❐   289

Iterative improved learning algorithm for petrographic image classification accuracy enhancement

Ashutosh Marathe, Priya Jain, Vibha Vyas
Department of Electronics & Tele-communication Engineering, College of Engineering, India

Article Info	ABSTRACT
	Rock image classification using image processing has been practiced to assist trained geologists in decision making. However, the study of microstructures of rocks and their use in geological investigations offer challenges in the areas of Image processing and Pattern Classification due to the stochastic nature of the mineral textures that is revealed at the microscopic level. Locally relevant Igneous Rock Microstructure images were classified from Volcanic and Plutonic Rock subtypes. The imaging method used mineral grain size as the key physical feature of classification. Three algorithms, namely Radial Basis Function (RBF) Support Vector Machine classifier; Improved (RBF) Support Vector Machine classifier; and AdaBoost algorithm with Improved RBF Support Vector Machine algorithm as base classifier, were used as a base classifier in a novel 'Iterative Improved Learning (IIL)' approach. Implementing the IIL approach in the chosen algorithm resulted in accurately classified images that were added to the training set to enhance the 'breadth and depth' of the learning knowledge. The algorithm iterated through all available classifier approaches and compared the inter-classifier performance and knowledge of the misclassified images accumulated during the execution of all algorithms.

Corresponding Author:

Ashutosh Marathe,
Department of Electronics & Tele-communication Engineering,
College of Engineering Pune,
5, Wellesley Road, Shivajinagar, Pune 411005, India.
Email: ashutosh.marathe@vit.edu

1. INTRODUCTION

As interdisciplinary applications, image processing and pattern classification techniques have been used for rock type identification. Most of this work is reported on handheld rock specimens [1], [2]. Studies of rock microstructures and applications of texture analysis-based approaches have been rare. This is mostly due to the stochastic nature of texture formed by minerals.

In this paper, microstructure image classification of Igneous rocks found in Volcanic and Plutonic subclasses are reported. Volcanic Igneous rocks are formed due to rapid cooling of lava on the Earth's surface. The fast cooling ensures that mineral grains do not grow to a large size. Plutonic rocks are formed below the Earth's crust due to slow cooling of the magma. Since the process of cooling is slow, the minerals tend to grow larger [3].

The basis for classification is the mineral grain size. The textures are analyzed using Haralick features [4] and a Laws' Mask-based [5] approach. Support Vector Machines (SVM) are used for classification. SVM based classification is successfully used in many applications ranging from medical applications such as MRI Brain image classification [6] to archeological applications such as monument classification [7].

For a Support Vector Machine, proper tuning its controlling parameters is important. Genetic Algorithm (GA) based approach is used to avoid trapping of SVM into local minimum [8], [9]. Rather than following such heuristic search based complex approaches, for the classification purpose in the chosen research work, A Radial Basis Function Support Vector Machine (RBFSVM) classifier is used. The controlling parameters C and Sigma are optimized using Grid search [10]. In addition, a classifier combination-based AdaBoost algorithm approach is used for classification.Boosting algorithm integrates weak classifiers. It has strong applicability for small size and high dimensionality data. It is easily programmable and highly adoptable [11].

The percentage accuracy for classification is calculated after each algorithm is implemented on the chosen image database. A progressive improvement is seen when an elaborate classification approach such as AdaBoost is used, after starting from a conventional RBFSVM approach. When the correctly classified images are added to the training image set, the 'Improved Learning' results in an improvement in % Accuracy.

The following sections provide the detais of the research work. Section 2 provides details of the database of images used. In Section 3, details of Haralick features and Laws' Mask details are provided, and the process of selection of the feature pair is presented. In Section 4, the concept of Improved Learning (IL) is detailed. Section 5 describes the implementation using a flow chart and outlines the execution of the algorithms. Section 6 provides the details of results obtained. Finally, Section 7 consolidates the various conclusions drawn from the IL-based approach.

2.   DATABASE OF IMAGES USED

In this study, locally relevant Igneous rock microstructure samples were used. Images belonging to the Basalt, Andesite, Spherulite, Pseudotachylites, Pegmatites, Dolerite, Rhyolite subfamilies were chosen. These varieties are abundant in the Western and Northern parts of the country in areas such as Goa, Maharashtra, Gujrat, Rajasthan, the Arawali Mountains, and Kumaun [12], [13], [14], [15], [16]. The structural analysis of these subfamilies is vital for construction and irrigation purposes. A total of 128 images, 64 each from the Volcanic and Plutonic categories, with a fair representation of each sub-family mentioned above, was considered.

3.   FEATURE SELECTION

Haralick Features and Laws' Masks were used to analyze the stochastic textures examined here due to diverse mineral combinations at the microstructure level. The Haralick features considered were
a.   Contrast
b.   Energy
c.   Entropy
d.   Homogeneity
e.   Correlation
And the Laws' features considered were
a.   Absolute Mean
b.   Standard Deviation

To improve the likelihood of generalization, a small set of features, usually from the original input variables, is generated via feature selection. In the feature selection step, redundant or meaningless features are discarded to achieve higher generalization performance and faster classification compared to the initial set of features.

The approach starts from an empty set, and features are added continuously while progressively checking the classification method's performance using a suitable classifier. This approach is called the forward selection approach [17].

A Radial Basis Function Support Vector Machine Classifier is used for classification by implementing the Forward Selection approach. Using all seven features, every possible feature pair was created, and the classification performance was evaluated.

Thirty-two Volcanic rock images and 32 Plutonic rock images were selected at random using the MATLAB function randperm(n) for each trial. One thousand such trials were carried out, and the Average False Rejection Ratio (AFRR) was calculated. The AFRR for all 21 duplets is as shown in Figure 1.
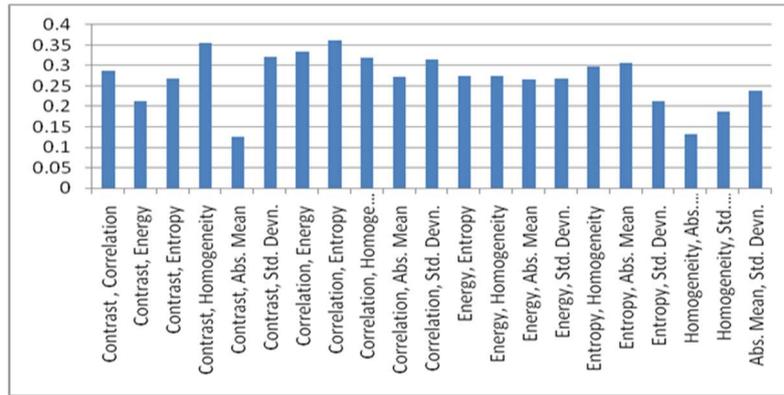
Figure 1. Average FRR for 21 duplets

The best AFRR of 0.125 was reported for Contrast and Abs. Mean (Feature Pair–1, 6). Consequently, triplets were formed with Pair (1, 6) as a base, adding a third feature and then checking AFRR. The performances of triplets, quadruplets are shown in Figure 2. Because there is no improvement in the False Rejection Ratio, compared to the cost of added computation time, the best performance is selected using the Contrast and Laws AM feature pair; this pair is then selected for improved learning-based experimentation.
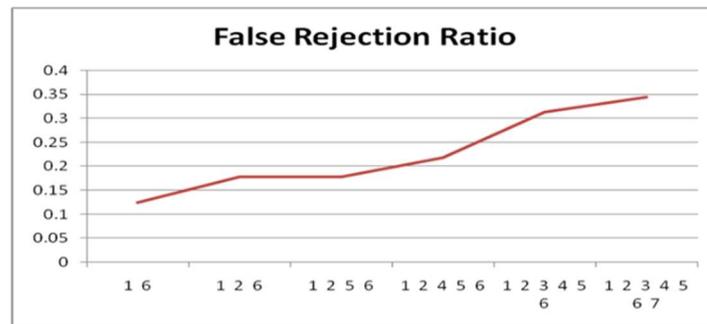

Figure 2. Average FRR variation with addition of features

4. IMPROVED LEARNING

Achieving the desired balance between stability and plasticity attributes in a classifier is critical. A classifier requires plasticity for the integration of any new knowledge, but it also requires stability to prevent the loss of previous knowledge [18].

Support Vector Machines are stable classifiers that exhibit better classification performance than many other machine learning methods [19]. However, stable SVMs suffer from a lack of plasticity and are prone to the catastrophic forgetting phenomenon. Catastrophic forgetting is defined as an event that occurs after a classifier is trained on the first task and then on another task, which results in the classifier forgetting to perform the first task.

Therefore, to fully benefit from the SVM Classifier performance, an Improved Learning method needs to be applied to the SVM, to retain its stability, while implementing plasticity [20].

A classifier can be eligible for improved learning by satisfying the following criteria:
a. It should be able to learn additional information from the input data.
b. It should not require access to the original data used to train the existing classifier
c. It should preserve previously acquired knowledge i.e., it should not succumb to catastrophic forgetting [21]

5. IMPLEMENTATION

For training purposes, a holdout-based training approach was used [22]. 64 Volcanic Rock images and 64 Plutonic Rock images, 32 from each class, were randomly selected from the image database. Overall, 10000 such training set combinations were attempted. The training set that gives the highest % Accuracy was considered. The Algorithm is as follows:

a. Compute Percentage Accuracy for Radial Basis Function SVM, Optimized Radial Basis SVM and AdaBoost Classifier with SVM as the base classifier. Identify the misclassified images in each case.
b. Add the classified images to the Training set to improve its 'Learning'.
c. Use the revised Training set and classify the remaining images, misclassified in an earlier iteration, using the 'Improved and more intelligent' training dataset.
d. Identify the misclassified images in each case again.
e. Repeat steps 1 to 4 if there is a reduction in the number of misclassified images.
f. Else, stop the algorithm because no further improvement is possible and no 'addition' in the training set is made for 'improved learning'.

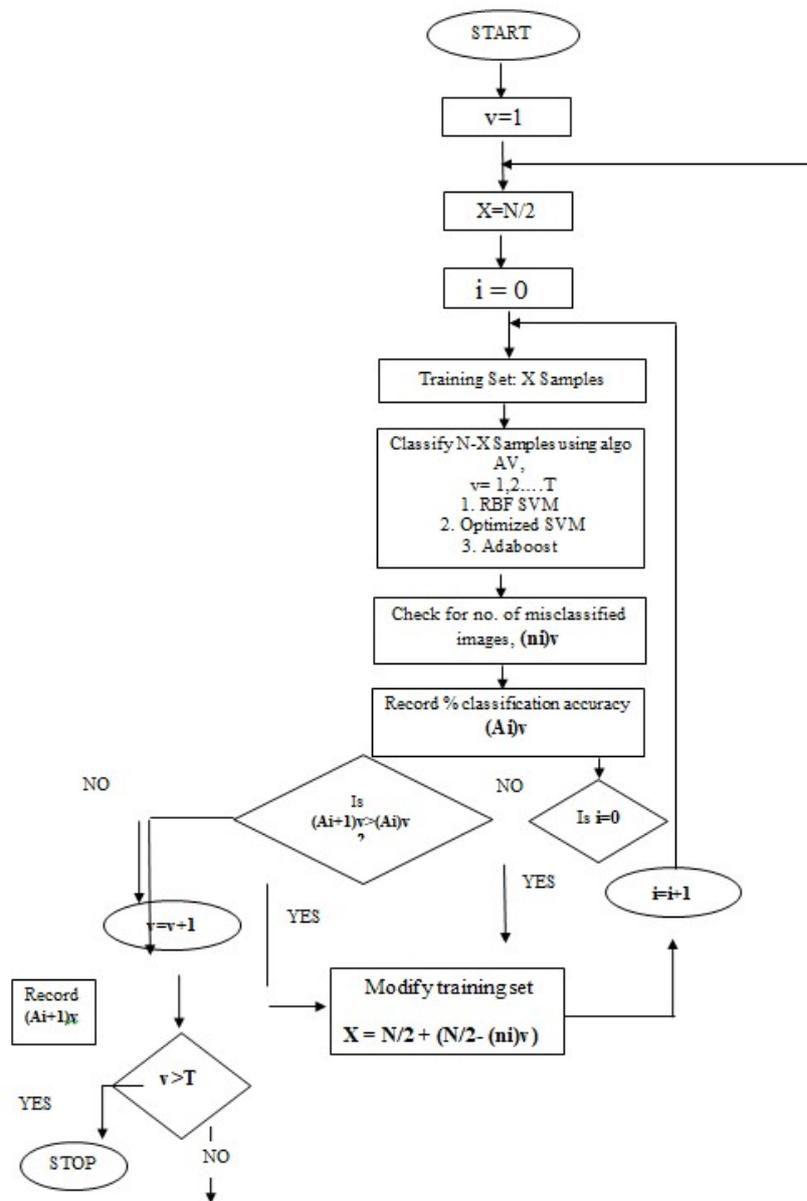The flow chart showing the implementation is as follows Figure 3:

Figure. 3 Flow chart of implementation

6. RESULTS & DISCUSSION

Table 1 precented about Algorithm-SVM, Table 2 precented Algorithm-Optimized SVM, Table 3 precented Algorithm–ADABOOST. The Improvement in percentage Accuracy for each algorithm is due to the Improved Learning approach used. Figure 4, highlights the details

Table 1. Algorithm-SVM

Iteration	1	2	3
No. of Training Images	64	64 + 56 = 120	120 + 1 = 121
No. of Classification Images	64	8	7
No. of Misclassified Images	8	7	7
Details of Misclassified Images	P = 8,29,30,34,40,47 V = 13,17	P = 8,29,30,40,47 V = 13,17	P = 8,29,30,40,47 V = 13,17
% Accuracy	87.5 %	89.06 %	89.06 %
Comment	Base accuracy	Improvement through learning	End because of no further improvement

Table 2. Algorithm-Optimized SVM

Iteration	1	2	3
No. of Training Images	64	64 + 57 = 121	121 + 2 = 123
No. of Classification Images	64	7	5
No. of Misclassified Images	7	5	5
Details of Misclassified Images	P = 8,29,30,34,40,47 V =13	P = 29,30,40,47 V = 13	P = 29,30,40,47 V = 13
% Accuracy	89.06 %	92.18 %	92.18 %
Comment	Base Accuracy	Improvement through Learning	End because of no further Improvement

Table 3. Algorithm-ADABOOST

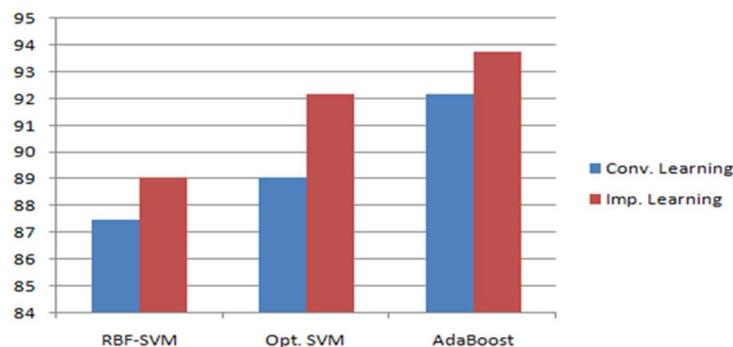| Iteration	1	2	3
No. of Training Images	64	64 + 59 = 123	123 + 1 = 124
No. of Classification Images	64	5	4
No. of Misclassified Images	5	4	4
Details of Misclassified Images	P = 29,30, 40,47 V = 13	P = 29,30,40 V = 13	P = 29,30,40 V = 13
% Accuracy	92.18 %	93.75 %	93.75 %
Comment	Base Accuracy	Improvement through Learning	End because of no further Improvement

Figure 4. Improvement in % accuracy with improved learning

7. DISCUSSIONS

a. An improvement in percentage accuracy is observed with improved learning
b. Classifier Ensemble approaches, i.e.the AdaBoost approach with an SVM as the base classifier performs the best among the algorithms used for this study.
c. The feature selection was implemented using a Forward Selection algorithm. Contrast and Laws Absolute Mean are the identified feature pair. Owing to the governing Pressure and Temperature conditions, Volcanic Rocks have smaller mineral Grain sizes, and Plutonic rocks develop grains of

larger size. The Volcanic and Plutonic Rock constellation on the 2-Dimensional feature plot is shown in Figure 5.
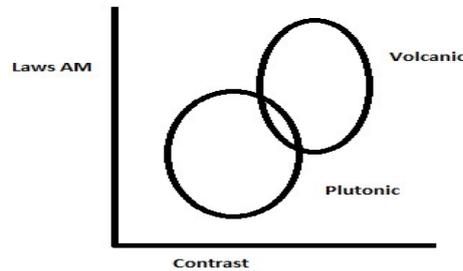

Figure 5. Typical data distribution for volcanic and plutonic rocks

d.   Some images remain misclassified even though the Improved Learning uses all of the algorithms. Some images that were misclassified by RBF-SVM and Optimized RBF-SVM were correctly classified by the AdaBoost approach. This confirms its relatively superior performance in terms of classification accuracy, as demonstrated by Figure 4.
e.   IILA is a comprehensive procedure that tests a database using all algorithms of interest. It also indicates the 'distinct outliers', which are not classified by any of the selected classifiers. The physical parameters governing the generated feature values can then be specifically investigated.
f.   Misclassification is attributed to the overlapping region shown in Figure 5. The misclassified images, even after applying the Improved Learning Based AdaBoost algorithm (Namely–Volcanic No. 13; Plutonic No. 29, 30, 40) demonstrated 'outlier' attributes. The outliers exhibit interim 'Hypabyssal' group feature attributes [23] and land in the overlapping region. Hence, these findings are consistent with the Geological properties.

A summary of research approaches used for Petrographic Image Analysis is shown in Table 4. It is seen that authors of this paper report the first of its kind use of Support Vector Machine based method and Classifier combination methods are practiced for Petrographic Image Classification. The Accuracy reported is around 94%. Majority of the authors have tried to classify Igneous, Metamorphic and Sedimentary rocks. The present paper is a unique research activity of classification of igneous rocks into its subfamilies. This study was carried out on an Intel i3 processor working at 2.10 GHz and used MATLAB version R 2012 b. A total 26.2 seconds was required to execute the IILA program. IILA had 2 iterations, as reported.

Table 4. Summary of Research Approaches used for Petrographic Image Analysis

Sr.No.	Authors	Approach	% Accuracy	Comments
1	Baykan, Yilmaz [24]	3-layer feedforward network with Minimum Square Error Correction Color Space Based classification	Variable between 81 to 98%	Study of total 5 minerals belonging to Igneous, Metamorphic and Sedimentary rocks
2	Marmo, Amodiao et. al [25]	Multilayer Perceptron (MLP) Network	93.3%	Ancient (Phanerozoic) Carbonates from marine environments were classified
3	N.Singh, T.N.Singh et. al [26]	Multilayer Perceptron (MLP) Network	92.22%	Study of Basalt rock
4	Bodziony et.al [27]	Tree Automata and Graph Automata theory	84.13%	Development of Expert System for Petrography and Mechanical analysis of Rocks
5	Ishikawa, Gulick [28]	Training of neural networks based on Raman Spectra	83%	Spectral mapping of Key minerals characterizing composition of igneous rocks
6	A. Marathe et al (in the present paper)	i. Support Vector Machine (SVM) based methods ii. Classifier Combinations – Adaptive Boosting	94%	Texture based Locally relevant Igneous rock petrographic image classification

8.   CONCLUSION

It can be concluded that the Classification of Igneous Rocks into Volcanic and Plutonic subfamilies can be carried out using Support Vector Machine based approach and classifier combinations. Fine tuning the penalty parameter and variance for the Radial Basis Function SVM, optimal performance can be extracted from the SVM Classifier. By using the optimized SVM as the base i.e. weak classifier, using Adaptive Boosting enhancement in classification accuracy can be achieved. Further improvement in classification accuracy can be achieved through the use of iterative Improved Learning approach. The iterative learning provided information about images those remain misclassified despite iteration. This information can be used for validation of outliers.

REFERENCES

[1]   Shankar V., Rodríguez J.J., Gettings M.E.,"Texture Analysis for Automated Classification of Geologic Structures,". *IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 81-85, 2006.
[2]   Hang Zhou, Monteiro S.T., Hatherly P., Ramos F, Nettleton E.Oppolzer F, "Automated rock recognition with wavelet feature space projection and Gaussian Process classification," *IEEE International Conference on Robotics and Automation (ICRA*, pp 4444-4450*)*, 2010.
[3]   Ron Vernon, "A Practical Guide to Rock Microstructure," *Cambridge University Press*, 2004.
[4]   Robert Haralick, K.Shanmugam, Its'hak Dinstein., "Textural Features for Image Classification,". *IEEE Transactions on Systems Man Cybernetics*, vol. 3, pp. 610-621, 1973.
[5]   Laws Kenneth.,"*Textured Image Segmentation*" PhD Thesis, University of Southern California. Los Angeles, 1980.
[6]   Madina Hamiane, Fateema Saeed. "SVM Classification of MRI Brain Images for Computer Assisted Diagnosis," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 7, pp. 2555-2564, October 2017.
[7]   Malay Bhatt, Tejas Patalia., "Indian Monuments Classification Using Support Vector Machine," *International Journal of Electrical and Computer Engineering (IJECE)*, vol.7, pp.1952-1963, August 2017.
[8]   Xiang Sheng, Zhou Yu, Xilong Qu., "Support Vector Machine Optimized by Improved Genetic Algorithm," *Telkomnika Indonesian Journal of Electrical Engineering*, vol.12, pp 631-636, 2014.
[9]   P.Liao, X.Zhang, K.Li, "Parameter Optimization for Support Vector Machine based on Nested Genetic Algorithms," *Journal of Automation and Control Engineering*, vol. 3, pp 507-511, 2015.
[10]   Zhigang yan, Y. Yang, Y,Ding, "An Experimental Study of the Hyper-Parameters Distribution Region and Its Optimization Method for Support Vector Machine with Gaussian Kernel," *IJSPIPPR*, vol.6, pp 437-446, 2013.
[11]   Wencheng Gu., "Application of Boosting Algorithm in Spam Filtration," *Telkomnika Indonesian Journal of Electrical Engineering*, vol.12, pp 5685-5692, 2014.
[12]   Rao Y.H. et al., "Magnetic and Petrographic Analysis of Andesite Rock Bodies, Khairmalia, South of Chittourgarh, Rajasthan," *International Journal of Science, Environment and Technology (IJSET),* vol. 1, pp 113-124, 2013.
[13]   Kshirsagar Pooja V. et al., "Spherulites and Thundereggs from Pitchstones of the Deccan Traps: Geology, Petrochemistry, and Emplacement Environments," *Bulletin of Volcanology*, vol. 74, pp 559-577, 2012.
[14]   Agarwal K.K. et al., "Occurrence of Pseudotachylites in the Vicinity of South Almora Thrust zone, Kumaun Lesser Himalaya," *Current Science*, vol. 101, pp. 431-434, 2011.
[15]   Sabale P.D., Meshram S.A., "Effect of Dyke Structure on Ground Water in between Sangamner and Sinnar area: A Case study of Bhokani Dyke," *International Journal of Computational Engineering Research (IJCER)*, vol. 2, pp. 1130-1136, 2012.
[16]   Sisodia M.S., "Malani Rhyolite: Highly Eroded Complex Crater," *Current Science*, vol. 101, pp 946-951, 2011.
[17]   Shigeo Abe. "Support Vector Machines for Pattern Classification". 2nd Edition. 2010*, Springer*, 2010.
[18]   Asaraf, M. Murty, S. Shevade, "Rough Set based Incremental Clustering of Interval Data," *Pattern Recognition Letters*, vol. 27. pp. 515-519, April 2006.
[19]   C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Springer Journal of Data mining and knowledge discovery*, vol. 2, pp. 121-167, 1998.
[20]   Parag Kulkarni., "Reinforcement and Systemic Machine Learning for decision making," Wiley – IEEE Press, 2012.
[21]   R. Polikar, L. Udpa, S. Udpaand, and V. Honavar, "Learn++: An Incremental Learning Algorithm for Supervised Neural Networks, "*IEEE Transactions on System, Man and Cybernetics, Special Issue on Knowledge Management*, vol. 31, pp. 497–508, 2001.
[22]   Kuncheva L. "Combining Pattern Classifiers, " 2nd edition ,John Wiley & Sons, 2014.
[23]   David Shelley, "Igneous and Metamorphic Rocks Under the Microscope," *Chapman and Hall*, 1993.
[24]   N. Baykan, N. Yilmaz., "Mineral Identification Using Color Spaces and Artificial Neural Networks," *Computers and Geosciences*, 2010, Vol. 36, No. 1, pp 91-97.
[25]   Marmo R., Amodio S., Tagliaferri R., Ferreri V., Longo G. "Textural Identification of Carbonate Rocks by Image Processing and Neural Network: Methodology, Proposal and Examples," *Computers and Geosciences*, vol. 31, pp. 649-659, 2005.
[26]   Singh N., Singh T.N., Tiwary A., Sarkar K.M., "Textural Identification of Basaltic Rock Mass Using Image Processing and Neural Networks," *Computers and Geosciences*, vol. 36, pp 301-310, 2010.
[27]   Bodziony J., Flasinski M., Mlynarczuk M., Baran M., "Towards Constructing Pattern Recognition-Based Expert System for Petrography and Rock Mechanics Analysis," *Proceedings of the 3rd Conference on Computer Recognition Systems* – KOSYR 2003.

[28] Ishikawa S., Gulick V., "An Automated Mineral Classifier Using Raman Spectra," *Computers and Geosciences*, vol. 39, pp 259-268, 2013.

BIOGRAPHIES OF AUTHORS

Ashutosh Marathe received B.E. Electronics and M.E. Electronics from Pune University, India in 1993 and 2000 respectively. He is pursuing Ph.D. from College of Engineering Pune. Ashutosh has 22 years of Teaching experience. He is Senior Member, IEEE since 2004. His Technical research interests include Image Processing, Pattern Recognition and Machine Learning with a focus on areas of interdisciplinary relevance. He has also been contributing in Academic research in the areas of Accreditation, Appraisals and Quality Improvement in higher education. He has publications in Technical as well as Pedagogy areas in many National and International Conferences in India and Abroad. The noteworthy publications being Frontiers in Education – 2013; ICSPCC – 2016 in Hong Kong and IConSIP – 2016 in India. Ashutosh also works as Data Auditor for the 'Technical Education Quality Improvement Program (TEQIP-II), a World Bank initiative. Besides IEEE, He is a life member of Indian Unit of Pattern Recognition and Artificial Intelligence (IUPRAI), a country affiliate of International Association of Pattern Recognition (IAPR); life member of IETE, India; life member of Institution of Engineers, India.

Priya Jain received the Bachelor of Engineering degree from the Department of Electronics and Communication, Samrat Ashok Technological Institute (SATI), Vidisha, MP in the year 2013. She is currently pursuing her Master of Technology under the stream Signal Processing from the Department of Electronics and Telecommunication, College Of Engineering, Pune (COEP). She has her interest in Image Processing, Computer Vision and Machine learning.

Vibha Vyas obtained her Bachelor's Degree from Nagpur University in 1995, her Masters and PhD Degrees from Pune University in 2002 and 2010 respectively. She has active research contribution in the areas of Signal & Image Processing. She has around 35 Quality Research Publications in International and National Journals as well as Conferences. She has filed two patents and continues to guide Doctoral and Masters students. Vibha has around 22 years of teaching experience. She is an active researcher in Centre of Excellence in Signal Processing, established in College of Engineering, by funding under 'Technical Education Quality Improvement Program (TEQIP-II), a World Bank initiative.