Biomedical Testing & Sensing
2024, VOL. 2, NO. 2, 14-18
DOI: 10.47297/wspbtsWSP2755-952103.20240202

Algorithm for Lightweight Mineral Recognition Based on Improved Neural Network Model

Wentao Chen

Nanjing University of Posts and Telecommunications, NJUPT, Nanjing, China 210023

ABSTRACT
This study explores the potential of deep learning in mineral identification and addresses the challenge of deploying large neural network models on embedded devices. To achieve this, I propose a lightweight algorithm that maintains accuracy while reducing the model's storage space and computational load.

KEYWORDS
Deep learning; Mineral identification; Lightweight algorithm; Embedded devices; Model optimization

1 Introduction

Traditional mineral extraction relies on manual labor and heavy machinery, leading to inefficiencies and safety risks. Identification primarily depends on expert judgment based on visual and chemical properties, making it time-consuming, specialist-dependent, and prone to bias [1]. Recently, intelligent mining technologies utilizing robotics and advanced equipment have improved efficiency, reducing errors and operational hazards. While deep learning models enhance identification accuracy, their deployment is constrained by the limited computational power of embedded devices. Therefore, developing a lightweight, efficient, and accurate neural network for mineral identification is crucial for practical implementation. With advancements in technology and social development, numerous scholars have begun applying deep learning to mineral identification. In 2019, Peng Weihang et al. increased data diversity through random image selection, establishing an InceptionV3 model and introducing the CenterLoss function, achieving an identification accuracy of 86% for 16 types of minerals [2]. In 2020, Guo Yanjun et al. selected the ResNet-18 framework in their design, achieving an accuracy rate of 89% for five types of minerals [3]. In the same year, Rong Gao et al. utilized the U-Net semantic segmentation model for image recognition, reporting an accuracy of 93% for sample detection [4]. In 2022, Yang Biao et al. constructed an intelligent recognition model based on depthwise separable convolution combined with attention mechanisms, reaching 90% accuracy for five types of minerals [5]. In 2023, Liu Bowen et al. enhanced rock image recognition performance by incorporating depthwise separable convolution layers and dynamic convolution layers into a VGG-16 neural network [6]. In 2024, Wang Lin et al. employed ensemble methods with average and weighted voting on ResNet, RegNet, EfficientNet, and Vision Transformer models, achieving an accuracy of 87.47% after dataset training [7]. That same year, Wan Chengzhou et al. proposed a progressive training mineral recognition model based on Next-ViT-Small, successfully achieving an accuracy of 86.5% for 36 types of minerals [8]. In summary, despite the demonstrated potential of deep learning technologies in the field of mineral identification.

2 Motivation

2.1 Data Set

This study used a mineral dataset with seven types: Biotite (67), Bornite (169), Chrysocolla (160), Malachite (230), Muscovite (77), Pyrite (96), and Quartz (142). To tackle the sample imbalance, I applied image augmentation techniques like random flips, rotations, cropping, and resizing. This increased data diversity, crucial for building a robust model to classify these minerals accurately.

Table 1 The dataset used for experiment

Order	Styles	Number
1	Biotite	67
2	bornite	169
3	chrysocolla	160
4	malachite	230
5	muscovite	77
6	pyrite	96
7	quartz	142

2.2 Data processing

In the code, PyTorch's transforms module is used for data preprocessing. For training images, transformations include resizing to 256x256, random flips, rotations up to 15 degrees, and cropping to 224x224. These augmentations boost sample diversity and make the model more robust. For test images, resizing to 224x224 and normalization are applied to ensure consistent model inputs. This preprocessing pipeline is vital for optimizing the mineral classification model's performance.

2.3 ResNet-50

ResNet, a deep convolutional neural network by Microsoft Research, tackles gradient vanishing and explosion issues through a "residual learning" mechanism. It uses skip connections to directly pass inputs to later layers, improving gradient and information flow. ResNet-50, a 50-layer variant, is widely used for image classification and object detection.

Its architecture includes convolutional layers, batch normalization, ReLU activations, and max pooling. Specifically, it begins with a 7×7 convolution (stride 2), followed by batch normalization and ReLU. The output is then downsampled by a 3×3 max-pooling layer (stride 2).

Residual blocks: ResNet-50 includes four residual blocks, each with multiple units. Each unit has two 3×3 convolutions with batch normalization and ReLU in between. Outputs are added to inputs via skip connections and passed through ReLU again. This enables learning residual mappings, boosting training efficiency and performance.

Global Average Pooling Layer: After the residual blocks, this layer reduces the feature map to a fixed-size vector.

Fully Connected Layer: The final layer performs classification, matching the number of output dimensions to the number of categories.

While the deep structure and residual learning mechanism of ResNet-50 enable it to perform well in complex image classification tasks, its computational complexity and large model size limit its deployment on resource-constrained embedded devices.

Figure 1 illustrates the detailed network architecture of ResNet-50, which clearly delineates five stages. Each stage contains a different number of residual blocks, and each residual block includes BTNK1 and BTNK2 units. These units implement feature transformation and fusion through 1×1 and 3×3 convolutional layers, and residual learning through jump connections.

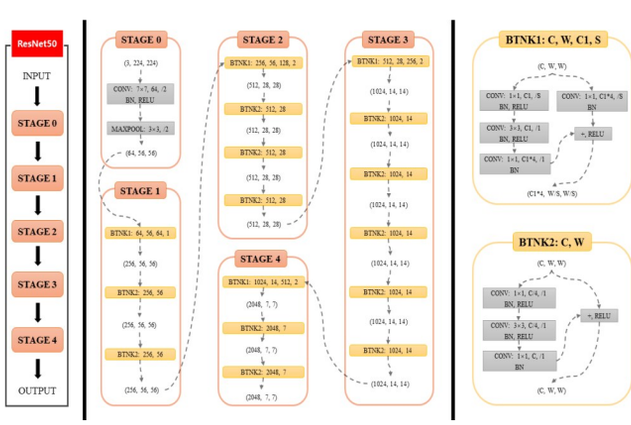

Figure 1 Schematic diagram of ResNet-50 network architecture

2.4 Introduction of SE modules

The Squeeze-and-Excitation (SE) module was integrated into the ResNet-50 model to enhance mineral recognition performance. This module leverages a channel attention mechanism to recalibrate feature responses, improving feature selection and generalization in complex tasks.

The SE module's core idea is to select important features for mineral recognition by weighting feature map channels through attention. It involves three steps:

Squeeze (compression): the feature map of each channel is first pooled globally on average to obtain the global description information of each channel. This operation compresses the spatial dimensions (width and height) of each channel into a single value, thus obtaining a global feature description for each channel.

Excitation: Next, the descriptive information for each channel is processed using two fully connected layers. Through the bottleneck layer (reduction layer), the number of channels is first reduced and then recovered using the activation function (ReLU) and the Sigmoid function, and a coefficient between 0 and 1 is generated. These coefficients represent the importance of each channel.

Recalibration: finally, the features of each channel are weighted by the weight coefficients generated by the Sigmoid function, enhancing the response to important features and suppressing the response to irrelevant features.

In this study, the SE module was introduced for layer3 and layer4 in the ResNet-50 model. To ensure the effectiveness of the module, I applied the SE module on these two layers because these layers usually extract higher-level features that are critical for mineral identification.

By introducing the SE module after layer3 and layer4 of ResNet-50, the model's ability to represent mineral features is effectively improved, especially in the complex feature extraction process of images. The SE module improves the accuracy of mineral classification by adaptively assigning weights to each channel, which enables the network to focus on important features and suppress redundant information.

2.5 Quantitative Awareness Training (QAT)

To enhance model efficiency on embedded devices, Quantization-Aware Training (QAT) is employed. QAT simulates quantization effects during training, enabling the model to adjust weights and activations while preserving accuracy, significantly reducing storage and computational costs.

The process begins with integrating a quantization module to mimic low-precision computation. PyTorch ' s QuantStub and DeQuantStub convert data between floating-point and integer formats, allowing the network to adapt to quantization constraints. A tailored quantization configuration (qconfig) is applied, defining quantized layers and methods optimized for embedded hardware to ensure minimal computational overhead.

Before training, the model undergoes preprocessing to support QAT. Essential network components, such as convolutional and fully connected layers, are embedded with quantization modules, ensuring adaptation to low-precision environments. During training, the model continuously learns through quantization and dequantization operations, allowing it to operate efficiently in reduced-precision settings. After training, a transformation step converts the model into its final quantized form while maintaining floating-point representation during training.

QAT optimizes both storage and computational efficiency, enabling practical deployment on embedded systems. The final quantized model significantly reduces parameters and memory footprint while maintaining classification performance, making it well-suited for real-world applications requiring lightweight and efficient deep learning solutions.

3 Conclusion and outlook

3.1 Experimental results and performance analysis

In the experiments, the improved lightweight model shows overwhelming advantages in terms of the number of parameters and storage space available:

The original ResNet-50 model has 25,636,712 parameters and 98.0 MB storage requirement. The improved lightweight model has only 52,992 parameters and 0.20 MB of storage. Compared with the original model, the number of parameters and the storage requirement of the lightweight model are reduced by 99.79% and 99.8%, respectively, which achieves great compression.

This achievement enables the improved model to run on resource-constrained embedded devices, providing great convenience and scalability for practical applications.

The compressed lightweight model maintains strong classification performance despite minor accuracy variations. As shown in Fig. 2, the original ResNet-50 model achieves 86.7% accuracy, while the lightweight model reaches 90.4%, reflecting a 3.7% improvement. Experimental results demonstrate that integrating the SE module and QAT enhances feature extraction while reducing computational complexity and storage needs. Additionally, training curve analysis indicates improved stability and reduced overfitting, further confirming the model's robustness and applicability.

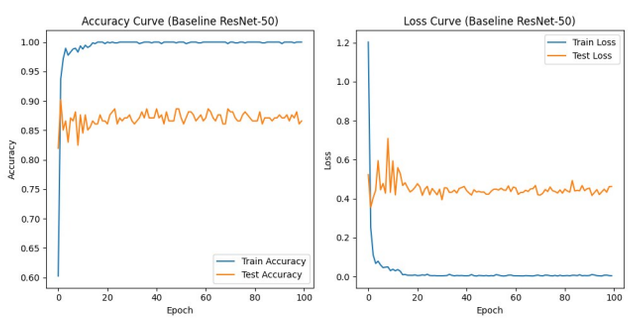

Figure 2 ResNet-50 model

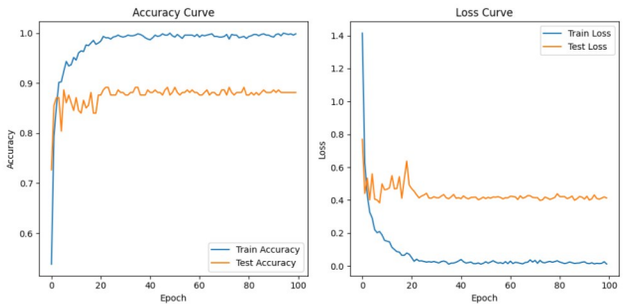

Figure 3 Improved ResNet-50 model

3.2 Strengths and improved features of the model

The proposed lightweight mineral identification model, enhanced by the SE module and QAT, shows significant advantages in resource-constrained environments. The detailed advantages and improved features of the model are analyzed below:

Model Lightweighting

The improved lightweight model achieves significant compression compared to the original ResNet-50. The number of parameters is reduced from 25.6 million to 53,000, a 99.79% decrease. Storage requirements drop from 98.0 MB to 0.20 MB, a 99.8% reduction. This makes the model highly suitable for deployment on resource-limited devices like embedded systems and mobile terminals, overcoming the challenges posed by the high resource demands of traditional deep learning models.

Efficient computational performance

The model employs low-precision quantization (e.g., INT8) along with QAT in the inference stage, which significantly reduces computational complexity and enhances hardware adaptability. Test results show that the quantized model on embedded devices achieves 3-5 times higher inference speed compared to the original model with FP32 precision, while also significantly reducing power consumption. This performance is crucial for industrial scenarios that require real-time performance and low power consumption.

Significantly Enhanced Feature Extraction Capability

The SE module enables dynamic feature weighting during forward propagation, prioritizing essential channels to enhance key mineral feature extraction while filtering out irrelevant information. This reduces background noise and improves classification accuracy, particularly in complex scenarios. Overall, it strengthens feature representation and maintains high performance under lightweight conditions.

More stable and robust training process

Experimental results show that the improved model achieves more stable convergence and effectively mitigates overfitting compared to the original ResNet-50. On a small-sample dataset, it demonstrates stronger generalization, increasing test accuracy by 3.7% (from 86.7% to 90.4%). Its consistent performance across varied data distributions highlights its robustness in identifying different mineral types and adapting to diverse image features

3.3 Limitations analysis and future directions

Despite the significant advantages of the proposed lightweight mineral identification model in terms of storage requirements, computational efficiency, and feature extraction capability, several limitations still need to be addressed:

Firstly, the dataset size and diversity are limiting factors. The training data, sourced mainly from a customized mineral image dataset, lacks sufficient scale and variety. Uneven sample distribution across mineral categories results in suboptimal classification performance and potential model bias. Additionally, the dataset's lab-based images lack real-world complexity, such as background interference and lighting variations, which may undermine the model's robustness in actual deployment.

Secondly, Quantization-Aware Training (QAT) minimizes quantization errors, extreme low-precision formats (e.g., INT4) can still impact model performance. In high-similarity mineral classes, information loss during quantization may blur classification boundaries. Additionally, hardware adaptation challenges arise due to weight distribution shifts, leading to performance fluctuations in the inference process caused by quantization method limitations.

Another limitation is the lack of cross-platform performance validation. While the model's quantization and lightweight design perform well on embedded devices, its performance on other hardware platforms (e. g., high-performance GPUs, FPGAs, etc.) has not been fully validated. Different hardware platforms may have different optimization strategies and computational efficiencies for the quantized model, necessitating further testing of the model's broad adaptability.

Finally, there is the computational overhead of the attention mechanism. Although the SE module significantly enhances the model's feature extraction capability, the additional attention computational overhead it introduces increases model complexity to some extent. In extreme resource-constrained embedded devices, this overhead may still be a bottleneck, especially when the model needs to process high-resolution images.

To address the limitations of the model in practical applications, this paper proposes the following key improvement directions to further enhance its performance and expand application scenarios:

Firstly, to enhance the model's generalization, focus on expanding and enriching the dataset. Introduce real-world mineral images from industrial and field settings to reflect practical applications. Use advanced data augmentation methods like CutMix and MixUp to create diverse training data, helping the model handle complex backgrounds and lighting. Additionally, apply self-supervised learning to extract features from unlabeled data, addressing the issue of limited samples for some mineral categories.

While Quantization-Aware Training (QAT) has significantly contributed to model lightweighting, further refinements are necessary. Exploring dynamic quantization and hybrid precision techniques can enhance performance. Dynamically adjusting quantization accuracy based on hardware or task requirements ensures a balance between efficiency and precision. Additionally, post-quantization calibration methods can fine-tune model weight distributions, minimizing low-precision quantization errors (e.g., INT4) and preserving classification accuracy in mineral identification tasks.

To enhance model applicability across various hardware platforms (e. g., embedded devices, NPUs, FPGAs), optimization strategies tailored to specific architectures are necessary. This involves adjusting model design to leverage hardware acceleration, such as optimizing ResNet-SE inference efficiency on NPUs. Additionally, compiler toolchains and frameworks like TensorRT or TVM can further compress the model and enhance computational efficiency, enabling streamlined end-to-end deployment.

Finally, there is the lightweight improvement of the attention mechanism. Although the SE module significantly improves the model's feature extraction capability, its computational overhead still needs to be optimized on resource-constrained devices. This can be achieved by lightweighting the attention module, such as replacing the existing SE module with a more efficient attention mechanism that reduces computational complexity while maintaining performance. Alternatively, hybrid attention can be incorporated. Using a hybrid mechanism of channel attention and spatial attention, the classification performance can be further improved with lower computational cost through hierarchical design.

References

[1] H. Huizhen, G. Qing, and H. Xiumian, "Research progress and prospects of intelligent mineral identification methods based on machine learning," Earth Science, vol. 46, no. 9, pp. 3091–3106, 2021.

[2] W. Peng, B. Lin, S. Shiwei, et al., "Intelligent identification of common minerals based on improved InceptionV3 model," Geological Bulletin of China, vol. 38, no. 12, pp. 2059–2066, 2019.

[3] G. Yanjun, Z. Zhe, L. Hexun, et al., "Research on intelligent mineral identification methods based on deep learning," Earth Science Frontiers, vol. 27, no. 5, pp. 39–47, 2020.

[4] Y. L. Gao, Z. Sun, W. Li, et al., "Automatic coal and gangue segmentation using U-Net based fully convolutional networks," Energies, vol. 13, no. 4, p. 829, 2020.

[5] Y. Biao, M. Yiji, N. Ruipu, et al., "Intelligent mineral image recognition based on multi-scale dense connection network," Journal of Yunnan University (Natural Science Edition), vol. 44, no. 6, pp. 1118–1126, 2022.

[6] B. Liu and X. Liu, "Application of optimized multi-scale convolutional neural network models in mineral identification," Mineralogy and Petrology, vol. 43, no. 3, pp. 10–19, 2023. DOI: 10.19719/j.cnki.1001-6872.2023.03.02.

[7] L. Wang, X. Ji, M. Yang, et al., "Miner identification based on data augmentation and ensemble learning," Earth Science Frontiers, vol. 31, no. 4, pp. 87–94, 2024.

[8] C. Wan, X. Ji, M. Yang, et al., "Mineral image recognition based on progressive multi-granularity training deep learning," Earth Science Frontiers, vol. 31, no. 4, pp. 112–118, 2024. DOI: 10.13745/j.esf.sf.2024.5.1.