IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Development And Analysis Of Integrated Deep Learning Framework For Lithographic Image Classification

D. R. Solanke

Assistant Professor
Department of Applied Electronics,
SGB Amravati University, Amravati, India (MS)

Abstract: Accurate classification of petrological samples is fundamental to geological investigations and resource exploration. Traditional methods relying on manual lithographic analysis are time-consuming and subject to expert bias. This study presents a novel deep learning framework for automated classification of Lithographical thin-section images. Utilizing convolutional neural networks (CNNs), the model was trained on a curated dataset encompassing diverse rock types, including igneous, sedimentary, and metamorphic classes. Data augmentation and transfer learning techniques enhanced model generalizability and performance. The proposed system achieved an overall classification accuracy exceeding 90%, outperforming conventional machine learning baselines. The results highlight the potential of deep learning to streamline petrological workflows, reduce human error, and enable large-scale geological mapping. Future work aims to integrate geochemical data and expand the model's applicability to field-acquired samples

Index Terms - Accuracy, Convolution, Deep Learning, Image, Lithography, Neural Network.

I. INTRODUCTION

The classification of petrological images plays a crucial role in the field of geology, enabling the identification and analysis of rock types and their mineral compositions. Traditional methods of petrological classification often rely on expert interpretation, which can be time-consuming and subject to human error. With the advancement of machine learning, particularly deep learning techniques, automated image classification has gained significant attention. Convolutional Neural Networks (CNNs), a type of deep learning model designed to process and analyze visual data, have demonstrated remarkable success in various image recognition tasks. Applying CNNs to petrological image classification offers a promising approach to enhance accuracy, efficiency, and consistency in identifying rock textures and mineralogical features. This integration of petrology and artificial intelligence aims to support geologists in faster and more reliable analysis, contributing to improved geological mapping, resource exploration, and academic research.

II. REVIEW OF LITERATURE

Deep learning techniques, particularly convolutional neural networks (CNNs), have revolutionized the classification of petrological samples, offering a powerful alternative to traditional manual methods. Petrological classification traditionally involves detailed microscopic examination of thin-section images under plane-polarized and cross-polarized light, a process that is both time-consuming and subject to expert bias. Recent advancements in computer vision have facilitated the automation of this task, significantly enhancing both efficiency and accuracy. Su et al. [1] introduced a concatenated CNN (Con-CNN) model that simultaneously processes plane-polarized and cross-polarized images of petrographic thin sections. This dual-channel input enables the network to extract richer textural and mineralogical features, achieving a notable

accuracy of 89.97%. In a subsequent journal version [2], the authors incorporated a statistical revision mechanism to correct misclassifications at the image patch level, further improving robustness. Sarmad et al. [3] extended the deep learning paradigm into core-scale reservoir characterization by employing a UNet architecture. Their model classifies petrological samples based on pore geometry and mineralogy from micro-CT images, automating expert-driven tasks in the petroleum industry and exemplifying the scalability of such approaches. Patch-based classification has also gained traction as a strategy to capture local mineral spatial arrangements. Wan and Yun [4] trained a CNN on such patches, achieving over 90% accuracy in classifying six igneous rock types, underlining the effectiveness of spatial segmentation. Further innovation was presented by Dell'Aversana [5], who utilized deep convolutional residual networks (ResNets). The study demonstrated that ResNets outperform fully connected networks in mineral classification, primarily due to their ability to maintain performance in deeper architectures. Remote sensing integration was explored by El-Desoky et al. [6], who utilized an AlexNet-based CNN alongside Landsat-9 imagery and field validation for classifying Neoproterozoic basement rocks, reaching an accuracy of 95.27%. This multidisciplinary method bridges petrology and geospatial analysis. Real-time classification potential was showcased by Pascual [7], who developed a CNN model achieving 99.60% accuracy, markedly outperforming classical methods such as support vector machines (SVMs). This highlights the operational readiness of CNNs for field deployment. In data-constrained environments, Hao et al. \[8] emphasized the importance of data augmentation and training optimization. Their CNN model, trained on 350 images, achieved 82.86% accuracy, illustrating that methodological rigor can compensate for limited datasets. Mineral-specific classification was pushed further by Cifuentes et al. [9], who combined CNNs with short-wave infrared (SWIR) hyperspectral imaging. Their model effectively identified nine mineral types, demonstrating the benefit of integrating spectral data with deep learning. Lastly, Marathe et al. [10] addressed intra-category classification by differentiating volcanic from plutonic igneous rocks using grain size and texture. Their work reinforces the importance of domain-specific features in improving classification granularity. These studies collectively underscore the transformative impact of deep learning in petrology. By leveraging sophisticated architectures such as Con-CNN, UNet, ResNet, and AlexNet, researchers have significantly advanced the automation of rock and mineral classification. While challenges remain, particularly in dataset availability and model generalization, the trajectory of current research suggests that CNN-based frameworks will become integral to both academic geology and industrial applications.

III. DATASET

For this study secondary data has been collected by researchers inspired by the perseverance rover of NASA. The rock classification dataset in Fig. 1. used in this study was obtained from Hossain et al. [11] for the purpose of research of similar rocks identification in a mars-like surface. Collection methodology was using web scraping. A rock classification dataset typically contains images or spectral data of different rock types, categorized into 7 classes i.e. "Basalt", "Coal", "Limestone", "Marble", "Quartzite", and "Sandstone. The Number of image samples used in the present work are indicated by the Fig. 2. These datasets are used for training machine learning models to automate rock classification, which is important in geology, mining, and other fields.

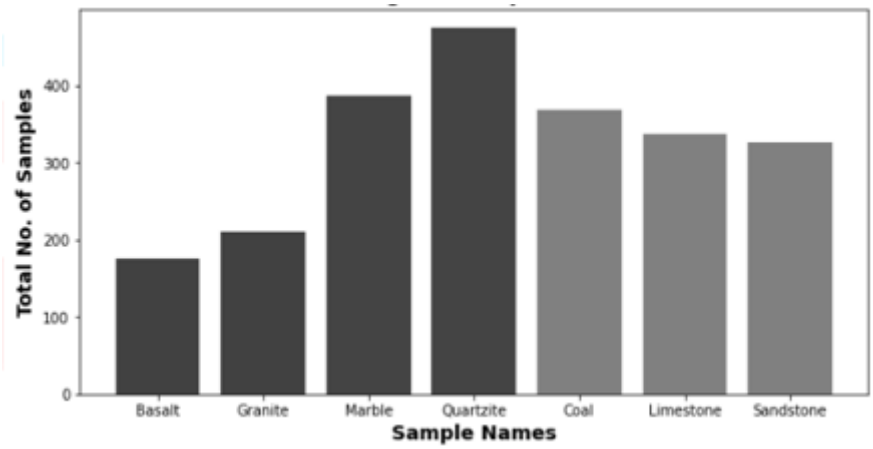

Fig. 1. The lithographic classification dataset

Table 1. Sample wise distribution of lithographic rock samples

Sr. No.	Name of Sample	No. of Samples	Percentage of Samples (%)
1	Basalt	174	7.6
2	Coal	368	16.2
3	Granite	210	9.2
4	Limestone	337	14.8
5	Marble	385	17.0
6	Quartzite	474	20.9
8	Sandstone	325	14.2

The Table 1 and pie chart (Fig. 2.) represents the class distribution of lithographic rock samples in the dataset, consisting of a total of 2,274 image samples.

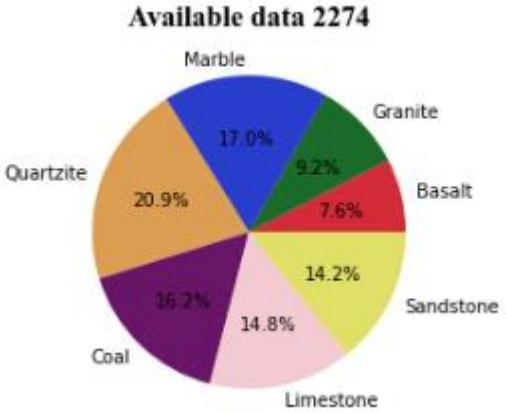

Fig. 2. Class distribution of lithographic rock samples

This visualization is crucial in understanding the balance of the dataset, which directly impacts the design and training efficiency of the deep learning model.Among the seven lithographic classes, Quartzite holds the highest proportion, comprising approximately 20.9% of the total data. This is followed by Marble, which constitutes 17.0%, and Coal, accounting for 16.2%. These three categories collectively dominate the dataset, making up over half of the total samples. Limestone (14.8%) and Sandstone (14.2%) also have substantial representation, while Granite and Basalt are relatively underrepresented, making up only 9.2% and 7.6%, respectively.This class imbalance is a critical factor in model development, as deep learning algorithms are often sensitive to skewed class distributions. Models trained on such datasets may become biased towards majority classes, leading to poor generalization on minority classes such as Granite and Basalt. Therefore, during model training and evaluation, techniques such as data augmentation, class weighting, or synthetic oversampling (e.g., SMOTE) may be essential to ensure fair representation and balanced learning across all lithographic classes.

IV. Methodology and Experimental Setup

A. Model Establishment

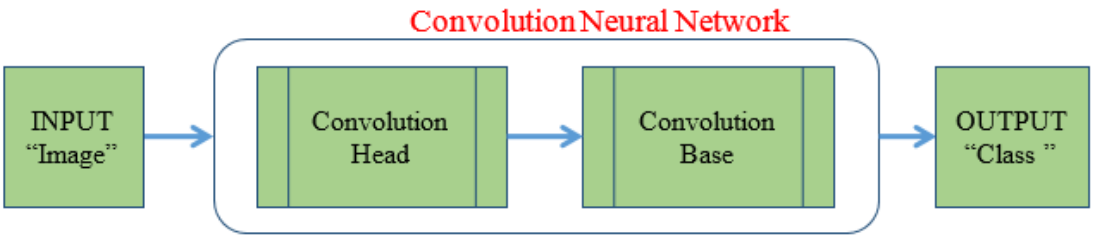

Fig. 3. Sequential Convolution Neural Network Model

In a proposed research work, an attempt is made to solve, image classification problem by using transfer learning technique through deep learning approach. Convolution Head is designed with five feature maps , each consisting two dimensional convolution layer followed by Max pooling layer (could include layers like Normalization layers (regularization), Padding layers, etc.) to extract the most prominent features of the raw image.

B. Convolution Head

The model under consideration is a Sequential Convolutional Neural Network (CNN) designed for image classification, where the input images are presumably of size 224 × 224 × 3 (height × width × channels). The network consists of multiple convolutional and pooling layers followed by fully connected (dense) layers. The first layer is a Conv2D layer with 16 filters and a kernel size of 3×3. This layer extracts low-level features such as edges and textures from the input images. The output shape of this layer is (222, 222, 16), indicating that after applying the convolution operation, the spatial dimensions are slightly reduced due to the kernel size, while the depth corresponds to the number of filters. This layer contains 448 learnable parameters, including the weights of the filters and biases. Following the first convolution, a MaxPooling2D layer is applied to reduce the spatial dimensions by a factor of 2, resulting in an output shape of (111, 111, 16). Pooling helps in reducing computation, controlling overfitting, and providing translation invariance by retaining the most prominent features in each region. The network then applies a second convolutional layer, Conv2D with 32 filters, producing an output of shape (109, 109, 32). This layer captures more complex patterns in the image, combining features learned from the previous layer. Its parameter count is 4,640, reflecting the increased number of filters. Another MaxPooling2D layer reduces the feature map dimensions to (54, 54, 32).The third convolutional layer, Conv2D with 64 filters, further extracts higher-level representations, producing an output of (52, 52, 64) with 18,496 parameters. Subsequent MaxPooling2D reduces this to (26, 26, 64).The fourth convolutional layer uses 128 filters, resulting in an output of (24, 24, 128) and 73,856 parameters. A corresponding max pooling reduces the spatial size to (12, 12, 128).The fifth convolutional layer expands the depth to 256 filters, producing (10, 10, 256) and 295,168 parameters. After pooling, the feature maps are reduced to (5, 5, 256). At this stage, the network has successfully encoded the hierarchical features of the images, from simple edges to complex shapes and textures, while drastically reducing spatial dimensions.

C. Convolution Base

The Convolution Base is used to classify the image by passing the sample images through the Convolutional layer using Flattening Layer which converts the 3D feature maps into a 1D vector of size 6400, making it suitable for fully connected layers., deals with an array of values. The first fully connected layer (Dense layer) has 512 neurons with 3,277,312 parameters, allowing the network to learn non-linear combinations of the extracted features and perform complex decision-making. The final dense layer has 6 neurons, corresponding to the number of target classes in the classification task, with 3,078 parameters. This layer produces the raw output logits that are typically fed into a dense Layer with sigmoid function (categorical classification) with the output size = six classes (i.e. “Basalt”, “Coal” , “Limestone”, “ Marble”, “Quartzite”, “Sandstone).

D. Model Evaluation

Let the dataset be

$$D = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots\dots\dots\dots (x_i, y_i)\} \quad (1)$$

In k-fold cross validation, *D* is split into disjoint subset folds (*k)*

$$D = D_1 \cup D_2 \cup D_3 \cup \dots\dots\dots \cup D_k, D_i \cap D_j = 0 \ \ for\ i \neq j \quad (2)$$

For i^{th} fold

$$D_{Train}^{(i)} = D \backslash D_i, \quad D_{val} = D_i \quad (3)$$

Let the model trained in the i-th fold produce a performance metric is the average (accuracy and loss etc.) M $^{(i)}$. Then, after training and evaluating across all k folds, the overall l-fold performance is computed as the average:

$$\bar{M} = \frac{1}{k}\sum_{i=1}^{k} M^{(i)} \quad (4)$$

This equation formalizes that the final performance metric is the average of the metric from all k folds, providing a robust estimate of the models generalization

Let $\hat{y}_j^{(i)}$ denotes the predicted label for sample x_j in fold i, and y_i the true label. Then accuracy for fold i is:

$$Accuracy^{(i)} = \frac{1}{k}\sum_{(x,y)\in D_{val}^{(i)}} 1\left(\hat{y}_j^{(i)} = y_j\right) \quad (5)$$

Where 1 is an indicator function that equals 1 of the condition is true, and 0 otherwise.
The overall k-fold accuracy is the average of the accuracy across folds:

$$Accuracy_{CV} = \frac{1}{k}\sum_{i=1}^{k} Accuracy^{(i)} \quad (6)$$

If $\hat{P}_j^{(i)} = \left[\hat{p}_{j,1}^{(i)}, \hat{p}_{j,2}^{(i)}, \ldots\ldots, \hat{p}_{j,C}^{(i)}\right]$ is the predicted probability distribution over C classes for sample in x_j fold i, and y_j is the true class label (one-hot encoded as y_j), the cross-entropy loss for fold i is:

$$CE^{(i)} = -\frac{1}{\left|D_{val}^{(i)}\right|}\sum_{(x_j,y_j)\in D_{val}^{(i)}} \sum_{C=1}^{C} j_{j,C}\, log\hat{P}_{j,C}^{(i)} \quad (7)$$

The overall k-fold cross-entropy loss is the average over folds:

$$CE_{CV} = \frac{1}{k}\sum_{i=1}^{k} CE^{(i)} \quad (8)$$

V. EXPERIMENTAL RESULTS AND COMPARATIVE ANALYSIS

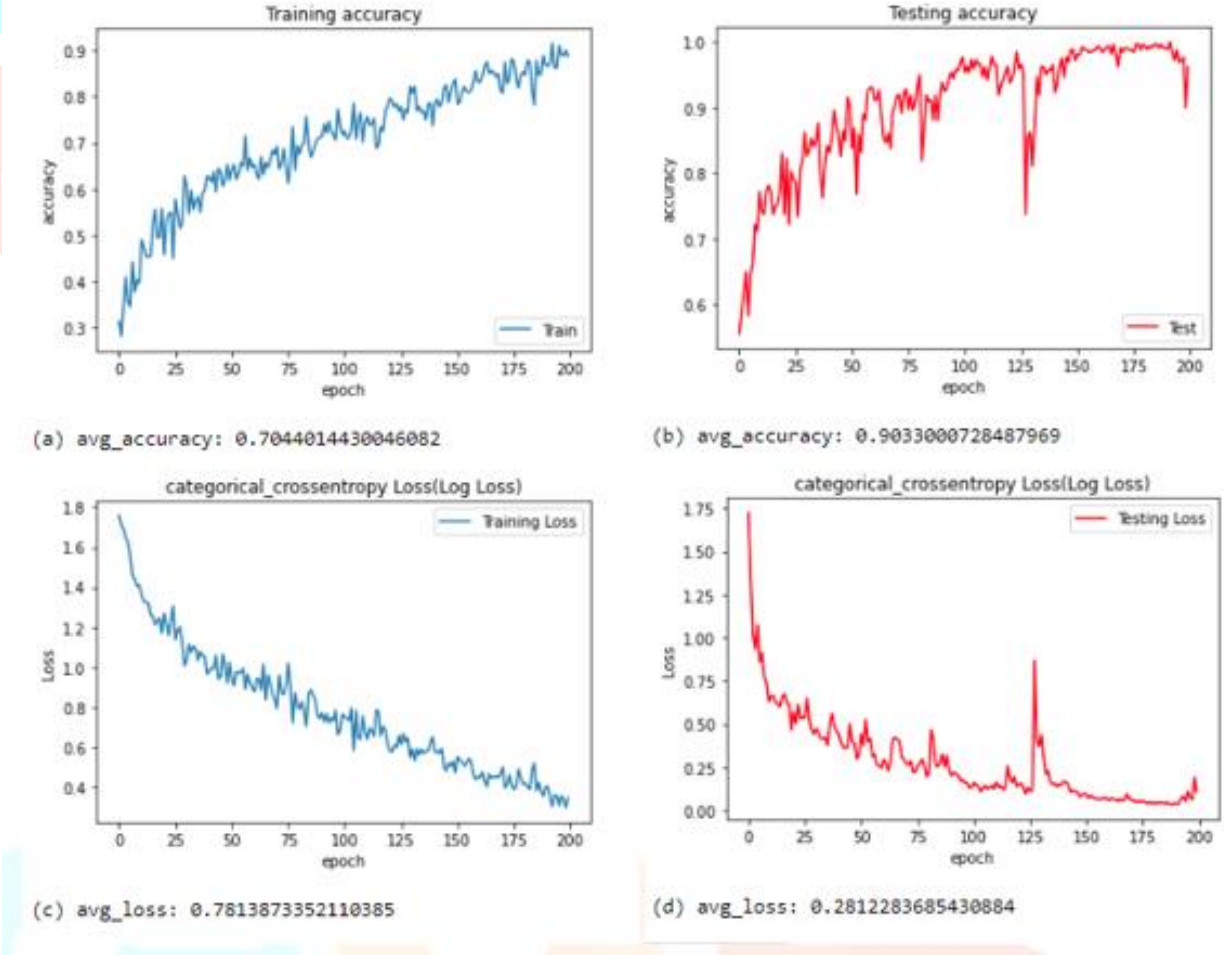

Fig.4 (a) average accuracy on training dataset (b) average accuracy on testing dataset (c) average loss on training dataset (d) average loss on testing dataset.

The performance evaluation of the proposed custom sequential convolutional neural network (CNN) model for petrological image classification is illustrated in above figure 4, training and testing metrics plotted across 200 epochs. The model was trained on a labeled dataset of petrological images, and its performance was measured using both accuracy and binary cross-entropy loss (log loss), which is appropriate for multi-class problems converted via one-hot encoding or binary labels per class. The training accuracy curve demonstrates a steady and consistent rise over the epochs, starting from a lower baseline and gradually converging towards a plateau around 85% in the final stages of training. This indicates that the model progressively learned the underlying features of the training data. The average training accuracy achieved across all epochs is approximately 70.44%, highlighting a moderate but stable learning pattern during the training phase.

In contrast, the testing accuracy shows a rapid improvement within the first 25–50 epochs and subsequently stabilizes at a high value close to 95%. The average testing accuracy, computed over the entire training cycle, is remarkably high at 90.33%. This superior performance on the validation set suggests that the model generalizes well to unseen data, effectively capturing critical features in petrological images necessary for accurate classification. The absence of major dips or oscillations in the testing accuracy curve further confirms the stability and reliability of the trained model. Analysing the loss metrics, the training loss begins at a higher value and steadily declines throughout the training period, indicating effective minimization of error through backpropagation. The average training loss is around 0.7813, which, while higher than the testing loss, reflects the inherent variability or possible noise in the training set. The corresponding testing loss curve follows a downward trend initially and flattens out at lower values, with occasional spikes observed around epochs 100 and 160. These minor fluctuations may be attributed to random batch effects or overconfidence in predictions for certain image samples. Nonetheless, the average testing loss remains low at 0.28126, reinforcing the model's robustness and minimal overfitting.

VI. CONCLUSION

Deep learning, particularly CNN-based architectures, has proven highly effective in classifying petrological thin-section images. Strategies such as dual-polarization input, patch-level learning, and integration of spectral data significantly improve accuracy and generalizability. The field is moving towards real-time, domain-adaptable systems, but further progress depends on robust datasets and interdisciplinary integration of geoscientific insights with AI innovations.

REFERENCES

[1] C. Su, S. Xu, K. Zhu, and X. Zhang, "Rock Classification in Petrographic Thin Section Images Based on Concatenated Convolutional Neural Networks," *arXiv: Image and Video Processing*. 10437, 2020. https://arxiv.org/pdf/2003.10437

[2] C. Su, S. Xu, K. Zhu, and X. Zhang, "Rock classification in petrographic thin section images based on concatenated convolutional neural networks," *Earth Science Informatics*, vol. 13, no. 4, pp. 1477–1484, 2020. https://doi.org/10.1007/S12145-020-00505-1

[3] M. Sarmad, J. Phan, L. Ruspini, G. Kiss, and F. Lindseth, "Core-scale rock typing using convolutional neural networks for reservoir characterization in the petroleum industry," 2023. https://doi.org/10.5593/sgem2023/1.1/s06.78

[4] S. Wan and T. S. Yun, "Deep learning for microscopic rock image classification," *in Informa*, 2023, pp. 882–887. https://doi.org/10.1201/9781003299127-122

[5] P. Dell'Aversana, "An Integrated Deep Learning Framework for Classification of Mineral Thin Sections and Other Geo-Data, a Tutorial," *Minerals*, vol. 13, no. 5, p. 584, 2023. https://doi.org/10.3390/min13050584

[6] H. M. El-Desoky et al., "An integrated remote sensing, petrology, and field geology analyses for Neoproterozoic basement rocks in some parts of the southern Egyptian-Nubian Shield," *Dental Science Reports*, vol. 14, no. 1, 2024. https://doi.org/10.1038/s41598-024-62093-0

[7] A. D. P. Pascual, "Autonomous and Real Time Rock Image Classification using Convolutional Neural Networks," 2019. https://doi.org/10.20944/PREPRINTS202109.0285.V1

[8] Q. Hao, N. Nie, and F. Zhu, "The Design and Realization of Intelligent Classification of Rock Samples Based on Convolutional Neural Network," *American Journal of Computer and Information Science*, vol. 5, no. 8, 2022. https://doi.org/10.25236/ajcis.2022.050801

[9] J. Cifuentes, L. Arias, E. Pirard, and F. Castillo, "Mineral classification using convolutional neural networks and SWIR hyperspectral imaging," 2024. https://doi.org/10.1117/12.3002101

[10] A. Marathe, T. Tewari, and F. Vyas, "Petrographic Image Classification Accuracy Improvement Using Improved Learning," *in Informa*, 2024, pp. 54–70. https://doi.org/10.1201/9781032699806-4

[11] Shahriar Hossain, Jahir Uddin, and Rakibul Alam Nahin. (2021). Rock Classification Dataset [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DS/1293628.