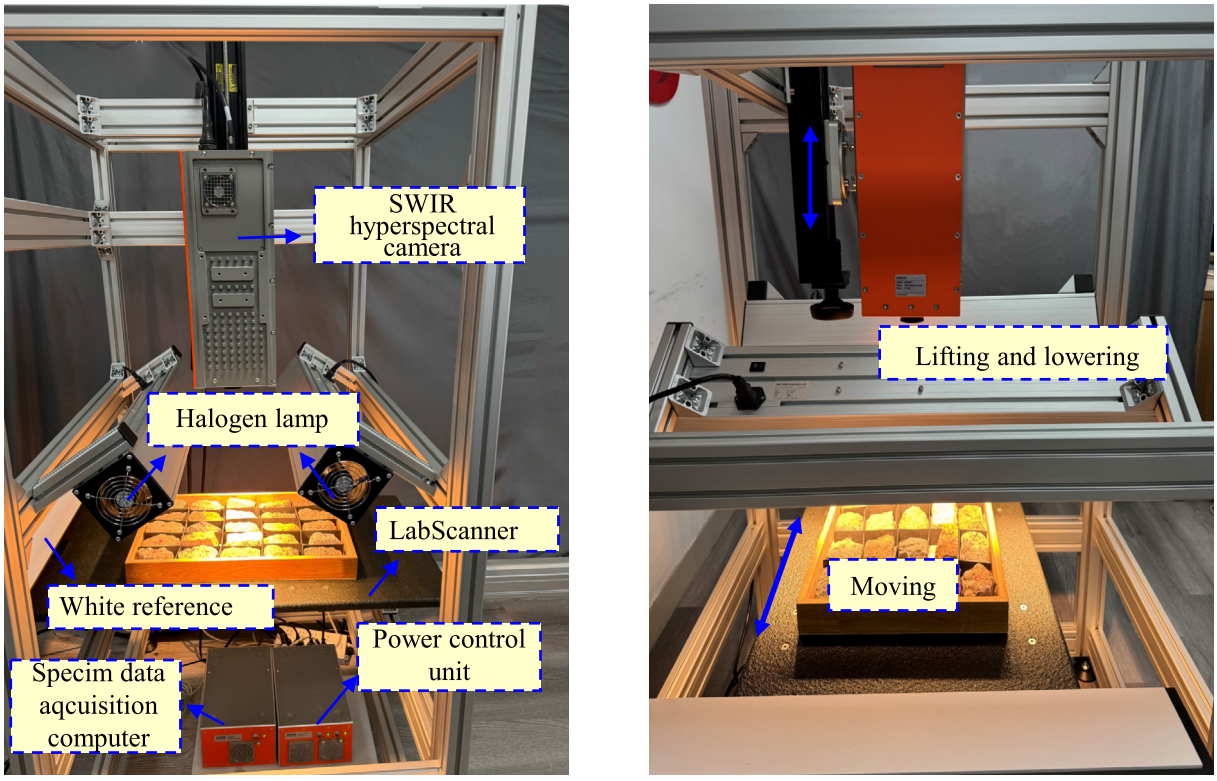


Fig. 5. Hyperspectral imaging data acquisition system. (a) Front, and (b) side. (Li et al., 2025)

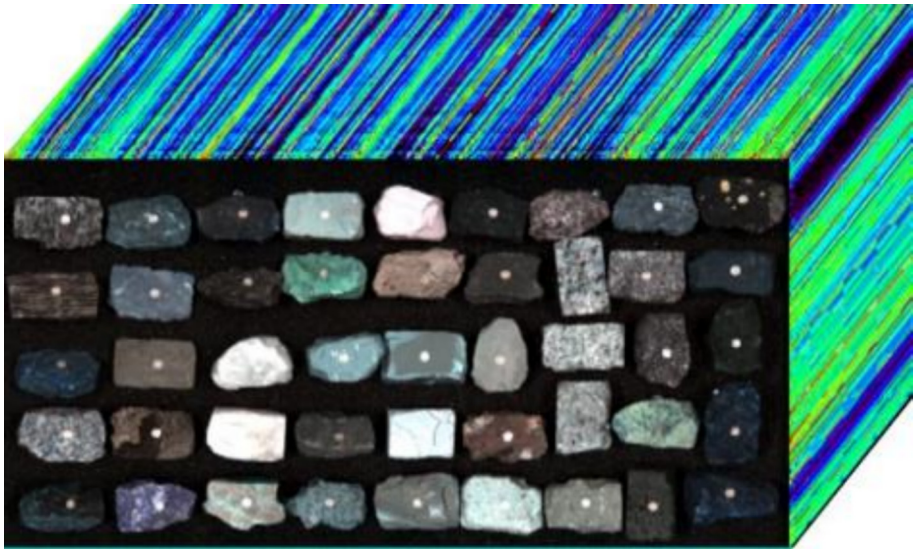


Fig. 6. Three-dimensional data volume.

of rocks in the range of 1000–2500 nm bands were collected, as shown in Fig. 6.

Before processing and analyzing the acquired rock hyperspectral images, a black and white correction of the reflectance of all raw spectral images was required due to dark current and light inhomogeneity in the hyperspectral camera. A white reference image is obtained by acquiring a white calibration plate through the camera, and the image when the lens is completely obscured is acquired as the black reference image. Correction is performed using the acquired black and white images, and the corrected image is calculated by the formula below:

$$I = \frac{I_r - I_d}{I_w - I_d}, \tag{12}$$

where I_r is the raw hyperspectral image, I_d is the dark reflection image, and I_w is the white reflection image. After all images are calibrated, the next step of data processing can be performed.

3.2 Feature extraction

In this study, the extraction of spectral and textural features is based on the established short-wave infrared imaging spectral database of 45 rocks. Firstly, the ROIs were extracted from the original short-wave infrared (SWIR) imaging spectral data, and the spectral data within the ROIs were extracted and reconstructed to obtain a high-quality dataset; then, the spectral data within the ROIs were normalized to further reduce the variations of the spectral data of the rocks due to the lighting conditions, instrumental differences, etc. Subsequently, in the spectral dimensions, the PCA was performed on the normalized rock imaging spectral data along the spectral dimension, selecting 10 principal components, which accounted for 3.68% of the original feature dimensions. Through these steps, the spectral features of the short-wave infrared imaging spectral data of the rocks were obtained.

The hyperspectral imaging system used in this study covers a spectral range of 1000–2500 nm, with a total of 272 bands. Each band can be treated as an individual image. Therefore, PCA was first applied to the raw spectral images for dimensionality reduction, where images corresponding to each band were linearly combined to form principal component images, eliminating redundant information in the hyperspectral images. The first few principal component images contained the majority of the information from the original spectral images. After principal component analysis, the first three principal component images obtained are shown in Fig. 7. Among them, the PCA1 image has 96.69% of interpretable variance, which retains most of the real information of the rock samples, while the PCA2 and PCA3 images have only 2.02% and 0.73% of interpretable variance, respectively, which basically do not retain the information of the rock samples. Therefore, the PCA1 image is used as the principal component image for extracting the texture feature information, and the local grey scale covariance matrix is used to extract the local texture feature information of the imaging spectral data. Specifically, a local window size of 25 × 25 pixels was utilized to extract four texture features: contrast, energy, homogeneity, and correlation, all along the 0° direction. These features were then normalized to the range [0, 1], facilitating more effective learning and identification of texture features by the model.

3.3 Model training

In the field of deep learning-based rock hyperspectral image identification, different machine learning algorithms are usually selected according to different tasks. This article uses ResNet's dual-channel residual neural network as a feature fusion framework and trains it under the PyTorch framework. In lithology identification tasks, ResNet-50/101 networks are widely used. In order to better evaluate the deep transfer learning model of rock hyperspectral

(a) (b) (c)

Fig. 7. First three principal component images. (a) PCA1, (b) PCA2, and (c) PCA3.

images, ResNetX2-50, ResNetX2-101, ResNet50-Spectral, ResNet101-Spectral, ResNet50-Image, and ResNet101-Image were trained. The specific description of the model is shown in Table 2.

Firstly, conduct model pre-training to enable the model to learn the general features of rock hyperspectral image data. By conducting multiple experiments on the model, the model was set to select epoch as 40 and batch size as 256, and the input size of spectral data used for pre-training was $10 \times 25 \times 25$, and the input size of texture data was $4 \times 25 \times 25$. Then, load the pre-trained model for transfer learning, compare the different models that require experiments to be conducted under the same hardware conditions, and adjust the model parameters to be the same. After multiple experiments, the transfer learning model was set with an epoch of 30 and a batch size of 64.

Table 2
Six model descriptions.

Model	Description
ResNetX2-50	Data: Spectral data and texture data Model: ResNet50 and ResNet50 fusion dual branch network
ResNetX2-101	Data: Spectral data and texture data Model: ResNet101 and ResNet101 fusion dual branch network
ResNet50-Spectral	Data: Spectral data Model: ResNet50
ResNet101-Spectral	Data: Spectral data Model: ResNet101
ResNet50-Image	Data: Texture data Model: ResNet50
ResNet101-Image	Data: Texture data Model: ResNet101

The accuracy curves of several classification models on the training and testing sets during the transfer learning phase are shown in Fig. 8. After 10 iterations, the accuracy of the models tends to stabilize, indicating that they have basically acquired the ability to extract rock features. On the training set, the model based on ResNetX2-50 has an accuracy rate of up to 98.98%, the ResNetX2-101 model has an accuracy rate of up to 97.71%, and the ResNet101-Spectral model has an accuracy rate of up to 94.95%. On the test set, the model based on ResNetX2-50 achieved the highest accuracy of 99.78%, the model based on ResNetX2-101 achieved the highest accuracy of 98.32%, the ResNet101-Spectral model achieved the highest accuracy of 97.41%, and the ResNet101-Image model achieved the highest accuracy of 79.21%. The convergence loss curves of different models in the training dataset are shown in Fig. 9.

In addition, the values of TP, TN, FP, and FN were calculated to provide a more comprehensive evaluation of the model performance, as shown in Fig. 10. A high TP value indicates that the model performs well in discovering positive samples and can effectively filter out true positive samples from all samples. A high TN value means that the model has high accuracy in identifying negative class samples, reducing the occurrence of misclassifying negative classes. The model based on ResNetX2-50 has the best classification performance for rock samples.

In summary, the ResNetX2-50, ResNetX2-101, and ResNet101-Spectral models have better performance, with the ResNetX2-50 based spectral-image fusion model achieving the best classification results in the transfer learning process.

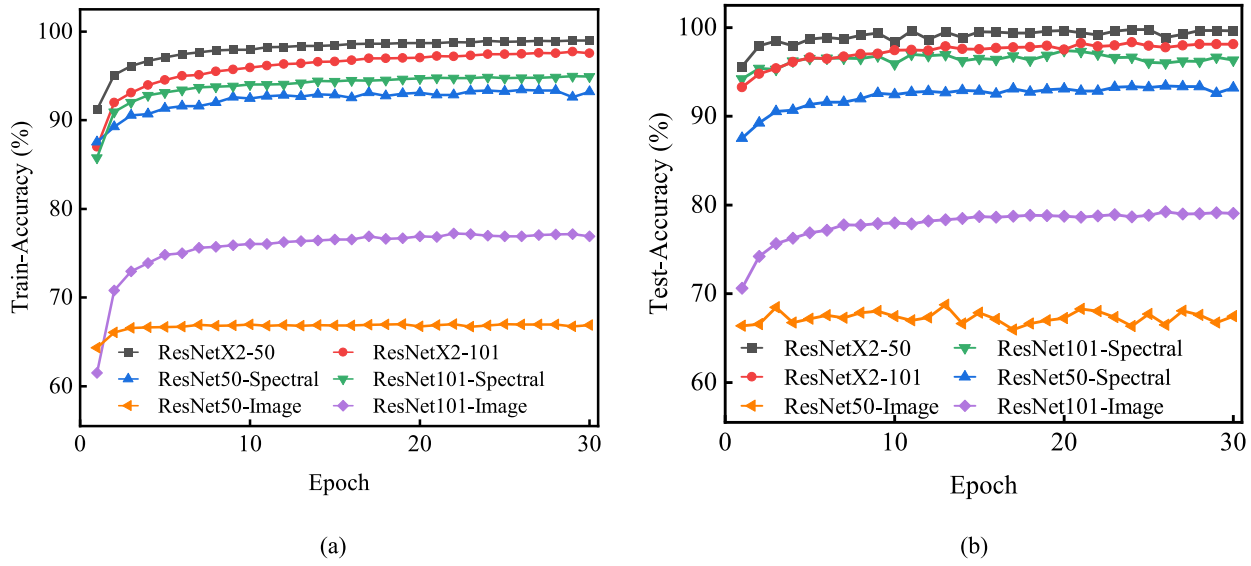


Fig. 8. Accuracy curves of training and testing sets during the transfer learning phase. (a) Training set, and (b) testing set.

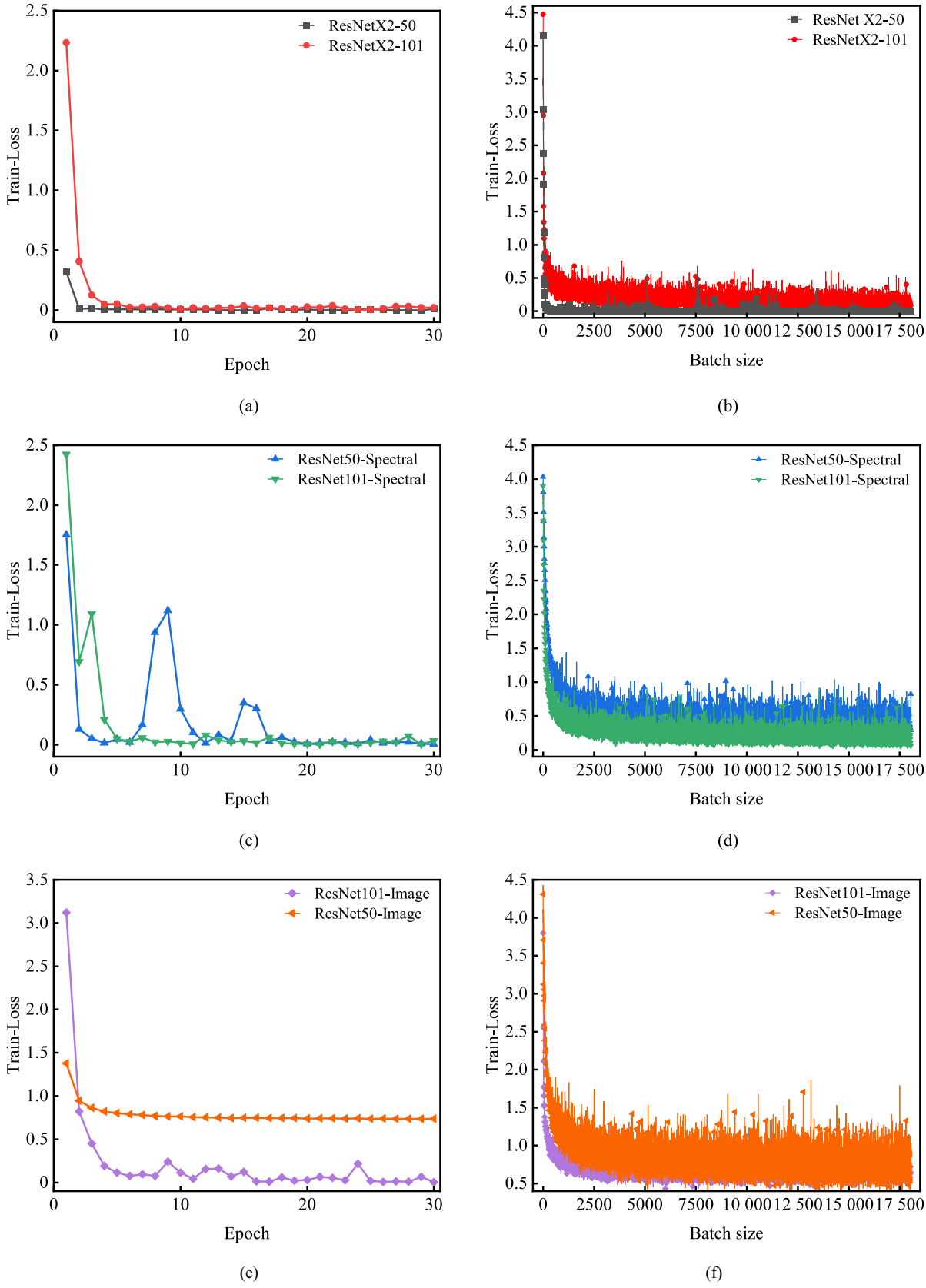


Fig. 9. Convergence loss curves in the training dataset. (a) ResNetX2-50 and ResNetX2-101 models at different epochs, (b) ResNetX2-50 and ResNetX2-101 models at different batch sizes, (c) ResNet50-Spectral and ResNet101-Spectral models at different epochs, (d) ResNet50-Spectral and ResNet101-Spectral models at different batch sizes, (e) ResNet101-Image and ResNet50-Image models at different epochs, and (f) ResNet101-Image and ResNet50-Image models at different batch sizes.

3.4 Identification results

3.4.1 Model validation

To assess the spectral-image fusion model's effectiveness in lithology identification, evaluation metrics P, R, F_1-score, and a confusion matrix were employed. The P of 45 types of rocks is shown in Fig. 11(a)–(c), in the ResNetX2-50-based model, 44 out of 45 rock types had a precision above 98%, with 29 reaching 100%, and the lowest (33-coarse-crystalline marble) at 96.93%. In the

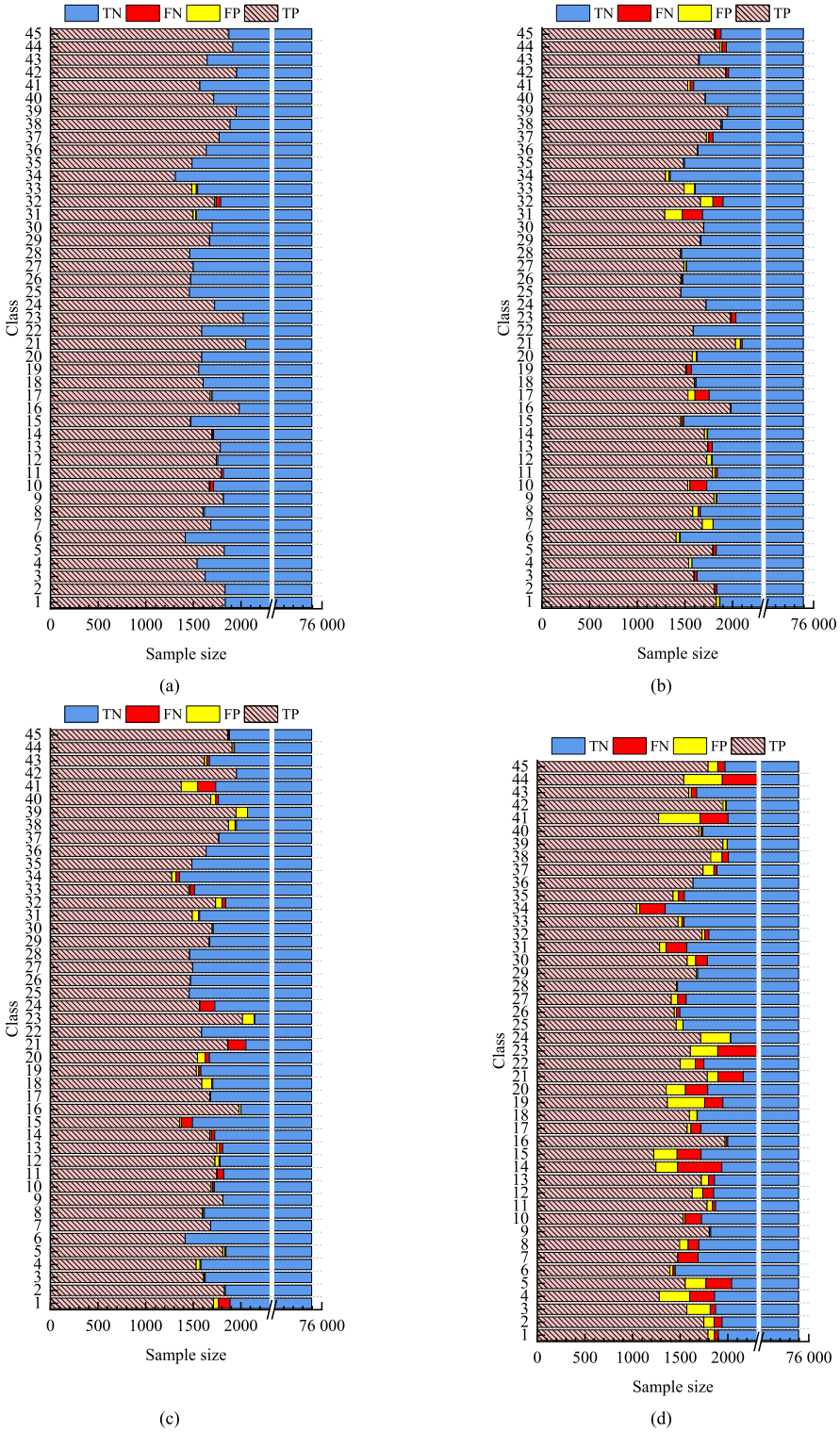


Fig. 10. Performance evaluation of test set. (a) ResNetX2-50, (b) ResNetX2-101, (c) ResNet101-Spectral, (d) ResNet50-Spectral, (e) ResNet101-Image, and (f) ResNet50-Image.

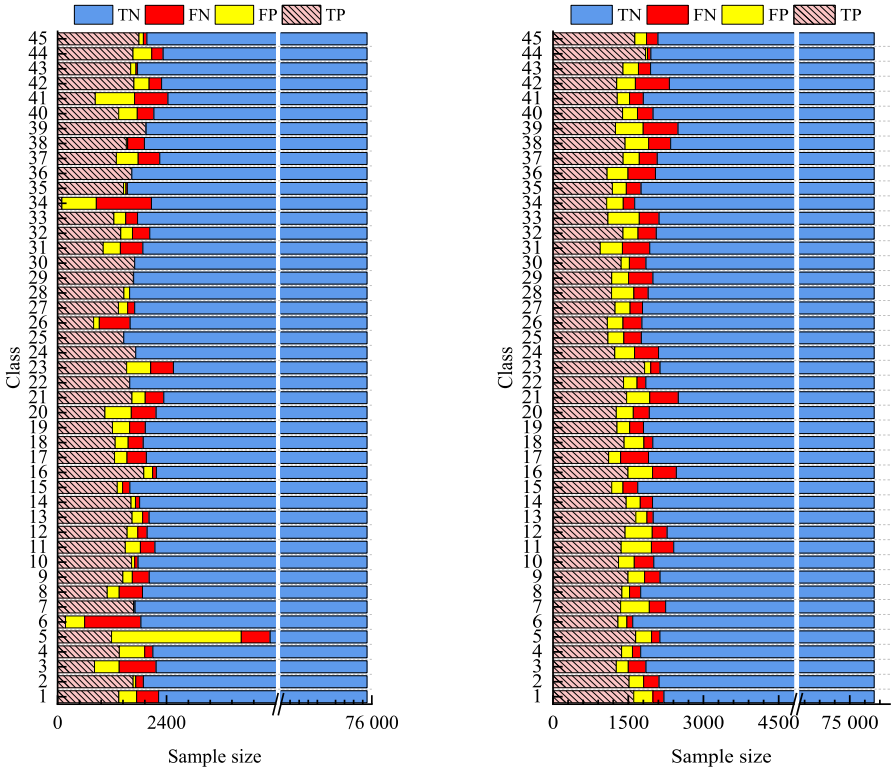


Fig 10. *(continued)*

ResNetX2-101 model, 44 types had precision over 90% except 31-quartzite (87.51%), with 29-quartz chert and 39-sericite slate at 100%. The ResNetX2-50 model showed higher overall precision and better differentiation for positive samples.

As shown in Fig. 11(d)–(f), the R for 45 types of rocks was also analyzed. The ResNetX2-50 model had all 45 rock types with recall above 97%, 28 at 100%, and the lowest (32-white marble) at 97.34%. In the ResNetX2-101 model, 42 types had recall over 90%, except 10-syenite (89.68%) and 31-quartzite (85.88%), with 6 types at 100%. The ResNetX2-50 model had a higher overall recall.

For the F_1-score (Fig. 11(g)–(i)), 21 rock types in the ResNetX2-50 model achieved 100%, compared to only one in the ResNetX2-101 model. Overall, the ResNetX2-50 model excelled, with P, R, and F_1-score for all 45 rock types above 95%, indicating strong sample differentiation and stable identification, even with imbalanced samples. This shows that increasing network depth (ResNetX2-101 has 101 layers per channel vs. 50 in ResNetX2-50) does not guarantee better performance, as it may cause overfitting.

The confusion matrix comparison revealed that the ResNetX2-50 model achieved over 95% accuracy for all 45 rock types, with 29 at 100%. Some misclassifications occurred, e.g., 10-syenite had small probabilities of being misclassified as other types, as shown in Fig. 12(a). In Fig. 12(b), the ResNetX2-101 model had most types above 90% accuracy, except 10-syenite and 31-quartzite, with notable misclassification probabilities for these two types.

The overall P, R, and F_1-score of the model were further used to compare the spectral-image fusion feature classification models constructed based on ResNetX2-50 and ResNetX2-101. As can be seen from Table 3, overall, the evaluation metrics for the feature fusion classification model based on ResNetX2-50 were higher than those for the ResNetX2-101 model. Specifically, the ResNetX2-50 model achieved a P of 99.78%, which was 1.44 percentage points higher than the ResNetX2-101 model, indicating stronger differentiation ability for negative samples. The R of the ResNetX2-50 model was 99.78%, which was 1.46 percentage points higher than the ResNetX2-101 model, indicating stronger differentiation ability for positive samples. The combined P and R metrics produced an F_1-score of 99.78% for the ResNetX2-50 model, 1.45 percentage points higher than the ResNetX2-101 model.

In conclusion, the feature fusion classification model based on ResNetX2-50 achieved superior identification performance, with classification accuracy for most rocks reaching 100%. The P, R, and F_1-score were all excellent,

demonstrating strong generalization capability. Therefore, by constructing an intelligent lithology identification model based on feature fusion and spectral-spatial attention mechanisms, using the ResNetX2-50 network yielded better identification results, enabling the extraction of fine details from spectral and texture data for more stable rock classification, with outstanding generalization and broad applicability.

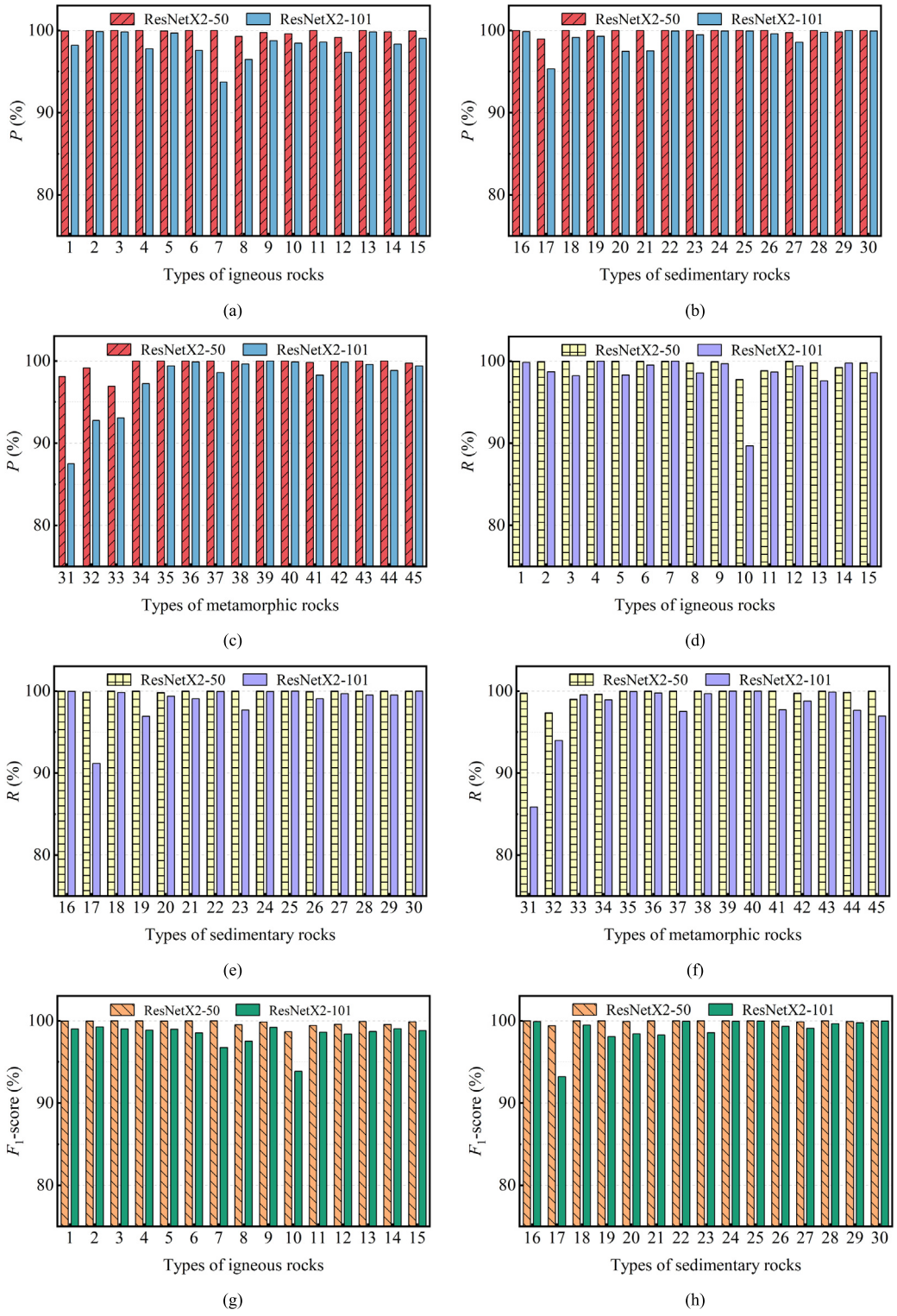
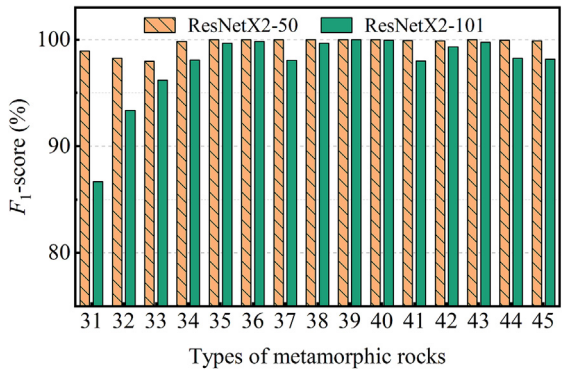


Fig. 11. Model accuracy of P, R, and F_1-score chart.



(i)

Fig 11. *(continued)*

3.4.2 Comparison of prediction results

This study employs transfer learning fine-tuned models to predict 45 rock types, comparing lithology identification performance between fused and individual features. Texture data processing and model training parameters mirrored those for spectral data. For pixel-wise predictions, the transfer-learned model was applied directly without dataset splitting; the entire dataset underwent classification. Table 4 presents evaluation metrics: the ResNetX2-50 fused feature model achieved the highest scores (all exceeding 99%), while the texture-only model scored lowest (approximately 80%). Spectral-only models used PCA to reduce normalized data to 10 principal components, generating 10 channels. Texture-only models extracted four features (contrast, energy, homogeneity, correlation) via a 25 × 25 local gray-level co-occurrence matrix in the 0° direction, producing four channels. Additional layers improved single-feature accuracy by enhancing feature exploration without excessive complexity, especially with abundant representative data. For fused features, however, excessive layers induced overfitting: combined spectral-texture information made extra layers prone to memorizing training noise, degrading generalization. Complex interactions in fused data caused overinterpretation in deep models, losing the simplicity that aids in capturing feature interactions.

Figure 13 shows pixel-wise results for three models: ResNet-101 (spectral), ResNet-101 (texture), and ResNetX2-50 (fused). In Fig. 13(a), the results for the ResNet-101 spectral feature classification model indicate that although the identification performance was excellent for most rock types, certain rocks, such as 15-granite porphyry, 21-phyllitic shale, 24-montmorillonite claystone, and 41-chlorite schist, exhibited identification accuracies of around 90%. Taking 21-phyllitic shale and 41-chlorite schist as examples, the comparison between their rock sample photographs and predicted results is shown in Fig. 14. These rocks might have undergone weathering or metamorphism due to prolonged exposure to external geological conditions, which altered their chemical composition and optical characteristics, thereby affecting the spectral feature classification model's predictive performance.

Figure 13(b) depicts the prediction results of the ResNet-101 texture feature classification model using local gray-level co-occurrence matrices, which quantify texture by analyzing spatial relationships between pixel gray levels and their neighbors. Despite significant spectral differences among the 45 rock types in the hyperspectral database, texture features—independent of chemical composition and optical properties—may exhibit similar patterns. Consequently, the model's predictions showed substantial noise for each rock type, resulting in low overall accuracy. Texture-based classification, akin to traditional image classification, relies heavily on dataset quality/quantity; however, this study extracted only four texture types (contrast, energy, homogeneity, correlation in the 0° direction), which may have contributed to poor performance.

Figure 13(c) shows the ResNetX2-50 fused feature classification model achieving an accuracy of over 99% accuracy for all 45 rock types, with stable predictions and minimal noise. This confirms that integrating spectral and texture features in ResNetX2-50 enhances classification accuracy and stability by comprehensively capturing rock characteristics, overcoming single-feature limitations, and providing robust data for geological research and exploration.

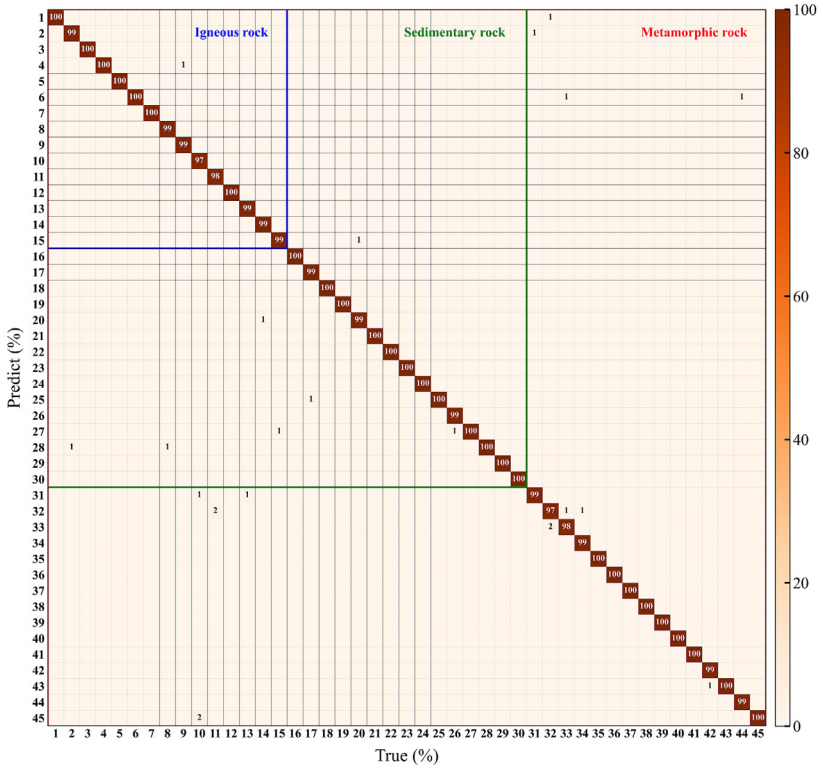
In conclusion, the fused feature classification model based on ResNetX2-50 outperformed single-feature classification models based on ResNet-101 across multiple machine learning evaluation metrics, and it also performed better in identifying weathered and metamorphosed regions of rocks.

4 Engineering application and validation

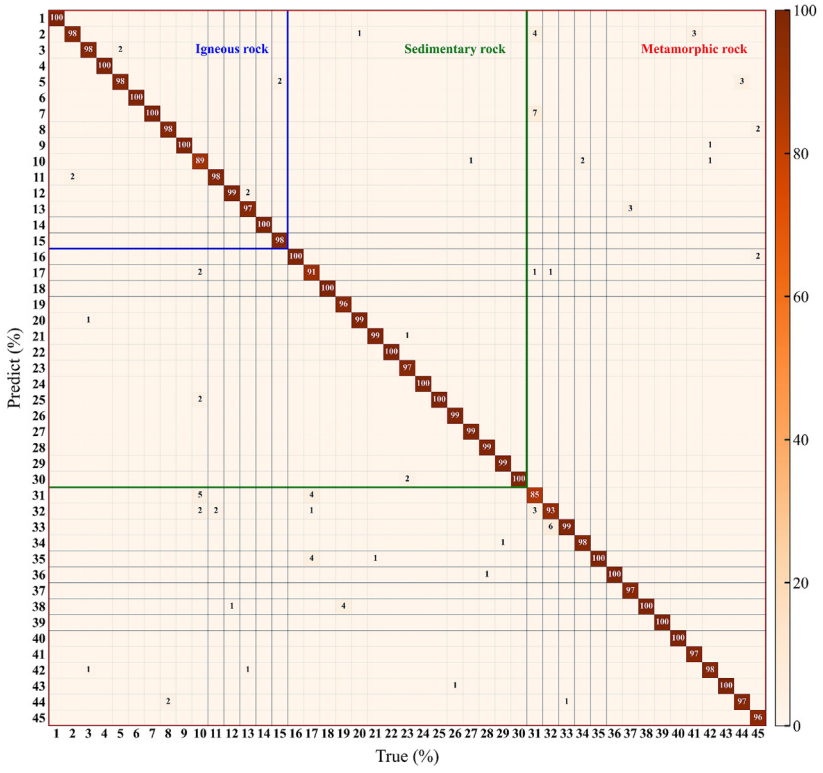
4.1 Hyperspectral imaging database for tunnel site

For application and validation, a tunnel project using full-face excavation was chosen. The geological sketch of the studied tunnel section showed that the surrounding rock was mainly Grade III, predominantly granite with some limestone. The rocks were generally grayish-white, with yellowish-brown rust stains on some surfaces, exhibiting a blocky texture, low weathering, and high rock mass integrity. Seventeen rock samples (granite and limestone) were collected from multiple tunnel sections.

In-situ hyperspectral image acquisition of the tunnel was carried out (Fig. 15). The hyperspectral imaging system setup in the tunnel included: (1) placing the tripod on flat ground and mounting the swing scanning gimbal and hyperspectral camera; (2) adjusting the tripod height and level to capture the study area's surrounding rock; (3) powering the swing scanning platform, camera, and halogen



(a)



(b)

Fig. 12. Model confusion matrix. (a) ResNetX2-50, and (b) ResNetX2-101.

Table 3
Comparison of overall model evaluation indicators.

Model	P (%)	R (%)	F_1-score (%)
ResNetX2-50	99.78	99.78	99.78
ResNetX2-101	98.34	98.32	98.33

Table 4
Comparison of model prediction evaluation indicators based on experimental analysis.

Type	Model	P (%)	R (%)	F_1-score (%)
Spectral feature	ResNet-50	93.23	93.10	93.16
	ResNet-101	98.33	98.30	98.31
Texture feature	ResNet-50	79.92	79.81	79.86
	ResNet-101	82.40	81.73	82.06
Feature fusion	ResNetX2-101	98.44	98.32	98.38
	ResNetX2-50	**99.83**	**99.83**	**99.83**

light source with the tunnel power supply; (4) positioning the light source brackets behind the camera tripod (irradiance decreases with distance and field of view); (5) aiming the light source at the rock in front of the camera for even lighting, keeping it as close as possible to the rock surface and concentrated within the geometric constraints.

In the laboratory, a push-broom SWIR hyperspectral imaging system was used to acquire data for the rock samples, establishing a hyperspectral imaging database with data and raw spectral curves of granite and limestone. The raw spectral curves, averaged from all pixels in the ROI, covered the SWIR bands with 272 bands.

4.2 Feature extraction from field data

After creating the tunnel rock hyperspectral imaging database, raw data were preprocessed to extract rock spectral and texture features. Spectrally, normalized data underwent PCA for dimensionality reduction, setting 10 principal components. In the image domain, the PCA1 image was selected for texture extraction. Local-GLCM was applied to extract local texture features from hyperspectral images. Four texture features (contrast, energy, homogeneity, and correlation) were extracted in the 0° direction for each Local-GLCM.

4.3 Transfer learning and model training

The pre-trained model used in laboratory analysis was first loaded, and transfer learning was performed on the hyperspectral features of the rock samples collected from the project site. The model was trained with an epoch setting of 30 and a batch size of 64. The input size for the spectral data during transfer learning was $10 \times 25 \times 25$, while the input size for the texture data was $4 \times 25 \times 25$. The accuracy curves of several models on the training and testing sets during transfer learning are shown in Fig. 16. On the training set, the ResNetX2-50 model achieved a maximum accuracy of 99.93%. On the testing set, the ResNetX2-50 model achieved a maximum accuracy of 99.997%. These results indicate that during the transfer learning phase, ResNetX2-50 achieved better classification results overall.

4.4 Identification results

To further compare the performance of the fused spectral-texture classification models built using ResNetX2-50 and ResNetX2-101, P, R, and F_1-score were evaluated for both models. As shown in Table 5, overall, the fused feature classification model based on ResNetX2-50 exhibited higher evaluation metrics than the model based on ResNetX2-101. Specifically, the ResNetX2-50 model achieved a P of 99.99%, which was an improvement of 0.6 percentage points compared to the ResNetX2-101 model. The R of the ResNetX2-50 model was also 99.99%, representing a 0.66 percentage points improvement over the ResNetX2-101 model. Combining the P and R, the F_1-score for the ResNetX2-50 model reached 99.99%, an increase of 0.66 percentage points over the ResNetX2-101 model.

Figure 17 shows the pixel-by-pixel prediction results of three laboratory models: the ResNet-101-based spectral feature classification model, the ResNet-101-based texture feature classification model, and the ResNetX2-50-based map fusion feature classification model. The ResNet-101-based spectral model correctly identified most granite and limestone regions but had some misclassifications (Fig. 17 (a)). The ResNet-101-based texture model produced noisy predictions for limestone (Fig. 17(b)), because similar texture features of granite and limestone led to low classification accuracy. In contrast, the ResNetX2-50-based fused model achieved over 99.9% accuracy for both rock types, with stable predictions and minimal noise (Fig. 17(c)).

These results confirm that fusing spectral and texture features in ResNetX2-50 enhance lithological classification accuracy and stability. By integrating both types of features, the model captures more comprehensive rock characteristics, overcomes single-feature limitations, and supports geological research and exploration.

5 Discussion

The spectral-image fusion-based lithological identification method using ResNetX2-50 achieved optimal performance, with over 99% accuracy for 45 rock types in the short-wave infrared database (29 types at 100%). Compared to the ResNetX2-101 model, ResNetX2-50 demonstrated superior stability and robustness for individual rock classifications. While hyperspectral-based lithology identification is promising, several research directions merit attention:

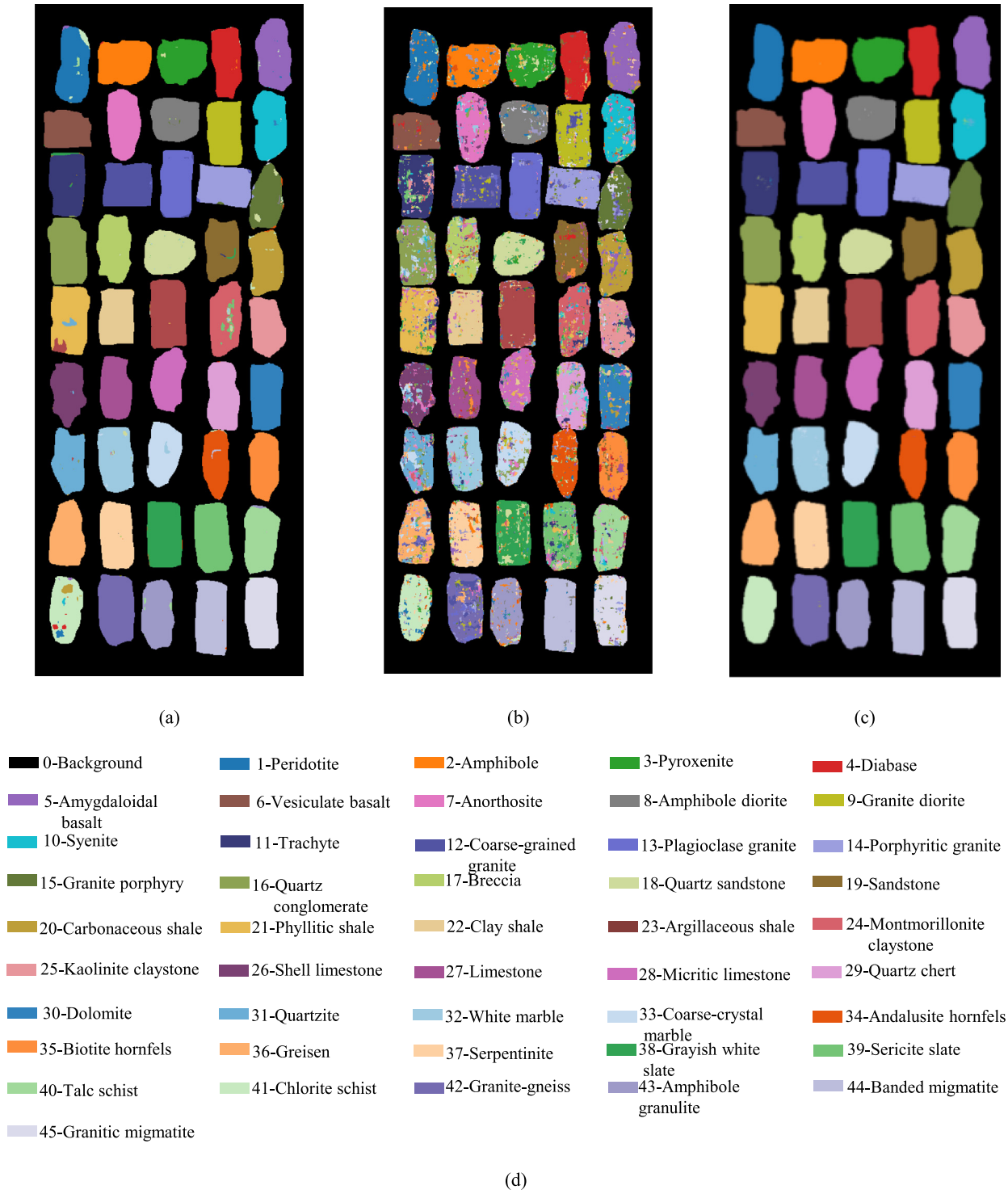


Fig. 13. Prediction results chart of different classification models. (a) Spectral features, (b) texture features, (c) fusion features, and (d) legend.

(1) Future work should enlarge the rock spectral library to include more lithologies, addressing class imbalance via oversampling minority classes in tunnel datasets to rebalance training data.

(2) Research should focus on acquiring high-quality hyperspectral data in challenging tunnel conditions (e.g., low-light, dusty environments) and assessing noise impacts on spectral features to inform model design.

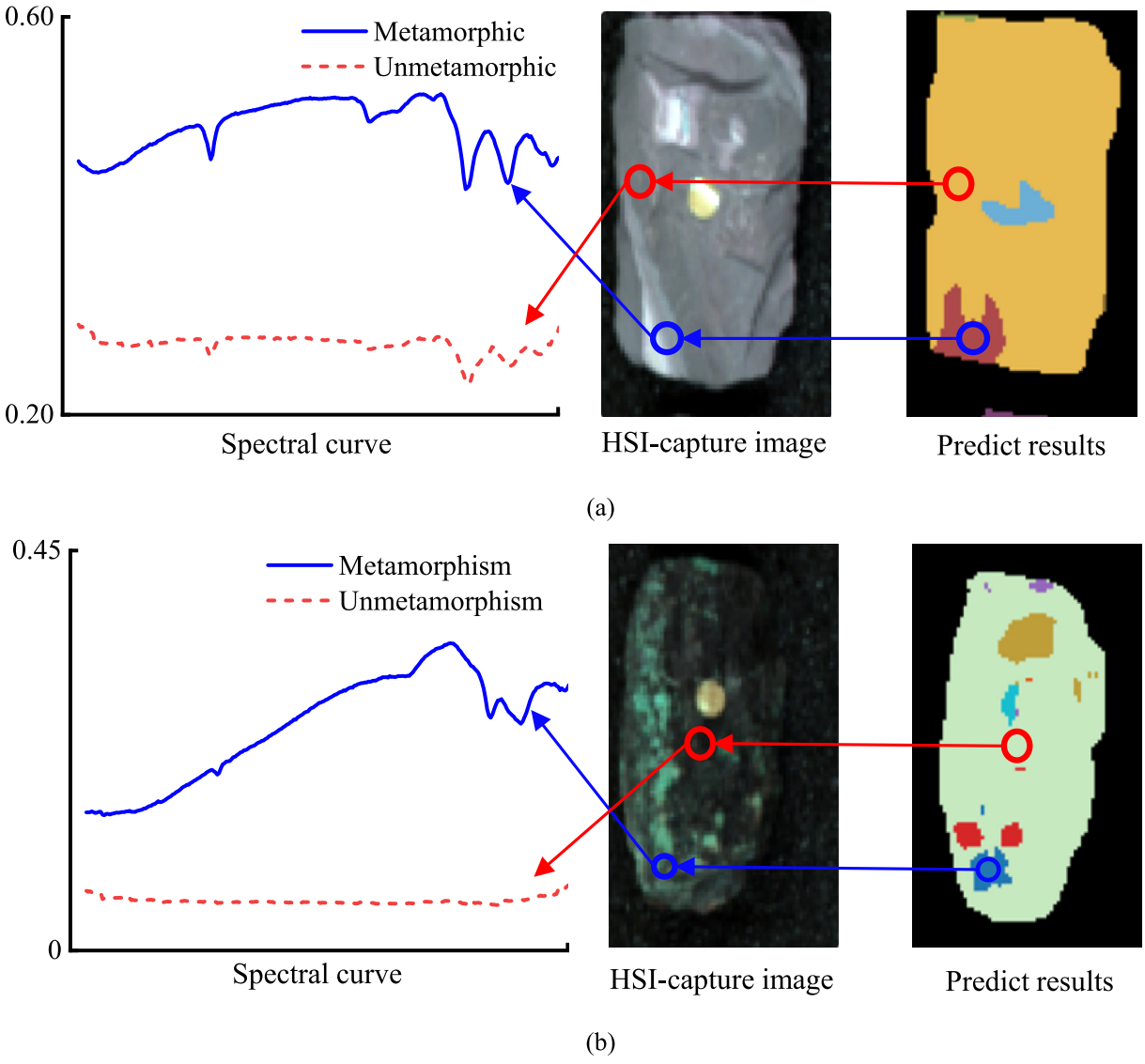


Fig. 14. Comparison of rock HSI and predicted results. (a) 21-phyllitic shale, and (b) 41-chlorite schist.

(3) Given limitations in shortwave infrared bands for some minerals (e.g., quartz, feldspar), incorporating thermal infrared spectral data into the existing database would expand its utility.
(4) Deepening analysis of spectral feature bands' relationships with mineral composition—by integrating geological mechanisms into the model and developing advanced feature selection—can enhance interpretability beyond PCA-based dimensionality reduction.
(5) Future efforts will explore efficient network architectures that balance global and local feature extraction, including optimizing spatial-spectral modules and adopting lightweight convolutions (e.g., depthwise separable convolutions) to reduce computational complexity and model parameters (Chen et al., 2025).

6 Conclusions

This article integrates deep learning with hyperspectral imaging technology and proposes a spectral-image fusion method for intelligent lithology identification. The main conclusions are as follows:

(1) A spectral-image feature fusion model was developed using a dual-channel residual neural network (ResNetX2-50/101). Feature-level fusion was performed at the fully connected layer, which improved the model's ability to handle different data sources and achieved a balance between training efficiency and model performance when applying transfer learning.

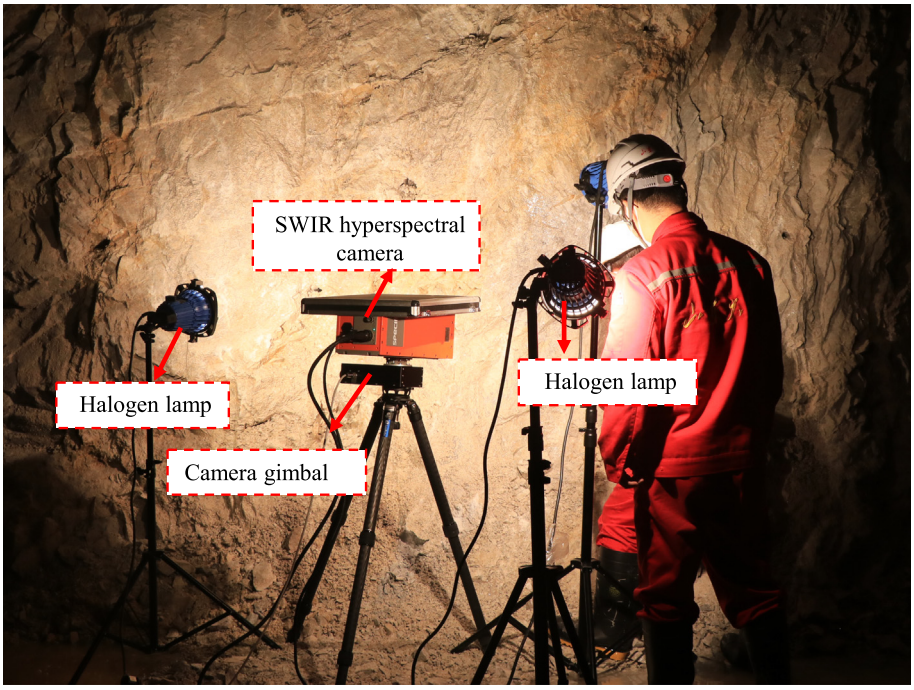


Fig. 15. In-situ hyperspectral image acquisition of the tunnel.

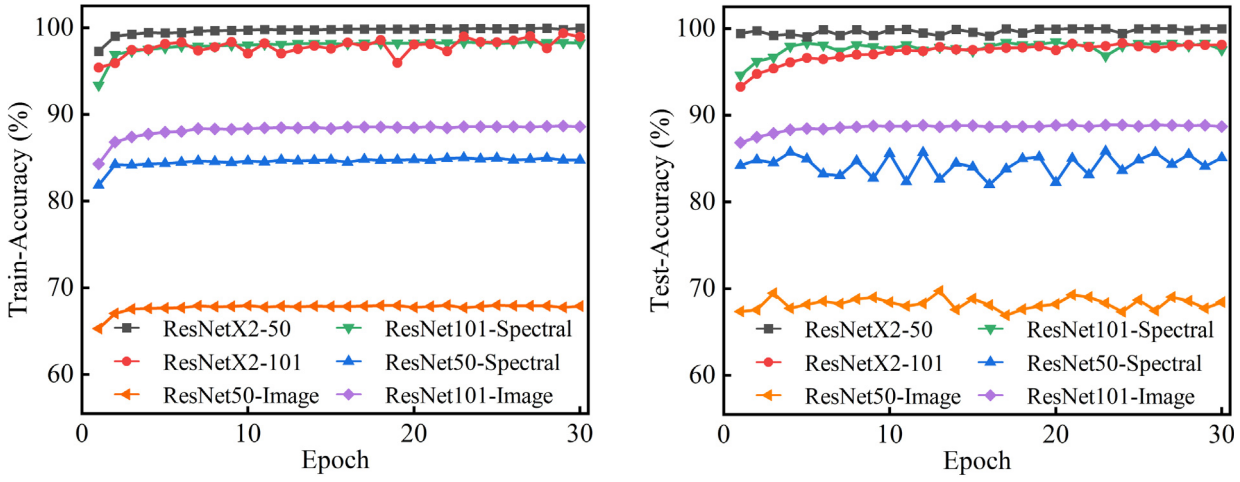


Fig. 16. Accuracy curve of engineering verification. (a) Training set, and (b) testing set.

Table 5
Comparison of overall evaluation indicators for different models based on engineering applications.

Model	P (%)	R (%)	F_1-score (%)
ResNetX2-50	99.99	99.99	99.99
ResNetX2-101	99.33	99.33	99.33

(2) Key features were effectively captured through spectral-spatial attention mechanisms, including STN and SAFE. Furthermore, the issue of limited sample sizes in engineering applications was addressed by customized transfer learning strategies.

(3) Compared to methods relying on single-feature lithology identification, the feature fusion-based model demonstrated superior accuracy and improved stability in identification performance.

(4) The ResNetX2-50 model based on feature fusion achieved the highest accuracy of 99.78%, while ResNetX2-101 attained an accuracy of 98.32% on the test set. The ResNetX2-50-based model, which incorporates feature fusion and spectral-spatial attention mechanisms, exhibited superior performance in lithology identification.

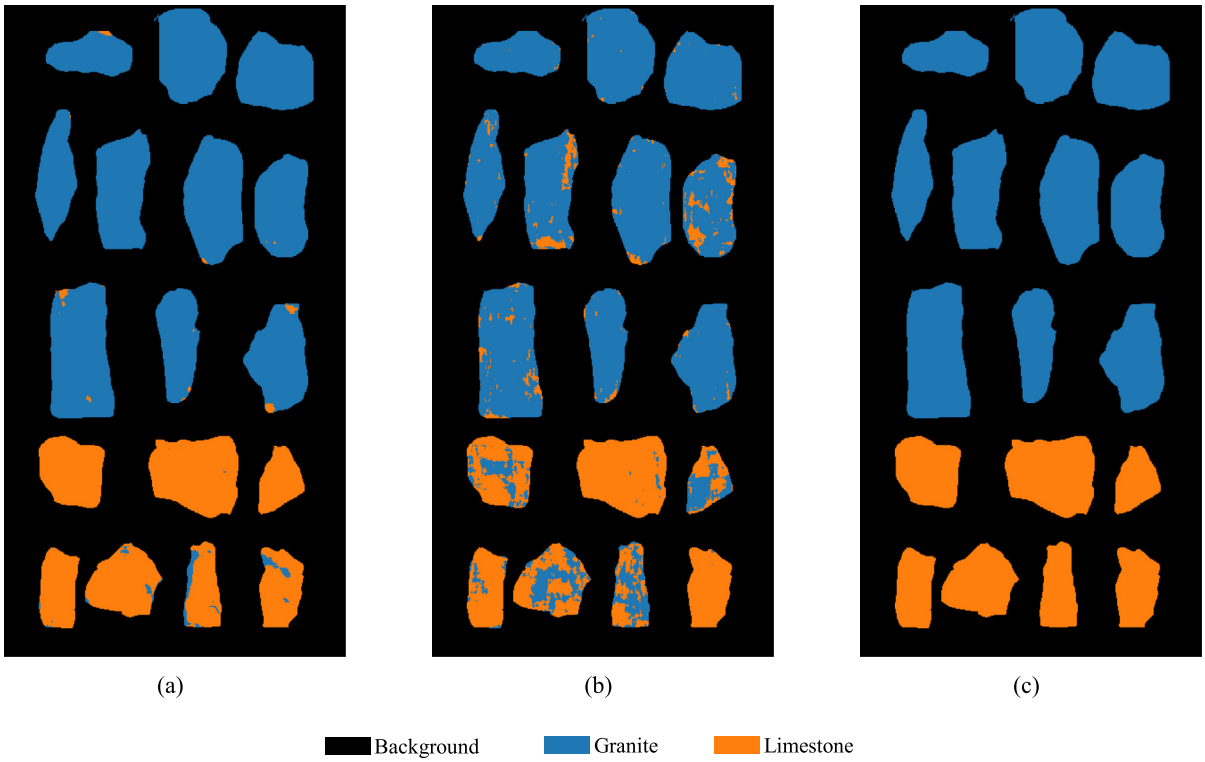


Fig. 17. Prediction results chart of different classification models based on tunnel engineering applications. (a) Spectral features, (b) texture features, and (c) fusion features.

Data availability

The data and code that support the findings of this study are available on https://github.com/haha1583091/Spectral-Image-fusion-method-for-lithology-identification.

CRediT authorship contribution statement

Zhenhao Xu: Funding acquisition, Writing – review & editing, Conceptualization, Methodology, Supervision. **Shan Li:** Visualization, Formal analysis, Writing – original draft, Data curation. **Peng Lin:** Validation, Funding acquisition. **Qianji Li:** Validation, Investigation.

Declaration of competing interest

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.

Acknowledgement

We would like to acknowledge the support from the National Natural Science Foundation of China (Grant Nos. 52379103 and 52279103) and the Natural Science Foundation of Shandong Province (Grant No. ZR2023YQ049).

References

Ahmad, M., Shabbir, S., Roy, S. K., Hong, D. F., Wu, X., Yao, J., Khan, A. M., Mazzara, M., Distefano, S., & Chanussot, J. (2022). Hyperspectral image classification—Traditional to deep models: A survey for future prospects. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15*, 968–999.

Audebert, N., Le Saux, B., & Lefèvre, S. (2019). Deep learning for classification of hyperspectral data: A comparative review. *IEEE Geoscience and Remote Sensing Magazine, 7*(2), 159–173.

Barker, R. D., Barker, S. L. L., Wilson, S., & Stock, E. D. (2021). Quantitative mineral mapping of drill core surfaces I: A method for μ XRF mineral calculation and mapping of hydrothermally altered, fine-grained sedimentary rocks from a carlin-type gold deposit. *Economic Geology, 116*(4), 803–819.

Chen, H., Li, W., & Zhu, Y. Y. (2021). Improved window adaptive gray level co-occurrence matrix for extraction and analysis of texture characteristics of pulmonary nodules. *Computer Methods and Programs in Biomedicine, 208*, 106263.

Chen, W. Z., Tan, X. Y., & Yang, J. P. (2025). Review of state-of-the-art in structural health monitoring of tunnel engineering. *Smart Underground Engineering, 1*(1), 40–50.

Ferrari, C., Foca, G., & Ulrici, A. (2013). Handling large datasets of hyperspectral images: Reducing data size without loss of useful information. *Analytica Chimica Acta, 802*, 29–39.

Gao, K. P., & Jiao, S. J. (2022). Research on lithology identification based on multi-sensor hybrid domain information fusion and support vector machine. *Earth Science Informatics, 15*(2), 1101–1113.

He, L., Li, J., Liu, C. Y., & Li, S. T. (2018). Recent advances on spectral–spatial hyperspectral image classification: an overview and new guidelines. *IEEE Transactions on Geoscience and Remote Sensing, 56*(3), 1579–1597.

Hong, D. F., Han, Z., Yao, J., Gao, L. R., Zhang, B., Plaza, A., & Chanussot, J. (2022). SpectralFormer: Rethinking hyperspectral image classification with transformers. *IEEE Transactions on Geoscience and Remote Sensing, 60*, 5518615.

Li, F., Wang, J., Lan, R. S., Liu, Z. B., & Luo, X. N. (2019a). Hyperspectral image classification using multi-feature fusion. *Optics & Laser Technology, 110*, 176–183.

Li, S., Lin, P., Xu, Z. H., Xiang, H., & Li, Q. J. (2025). Intelligent lithology identification based on transfer learning of rock hyperspectral images. *Spectroscopy and Spectral Analysis, 45*(8), 2289–2301 (in Chinese).

Li, S. T., Song, W. W., Fang, L. Y., Chen, Y. S., Ghamisi, P., & Benediktsson, J. A. (2019b). Deep learning for hyperspectral image classification: an overview. *IEEE Transactions on Geoscience and Remote Sensing, 57*(9), 6690–6709.

Li, W., Zhang, C. P., Tu, S. Q., Chen, W., & Ma, M. S. (2023). Face stability analysis of a shield tunnel excavated along inclined strata. *Underground Space, 13*, 183–204.

Li, X. Y., Li, Z. M., Qiu, H. M., Chen, G. Y., Fan, P. P., & Liu, Y. (2024a). Multi-scale spatial and spectral feature fusion for soil carbon content prediction based on hyperspectral images. *Ecological Indicators, 160*, 111843.

Li, Z. Y., Xue, Z. H., Xu, Q., Zhang, L., Zhu, T. Z., & Zhang, M. X. (2024b). SPFormer: Self-pooling transformer for few-shot hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing, 62*, 5502019.

Lin, N., Fu, J. W., Jiang, R. Z., Li, G. J., & Yang, Q. (2023). Lithological classification by hyperspectral images based on a two-layer XGBoost model, combined with a greedy algorithm. *Remote Sensing, 15*(15), 3764.

Liu, T. X., Zhang, C. P., Li, W., Tu, S. Q., Wang, L. B., & Jin, Z. X. (2025). Face failure mechanism of fault tunnels under high-temperature and high-pressure conditions using the discrete element method. *Computers and Geotechnics, 179*, 107059.

Lu, M., Zhang, J., Lyu, Q., & Zhang, L. L. (2023). Assessing the annual probability of rainfall-induced slope failure based on intensity–duration–frequency (IDF) curves. *Natural Hazards, 117*(1), 763–778.

Muktadir, G., Amro, M., Kummer, N., Freese, C., & Abid, K. (2021). Application of X-ray diffraction (XRD) and rock–eval analysis for the evaluation of middle eastern petroleum source rock. *Energies, 14*(20), 6672.

Rashid, M., Luo, M., Ashraf, U., Hussain, W., Ali, N., Rahman, N., Hussain, S., Martyushev, D. A., Vo Thanh, H., & Anees, A. (2023). Reservoir quality prediction of gas-bearing carbonate sediments in the qadirpur field: insights from advanced machine learning approaches of SOM and cluster analysis. *Minerals, 13*(1), 29.

Tan, K., Ma, W. B., Chen, L. H., Wang, H. M., Du, Q., Du, P. J., Yan, B. K., Liu, R. Y., & Li, H. D. (2021). Estimating the distribution trend of soil heavy metals in mining area from HyMap airborne hyperspectral imagery based on ensemble learning. *Journal of Hazardous Materials, 401*, 123288.

Tejasree, G., & Agilandeeswari, L. (2024). Land use/land cover (LULC) classification using deep-LSTM for hyperspectral images. *The Egyptian Journal of Remote Sensing and Space Sciences, 27*(1), 52–68.

Wang, S., Guan, K. Y., Zhang, C. H., Lee, D., Margenot, A. J., Ge, Y. F., Peng, J., Zhou, W., Zhou, Q., & Huang, Y. Z. (2022). Using soil library hyperspectral reflectance and machine learning to predict soil organic carbon: Assessing potential of airborne and spaceborne optical soil sensing. *Remote Sensing of Environment, 271*, 112914.

Wang, X. F., Wei, Y. Y., Jiang, T., Hao, F. X., & Xu, H. F. (2024). Elastic–plastic criterion solution of deep roadway surrounding rock based on intermediate principal stress and Drucker–Prager criterion. *Energy Science & Engineering, 12*(6), 2472–2492.

Wang, Z. Y., & Zuo, R. G. (2024). An evaluation of convolutional neural networks for lithological mapping based on hyperspectral images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17*, 6414–6425.

Wang, Z. Y., Zuo, R. G., & Jing, L. H. (2021). Fusion of geochemical and remote-sensing data for lithological mapping using random forest metric learning. *Mathematical Geosciences, 53*(6), 1125–1145.

Xu, Z. H., Li, S., Lin, P., Xiang, H., & Li, Q. J. (2025). Bayesian-optimized lithology identification via visible and near-infrared spectral data analysis. *Intelligent Geoengineering, 2*(1), 1–13.

Xu, Z. H., Liu, F. M., Lin, P., Shao, R. Q., & Shi, X. S. (2021). Non-destructive, in-situ, fast identification of adverse geology in tunnels based on anomalies analysis of element content. *Tunnelling and Underground Space Technology, 118*, 104146.

Xu, Z. H., Ma, W., Lin, P., & Hua, Y. L. (2022). Deep learning of rock microscopic images for intelligent lithology identification: Neural network comparison and selection. *Journal of Rock Mechanics and Geotechnical Engineering, 14*(4), 1140–1152.

Xu, Z. H., Yu, T. F., Lin, P., & Li, S. C. (2023). Anomalous patterns of clay minerals in fault zones. *Engineering Geology, 325*, 107279.

Xu, Z. H., Shi, H., Lin, P., & Li, S. (2024). LithoSegNet: Regional attention-based deep fusion of multi-scale and cross-stage features for real-time lithology segmentation. *International Journal of Rock Mechanics and Mining Sciences, 180*, 105814.

Yang, H. Q., Chen, C. W., Ni, J. H., & Karekal, S. (2023). A hyperspectral evaluation approach for quantifying salt-induced weathering of sandstone. *Science of The Total Environment, 885*, 163886.

You, M. L., Hong, Z. K., Tan, F., Wen, H., Zhang, Z. R., & Lyu, J. H. (2024). Stratigraphic identification using real-time drilling data. *Journal of Rock Mechanics and Geotechnical Engineering, 16*(9), 3452–3464.

Zhang, H., Liu, H. H., Yang, R. H., Wang, W., Luo, Q. Q., & Tu, C. D. (2024). Hyperspectral image classification based on double-branch multi-scale dual-attention network. *Remote Sensing, 16*(12), 2051.

Zhang, J., Sun, Y., Hu, J. Z., & Huang, H. W. (2023a). Assessing site investigation program for design of shield tunnels. *Underground Space, 9*, 31–42.

Zhang, Y., Huynh, C. P., & Ngan, K. N. (2019). Feature fusion with predictive weighting for spectral image classification and segmentation. *IEEE Transactions on Geoscience and Remote Sensing, 57*(9), 6792–6807.

Zhang, Z. Q., Wang, G. W., Carranza, E. J. M., Liu, C., Li, J. J., Fu, C., Liu, X. X., Chen, C., Fan, J. J., & Dong, Y. L. (2023b). An integrated machine learning framework with uncertainty quantification for three-dimensional lithological modeling from multi-source geophysical data and drilling data. *Engineering Geology, 324*, 107255.

Zhong, Z. L., Li, J., Luo, Z. M., & Chapman, M. (2018). Spectral–spatial residual network for hyperspectral image classification: A 3-D deep learning framework. *IEEE Transactions on Geoscience and Remote Sensing, 56*(2), 847–858.