Shahrood University of Technology

Journal of Mining and Environment (JME)

Journal homepage: www.jme.shahroodut.ac.ir

Iranian Society of Mining Engineering (IRSME)

Evaluation of Band Ratio Technique for Prediction of Iron-Titanium Mineralization Using Ensemble Machine Learning Model: A Case Study from Khamal area, Western Saudi Arabia

Ahmed Ali Madani

Department of Geology, Faculty of Science, Cairo University, Giza, Egypt

Article Info

Received 24 April 2024

Received in Revised form 15 May 2024

Accepted 11 June 2024

Published online 11 June 2024

DOI: 10.22044/jme.2024.14451.2711

Keywords

AI-based Predictive Model

Random Forest Algorithm

SPOT-5 Data

Fe-Ti Mineralization

Western Saudi Arabia

Abstract

Innovation in mineral exploration occurs either in the construction of new ore deposit models or the development of new techniques used to locate the ore deposits. Band ratio is the image processing technique developed for mineral exploration. The present study presents a new approach used to evaluate the band ratio technique for discrimination and prediction of the Iron-Titanium mineralization exposed in the Khamal area, Western Saudi Arabia using the ensemble Random Forest model (RF) and SPOT-5 satellite data. SPOT-5 band ratio images are prepared and used as the explanatory variables. The target variable is prepared in which (70%) of the target locations are used for training and the rest are for validation. A confusion matrix and the precision-recall curves are constructed to evaluate the RF model performance and the Receiver Operating Characteristics curves (ROC) are used to rank the band ratio images. Results revealed that the 3/1, 2/1 & 3/2 band ratio images show excellent discrimination with AUC values of 0.986, 0.980 & 0.919 respectively. The present study successfully selects the 3/1 band ratio image as the best classifier and presents a new Fe-Ti mineralization image map. The present study proved the usefulness of the Random Forest classifier for the prediction of the Fe-Ti mineralization with an accuracy of 97%.

1. Introduction

Band ratio is one of the most important remote sensing techniques used to locate the ore deposits. Selection of the best band ratio classifier can be performed either visually or statistically [1-22]. The present study presents a new approach used to predict the Iron-Titanium Mineralization exposed in the Khamal area, Western Saudi Arabia (Figure 1) using the RF machine learning algorithm and the ROC curves. Data-driven machine learning (ML) algorithms are capable of learning and modeling complex patterns in a large dataset [23]. The capabilities of machine learning predictive models have emerged as a powerful decision-making tool for mineral exploration. Machine learning algorithms provide a valuable prediction using several geo-datasets. The Support Vector Machine (SVM) and the Random Forest (RF) are the most common ML models used to predict mineral occurrences [24-45].

Multi-criteria decision-making GIS predictive models (Knowledge-driven and data-driven) are developed to generate predictive favorability maps showing promising sites for mineral occurrences [29, 46-64]. The Random Forest regression is used to locate the gold deposits in the Rodalquilar mining district, Southern Spain [29]. Results indicated that the use of the RF offers several advantages over existing methods. Two individual regression models (KNN and SVM) and two robust ensemble methods (RFR and GBR) are developed to predict ore grades (Pb and Zn) in the Irankuh area of Central Iran [43]. Results revealed that the hybrid method is promising for predicting the ore elemental distribution. The ensemble learning models,

Corresponding author: aamadani@sci.cu.edu.eg (A.A. Madani)

logistic regression (LR), and support vector machines (SVM) are compared to predict the lithological classes using geological and geochemical data [45]. Results emphasize the potential of ensemble learning models for lithologic classification. A decision support system based on remote sensing and GIS techniques is developed for gold-rich area mapping in SE Spain [50]. A Knowledge-driven predictive model is used to generate a favorability map for gold mineralization at the Bulghah gold mine area, in Saudi Arabia using the integration of several geo-datasets [56]. Also, a data-driven predictive mapping using the Random Forest algorithm is used to map mineral occurrences in the Baguio gold district, Philippines using different datasets [59]. The integration of Landsat ETM imagery, geochemistry, airborne radiometric data, and aeromagnetic data is performed to produce a mineral potential exploration model of the Riruwai Complex, Nigeria [62]. A knowledge-driven GIS predictive model is used to generate the favorability maps for gold and copper mineralization for south Gabal Um Monqul and the Gabal Al Kharaza prospects, the northern Eastern Desert of Egypt [63]. A processed multispectral satellite data is used to map the gold deposits in the Hamissana area, NE Sudan based on a Random Forest predictive model [64]. Results revealed that ASTER and Sentinel-2 datasets achieved very similar accuracy.

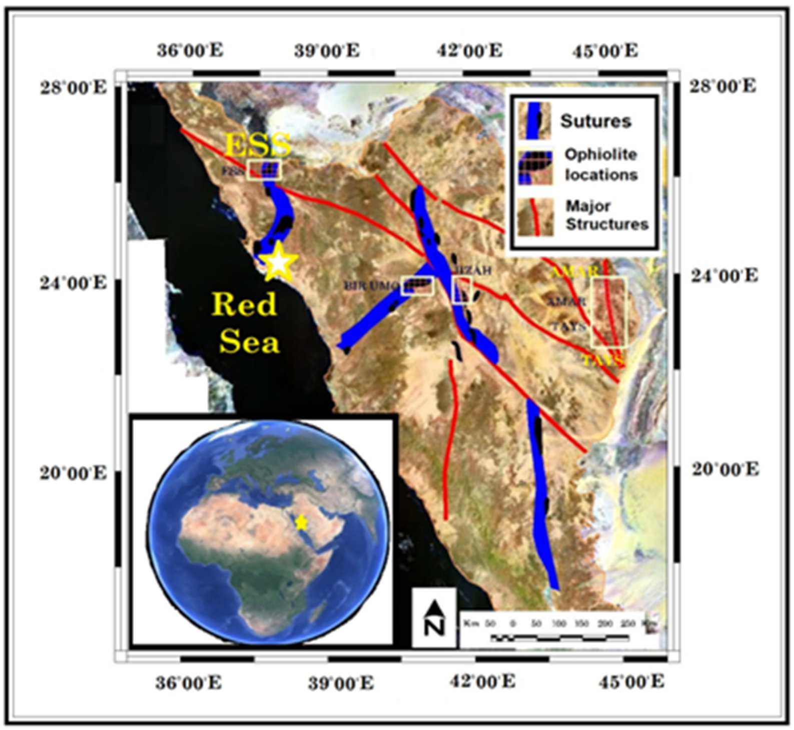


Figure 1. Location map for the study area (modified after [20]). The yellow asterisk indicates the location of the study area.

2. Study Area and the Fe-Ti Mineralization

The study area is located in the northwestern part of the Arabian Shield (Figure 1). It is covered by a sequence of ultramafic-mafic rocks that belong to the Wadi Khamal-Wadi Murattijah complex [65]. It represents a well-known example of Neoproterozoic post-collisional layered mafic intrusions [66]. It hosts ore deposits characteristic of layered intrusions, including Fe-Ti oxides and Fe-Ti-apatite-rich nelsonite. Several authors studied the geology, mineralogy, stratigraphy, and mineralization of the Wadi Khamal Complex [67-72]. The study area is dominated by a sequence of ultramafic-mafic rocks intruded on the metamorphic rocks of the Farri group and post-Al Ays granitoid rocks (Figure2). This rock sequence was intruded by younger granitic batholiths, and mafic dykes, and is locally overlain by Tertiary to Quaternary basalt flows, marine sediments, and alluvial terraces. The Khamal complex can be classified into four main units namely; 1) Marginal gabbro unit (KU1), 2) Anorthosites (KU2), 3) Central gabbros (KU3), and 4) Northern gabbronorite (KU4) [66]. The RF model was applied to the SPOT 5 subsets covering the KU2 and KU3 units. The Fe-Ti mineralizations are commonly formed during the late stage of fractionation of basic magma and are mainly associated with the Gabbro-Anorthosite Complex. The main factor controlling the distribution of the mineralization in the study area is lithology. Two types of Fe-Ti mineralization are recorded [65]; (i) massive nelsonite bands of magnetite-ilmenite and apatite in different proportions, and (ii) massive magnetite-ilmenite ore found either as bands intercalated with nelsonite or as dike-like bodies hosted by anorthosite.

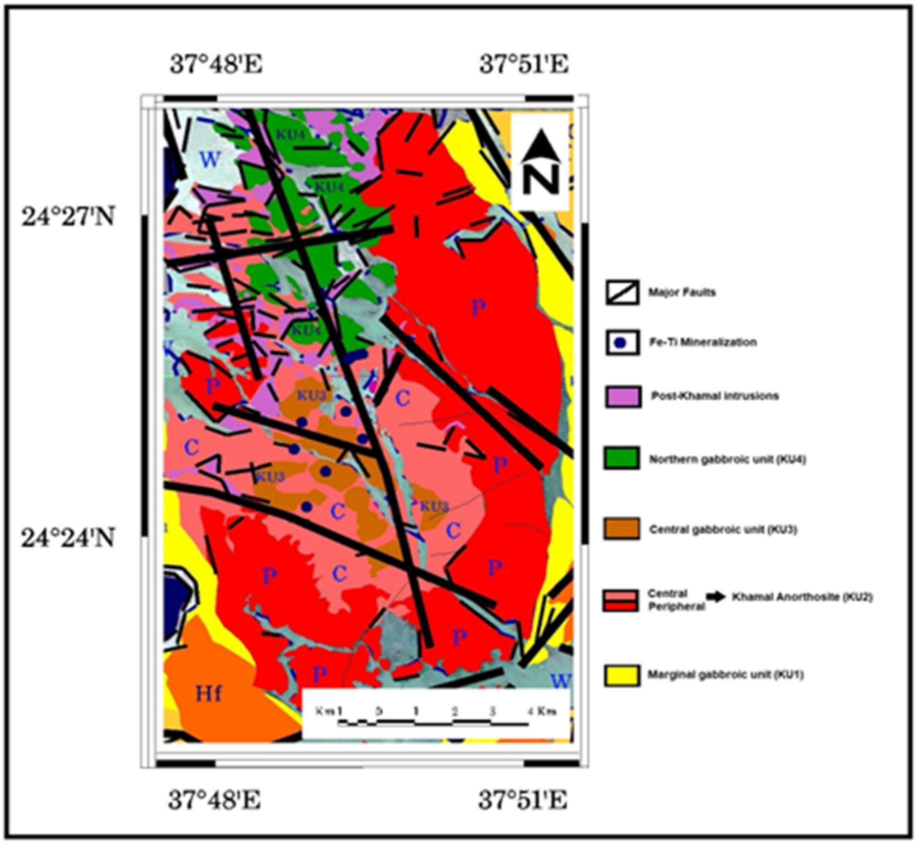


Figure 2. Part of the geological map for the Khamal area, (modified after [66]).

3. Methodology

Spatial Data Science (SDS) is a part of data science concerned with the analyses of spatial data. One of the main activities of SDS is the predictive modeling. It is a statistical technique that utilizes machine learning algorithms to predict future outcomes with the aid of historical data. Figure 3 shows the general methodology used to predict Fe-Ti-mineralization and the selection of the best band ratio classifier. SPOT 5 data are digitally processed using PCI image processing software. Image subsets from the original multispectral images covering the study area are prepared. The band ratio technique is the main image processing technique used to generate SPOT-5 band ratio images which are used as explanatory variables in the Random Forest ML model. A dependent variable consisting of 106 mineralized and non-mineralized locations is prepared. It is split into (70%) for training the model and (30%) for validation. A Random Forest model is performed using the ArcGIS Pro package. A confusion matrix and precision-recall curves are used for model performance and validation. The ROC curves are generated and evaluated to determine the most successful classifier for Fe-Ti-mineralization predication and mapping.

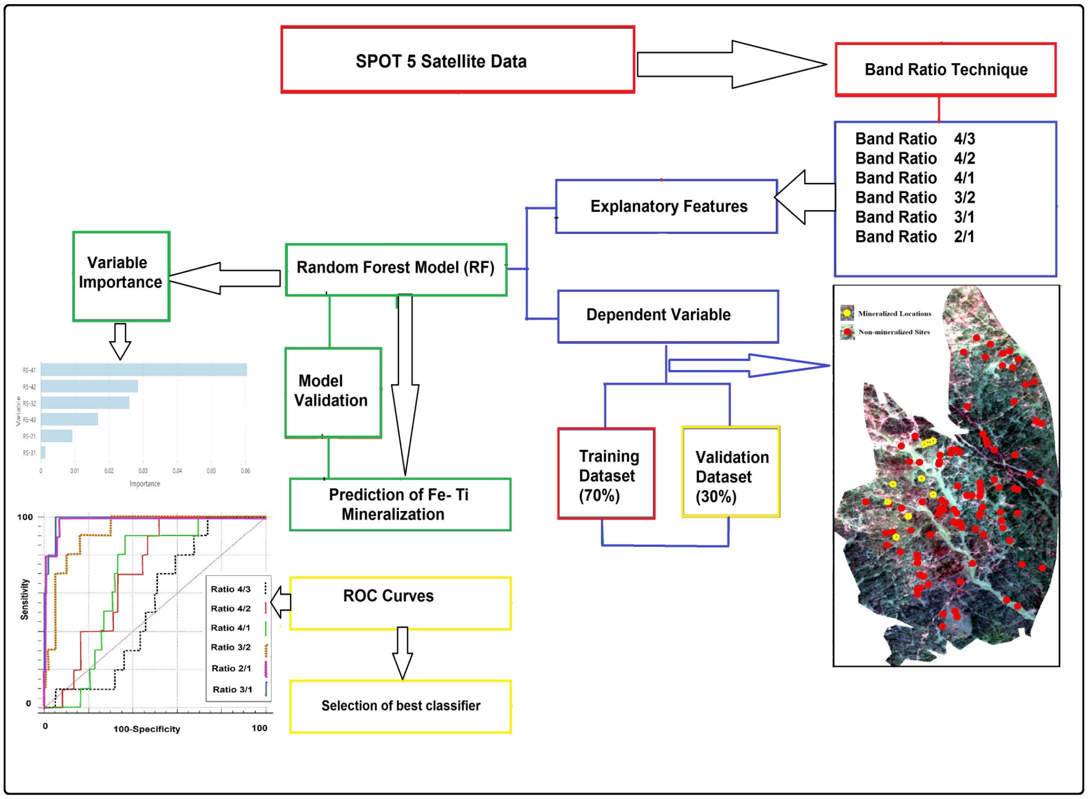


Figure 3. General methodology for prediction of Fe-Ti-mineralization using the RF ML model.

3.1. SPOT 5 Satellite Data

SPOT series are high spatial resolution Earth observation satellites operating in visible-near-infrared wavelength regions. The SPOT 5 satellite contains two identical High-Resolution cameras providing 2.5 m and 5 m resolution in a panchromatic mode and a 10 m resolution in a multi-spectral mode. The SPOT 5 has four spectral bands: Band1 (0.50-0.59 µm); Band2 (0.61-0.68 µm); Band3 (0.79-0.89 µm); and MIR-band (1.58-1.75 µm). Several authors utilized SPOT 5 data in land use/land cover mapping, Natural disasters, and urban planning applications [73-79]. Few studies utilized SPOT data for geological mapping [7]. In the present study, image subsets from the original SPOT 5 multispectral images are prepared and processed using the band ratio technique.

3.2. Preparation of Explanatory and Target Variables

The explanatory variables are represented by the band ratio images prepared from the multispectral SPOT-5 data. Band ratio is an important technique used for mineral identification and lithologic discrimination [1, 3, 4, 9, 18, 20]. SPOT 5 satellite data is utilized for mapping gold mineralized diorite-tonalite intrusion in the Bulghah gold mine area, Saudi Arabia [7]. Band ratio technique is used to discriminate the listvenite and serpentinite rocks along some shear zones, in Saudi Arabia [20]. Figure 4 shows the results of the band ratio technique applied to SPOT data. These images are used to discriminate gabbros from anorthosite rocks and to predict the Fe-Ti mineralization. Visual inspection revealed that: 1) anorthosites have dark grey image signatures on images 4/1, 4/2 & 4/3 whereas central gabbros have bright grey image signature on 2/1 band ratio image.

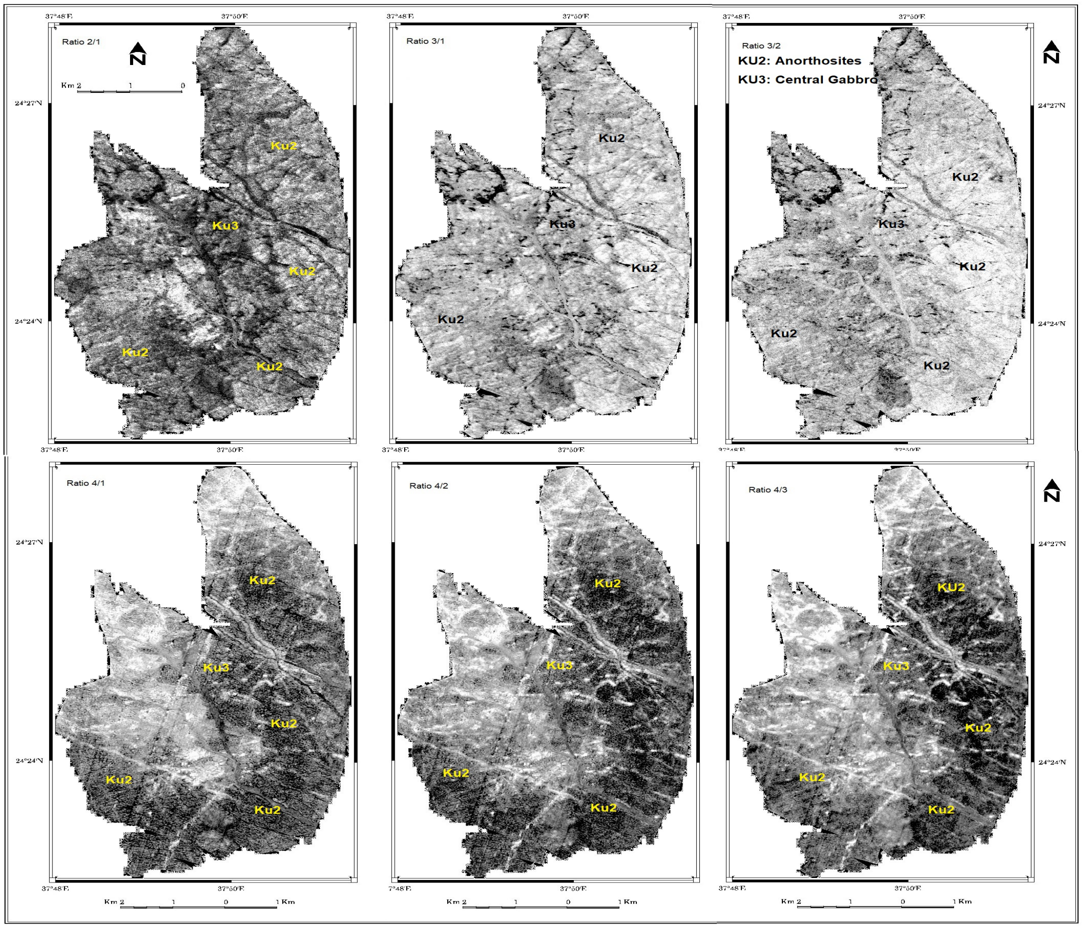


Figure 4. Explanatory variables: SPOT 5 band ratio images.

The target variable is represented by the mineralized and non-mineralized locations covered in the study area (Figure5). The mineralized locations are used after [66]. Two types of Fe-Ti mineralization are recorded in the study area; (i) massive nelsonite bands, and (ii) massive magnetite-ilmenite ore found as dike-like bodies hosted by anorthosite. The mode of occurrence of the mineralization and their dimensions in addition to the spatial resolution of SPOT 5 data are favorable conditions for the prediction of the Fe-Ti mineralization using a machine learning model. The last type of mineralization (massive magnetite-ilmenite dike-like bodies) is the main target of the RF model. The non-mineralized sites are represented by the host rocks such as gabbros and anorthosites. The RF model is applied to the SPOT 5 subsets covered by the anorthosites and central gabbros.

3.3. Random Forest (RF) Algorithm and Model Performance

The Random Forest machine learning model is used to predict the Fe-Ti mineralization. The model falls under the umbrella of ensemble classifiers and is characterized by bootstrap aggregation and randomization. The RF model allows the user to build optimal decision trees based on the aggregation of multiple iterative trees built from randomly selected samples of the training step [80]. Several authors demonstrated the ability of the RF model to show the variable importance during the training and prediction stages [30,33,81]. The number of trees and the explanatory variables are the main parameters required to implement the RF model. In the present work, the number of trees is set to the default (100) whereas the explanatory variable is represented by the SPOT 5 band ratio images. Several authors utilized the RF model and other machine-learning models for mineral exploration, among them [30,33,39,40, 43,44,45,59, 82,83]. In this study, confusion matrix and classification reports are generated to evaluate the model performance. Accuracy, precision, recall (sensitivity), and FI-score are generated and evaluated. In the present study, a precision-recall curve is evaluated for the model performance. It is a plot of the precision (y-axis) and the recall (x-axis) for different thresholds [84, 85]. A precision-recall curve is useful in the case of imbalanced data (as in the case of our study) in which there are many observations of class 0 and few of class 1. The Receiver Operating Characteristics (ROC) curves are evaluated to select the best classifier from the explanatory variables.

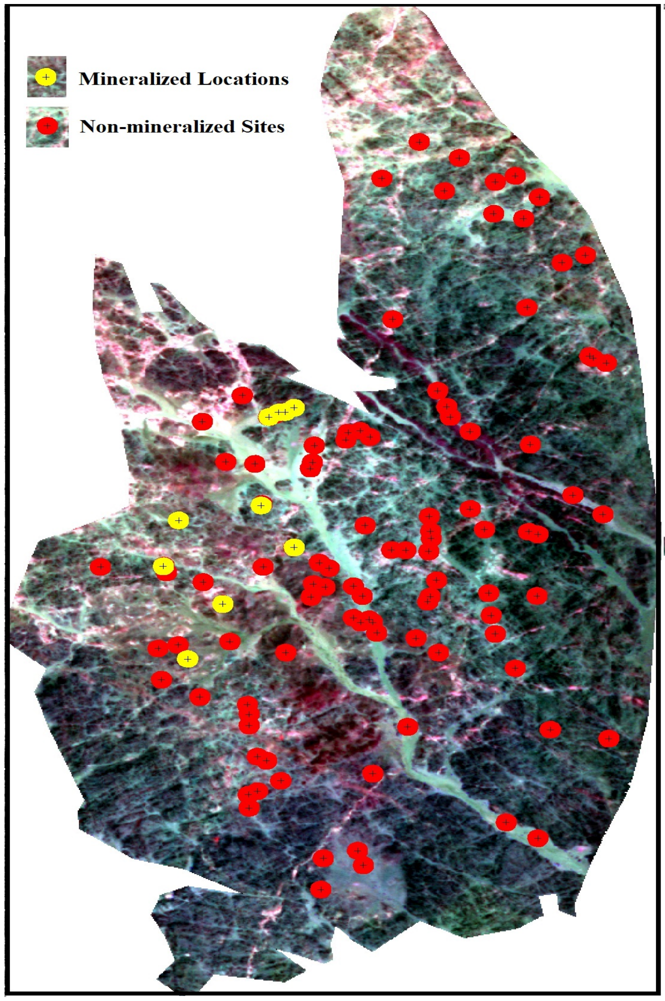


Figure 5. The target variable. Yellow dots represent the mineralized locations.

4. Results and Discussion

Collinear analysis is the study of the linear correlation between the independent variables. Results of this analysis revealed: 1) the presence of strong positive correlations between a) ratio 4/3 and ratio 4/1 (R2= 0.97); b) ratio 3/1 and ratio2/1 (R2= 0.86); c) ratio 4/2 and ratio4/1 (R2= 0.81); and d) ratio 3/2 and ratio 3/1 (R2= 0.81); 2) the presence of negative correlations between a) ratio 4/3 versus 3/2, 3/1 and 2/1 ratios; b) ratio 4/2 versus 3/2, 3/1 and 2/1; and c) ratio 4/1 versus 3/2, 3/1 and 2/1. In the present study, all the above ratios are used as the explanatory variables during the implementation of the RF model.

4.1. The Random Forest (RF) Model Evaluation

The Random Forest model is implemented using a reference variable containing about (106 points) of mineralized and non-mineralized locations. The model is trained using 70% of these points and is validated using the rest (30%). The six SPOT band ratio images are used as explanatory variables. Figure 6 shows the variable importance of SPOT band ratio variables during the training (Figure6a) and the prediction (Figure6b) stages.

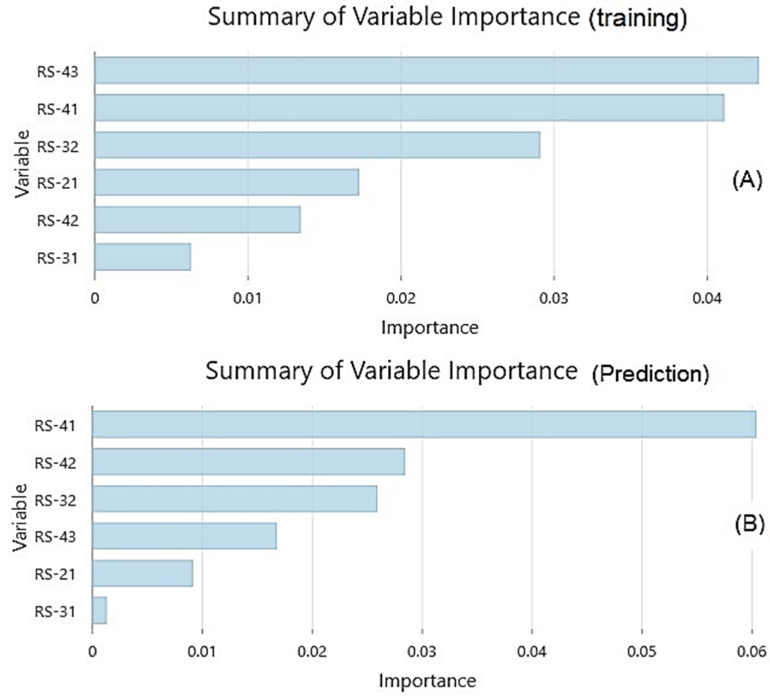


Figure 6. SPOT variable importance during, a) training and b) prediction stages.

The ratios 4/3, 4/1, and 3/2 are the most important variables in the training stage with values reaching 28%, 27%, and 19% respectively. During the prediction stage, the ratios 4/1, 4/2, and 3/2 are the important variables reaching 42%, 20%, and 18% respectively. Table 1 shows the importance percentage for each variable during the training and prediction stages.

Table 1. SPOT variable importance percentage.

Variables	I-train	%train	I-Pred	%Pred
RS-41	0.041082	27.328	0.06035	42.6
RS-42	0.013397	8.912	0.02837	20.026
RS-32	0.029056	19.328	0.02588	18.268
RS-43	0.043331	28.823	0.016726	11.807
RS-21	0.017216	11.452	0.009091	6.417
RS-31	0.00625	4.157	0.00125	0.882

The confusion matrix and the precision-recall curves are used for model performance and evaluation. The confusion matrix is organized to map the prediction classes to the original classes of the data. It reports the numbers of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). During the training stage, among 106 samples of actual data, 66 samples were classified as true positive, 3 samples as true negative, 8 samples as false positive, and 29 samples as false negative. Figure 7 shows the confusion matrix of the RF model applied for SPOT ratios. For the prediction, among 32 (30%) samples of actual data, 28 samples are classified as true positive, and 3 samples are classified as true negative. The overall accuracy reached up to 97%. Figure 8 shows the new locations of the Fe-Ti mineralization predicted using the RF model.

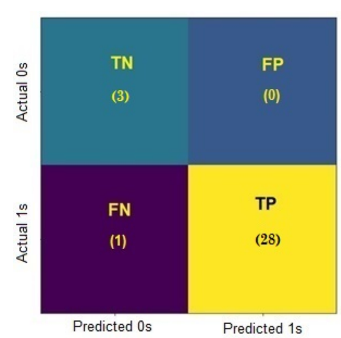


Figure 7. Confusion matrix of SPOT data.

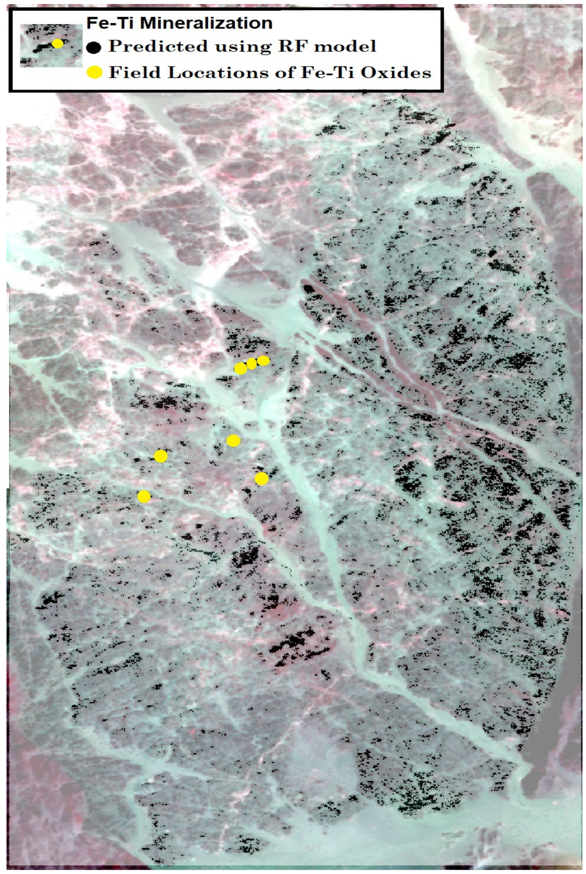


Figure 8. New locations of Fe-Ti mineralization (black spots) predicted using the RF model.

Due to the imbalanced observations of the dependent variable, the precision-recall curves have been constructed to show the model performance. Figure 9 shows the result of the precision-recall curves. Table 2 shows the variable performance items F1& AUC for each variable at the optimum threshold value. The ratios 3/1, 2/1 & 3/2 have the F1 & AUC values of 0.800, 0.842, 0.636 & 0.894, 0.806, 0.514 respectively.

Table 2. Variable performance items (F1& AUC).

Variable	Associated Criterion	F1	AUC
Ratio 3/1	1.582	0.800	0.894
Ratio 2/1	1.420	0.842	0.806
Ratio 3/2	1.129	0.636	0.514
Ratio 4/3	0.532	0.286	0.139
Ratio 4/1	0.835	0.333	0.131
Ratio 4/2	0.632	0.220	0.094

4.2. Evaluation of the Band Ratio technique using the ROC Curves

The Receiver Operating Characteristic (ROC) curves are graphs showing the performance of the classification model at different thresholds. They are useful for organizing classifiers and visualizing their performance [86]. It is commonly used to evaluate predictive models and is frequently used in machine learning models. The curve has two parameters namely true positive rate and false positive rate, expressed as the following equations: (True positive rate (TPR) = TP/TP+FN) and (False positive rate (FPR) = FP/FP+TN). Whereas TP is true positive; FN is false negative; FP is false positive & TN is true negative. In the present study, the ROC curves are used to select the best classifier used for the prediction of Fe-Ti-mineralization. Table 3 shows the variable performance items (sensitivity, specificity, accuracy & AUC) measured at the optimum threshold value. It shows that the 3/1 ratio is the best classifier with accuracy reaching 94.79% and AUC reaching 0.986. Band ratios 3/1, 2/1 & 3/2 exhibit good discrimination with AUC values of 0.986, 0.980 & 0.919 respectively. Band ratios 4/3, 4/1 & 4/2 show less discrimination and have AUC values of 0.703, 0.683 & 0.534 respectively. Figure 10 shows the ROC curves for the explanatory variables.

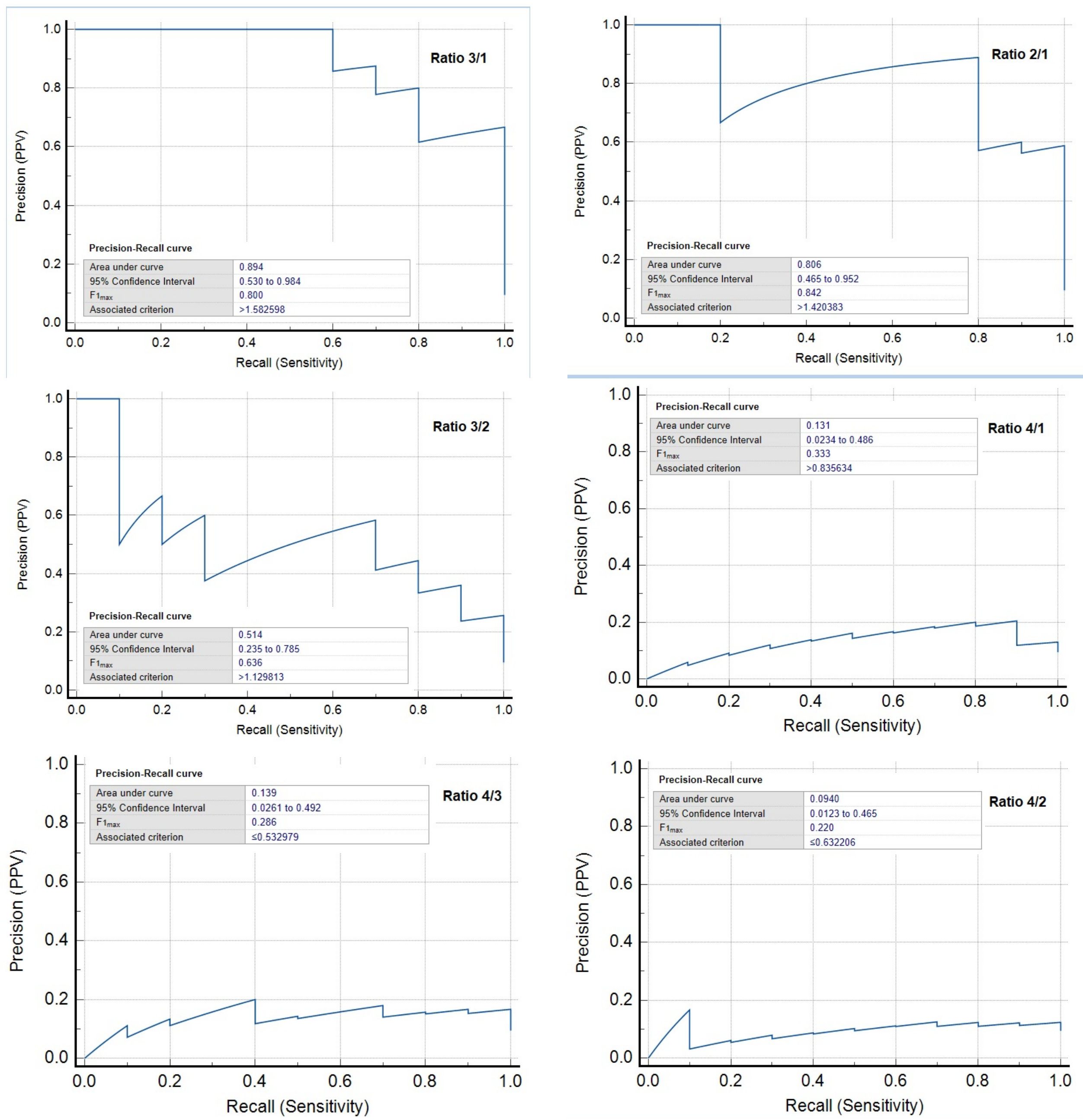
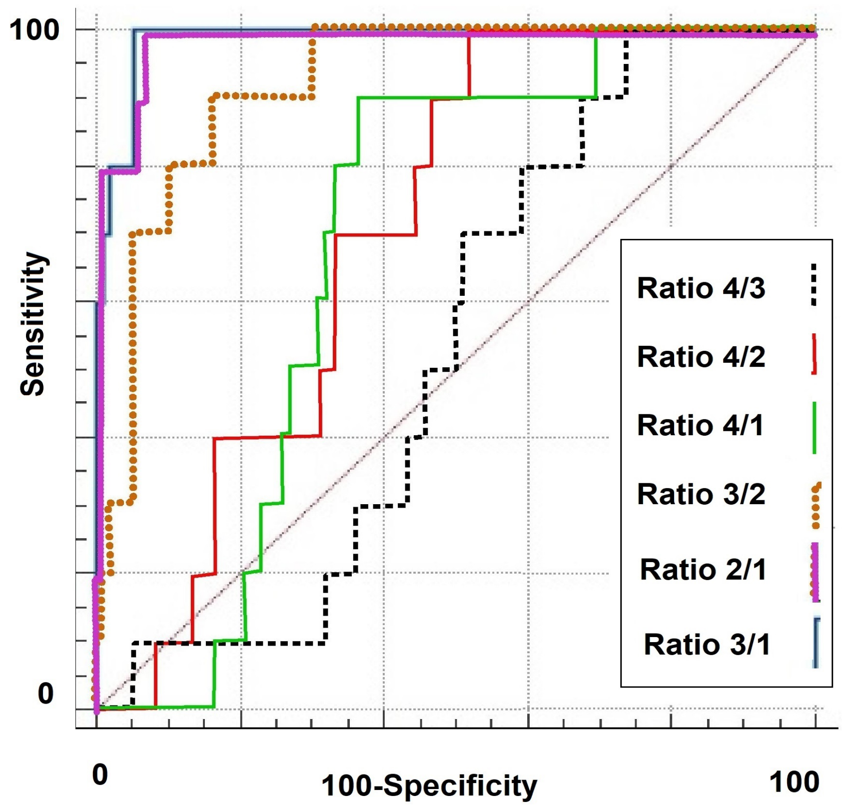


Figure 9. Precision-recall curves of the explanatory variables.

Table 3. Variable performance items (sensitivity, specificity, accuracy & AUC).

Variable	Cutoff value	Sensitivity	Specificity	Accuracy	AUC	Remarks
Ratio 3/1	1.568918	100.00	94.79	0.9479	0.986	excellent discrimination
Ratio 2/1	1.398237	100	92.71	0.9271	0.980	excellent discrimination
Ratio 3/2	1.117008	90.00	83.33	0.7333	0.919	Acceptable discrimination
Ratio 4/3	0.553021	100.00	47.92	0.4792	0.703	less discrimination
Ratio 4/1	0.835634	90.00	63.54	0.5354	0.683	less discrimination
Ratio 4/2	0.632206	100	26.04	0.2604	0.534	less discrimination



s

Figure 10. The ROC for SPOT band ratio classifiers.

5. Conclusions

The present study successfully evaluates the band ratio technique using the ROC curves. It utilized the SPOT 5 satellite data and the Random Forest (RF) model to predict the Fe-Ti mineralization. A set of band ratio images are prepared and used as input variables. The RF model was trained according to the target variable which contains mineralized and non-mineralized locations. The following conclusions are reached: 1) the 3/1 band ratio image shows excellent discrimination of Fe-Ti mineralization with AUC values of 0.986; 2) the prediction accuracy of the RF model reached up to 97%; 3) a new image map shows the distribution of the Fe-Ti mineralization is generated; 4) the present study proved the usefulness of the RF algorithm for the prediction of Fe-Ti mineralization.

References

[1]. Madani, A. (2001). Geological studies and remote sensing applications on Wadi Natash volcanic, Eastern Desert, Egypt. *PhD thesis, Faculty of Science, Cairo University.*

[2]. Tommaso, I., & Rubinstein, N. (2007). Hydrothermal alteration mapping using ASTER data in the Infiernillo porphyry deposit, Argentina. *Ore Geol. Rev., 32,*275-290.

[3]. Madani, A. (2009). Utilization of landsat ETM+ data for mapping gossans and iron rich zones exposed at Bahrah Area, Western Arabian Shield, Saudi Arabia. *JKAU Earth Sci 20(1),* 35–49.

[4]. Madani, A., & Emam, A. (2011). SWIR ASTER band ratios for lithological mapping and mineral exploration: a case study from El Hudi area, South Eastern Desert, Egypt. *Arab J Geosci 4,* 45–52.

[5]. Nouri, R., Jafari, MR., Arian, F., & Feizi, F. (2012). Hydrothermal alteration zone identification based on remote sensing data in the Mahn Area, West of Qazvin Province Iran. *World Acad Sci Eng Technol, 67,* 479-482.

[6]. Madani, A., & Harbi, H. (2012). Spectroscopy of the mineralized tonalite–diorite intrusions, Bulghah gold mine area, Saudi Arabia: effects of opaques and alteration products on FieldSpec data. *Ore Geol Rev 44,*148–157.

[7]. Harbi, H., & Madani, A. (2013). Utilization of SPOT 5 data for mapping gold mineralized diorite–tonalite intrusion, Bulghah gold mine area, Saudi Arabia. *Arab J Geosci,.*

[8]. Tayebi, MH., Hashemi, TM., & Keller, VR. (2014). Alteration mineral mapping with ASTER data by integration of coded spectral ratio imaging and SOM neural network model. *Turkish J. Earth Sci., 23,* 627-644,

[9]. Madani, A. (2015). Spectroscopy of olivine basalts using FieldSpec and ASTER data: a case study from Wadi Natash volcanic field, South Eastern Desert, Egypt. *J Earth Syst Sci 124(7),* 1475–1486.

[10]. Sadiya, TB., Halilu, AS., Asmah, TF., Agu, NV., Nsofor, CJ., Sanusi, M., Aliyu, I., & Ibrahim, I. (2015). Hydrothermal alteration mapping in Ijio, Oyo state, Nigeria using satellite imagery & remote sensing technique. *SSRG Int. J. Geo Inform. Geolog. Sci. (SSRG-IJGGS), 2,* 2.

[11]. Zhang, Z., Zuo, R., & Xiong, Y. (2016). A comparative study of fuzzy weights of evidence and random forests for mapping mineral prospectivity for skarn-type Fe deposits in the southwestern Fujian metallogenic belt, China. *Science China Earth Sciences 59,*556–572.

[12]. Rajendran, S., & Nasir, S. (2017). Characterization of STER spectral bands for mapping of alteration zones of volcanogenic massive sulphide deposits. *Ore Geology Reviews 88,* 317–335.

[13]. Zamyad, M., Afzal, P., Pourkermani, M., Nouri, R., & Jafari, MR. (2019). Determination of hydrothermal alteration zones using remote sensing methods in Tirka Area, Toroud, NE Iran. *J. Indian Soc. Remote Sens., 47(11),* 1817-1830.

[14]. Traore, M., Takodjou, Wambo, JD., Ndepete, CP., Tekin, S., Pour, AB., & Muslim, AM. (2020). Lithological and alteration mineral mapping for alluvial gold exploration in the south east of Birao area, Central African Republic using Landsat-8 Operational Land Imager (OLI) data. *J. Afr. Earth Sci.* p.170.

[15]. Shi, X., Al-Arifi, N., Abdelkareem, M., & Abdalla, F. (2020). Application of remote sensing and GIS techniques for exploring potential areas of hydrothermal mineralization in the central Eastern Desert of Egypt. *J. Taibah Univ. Sci. 14,* 1421–1432.

[16]. Sekandari, M., Masoumi, I., Pour, AB., Muslim, AM., Rahmani, O., Hashim, M., Zoheir, B., Pradhan, B., Misra, A., & Aminpour, SM. (2020). Application of Landsat-8, Sentinel-2, ASTER andWorldView-3 Spectral Imagery for Exploration of Carbonate-Hosted Pb-Zn Deposits in the Central Iranian Terrane (CIT). *Remote Sens. 12,* 1239.

[17]. El Sobky, MA., Madani, AA., & Surour, AA. (2020). Spectral characterization of the Batuga granite pluton, South Eastern Desert, Egypt: influence of lithological and mineralogical variation on ASD Terraspec data. *Arabian Journal of Geosciences 13,* 1246.

[18]. Sadek, MF., El-Kalioubi, BA., Ali-Bik, MW., El Hefnawi, MA., & Elnazer, AA. (2020). Utilizing Landsat-8 and ASTER data in geologic mapping of hyper-arid mountainous region: case of Gabal Batoga area, South Eastern Desert of Egypt. *Environ Earth Sci 79(5),* 1–14.

[19]. El-Din, GM., El-Noby, ME., Abdelkareem, ZM., & Hamimi, Z. (2021). Using multispectral and radar remote sensing data for geological investigation, Qena-Safaga Shear Zone, Eastern Desert, Egypt. *Arab. J. Geosci. 14,* 997.

[20]. Madani, AA., Harbi, HM., El-Dougdoug, AA., Surour, AA., & Ahmed, AH. (2021). Spectral Characteristics of Listvenites and Serpentineites Along Ophiolite-Decorated Megashears (Suture Zones) in the Arabian Shield Using ASD Fieldspec and Satellite Data *in:Hamimi (ed) The Geology of the Arabian-Nubian Shield, Springer, Cham,* 559-583.

[21]. Ige, O., Tende, A., Bale, R., Gajere, J., & Aminu, M. (2022). Spatial mapping of hydrothermal alterations and structural features for gold and cassiterite exploration. *Scientific African Volume 17, September 2022, e01307.*

[22]. Abd El-Fatah, A., Madani, A., Surour, A., & Azer, M. (2023). Integration of Landsat-8 and Reflectance Spectroscopy data for Mapping of Late Neoproterozoic Igneous Ring Complexes in an Arid Environment: a Case Study of Gebel El-Bakriyah Area, Eastern Desert, Egypt. *Journal of Mining and Environment (JME).*

[23]. Ali, D., & Frimpong, S. (2020). Artificial intelligence, machine learning and process automation: Existing knowledge frontier and way forward for mining sector. *Artif. Intell. Rev. 53,* 6025–6042.

[24]. Brown, WM., Gedeon, TD., Groves, D., & Barnes, RG. (2000). Artificial neural networks: A new method for mineral prospectivity mapping. *Australian Journal of Earth Sciences 47,* 757–770.

[25]. Carranza, EJ. (2014). Data-driven evidential belief modeling of mineral potential using few prospects and evidence with missing values. *Natural Resources Research 24,* 291–304.

[26]. Porwal, A., Carranza, EJ., Hale, M. (2003). Artificial neural networks for mineral-potential mapping: A case study from Aravalli Province, western India. *Nat Resour Res 12,* 155–171.

[27]. Fung, CC., Iyer, V., Brown, W., & Wong, KW. (2005). Comparing the performance of different neural networks architectures for the prediction of mineral prospectivity. *In Proceedings of the fourth international conference on machine learning and cybernetics, Guangzhou* 394–398.

[28]. Abedi, M., Norouzi, GH., & Bahroudi, A. (2012). Support vector machine for multi- classification of mineral prospectivity areas. *Computers Geosciences 46, 272*–283.

[29]. Rodriguez-Galiano, VF., Chica-Olmo, M., & Chica-Rivas, M. (2014). Predictive modelling of gold potential with the integration of multisource information based on random forest: A case study on the Rodalquilar area, Southern Spain. *Int J Geogr Inf Sci 28,* 1336–1354.

[30]. Carranza, EJ., & Laborte, AG. (2015b). Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). *Computers & Geosciences 74,* 60–70.

[31]. Zhang, Z., Zuo, R., & Xiong, Y. (2015). A comparative study of fuzzy weights of evidence and random forests for mapping mineral prospectivity for skarn-type fe deposits in the southwestern Fujian metallogenic belt, China. *Sci China Earth Sci 59,* 556–572.

[32]. Carranza, EJ., & Laborte, AG. (2016). Data-driven predictive modeling of mineral prospectivity using random forest: a case study in Catanduanes Island (Philippines). *Natural Resources Research 25,* 35-50.

[33]. Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geol Rev 71,* 804–818.

[34]. McKay, G., & Harris, JR. (2016). Comparison of the data-driven random forests model and a knowledge-driven method for mineral prospectivity mapping: A case study for gold deposits around the Huritz Group and Nueltin Suite, Nunavut, Canada. *Natural Resources Research 25,* 125–143.

[35]. Zhang, Z., Zuo, R., & Xiong, Y. (2016). A comparative study of fuzzy weights of evidence and random forests for mapping mineral prospectivity for skarn-type Fe deposits in the southwestern Fujian metallogenic belt, China. *Science China Earth Sciences 59,* 556–572.

[36]. Hariharan, S., Tirodkar, S., Porwal, A., Bhattacharya, A., & Joly, A. (2017). Random forest-based prospectivity modelling of greenfieldterrains using sparse deposit data: An example from the Tanami Region,Western Australia. *Nat Resour Res 26,* 489–507.

[37]. Wang,Y., Fang, Z., & Hong, H. (2019). Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Sci Total Environ 666,* 975–993.

[38]. Sun, T., Chen, F., Zhong, L., Liu, W., & Wang, Y. (2019). GIS-based mineral prospectivity mapping using machine learning methods: A case study from Tongling ore district, eastern China. *Ore Geol Rev 109,* 26–49.

[39]. Chen, C., He, B., & Zeng, Z. (2014). A method for mineral prospectivity mapping integrating C4. 5 decision tree, weights-of-evidence and m-branch smoothing techniques: a case study in the eastern Kunlun Mountains, China. *Earth Science Informatics 7,* 13–24.

[40]. Sun, T., Hui, Li., Kaixing, Wu., Fei Chen, Zhong Zhu, & Zijuan Hu (2020). Data-Driven Predictive Modelling of Mineral Prospectivity Using Machine Learning and Deep Learning Methods: A Case Study from Southern Jiangxi Province, China. *Minerals, 10(2),*102.

[41]. Jooshaki, M., Nad, A., & Michaux, S. (2021). A Systematic Review on the Application of Machine Learning in Exploiting Mineralogical Data in Mining and Mineral Industry. *Minerals 11,* 816.

[42]. Li, S., Chen, J., & Liu, C. (2022). Overview on the Development of Intelligent Methods for Mineral Resource Prediction under the Background of Geological Big Data. *Minerals 12,* 616.

[43]. Farhadi, S., Afzal, P., Boveiri Konari, M., Daneshvar Saein, L., & Sadeghi, B., (2022). Combination of machine learning algorithms with concentration-area fractal method for soil geochemical anomaly detection in sediment-hosted Irankuh Pb-Zn deposit, Central Iran. *Minerals 12(6),* 689.

[44]. Afzal, P., Farhadi, S., Konari, MB., Meigooni, MS., & Saein, LD., (2022). Geochemical anomaly detection in the Irankuh District using Hybrid Machine learning technique and fractal modeling. *Geopersia 12(1),* 191-199.

[45]. Farhadi, S., Tatullo, S., Konari, MB., & Afzal, P., (2024). Evaluating StackingC and ensemble models for enhanced lithological classification in geological mapping. *Journal of Geochemical Exploration 260,* 107441.

[46]. Bonham-Carter, G. (1994). Geographic information systems for geoscientists: Modelling with GIS. *Oxford:Pergamon Press.*

[47]. Carranza, EJ., Mangaoang, JC., & Hale, M. (1999). Application of mineral exploration models and GIS to generate mineral potential maps as input for optimum land-use planning in the Philippines. *Natural Resources Research 8,* 165–173.

[48]. Singer, DA., & Kouda, R. (1999). A Comparison of the weights of evidence method and probabilistic neural networks. *Natural Resources Research 8,* 87–298.

[49]. Asadi, HH., & Hale, M. (2001). A predictive GIS model for mapping potential gold and base metal mineralization in Takab area, Iran. *Comput Geosci 27,* 901–912.

[50]. Chica-Olmo, M., Abarca, F., & Rigol, JP. (2002). Development of a Decision Support System based on remote sensing and GIS techniques for gold-rich area identification in SE Spain. *International Journal of Remote Sensing, 23(22),* 4801-4814.

[51]. Harris, D., Zurcher, L., Stanley, M., Marlow, J., & Pan, G. (2003). A comparative analysis of favorability mappings by weights of evidence, probabilistic neural networks, discriminant analysis, and logistic regression. *Nat Res Res 12,* 241-255.

[52]. Carranza, EJ., Woldai, T., & Chikambwe, EM. (2005). Application of data-driven evidential belief functions to prospectivity mapping for aquamarine-bearing pegmatites, Lundazi District, Zambia. *Natural Resources Research 14,* 47–63.

[53]. Caumon, GJ., Ortiz, O., & Rabeau, (2006). A comparative study of three data-driven Mineral Potential Mapping techniques. *Int. Assoc. for Mathematical Geology XIth International Congress Université de Liège – Belgium Liège – September, 3rd - 8th S13-05.*

[54]. Carranza, EJ., van Ruitenbeek, FJ., Hecker, C., van der Meijde, M., & van der Meer, FD. (2008). Knowledge-guided data-driven evidential belief modeling of mineral prospectivity in Cabo de Gata, SE Spain. *Int J Appl Earth Obs 10,* 374–387.

[55]. Partington, G. (2010). Developing models using GIS to assess geological and economic risk: An example from VMS copper gold mineral exploration in Oman. *Ore Geology Reviews 38,* 197–207.

[56]. Madani, A. (2011). Knowledge-driven GIS modeling technique for gold exploration, Bulghah gold mine area, Saudi Arabia. *Egypt J Remote Sensing and Space Sci 14:91–97.*

[57]. Joly, A., Porwal, A., & McCuaig, TC. (2012). Exploration targeting for orogenic gold deposits in the Granites-Tanami Orogen: Mineral system analysis, targeting model and prospectivity analysis. *Ore Geology Reviews 48,* 349–383.

[58]. Ford, A., Miller, JM., & Mol, AG. (2015). A comparative analysis of weights of evidence, evidential belief functions, and fuzzy logic for mineral potential mapping using incomplete data at the scale of investigation. *Natural Resources Research 25,* 19–33.

[59]. Carranza, EJ., & Laborte, AG. (2015a). Data-driven predictive mapping of gold prospectivity, Baguio district, Philippines: application of Random Forest algorithm. *Ore Geology Reviews 71,* 777–787.

[60]. Harris, JR., Grunsky, E., Behnia, P., & Corrigan, D. (2015). Data- and knowledge-driven mineral prospectivity maps for Canada's North. *Ore Geology Reviews 71:788-803.*

[61]. Sun, T., Wu, K., Chen, L., Liu, W., Wang, Y., & Zhang, C. (2017). Joint application of fractal analysis and weights-of-evidence method for revealing the geological controls on regional-scale tungsten mineralization in southern Jiangxi Province, China. *Minerals 7,* 243.

[62]. Olasehinde, A., & Ashano, E. (2021). Data Driven Predictive Modelling of Mineral Prospectivity Using Principal Component Analysis: A Case Study of Riruwai Complex. *Advances in Applied Science Research 12(7),* 33.

[63]. Atef, A., Madani, A., Surour, A., & Azer, M. (2023). A predictive GIS model for mapping gold- and copper-bearing alteration zones at South Gabal Um Monqul and Gabal Al Kharaza prospects, north Eastern Desert, Egypt. *Journal of mining and environment.*

[64]. Taha, AM., Xi, Y., He, Q., Hu, A., Wang, S., & Liu, X. (2023), Investigating the Capabilities of Various Multispectral Remote Sensors Data to Map Mineral Prospectivity Based on Random Forest Predictive Model: A Case Study for Gold Deposits in Hamissana Area, NE Sudan. *Minerals 13,* 49.

[65]. Chevremont, P., & Johan, Z., (1981). Wadi Khamal-Wadi Murattijah Ultramafic-Mafic Layered Complex, Saudi Arabian. *Deputy Ministry for Mineral Resources: Open File Report BRGMOF-01,36,* 143.

[66]. Harbi, HM. (2008). Geology and Lithostratigraphy of the Ultramafic-Mafic Rocks and Associated Mineralizations, Wadi Khamal Area, West-Central Arabian Shield. *JKAU: Earth Sci 19,* 119-157.

[67]. Bache, J., & Chevermont, P. (1976). Mineral Investigations for Nickel and Copper in the Wadi Khamal Region: *Bureau de Recherches Geologiques et Minieres, Open-File-Report-JED-OR-79(8),*37.

[68]. Pellaton, C. (1979). Geologic Map of the Yanbu Al Bahr Quadrangle, Sheet 24C, Kingdom of Saudi Arabia, *Saudi Dir. Gen. Miner. Resour. Geologic Map GM-48-A,* 16.

[69]. Hashem, WB. (1981). The Geology of the Wadi Khamal Basic Layered Intrusion, Yanbou Al Bahr, Saudi Arabia*, Unpublished Ph.D. Thesis, University of Bristol UK.*

[70]. Al Ghamdi, AM. (1994). Mineralization and Associated Platinum Group Elements in mafic-ultramafic rocks, Northwestern Arabian Shield, K.S.A: *Unpublished Ph.D. Thesis, King Abdulaziz University, Faculty of Earth Sciences.*

[71]. Eldougdoug, A., Abd El-Rahman, Y., & Harbi, H. (2020). The Ediacaran Post-Collisional Khamal Gabbro-Anorthosite Complex from the Arabian Shield and Its Fe-Ti-P Ore: An Analogy to Proterozoic Massif-Type Anorthosites. *Lithos* 372–373, 105674.

[72]. Abuamarah, B A., Alshehri, F., Azer, MK., & Asimow, PD. (2023). Loveringite from the Khamal Layered Mafic Intrusion: The First Occurrence in the Arabian Shield, Northwest Saudi Arabia. *Minerals 13,*172.

[73]. Galaup, M., & Dupuy, S. (2003). Benefits of SPOT 5 imagery for town planning with new adapted processing techniques. *Proceedings of IGARSS 2003, IEEE International Geoscience and Remote Sensing symposium.*

[74]. Clandillon, S., Yesou, H., & Meyer, C. (2003). Benefits of SPOT 5 HR and VHR data for forest management and windfall damage mapping. *Proceedings of IGARSS, IEEE International Geoscience and Remote Sensing symposium.*

[75]. Yésou, H, Clandillon, S., Allenbach, B., Bestault, C., De Fraipont, P., Inglada, J., & Favard, C. (2003). A constellation of advantages with SPOT SWIR and VHR SPOT 5 data for flood extent mapping during the September 2002 Gard event (France). *Proceedings of IGARSS 2003, IEEE International Geoscience and Remote Sensing symposium.*

[76]. Kakiuchi, H., Onaka, M., Asai, M., & Itoh, F. (2004). Topographic mapping at scale of 1:25,000 using SPOT 5 satellite imagery. *ISPRS 2004, Istanbul, Turkey.*

[77]. Ferreira, F. (2004). Using SPOT 5 to improve census cartography. *ISPRS, Istanbul, Turkey.*

[78]. Retière, A., Senegas, O., Parriaux, A., Haeberlin, Y., & Turberg, P. (2004). Validation of SPOT 5 satellite imagery for geological hazard identification and risk assessment for landslides, mud- and debris flows in Matagalpa, Nicaragua. *ISPRS 2004, Istanbul, Turkey.*

[79]. Fajji, NG., Palamuleni, LG., & Mlambo, V. (2018). Application of SPOT Imagery for Landcover Mapping and Assessing Indicators of Erosion and Proportion of Bareground in Arid and Semi-arid Environment. *J Remote Sens GIS 7,* 240.

[80]. Breiman, L. (2001). Random forests. *Machine learning 45,* 5-32.

[81]. Madani, A., & Niyazi, B. (2023). Groundwater Potential Mapping Using Remote Sensing and Random Forest Machine Learning Model: A Case Study from Lower Part of Wadi Yalamlam,Western Saudi Arabia. *Sustainability 2023, 15,* 2772.

[82]. O'Brien, JJ., Spry, PG., Nettleton, D., Xu, R., & Teale, GS. (2015). Using Random Forests to distinguish gahnite compositions as an exploration guide to Broken Hill-type Pb–Zn–Ag deposits in the Broken Hill domain, Australia. *J Geochem Explor 149,* 74–86.

[83]. Long, T., Zhou, Z., Hancke, G., Bai, Y., & Gao, Q. (2022). A Review of Artificial Intelligence Technologies in Mineral Identification: Classification and Visualization. *J. Sens. Actuator Netw. 2022, 11,* 50.

[84]. Davis, J., & Goadrich, M., (2006). The relationship between Precision-Recall and ROC curves. *Proceedings of the 23rd international conference on Machine learning.* 233–240.

[85]. Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE 10(3): e0118432.*

[86]. Fawcett, T. (2005). An introduction to ROC analysis. *Pattern Recognition Letters 27,* 861–874.

ارزیابی تکنیک نسبت نواری برای پیش‌بینی کانی‌سازی آهن-تیتانیوم با استفاده از مدل یادگیری ماشینی گروهی: مطالعه موردی از منطقه خمال، عربستان غربی

احمد علی مدنی

گروه زمین شناسی، دانشکده علوم، دانشگاه قاهره، جیزه، مصر

ارسال ۲۰۲۴/۰۴/۲۴، پذیرش ۲۰۲۴/۰۵/۱۱

* نویسنده مسئول مکاتبات: aamadani@sci.cu.edu.eg

چکیده:

نوآوری در اکتشاف معدنی یا در ساخت مدل‌های جدید کانسار یا توسعه تکنیک‌های جدید مورد استفاده برای مکان‌یابی ذخایر معدنی رخ می‌دهد. نسبت باند تکنیک پردازش تصویر است که برای اکتشاف مواد معدنی توسعه یافته است. مطالعه حاضر رویکرد جدیدی را ارائه می‌کند که برای ارزیابی تکنیک نسبت نواری برای تشخیص و پیش‌بینی کانی‌سازی آهن-تیتانیوم در معرض در منطقه خمال، غرب عربستان سعودی با استفاده از مدل جنگل تصادفی (RF) و داده‌های ماهواره‌ای SPOT-5 استفاده می‌شود. تصاویر نسبت باند SPOT-5 تهیه شده و به عنوان متغیرهای توضیحی استفاده می شود. متغیر هدف تهیه شده است که در آن (٪۷۰) از مکان های هدف برای آموزش و بقیه برای اعتبار سنجی استفاده می شود. یک ماتریس سردرگمی و منحنی‌های فراخوان دقیق برای ارزیابی عملکرد مدل RF و منحنی‌های ویژگی‌های عملیاتی گیرنده (ROC) برای رتبه‌بندی تصاویر نسبت باند استفاده می‌شوند. نتایج نشان داد که تصاویر نسبت باند ۱/۳، ۱/۲ و ۲/۳ تمایز بسیار خوبی را با مقادیر AUC به ترتیب ۰.۹۸۶، ۰.۹۸۰ و ۰.۹۱۹ نشان می‌دهند. مطالعه حاضر با موفقیت تصویر نسبت باند ۱/۳ را به عنوان بهترین طبقه‌بندی کننده انتخاب کرده و یک نقشه تصویر کانی‌سازی Fe-Ti جدید ارائه می‌کند. مطالعه حاضر سودمندی طبقه‌بندی‌کننده جنگل تصادفی را برای پیش‌بینی کانی‌سازی Fe-Ti با دقت ۹۷ درصد اثبات کرد.

کلمات کلیدی: مدل پیش‌بینی مبتنی بر هوش مصنوعی، الگوریتم جنگل تصادفی، داده‌های SPOT-5، کانی‌سازی Fe-Ti، عربستان سعودی غربی.