

Optimal Classification of Minerals by Microscopic Image Analysis Based on Seven-State “Deep Learning” Combined with Optimizers

Kouadio Krah^{1,2}, Sie Ouattara^{2,3,4}, Gbele Ouattara^{1,2}, Alain Clement⁵, Joseph Vangah⁶

¹Laboratoire de Génie Civil, Géosciences et Sciences Géographiques, Institut National Polytechnique Houphouët Boigny (INPHB), Yamoussoukro, Côte d'Ivoire

²Institut National Polytechnique Houphouët Boigny (INPHB), Ecole Doctorale Polytechnique (EDP), Yamoussoukro, Côte d'Ivoire

³Laboratoire des Sciences et Technologies de la Communication et de l'Information (LSTCI), Institut National Polytechnique Houphouët Boigny (INP-HB), Yamoussoukro, Côte d'Ivoire

⁴Ecole de Géomatique et du Territoire (EGT), Abidjan, Côte d'Ivoire

⁵LARIS, SFR MATHSTIC, Université d'Angers, Angers, France

⁶Université Virtuelle de Côte d'Ivoire, Abidjan, Côte d'Ivoire

Email: sie_ouat@yahoo.fr

How to cite this paper: Krah, K., Ouattara, S., Ouattara, G., Clement, A. and Vangah, J. (2024) Optimal Classification of Minerals by Microscopic Image Analysis Based on Seven-State “Deep Learning” Combined with Optimizers. *Open Journal of Applied Sciences*, **14**, 1550-1572.

<https://doi.org/10.4236/ojapps.2024.146103>

Received: May 4, 2024

Accepted: June 21, 2024

Published: June 24, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The development of artificial intelligence (AI), particularly deep learning, has made it possible to accelerate and improve the processing of data collected in different fields (commerce, medicine, surveillance or security, agriculture, etc.). Most related works use open source consistent image databases. This is the case for ImageNet reference data such as coco data, IP102, CIFAR-10, STL-10 and many others with variability representatives. The consistency of its images contributes to the spectacular results observed in its fields with *deep learning*. The application of deep learning which is making its debut in geology does not, to our knowledge, include a database of microscopic images of thin sections of open source rock minerals. In this paper, we evaluate three optimizers under the *AlexNet* architecture to check whether our acquired mineral images have object features or patterns that are clear and distinct to be extracted by a neural network. These are thin sections of magmatic rocks (biotite and 2-mica granite, granodiorite, simple granite, dolerite, charnokite and gabbros, etc.) which served as support. We use two hyper-parameters: the number of *epochs* to perform complete rounds on the entire data set and the “*learning rate*” to indicate how quickly the weights in the network will be modified during optimization. Using *Transfer Learning*, the three (3) opti-

mizers all based on the gradient descent methods of Stochastic Momentum Gradient Descent (*sgdm*), Root Mean Square Propagation (*RMSprop*) algorithm and Adaptive Estimation of moment (*Adam*) achieved better performance. The recorded results indicate that the *Momentum* optimizer achieved the best scores respectively of 96.2% with a learning step set to 10^{-3} for a fixed choice of 350 *epochs* during this variation and 96, 7% over 300 *epochs* for the same value of the learning step. This performance is expected to provide excellent insight into image quality for future studies. Then they participate in the development of an intelligent system for the identification and classification of minerals, seven (7) in total (quartz, biotite, amphibole, plagioclase, feldspar, muscovite, pyroxene) and rocks.

Keywords

Classification, Convolutional Neural Network, Deep Learning, Optimizers, Transfer Learning, Rock Mineral Images

1. Introduction

The cooling and crystallization of molten rock, often called magma, is the origin of igneous rocks [1] [2]. The planet's primarily molten origins can use magma as marking the start of the rock cycle and the origins of igneous rocks are documented in their composition [3]. Analyzing, examining and interpreting the information that its rocks carry will allow us to deduce the processes occurring in the Earth. The same information will help to understand volcanic activities on the earth's surface, as well as their socio-economic importance in mineral and/or water resources [4]. However, it would be necessary to identify the interior of these rocks. Indeed, they are made up of minerals. These minerals each have a specific crystal structure produced naturally in pure and ordered form [5] [6]. Each structure defines the texture of a rock as a whole. Physically identifying igneous rocks by texture, grain size, colors, defects and patterns is a difficult process. This task requires the use of an informed geologist with a background in rocks and also constituent minerals. The precise characterization of minerals present in rocks constitutes a fundamental step in the in-depth understanding of geological processes (identification, geological prospecting, materials science and engineering) and conditions of formation of the earth's crust [7] [8]. The act of identifying and classifying minerals and rocks can be arduous and time-consuming for people from other fields. To overcome this difficulty, geology uses the optical properties of these minerals. Said properties are observed when the image of a thin section of rock is visualized under a polarizing microscope [9] [10]. [11] used these properties to automate the identification and classification procedure of minerals by efficiently integrating color variations under plane polarization (*PPL*) and cross polarization (*XPL*) illumination modes into the *CieLab* space. This technique has the advantage of not ignoring the pixel path of mineral im-

ages rotated in this space. In recent decades, the use of machine learning methods from microscopic images boosts the limits of old practices for identifying, naming and classifying rocks [12] [13] [14]. This new approach makes the classification of rock types a booming subject in geosciences. It uses computer science and pattern recognition techniques [15]. The most characteristic visual properties for identifying minerals in rocks concern color and then texture. Since computer technologies provided consistent definitions of colors, colors have been used for several identification and classification purposes [16]. In the case of minerals, color, expression of birefringence, is of capital importance during the formation of their images. The work of [17] indicates that minerals can be identified with high accuracy if their colors and birefringence colors are not changed depending on slide thickness or lighting conditions. The authors used an artificial neural network (ANN) for mineral recognition (quartz, muscovite, biotite, chlorite and opaque minerals) via image processing. Thus, computers perform operations such as recognition [8] [18] and measurement of target objects on images using color and/or pixels. These operations replace the eyes and provide a more adapted form of object detection unlike human observation [19]. Image classification appears to be a means that helps distinguish different types of images based on the characteristics extracted from them [20] [21]. [22] proposed a CNN that chooses and extracts features from image samples to recognize the granularity of a rock. The accuracy of the model reached 98.5% but deviations were revealed and are linked to the use of single polarization images. The redundancy of information and the lack of differentiation of textural features extracted from a rock image slow down the training of a CNN and give it poor classification accuracy. [23] proposed a hybrid *DCT-CNN* method to overcome the problem. The technique reveals short training time with more accurate classification result. The study we are carrying out aims to verify the quality of the images of minerals acquired with a view to using their content through deep learning. A coherent image of rock and minerals can be fed into intelligent systems that use image processing and machine learning techniques. These techniques aim to do recognition and classification. Specifically, we must: 1) update the network weights using the *sgdm*, *Adam* and *RMSprop* optimizers to minimize the loss function, 2) study their convergence by the transfer learning method with the pre-trained parameters of the *AlexNet* model, 3) evaluate its optimizers by the performance results (score) recorded. To do this, variations in learning steps and the number of *epochs* (recursions) were used as parameters.

2. State of the Art

The classification and identification of igneous rocks and rocks in general, remains an inherent subject for geologist and geology as they contribute to the understanding of the relative mineral wealth of the rock [15]. The classification of minerals in rocks has evolved significantly over the decades. It has moved

from classical methods to more advanced approaches and is based on emerging technologies [24]. In this section, we will explore previous work. We will highlight classic methods, the first applications of machine learning, and finally, recent advances through the use of deep learning.

Classical methods of classifying minerals from rocks relied largely on manual analyzes through optical microscopy. It begins by producing a thin section in the laboratory [25]. Geologists, with the naked eye or under a microscope, observed the samples for identification of minerals on the basis of their optical properties and/or their morphology [26] [27]. With a magnifying glass (approximately $\times 10$), a texture is determined by estimation [28]. Although these methods were invaluable for their time, they sometimes had errors due to the complexity of the rock samples. It was also necessary to give yourself time, and sometimes even refer cases to an expert in order to have a more reassured decision [29].

The emergence of Machine Learning (*ML*) has made it possible to introduce automated approaches. The classification of minerals with image processing and computer vision has improved the visual observation of thin sections and the description of rocks using their images. [30] successfully classified minerals using supervised learning by combining data from three spectroscopic methods: vibrational Raman scattering, reflective visible-near infrared (*VNIR*), and laser-induced breakdown spectroscopy (*LIBS*). The results demonstrate that multi-scale spectroscopy associated with *ML* leads to rapid and precise characterization of rocks and minerals. However, the authors note that differences in the spectral data set may affect some results. [31] conducted a comparative classification study between various combinations of manual features such as first and second order statistics. These characteristics respectively describe the distribution of the intensity of the pixels in the image and the information relating to the position of the different intensities. Unsupervised learning based on the K-Means algorithm is subsequently used, in order to understand the way in which the gray levels are distributed across the pixels. They claim that while there are no guarantees of the model, the method outperformed any manual feature setup. They also learned the feature representation through self-learning based on unsupervised learning from a dataset of unlabeled rocks. Under polarized reflected light microscopy, [32] carried out the recognition of hematite grains. Good results are recorded with a better adjustment of the parameter controlling the sensitivity of the Euclidean distance of pixels in *RGB* space. [33], developed a correspondence relationship between the characteristics of an image and the type of rocks. The goal is to automatically classify rocks on the basis of their different images. This was done in different color spaces using 1000 images of rock from the Odors Basin in China in the Shaanxi province. The accuracy of the method is estimated at 95.0%. However, the authors are planning a study on the influence of the types and number of images which would make it possible to better train the network in order to improve this score. [34] following the results of their work, stated this: “this is the first time that computer vision based on machine learning algorithms has succeeded in the automated recognition of mineral

grains from digital images acquired with a simple optical microscope”. The model uses a simple linear iterative Clustering segmentation to generate super-pixels to isolate sand grains. An action impossible with traditional segmentation techniques. The limits of the model highlighted by the authors lie in the origin (regions) of the mineral grains and the light sources (plane, circular polarized, infrared, etc.) during the acquisition. For simplicity, some researchers often consider rock images as samples with two phases: porosity and minerals [35] [36] [37]. The application of this simplification is not universal and significantly influences subsequent calculations of rock physical parameters. In particular, the velocities of *P* waves (compression waves or primary waves) and *S* waves (shear waves or secondary waves) ([38] Andrä *et al.*, 2013). Some studies have used discrete cosine transform (*DCT*), local binary model (*LBP*) to index and calculate the signature of minerals by macroscopic view, as well as wide margin support (*SVM*) [39] [40]. However, these machine learning automatons are often dependent on manually extracted features. This limits their adaptability because of the diversity of minerals, their complexity of formation, the fineness of their modal structure, the similarities and even their life process in rocks. It is therefore necessary to innovate in the processing of information for complex and very varied cases or because the characteristics extracted from the image are independent of the classifier.

Recent advances in the field of artificial intelligence, in particular deep learning, have considerably transformed the scientific field of classification in medicine, security, agriculture [41] [42] [43]. Based on deep neural networks, its algorithms can learn complex features directly from microscopic images. They thus eliminate the need for manual extractions. A deep neural network (CNN) has the advantage of extracting features from the image without the need for a human to intervene manually [44]. CNNs remain the most efficient neural networks in terms of pattern recognition [45].

Regarding geology, for the first time, [21] implemented transfer learning technique to automate lithology identification and classification of rock images. They were able to effectively distinguish graphite, phyllite and breccia (various minerals from several fragments). [46], using AI, used the *Transfer Learning* method to identify minerals in the field from smartphones. The basic application is modeled on the architecture of the CNN *ShuffleNet* network. [8], thanks to an intelligent lithology identification method, based on the Faster R-CNN architecture, managed to predict lithological information and detect rock targets. [47] propose multi-class classification beyond universal CNN models. They combine deep learning and transfer learning using *VGGNet*, *InceptionNet* and *ResNet* architectures for online multi-coal and multi-class sorting. The size of images influences recognition during the learning process of certain models. The authors invite to diversify the components of the dataset in terms of rock types in order to better evaluate the scalability capacity of the classifier. Also overcoming the problem of estimating the background and sample points with transfer learning is another challenge.

3. Material and Methods

With the improvement in the performance of calculation units, CNNs with convolutional filter layers linked to an artificial neural network make it possible to identify the content of images by extraction of characteristics. These operations are carried out at the same time as the learning stage of the algorithm. This algorithm requires a large number of parameters which must be defined by the user as well as a large number of data. Fortunately, there are CNN architectures such as *AlexNet*, *GoogleNet*, *VGGNet* and others [48]. Knowing the quality of mineral images acquired using the *AlexNet* architecture and *Transfer Learning* is the major objective of this study.

In this part, we will detail how to put this approach into practice to achieve this.

3.1. Diagram of the Experimental Approach

The structure in **Figure 1** labels the main procedure of our experimental approach. First of all, we have a database of various images of minerals. These images will be preprocessed with a certain number of operations (cutting, resizing, etc.). Then divided into a training and validation data set, the images are used to train a convolutional neural network via *AlexNet*. Finally, the membership classification of each mineral image is calculated and the precision rate is recorded serving as the performance of the chosen optimizers.

3.2. Data Set and Protocol

For the needs of the project and in order to give it a particular meaning, the creation of the database began with the collection of 15 samples of magmatic rocks in the field. These samples were used to prepare 15 thin sections for each sample. Each slide was admitted under homogeneous polarized light from an MD500 microscope (using Amscope 3.7 software) on which a camera is mounted and transmitted to a computer (Intel(R) Core(TM) i5-8265U CPU @ 1.60GHz 1.80 GHz, RAM size 16.0 GB (15.8 GB usable)) to collect slide images.

3.2.1. Acquisition of the Image Database

For our research, we needed a fairly large number of images. By capture, the microscope scene images are acquired following the rotation of its stage ranging from 0 to 315 degrees in increments of 45 degrees. Images are observed at magnification ($\times 40$). The flowchart for this step is illustrated in **Figure 1**. All images (*RGB*) were stored at the size of 2592×1944 pixels with a resolution of 120 dpi, and saved in "JPG" format. The increment made is due to the fact that different minerals have different quenching properties.

3.2.2. Preprocessing and Organization of Data

The images acquired include one or more minerals at a time with different hues. The pretreatment began with the isolation of each mineral from its matrix, resulting in a resizing. The operation was done by screenshot with the *XnView*

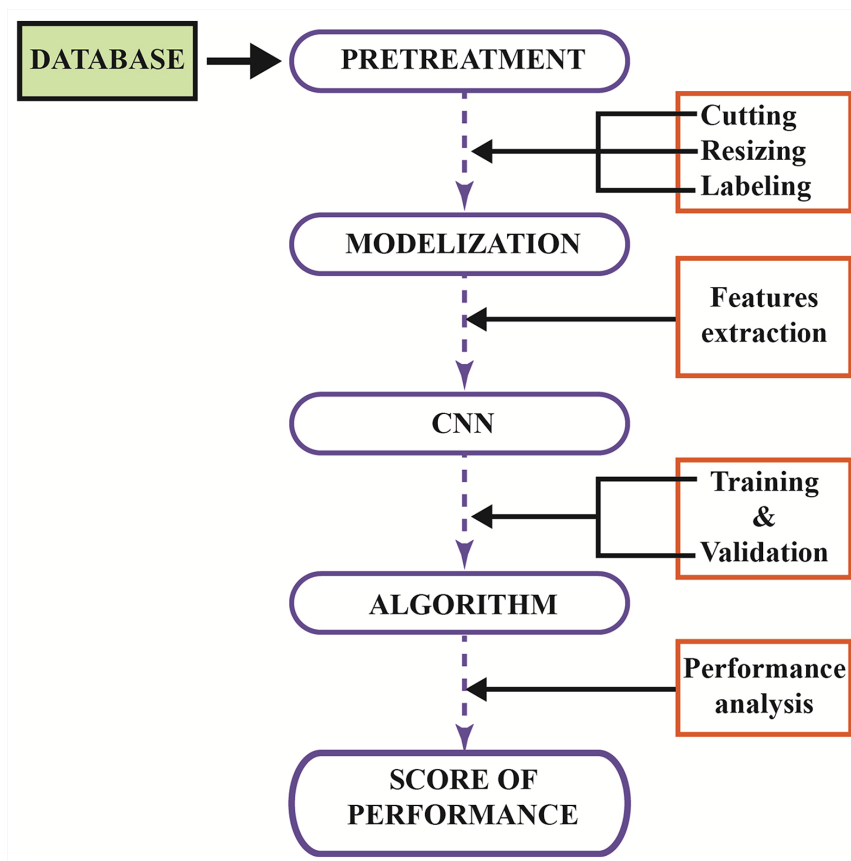


Figure 1. Structure of our work plan.

software. The images occupy on average a storage space of between 3.5 and 8 KB with each a resolution of 144 ppi (pixels per inch). The minerals dataset in this experiment is a new dataset. A total of 700 images (due to 100 images per class) of single species minerals with different hues constituted our database (**Figure 2**). The classes were chosen because of their abundance in igneous rocks and whose proportions are used to name the rocks based on the Streckeisen diagram. We distinguish: amphibole, biotite, alkali feldspar, muscovite, plagioclase, pyroxene and quartz. Feldspathoids were not taken into account in this study because they are very rare in rock formations in Ivory Coast. To reduce the model parameters, the original images were compressed to 224×224 pixels. However, each image was labeled according to its class.

The exact distribution of data by classes and types is described in **Table 1**. The number of images in the test folder has no impact on that of the training or learning divided into two (training images or train, and validation images) with the respective proportion (70% - 30%) [49]. We just need a few different images to test the prediction once processing is complete.

3.2.3. Training and Validation

After acquiring and organizing the data, it's time for training which consists of different processes.

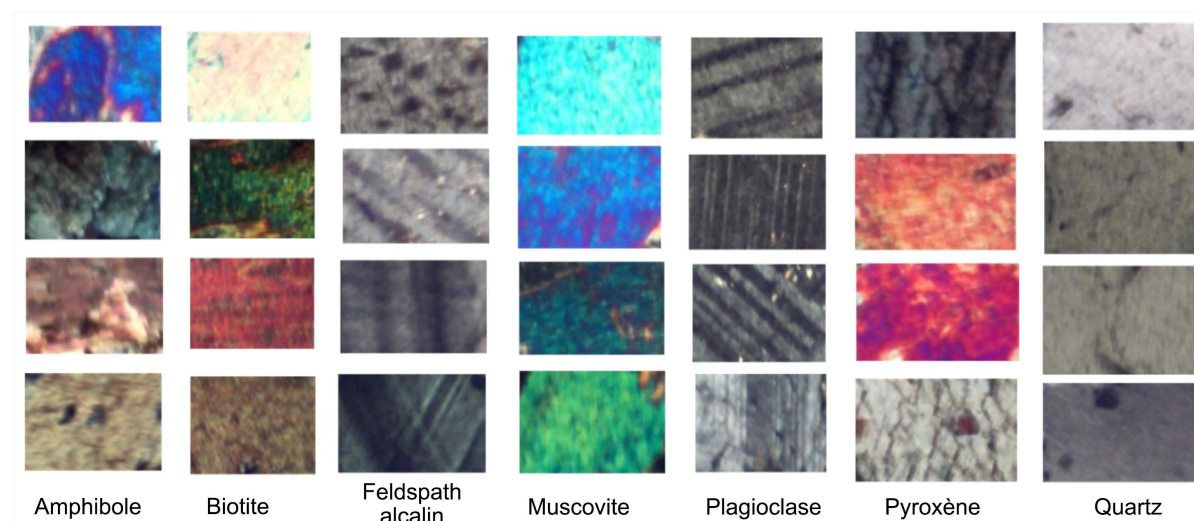


Figure 2. Images of different shades of the same mineral in analyzed polarized light (LPA).

Table 1. Distribution of mineral data by class.

	Train	Validation	Test
Amphibole	70	30	25
Biotite	70	30	25
Alkali feldspar	70	30	25
Muscovite	70	30	25
Plagioclase	70	30	25
Pyroxene	70	30	25
Quartz	70	30	25

1) Choice of model

In this work, mineral classification is implemented using a convolutional neural network with the *AlexNet* model. The choice of *AlexNet* is both a function of the quantity of our images (700 in total for the 7 classes or 100/class) and for its convolutional and fully connected layers. The latter mentioned are adjusted so as to maximize the extraction of descriptor details at the intermediate level of features. The network was developed by Alex Krizhevsky in 2012, hence its name with his collaborators Ilya Sutskever and Geoffrey Hinton [50]. Winner of the *ILSRC2012* ImageNet challenge, *AlexNet* is a CNN model with more depth and width allowing it to adapt to graphics processing units (*GPUs*) with their great potential for parallel calculations. The reasons for our choice of *AlexNet* are as follows:

- To our knowledge, no literature mentions the application of transfer learning based on *AlexNet* for the recognition of rock mineral images.
- The application of *AlexNet* in various deep learning problems shows promising results [51] [52].
- *AlexNet* is considered the first deep CNN architecture which has shown satisfactory results in image recognition and classification tasks [53].

Unlike the all-connected layers of CNNs, its convolutional layers maintain spatial coherence of information. This is the first work that propelled convolutional neural networks (CNN). It takes as input a 224×224 pixels' image with 3 color channels. Its architecture includes 8 layers including 5 convolutional and 3 fully connected (**Figure 3**).

The block of the first convolutional layers filters the input image ($224 \times 224 \times 3$) with 96 kernels of size $(11 \times 11 \times 3)$ with a step of 4 pixels. The second layer, which has the output (normalized and grouped) of the first as input, filters with 256 kernels of size $(5 \times 5 \times 48)$. The next 3 layers are connected to each other without any intermediate normalization or pooling layers. However, the third layer has 384 cores of size $(3 \times 3 \times 256)$ connected to the outputs (normalized and grouped) of the second. The fourth also has 384 cores but of size $(3 \times 3 \times 192)$. The fifth has 256 cores of size $(3 \times 3 \times 192)$. The fully connected layers have 4096 neurons each and perform global classification [55].

AlexNet promotes a fast *GPU* implementation of CNNs for image recognition with the “*sigmoid*” or *softmax* function as activation function [56]. This function acts as an output layer and calculates the output probability.

2) Optimization algorithms

Typically, in machine learning, the goal of a model is to create a prediction function $f(x)$ from a data set \mathcal{D} , also known as a set of training. In the context of supervised learning, \mathcal{D} consists of pairs of examples (x, y) , where x represents an input vector to the model and y is a target vector indicating what we are trying to predict. Regardless of how the data was acquired, it is then put through a learning algorithm that aims to model the relationship between the inputs and the targets. The inputs to a network are generally denoted x_1, x_2, \dots, x_p , acting as the explanatory variables of the model. The weights associated with these inputs are represented by the parameters α and also β which must be estimated during the learning procedure. The output of the model y , represents the variable to be explained or the target of the model, formulated as:

$$y = f(x_1, x_2, \dots, x_p; \alpha). \quad (1)$$

Learning therefore consists of estimating these parameters (α, β) by minimizing the prediction error. To measure the prediction error, we use a function called the loss function which defines the learning rate and gradient descent is a technique to do this. It updates the parameters for each example of data x and labels y ; we speak of an iterative algorithm. The lower its value, the more robust the model. If the model has correctly learned on all the data, the value of the cost function becomes zero.

a) Stochastic gradient descent momentum (sgdm) optimizers

In order to improve the efficiency of stochastic gradient descent (*SGD*) [57], the concept of momentum [58] (combination of the current direction of the gradient and from the previous direction) which in the case of optimization gives a certain “inertia” to the updating of the model parameters. Therefore, rather than moving only in the direction of the instantaneous gradient at each ite-

ration, the momentum allows a sort of “memory” of previous directions to be retained. Thus, the algorithm keeps track of the direction in which the model parameters move. This helps accelerate the convergence of the optimization by attenuating unwanted oscillations or noisy variations in the gradient descent (Figure 4).

For a loss function $J(\theta)$, the network escapes from traps using the relation using the *momentum* coefficient [58] defined by:

$$v_{t+1} = \gamma - \eta \nabla_{\theta} J(\theta_t) \quad (2)$$

where:

$\gamma \in [0,1]$ represents the *momentum* coefficient,

v_t the *momentum* vector at iteration t ,

η , the learning rate,

$\nabla_{\theta} J(\theta_t)$, the gradient of the cost function with respect to the parameters at iteration t

The complete formula for updating the parameters with *SGDM* (stochastic gradient *momentum* descent) is then:

$$\theta_{t+1} = \theta_t - v_{t+1} \quad (3)$$

where : θ represents the model parameters. Using the *momentum* coefficient can help speed up the convergence of the optimization algorithm and smooth out oscillations when training the model.

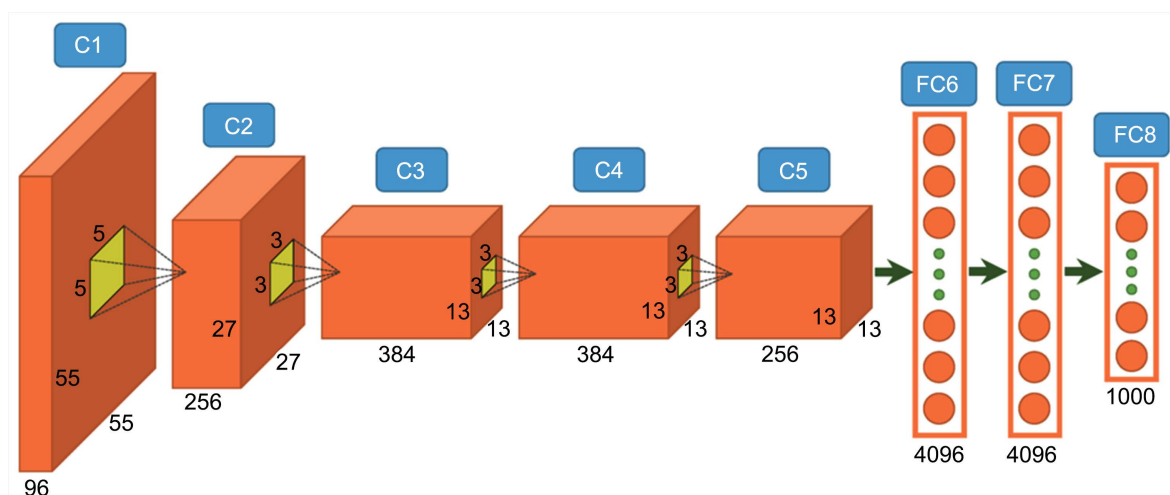


Figure 3. Example of the CNN called *AlexNet* [54].

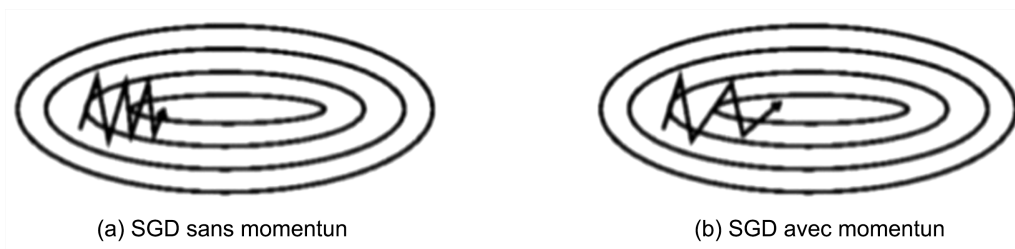


Figure 4. Acceleration and reduction of *SGD* oscillations by the *momentum* method.

b) Optimizers Root Mean Square Propagation (RMSprop) Algorithm

This optimizer was designed to address some issues related to updating *Learning rate* in gradient descent. Instead of accumulating all the squares of the previous gradients, we restrict the window of accumulated gradients to a fixed size. Therefore, we apply an exponential moving average of the squares of the gradients to automatically adapt the learning rate for each parameter of the model rather than storing the squares of the previous gradients. The parameter update (θ) with *RMSprop* is calculated as follows at each iteration t .

$$E[g^2]_t = \beta E[g^2]_{t-1} + (1 - \beta) g_t^2 \quad (4)$$

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{E[g^2]_t + \varepsilon}} \cdot g_t \quad (5)$$

where:

g_t , the gradient with respect to the parameters at iteration t

$[g^2]_t$, the exponential moving average of the squares of the gradients at iteration t

α , the *learning rate*;

β , attenuation coefficient, generally close to 1 (for example, 0.9);

ε , small constant added to avoid division by zero.

[59], propose to set β at 0.9 while a good default value for a better learning step α is 0.001.

c) Optimizers Adaptive moment estimation (Adam)

Adam, is a method which calculates a learning step qualified as adaptive for each parameter [60]. During training, in addition to storing an exponentially decreasing average of the squares of the previous gradients v_t as with *RMSprop*, it also keeps an exponentially decreasing average of the previous gradients m_t following the equations:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (6)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (7)$$

It should be noted that m_t makes it possible to update the exponential average moments of the gradient and is qualified as an order 1 estimation while v_t updates the exponential average moments of the squares of the gradient and is qualified as an order estimation 2. According to these authors, m_t and v_t are initialized as vectors of 0 and are biased around this point during the first steps, specifically when the decay coefficients (β_1 , β_2) are small (close to 1). They propose a bias correction for the estimates with the relationships:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (8)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (9)$$

Then comes the settings update (θ):

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \varepsilon}} \cdot g_t \quad (10)$$

where:

η , the learning rate;

β_1 et β_2 , attenuation coefficients (typically close to 1, for example, 0.9 and 0.999);

ε , small constant added to avoid division by zero.

For the parameters β_1 , β_2 and ε , the authors propose default values respectively 0.9, 0.999 and 10^{-8} .

3) AlexNet and Transfer Learning

Lack of knowledge of open source reference images of minerals is a disadvantage in having a sufficient quantity of images. This leaves us unable to train the AlexNet network from scratch. Our solution to this deficit takes us to work through *Transfer Learning* [61] [62] [63]. This method refers to the ability of a system to recognize and use knowledge and skills acquired in previous tasks, to apply them to new tasks, often in different domains. Thus, the influence of transfer learning follows the principle of **Figure 5**.

This change consists of modifying the classification of the 1000 classes of ImageNet into a classification of seven (7) classes. This will involve determining as outputs the presence of images of quartz, biotite, amphibole, plagioclase, feldspar, muscovite, pyroxene. The new network to obtain follows the following steps: This principle works as follows:

- First step: Initialize AlexNet settings.
- Second step: Train good performance in image recognition and classification. This by removing the last fully connected layers of AlexNet which were used to classify the 1000 classes of ImageNet.

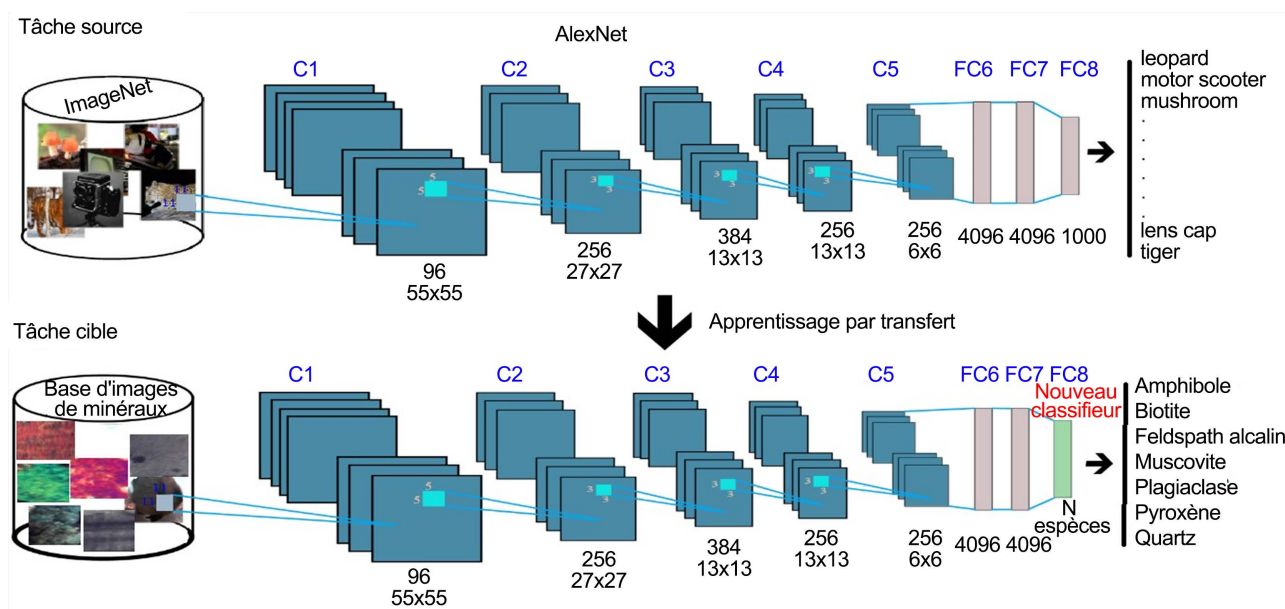


Figure 5. Principle of the transfer learning approach.

- Third step: Freeze all the weights of the pre-trained layers in order to perform the adaptation used to solve the new classification problem.
- Fourth step: The extension of *AlexNet* is done on the new model by adding a new classifier next to the output layer. Thus, all layers of the new model are trained on the target data and then the sigmoid function finds the probability that the data belongs to a class.

In each configuration, the parameters of the new model are optimized using the grid search method. This method uses two hyper-parameters: the number of *epochs* and the “Learning rate”. The grid method consists of defining a certain number of values (or grid of values) for each hyper-parameter [64]. The metric chosen to evaluate the performance of the said model is precision. However, the “loss” curve can be used (Figure 6 (curve in orange)). The combination that provides the best performance is then selected to certify the quality of the images.

4. Experimentation and Analysis of Results

4.1. Identification of Rock Minerals Using *AlexNet*

Trained to classify images into 1000 ImageNet categories, we need to adjust the output layer of the *AlexNet* pre-trained network to match the number of classes in our work. This leads to replacing the last dense layer with a new dense layer with the appropriate number of neurons for our case. Through transfer learning and data augmentation operations, the *AlexNet* pre-trained network will optimize its parameters to solve the classification of rock mineral images in the database. In this work, to prevent the size of the images from affecting the classification results [65], each image is compressed to [224 × 224 pixels] in order to fully assess the accuracy. The *ReLU* activation function is used to introduce non-linearity into the model, facilitating the learning of complex relationships between rock minerals. At the output of the network, *Softmax* is used specifically to convert scores into class probabilities.

4.2. Model Evaluation

Deep learning with its networks with complex architectures as well as its important training parameters, which by using transfer learning freezes some of its layers to a certain point requires evaluation. The evaluation phase will therefore determine what is correct in the model. A way to quantify which predictions is correct. The performance of a model results from the confusion matrix (Table 2) presented by the model including the metrics: *precision*, *recall*, *accuracy* and *trade-off* between *precision* and *recall* (*F1-score*).

For a confusion matrix the metrics raised are determined as follows:

$$Précision = \frac{TP}{TP + FP} \quad (11)$$

$$Rappel = \frac{TP}{TP + FN} \quad (12)$$

$$Accuracy = \frac{TP + TN}{TP + TN + PF + FN} \quad (13)$$

$$F1 - mesure = \frac{2 * (Précision * Rappel)}{Précision + Rappel} \quad (14)$$

With TP , TN , FP , FN respectively true positive, true negative, false positive and false negative.

4.3. Training and Results

This section of this article details the results after training our *AlexNet* model on image of rock minerals. Three optimizers were selected in order to understand and know their mode of operation by varying the *Learning rate* for a series of experiments and then doing the same for different *epochs* on our dataset. For a method, the learning step remains a very sensitive parameter for its convergence [66]. **Figure 6** shows an example of the curve obtained after training. At this step, the learning step is executed for N iterations, thus updating the parameters N times using N examples from the dataset. However, an epoch constitutes T updates, that is to say the equivalent of the entire dataset intended for this phase.

The results of each variation linked to each optimizer are recorded in **Table 3(a)**, **Table 3(b)**, **Table 4(a)**, **Table 4(b)**, **Table 5(a)** and **Table 5(b)**. The proportions of data for each training and validation set are respectively 70% - 30% and remain the same in each case.

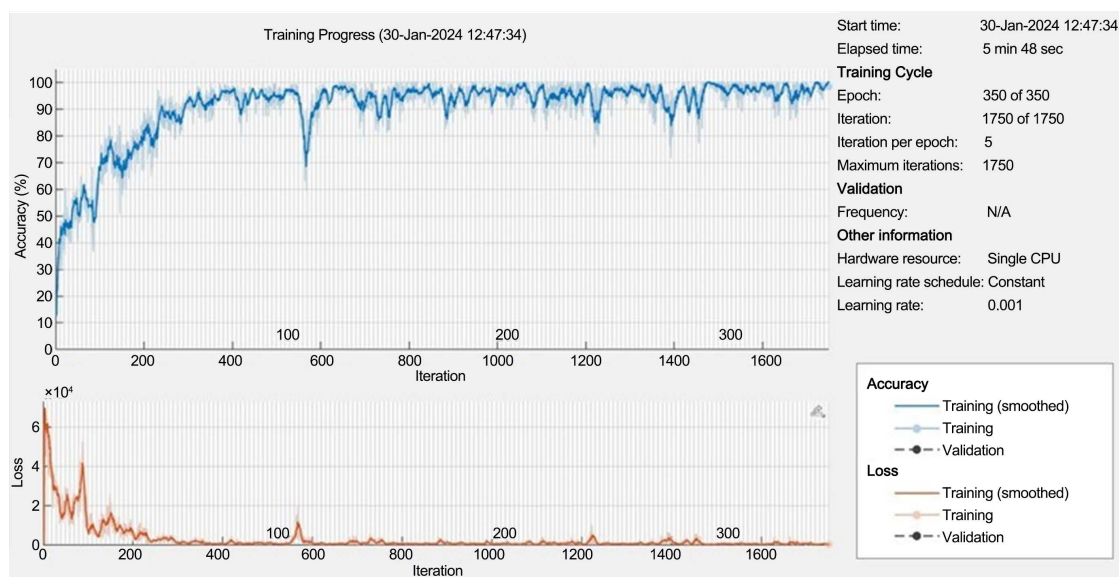


Figure 6. Training to test the network on the database.

Table 2. Matrix of confusion.

		Classes prédites	
		Classe 0	Classe 1
Classes réelles	Classe 0	TP	TN
	Classe 1	FP	FN

Table 3. (a) Results of variations in the *Learning rate* (LR), with *Epoch* set to 350; (b) results of *Epoch* variations where the *Learning rate* (LR) is set at 10^{-3} .

(a)								
<i>Learning rate</i> (LR)	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}		
Score (%)	93.3	94.4	95.2	96.2	92.9	82.4		
(b)								
<i>Epoch</i>	50	100	150	200	250	300	350	400
Score (%)	84.8	90	91.4	92.9	96.2	96.7	96.2	91.9

Table 4. (a) Results of *Learning rate* (LR) variations, with *Epoch* set to 350; (b) results of *Epoch* variations where the *Learning rate* (LR) is set at 10^{-1} .

(a)								
<i>Learning rate</i> (LR)	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}		
Score (%)	86.7	92.4	89.5	93.8	94.4	95.2		
(b)								
<i>Epoch</i>	50	100	150	200	250	300	350	400
Score (%)	76.2	82.9	89.0	93.8	89.0	92.4	82.4	81.4

Table 5. (a) Results of *Learning rate* (LR) variations, with *Epoch* set to 350; (b) results of *Epoch* variations where the *Learning rate* (LR) is set at 10^{-1} .

(a)								
<i>Learning rate</i> (LR)	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}		
Score (%)	78.1	79.0	89.5	74.3	76.7	75.2		
(b)								
<i>Epoch</i>	50	100	150	200	250	300	350	400
Score (%)	41.9	55.7	66.2	70.5	71.9	75.2	84.3	74.3

4.3.1. Results of Variations of Learning Rate and Epoch with Sgdm

Table 3(a) and **Table 3(b)** describe the results obtained with the “sgdm” optimizer with the *AlexNet* architecture using images of rock minerals using *Transfer Learning*. Part **(3a)** describes the accuracy scores for different *Learning rate* values. The highest score is 96.2% at the learning rate of value 10^{-3} . As for **Table 3(b)**, we record scores in two phases. The first phase has the same evolution as the values of the grid *epochs*. The step is set to 10^{-3} during the experiments. The highest score recorded is 96.7%. Subsequently, the second phase is positioned with a drop in precision values.

The stochastic gradient descent with momentum (*sgdm*) optimizer recorded its lowest score value (82.4%) for an estimated learning rate of 10^{-1} and 84.8% when the algorithm went through the entire training data set 50 times. During the different types of training to evaluate *sgdm*, with the different Learning Rate values. The curve of the loss function or “Loss” (curve shown in orange) com-

pletely begins to converge around 100 *epochs*, or around 400 iterations.

4.3.2. Results of Learning Rate and Epoch Variations with Adam

Table 4(a) and **Table 4(b)** show the results obtained with the “*Adam*” optimizer, still using *Transfer Learning*. Part (4a), however, describes the accuracy scores for different *learning rate* values. The highest score value is 95.2% at the *learning rate* value 10^{-1} .

As for **Table 4(b)**, we note the scores also corresponding to different *epochs* with a *learning rate*, this time maintained at 10^{-1} during the experimental phase. The highest score is 93.8% after 200 *epochs*.

The *Adam* optimizer recorded its lowest score values, which are 86.7% with the lowest learning rate value set to 10^{-6} and 76.2% when the algorithm performed 50 *epochs* on the entire training data.

4.3.3. Results of “Learning Rate” and “Epoch” Variations with RMSprop

Table 5(a) and **Table 5(b)** show the results obtained with the “*RMSprop*” optimizer. Part (5a) shows the model accuracy scores for different learning step values. The highest score with *RMSprop* is 89.5% for 10^{-4} as the optimal value of the learning step. As for **Table 5(b)**, the learning rate was maintained at 10^{-4} and we record the scores corresponding to the *epochs* proposed for the experiment. The highest score is 84.3% at 350 *epochs*.

This optimizer presents its lowest score value which is 75.2% with the high value of the *Learning rate* estimated at 10^{-3} . Unlike the other *epochs*, the logic is respected, *i.e.* a low score of 41.9% at the start of training and which increases as the *epochs* also increase before making a learning jump after 350.

5. Discussion

The performance results of the three optimizers from the point of view of accuracies are approximately more than 80% with regard to the “*Learning rate*” hyper-parameter. This shows that optimization algorithms such as gradient descent use the “*Learning rate*” as a scalar to determine the step size (modification made to the model parameters) at each iteration in order to tend towards the minimum of the loss function in order to achieve a good score [67] (Murphy, 2012). It is defined as an adjustable hyper-parameter and influences the result of a learning model. It metaphorically symbolizes how quickly a pre-trained learning model assimilates new information [68]. Thus, the average precision value achieved is considered a significant performance in image classification tasks. This observation reveals that all of the image data acquired has lower noise levels [69]. However, the best accuracy is obtained by setting the optimizer to the Momentum algorithm. The step value (10^{-3}) of this optimizer that resulted in its highest performance value also achieved the best performance at the expense of the performance values of the other two optimizers.

However, setting a learning rate is sometimes problematic, in the sense that for a high step value, the learning will undergo a jump and will be placed above

the minimums (unstable convergence). This hypothesis is verified in our case with the learning rate values at 10^{-2} and 10^{-1} of respective scores 92.9% and 82.4% for the optimizer “*sgdm*” after its limit at 10^{-3} . Conversely, for a low learning step value, learning will take too long to converge since the risks of being stuck in a *local minimum* are significant [70]. [71] points out that for these cases, there is a limit from which the error stops decreasing and begins to increase inversely to the score. Concerning the “Adam” optimizer which is a combination of *RMSprop* and *Momentum (sgdm)*, its highest score is in line with the highest value of the “*Learning rate*”. And since a higher *Learning rate* value can allow a model to move more quickly through the parameter space and explore different regions of the search space for better model generalization, then images contain functionalities usable by CNNs.

For the results related to the number of “*Epochs*”, the “*sgdm*” and “*Adam*” optimizers recorded precision performances around an average of more than 80% as well. These scores confirm the performance of CNNs for Image Classification tasks and giving credit a second time to the quality of our images. It is again “*sgdm*” which again achieves the best precision score (96.60%) after each image in the database has been observed by the model 300 times (300 “*Epochs*”). However, a growth in scores is observed for each optimizer. This is justified by the progressive initialization of the training weights from the pre-trained model.

Unlike the other two optimizers, “*RMSprop*” starts its precision performance around 40%, or half for the start of the other two. This weakness in precision demonstrates the idea that “*RMSprop*” does not accumulate all the squares of the previous gradients. The algorithm operates by applying an exponential moving average of the squares of the previous gradients. The current average on the precision at step t depends only on the previous average and the current gradient. This means that the initialization of the network weights makes the model ineffective while it takes time to update the new data. Some authors have worked to improve this aspect of initialization [72] [73]. However, the algorithm ended up reaching the significant value of CNNs for image classification at 350 “*Epochs*” [74]. This result reassures us of the quality of the images.

6. Conclusions

Image quality is of great importance in developing an intelligent identification and classification system. The consistency of mineral images allows better features to be extracted to analyze, examine and interpret certain information that its rocks carry. Which information could be useful to industry, engineering, academics and many others dealing with the field of geoscience.

In this article, we compared the accuracy-related performance results of 3 optimizers for rock mineral image quality. The approach applied *Transfer Learning* based on the architecture of the CNN *AlexNet* model. The idea is to show that at the limit of our data, the simplest architecture of *AlexNet* can perform better when implemented with *Transfer Learning* and optimized model parameters.

The results of this study show that *Transfer Learning* can be useful to improve the robustness of the model on new data. Furthermore, it is beneficial when the number of training images is small. In our case, this was possible thanks to the use of feature extractors from the pre-trained model and the limitation of the training to only the classification layers (fully connected layers). This approach was beneficial to us instead of using all the parameters of the pre-trained model.

The proposed method can therefore provide satisfactory quality of images of rock minerals. The best precision is recorded with the optimizer using the “*sgdm*” algorithm applied to the two hyper-parameters. The scores are 96.7% and 96.20% respectively for the number of “*Epochs*” and the “*Learning rate*”.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this article.

References

- [1] MacKenzie, W.S., Donaldson, C.H. and Guilford, C. (1982) Atlas of Igneous Rocks and Their Textures, vol. 148. Longman Harlow.
- [2] Cox, K.G. (2013) The Interpretation of Igneous Rocks. Springer Science & Business Media.
- [3] Egger, A.E. (2005) “Defining Minerals”. Vision Learning, Recherche Google.
- [4] Gill, R. and Fitton, G. (2022) Igneous Rocks and Processes: A Practical Guide. John Wiley & Sons.
- [5] Rafferty, J. (2012) Geology: Landform, Minerals and Rocks. Britannica Educational Publishing, p. 358.
- [6] Pellant, C. and Pellant, H. (2021) Rocks and Minerals. Dorling Kindersley Ltd.
- [7] Lin, P., Yu, T., Xu, Z., Shao, R. and Wang, W. (2022) Geochemical, Mineralogical, and Microstructural Characteristics of Fault Rocks and Their Impact on TBM Jamming: A Case Study. *Bulletin of Engineering Geology and the Environment*, **81**, Article No. 64. <https://doi.org/10.1007/s10064-021-02548-0>
- [8] Xu, Z., Ma, W., Lin, P., Shi, H., Pan, D. and Liu, T. (2021) Deep Learning of Rock Images for Intelligent Lithology Identification. *Computers & Geosciences*, **154**, Article 104799. <https://doi.org/10.1016/j.cageo.2021.104799>
- [9] Izadi, H., Sadri, J. and Bayati, M. (2017) An Intelligent System for Mineral Identification in Thin Sections Based on a Cascade Approach. *Computers & Geosciences*, **99**, 37-49. <https://doi.org/10.1016/j.cageo.2016.10.010>
- [10] Liu, H., et al. (2022) Rock Thin-Section Analysis and Identification Based on Artificial Intelligent Technique. *Petroleum Science*, **19**, 1605-1621. <https://doi.org/10.1016/j.petsci.2022.03.011>
- [11] Aligholi, S., Lashkaripour, G.R., Khajavi, R. and Razmara, M. (2017) Automatic Mineral Identification Using Color Tracking. *Pattern Recognition*, **65**, 164-174. <https://doi.org/10.1016/j.patcog.2016.12.012>
- [12] Marmo, R., Amodio, S., Tagliaferri, R., Ferreri, V. and Longo, G. (2005) Textural Identification of Carbonate Rocks by Image Processing and Neural Network: Methodology Proposal and Examples. *Computers & Geosciences*, **31**, 649-659. <https://doi.org/10.1016/j.cageo.2004.11.016>

- [13] Singh, N., Singh, T.N., Tiwary, A. and Sarkar, K.M. (2010) Textural Identification of Basaltic Rock Mass Using Image Processing and Neural Network. *Computers & Geosciences*, **14**, 301-310. <https://doi.org/10.1007/s10596-009-9154-x>
- [14] Harinie, T., Janani Chellam, I., Sathya Bama, S.B., Raju, S. and Abhaikumar, V. (2012) Classification of Rock Textures. *Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2012 (India 2012) Held in Visakhapatnam, India, January 2012*, Visakhapatnam, January 2012, 887-895. https://doi.org/10.1007/978-3-642-27443-5_102
- [15] Młynarczyk, M., Górszczyk, A. and Ślipek, B. (2013) The Application of Pattern Recognition in the Automatic Classification of Microscopic Rock Images. *Computers & Geosciences*, **60**, 126-133. <https://doi.org/10.1016/j.cageo.2013.07.015>
- [16] Gökyay, M.K. and Gundogdu, I.B. (2008) Color Identification of Some Turkish Marbles. *Construction and Building Materials*, **22**, 1342-1349. <https://doi.org/10.1016/j.conbuildmat.2007.04.016>
- [17] Baykan, N.A. and Yılmaz, N. (2010) Mineral Identification Using Color Spaces and Artificial Neural Networks. *Computers & Geosciences*, **36**, 91-97. <https://doi.org/10.1016/j.cageo.2009.04.009>
- [18] Perez, C.A., Saravia, J.A., Navarro, C.F., Schulz, D.A., Aravena, C.M. and Galdames, F.J. (2015) Rock Lithological Classification Using Multi-Scale Gabor Features from Sub-Images, and Voting with Rock Contour Information. *International Journal of Mineral Processing*, **144**, 56-64. <https://doi.org/10.1016/j.minpro.2015.09.015>
- [19] Lepistö, L., Kunttu, I. and Visa, A.J.E. (2005) Rock Image Classification Using Color Features in Gabor Space. *Journal of Electronic Imaging*, **14**, Article 040503. <https://doi.org/10.1117/1.2149872>
- [20] Foody, G.M. and Mathur, A. (2004) A Relative Evaluation of Multiclass Image Classification by Support Vector Machines. *IEEE Transactions on Geoscience and Remote Sensing*, **42**, 1335-1343. <https://doi.org/10.1109/TGRS.2004.827257>
- [21] Zhang, Y., Li, M. and Han, S. (2018) Automatic Identification and Classification in Lithology Based on Deep Learning in Rock Images. *Yanshi Xuebao Acta Petrologica Sinica*, **34**, 333-342.
- [22] Cheng, G. and Guo, W. (2017) Rock Images Classification by Using Deep Convolution Neural Network. *Journal of Physics: Conference Series*, **887**, Article 012089. <https://doi.org/10.1088/1742-6596/887/1/012089>
- [23] Li, Y., Shi, D. and Bu, F. (2019) Automatic Recognition of Rock Images Based on Convolutional Neural Network and Discrete Cosine Transform. *Traitement du Signal*, **36**, 463-469. <https://doi.org/10.18280/ts.360512>
- [24] Fan, G., Chen, F., Chen, D. and Dong, Y. (2020) Recognizing Multiple Types of Rocks Quickly and Accurately Based on Lightweight CNNs Model. *IEEE Access*, **8**, 55269-55278. <https://doi.org/10.1109/ACCESS.2020.2982017>
- [25] Karimpouli, S. and Tahmasebi, P. (2019) Segmentation of Digital Rock Images Using Deep Convolutional Autoencoder Networks. *Computers & Geosciences*, **126**, 142-150. <https://doi.org/10.1016/j.cageo.2019.02.003>
- [26] Pichler, H. and Schmitt-Riegraf, C. (1997) Rock-Forming Minerals in Thin Section. Springer Science & Business Media. <https://doi.org/10.1007/978-94-009-1443-8>
- [27] Massonne, H.-J., Bernhardt, H.-J., Dettmar, D., Kessler, E., Medenbach, O. and Westphal, T. (1998) Simple Identification and Quantification of Microdiamonds in Rock Thin-Sections. *European Journal of Mineralogy*, **10**, 497-504. <https://doi.org/10.1127/ejm/10/3/0497>

- [28] Pappalardo, G., Punturo, R., Mineo, S., Ortolano, G. and Castelli, F. (2015) Engineering Geological and Petrographic Characterization of Migmatites Belonging to the Calabria-Peloritani Orogen (Southern Italy). *Rock Mechanics and Rock Engineering*, **49**, 1143-1160. <https://doi.org/10.1007/s00603-015-0808-9>
- [29] Song, Y.-Q., Ryu, S. and Sen, P.N. (2000) Determining Multiple Length Scales in Rocks. *Nature*, **406**, 178-181. <https://doi.org/10.1038/35018057>
- [30] Jahoda, P., Drozdovskiy, I., Payler, S.J., Turchi, L., Bessone, L. and Sauro, F. (2021) Machine Learning for Recognizing Minerals from Multispectral Data. *The Analyst*, **146**, 184-195. <https://doi.org/10.1039/d0an01483d>
- [31] Shu, L., McIsaac, K., Osinski, G.R. and Francis, R. (2017) Unsupervised Feature Learning for Autonomous Rock Image Classification. *Computers & Geosciences*, **106**, 10-17. <https://doi.org/10.1016/j.cageo.2017.05.010>
- [32] Iglesias, J.C.A., da Fonseca Martins Gomes, O. and Paciornik, S. (2011) Automatic Recognition of Hematite Grains under Polarized Reflected Light Microscopy through Image Analysis. *Minerals Engineering*, **24**, 1264-1270. <https://doi.org/10.1016/j.mineng.2011.04.015>
- [33] Ye, L., Chao, G. and Guojian, C. (2014) Rock Classification Based on Images Color Spaces and Artificial Neural Network. 2014 *Fifth International Conference on Intelligent Systems Design and Engineering Applications*, Hunan, 15-16 June 2014, 897-900. <https://doi.org/10.1109/isdea.2014.199>
- [34] Maitre, J., Bouchard, K. and Bédard, L.P. (2019) Mineral Grains Recognition Using Computer Vision and Machine Learning. *Computers & Geosciences*, **130**, 84-93. <https://doi.org/10.1016/j.cageo.2019.05.009>
- [35] Fattahi, H. and Karimpouli, S. (2016) Prediction of Porosity and Water Saturation Using Pre-Stack Seismic Attributes: A Comparison of Bayesian Inversion and Computational Intelligence Methods. *Computational Geosciences*, **20**, 1075-1094. <https://doi.org/10.1007/s10596-016-9577-0>
- [36] Karimpouli, S., Khoshlesan, S., Saenger, E.H. and Koochi, H.H. (2018) Application of Alternative Digital Rock Physics Methods in a Real Case Study: A Challenge between Clean and Cemented Samples. *Geophysical Prospecting*, **66**, 767-783. <https://doi.org/10.1111/1365-2478.12611>
- [37] Karimpouli, S. and Fattahi, H. (2016) Estimation of P- and S-Wave Impedances Using Bayesian Inversion and Adaptive Neuro-Fuzzy Inference System from a Carbonate Reservoir in Iran. *Neural Computing and Applications*, **29**, 1059-1072. <https://doi.org/10.1007/s00521-016-2636-6>
- [38] Andrä, H., Combaret, N., Dvorkin, J., Glatt, E., Han, J., Kabel, M., et al. (2013) Digital Rock Physics Benchmarks—Part I: Imaging and Segmentation. *Computers & Geosciences*, **50**, 25-32. <https://doi.org/10.1016/j.cageo.2012.09.005>
- [39] Chang, C.-C. and Lin, C.-J. (2011) LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*, **2**, Article No. 27. <https://doi.org/10.1145/1961189.1961199>
- [40] Vangah, W.J. (2020) Reconnaissance automatique des roches basée sur la représentation parcimonieuse des signaux combinée aux descripteurs statistiques et fréquents de texture couleur. Institut de National Polytechnique Félix Houphouët Boigny de Yamoussoukro.
- [41] Alico, N.J. (2021) Caractérisation automatique des anémies basée sur la morphologie et la couleur des érythrocytes par traitement d'images microscopiques. Institut National Polytechnique Félix Houphouët-Boigny de Yamoussoukro.
- [42] Fofana, T., Ouattara, S. and Clement, A. (2021) Optimal Flame Detection of Fires in

- Videos Based on Deep Learning and the Use of Various Optimizers. *Open Journal of Applied Sciences*, **11**, 1240-1255. <https://doi.org/10.4236/ojapps.2021.1111094>
- [43] Coulibaly, S. (2021) Analyse intelligente des images pour la surveillance dans une agriculture de precision. Ph.D. Thesis, Institut National Polytechnique de Toulouse.
- [44] Kriegeskorte, N. (2015) Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing. *Annual Review of Vision Science*, **1**, 417-446. <https://doi.org/10.1146/annurev-vision-082114-035447>
- [45] Albawi, S., Mohammed, T.A. and Al-Zawi, S. (2017) Understanding of a Convolutional Neural Network. 2017 *International Conference on Engineering and Technology (ICET)*, Antalya, 21-23 August 2017, 1-6. <https://doi.org/10.1109/icengtechnol.2017.8308186>
- [46] Fan, G., Chen, F., Chen, D., Li, Y. and Dong, Y. (2020) A Deep Learning Model for Quick and Accurate Rock Recognition with Smartphones. *Mobile Information Systems*, **2020**, Article 7462524. <https://doi.org/10.1155/2020/7462524>
- [47] Liu, Y., Zhang, Z., Liu, X., Wang, L. and Xia, X. (2021) Deep Learning-Based Image Classification for Online Multi-Coal and Multi-Class Sorting. *Computers & Geosciences*, **157**, Article 104922. <https://doi.org/10.1016/j.cageo.2021.104922>
- [48] Long, J., Shelhamer, E. and Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 3431-3440. <https://doi.org/10.1109/cvpr.2015.7298965>
- [49] Ivanov, S. (2017) Reasons Why Your Neural Network Is not Working. Retrieved June, 20, 2020. https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides lec6.pdf
- [50] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90. <https://doi.org/10.1145/3065386>
- [51] Almisreb, A.A., Jamil, N. and Din, N.M. (2018). Utilizing *Alexnet* Deep Transfer Learning for Ear Recognition. 2018 *Fourth International Conference on Information Retrieval and Knowledge Management (CAMP)*, Kota Kinabalu, 26-28 March 2018, 1-5. <https://doi.org/10.1109/infrkm.2018.8464769>
- [52] Alom, M.Z., et al. (2018) The History Began from *AlexNet*: A Comprehensive Survey on Deep Learning Approaches. arXiv: 1803.01164. <https://doi.org/10.48550/arXiv.1803.01164>
- [53] Khan, A., Sohail, A., Zahoora, U. and Qureshi, A.S. (2020) A Survey of the Recent Architectures of Deep Convolutional Neural Networks. *Artificial Intelligence Review*, **53**, 5455-5516. <https://doi.org/10.1007/s10462-020-09825-6>
- [54] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) Imagenet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, Curran Associates, Inc.
- [55] Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R.R. (2012) Improving Neural Networks by Preventing Co-Adaptation of Feature Detectors. arXiv: 1207.0580. <https://doi.org/10.48550/arXiv.1207.0580>
- [56] Deng, J., Dong, W., Socher, R., Li, L., Li, K. and Li, F.-F. (2009) Imagenet: A Large-Scale Hierarchical Image Database. 2009 *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, 20-25 June 2009, 248-255. <https://doi.org/10.1109/cvpr.2009.5206848>
- [57] Rumelhart, D.E. and McClelland, J.L. (1986) Parallel Distributed Processing. The

- MIT Press. <https://doi.org/10.7551/mitpress/5236.001.0001>
- [58] Qian, N. (1999) On the Momentum Term in Gradient Descent Learning Algorithms. *Neural Networks*, **12**, 145-151. [https://doi.org/10.1016/s0893-6080\(98\)00116-6](https://doi.org/10.1016/s0893-6080(98)00116-6)
- [59] Hinton, G., Srivastava, N. and Swersky, K. (2012) Neural Networks for Machine Learning Lecture 6a Overview of Mini-Batch Gradient Descent. https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf
- [60] Kingma, D.P. and Ba, J. (2017) Adam: A Method for Stochastic Optimization. arXiv: 1412.6980. <https://doi.org/10.48550/arXiv.1412.6980>
- [61] Pan, S.J. and Yang, Q. (2010) A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, **22**, 1345-1359. <https://doi.org/10.1109/tkde.2009.191>
- [62] Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C. and Liu, C. (2018) A Survey on Deep Transfer Learning. *Artificial Neural Networks and Machine Learning-ICANN 2018*, Rhodes, 4-7 October 2018, 270-279. https://doi.org/10.1007/978-3-030-01424-7_27
- [63] Agarwal, N., Sondhi, A., Chopra, K. and Singh, G. (2021) Transfer Learning: Survey and Classification. In: Tiwari, S., Trivedi, M., Mishra, K., Misra, A., Kumar, K. and Suryani, E., Eds., *Smart Innovations in Communication and Computational Sciences*, Springer, 145-155. https://doi.org/10.1007/978-981-15-5345-5_13
- [64] Thalagala, S. and Walgampaya, C. (2021) Application of Alexnet Convolutional Neural Network Architecture-Based Transfer Learning for Automated Recognition of Casting Surface Defects. 2021 *International Research Conference on Smart Computing and Systems Engineering (SCSE)*, Colombo, 16 September 2021, 129-136. <https://doi.org/10.1109/scse53661.2021.9568315>
- [65] Zawadzka-Gosk, E., Wołk, K. and Czarnowski, W. (2019) Deep Learning in State-of-the-Art Image Classification Exceeding 99% Accuracy. In: Rocha, Á., Adeli, H., Reis, L. and Costanzo, S., Eds., *New Knowledge in Information Systems and Technologies*, Springer, 946-957. https://doi.org/10.1007/978-3-030-16181-1_89
- [66] Bottou, L. (2012) Stochastic Gradient Descent Tricks. In: Montavon, G., Orr, G.B. and Müller, KR., Eds., *Neural Networks. Tricks of the Trade*, Springer, 421-436. https://doi.org/10.1007/978-3-642-35289-8_25
- [67] Murphy, K.P. (2012) *Machine Learning: A Probabilistic Perspective*. MIT Press.
- [68] Zulkifli, H. (2018) Understanding Learning Rates and How It Improves Performance in Deep Learning. *Towards Data Science*, **21**. <https://towardsdatascience.com/understanding-learning-rates-and-how-it-improves-performance-in-deep-learning-d0d4059c1c10>
- [69] Abbasi Aghamaleki, J. and Moayed Baharlou, S. (2018) Transfer Learning Approach for Classification and Noise Reduction on Noisy Web Data. *Expert Systems with Applications*, **105**, 221-232. <https://doi.org/10.1016/j.eswa.2018.03.042>
- [70] Buduma, N. and Locascio, N. (2017) *Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms*. O'Reilly Media.
- [71] Smith, L.N. (2017) Cyclical Learning Rates for Training Neural Networks. 2017 *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Santa Rosa, 24-31 March 2017, 464-472. <https://doi.org/10.1109/WACV.2017.58>
- [72] Glorot, X. and Bengio, Y. (2010) Understanding the Difficulty of Training Deep Feedforward Neural Networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, Sardinia, 13-15 May 2010, 249-256.
- [73] He, K., Zhang, X., Ren, S. and Sun, J. (2015) Delving Deep into Rectifiers: Surpass-

ing Human-Level Performance on Imagenet Classification. 2015 *IEEE International Conference on Computer Vision (ICCV)*, Santiago, 7-13 December 2015, 1026-1034. <https://doi.org/10.1109/iccv.2015.123>

- [74] Bergstra, J. and Bengio, Y. (2012) Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research*, **13**, 281-305.