Computers & Geosciences 193 (2024) 105738

Contents lists available at ScienceDirect

Computers and Geosciences

journal homepage: www.elsevier.com/locate/cageo

Enhanced lithological mapping in arid crystalline regions using explainable AI and multi-spectral remote sensing data

Hesham Morgan [a,b,c], Ali Elgendy [c], Amir Said [c], Mostafa Hashem [d], Wenzhao Li [a,b], Surendra Maharjan [a,b], Hesham El-Askary [a,b,e,*]

[a] Schmid College of Science and Technology, Chapman University, One University Drive, Orange, CA, 92866, USA
[b] Earth Systems Science and Data Solutions Lab, Chapman University, Orange, CA, 92866, USA
[c] Department of Geology, Faculty of Science, Cairo University, P.O.B. 12613, Giza, Egypt
[d] Dell Technologies, Cairo, Egypt
[e] Department of Environmental Sciences, Faculty of Science, Alexandria University, Moharem Bek, Alexandria, 21522, Egypt

ARTICLE INFO

Keywords:
Imbalanced datasets
Lithological mapping
Random forest classifier
Variable importance
Explainable artificial intelligence
SHapley additive exPlanations

ABSTRACT

Lithological classification is essential for understanding the spatial distribution of rocks, especially in arid crystalline areas. Artificial intelligence (AI) recent advancements with multi-spectral satellite imagery have been utilized to enhance lithological mapping in these areas. Here we employed different AI models namely, Support Vector Machine (SVM), Random Forest Classification (RFC), Logistic Regression, XGBoost, and K-nearest neighbors (KNN) for lithological mapping. This was followed by the application of explainable AI (XAI) for lithological discrimination (LD) which is still not widely explored. Based on the highest accuracy and F1 score of the previously mentioned models, RFC model outperformed all of them, and hence, it was integrated with XAI, using the SHapley Additive exPlanations (SHAP) method.

This approach successfully identified critical multi-spectral features for LD in arid crystalline zones when applied on the Landsat-8, Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), and SRTM-DEM datasets covering the Hammash and the Wadi Fatimah areas in Egypt and the Kingdom of Saudi Arabia, respectively. Field validation in the Hammash area confirmed the RFC model's efficacy, achieving a satisfactory 94% overall accuracy for 18 features. SHAP was able to identify the top ten features for proper LD over the Hammash area with 90.3% accuracy despite the complex nature of the ophiolitic mélange. For validation purposes, RCF was then utilized in the Wadi Fatimah region, using only the top 10 critical features rendered from the SHAP analysis. It performed well and had 93% accuracy. Notably, XAI/SHAP results indicated that elevation data, Landsat-8's Green Band (B3), and the two ASTER SWIR bands (B5 and B6) were essential and significant for identifying island arc rocks. Moreover, the SHAP model effectively delineated complex mélange matrices, primarily using ASTER SWIR band (B8). Our findings highlight the successful combination of RFC with XAI for LD and its potential utilization in similar arid crystalline environments worldwide.

1. Introduction

Remote sensing (RS) plays a crucial role in lithological mapping, as the use of satellite data has significantly advanced geoscientists' ability to visualize and understand Earth's geological features. RS instruments capture reflected or absorbed electromagnetic spectra across different wavelengths, enabling the identification of objects based on their characteristic responses influenced by various chemical and physical properties (Clark and Roush, 1984; Cloutis, 1996). Until recently, several spectral-driven methods have been developed for lithological discrimination based on pixel-oriented concepts such as band rationing (BR) (Alifu et al., 2015; Pal et al., 2020; Sultan et al., 1986), spectral angle mapper (SAM) (Validabadi Bozcheloei and Tangestani, 2019; Zhang and Li, 2014), match filtering (MF) (Mehr et al., 2013; Rowan and Mars, 2003), and principal component analysis (PCA) (Amer et al., 2010; Khalifa et al., 2021). In practical applications, spectral responses are influenced by factors such as spatial and spectral resolutions, atmospheric effects, soil presence, weathering, sub-pixel variations in

* Corresponding author. Schmid College of Science and Technology, Chapman University, One University Drive, Orange, CA, 92866, USA.
E-mail address: elaskary@chapman.edu (H. El-Askary).

https://doi.org/10.1016/j.cageo.2024.105738
Received 2 July 2024; Received in revised form 11 October 2024; Accepted 13 October 2024
Available online 16 October 2024

mineralogical and chemical composition, and vegetation cover. Even with extensive pre-processing steps like calibration, radiometric and geometric corrections, the spectral properties captured by RS for a particular rock type can differ markedly from the laboratory spectra of pure samples (Girija and Mayappan, 2019; Gupta, 2017).

In the current era of artificial intelligence (AI), the exponential growth in computing power and advancements in algorithms have made machine learning (ML) a widely adopted solution for tackling a range of geoscience challenges, including lithological discrimination (LD) (Karpatne et al., 2019; Morgan et al., 2023; Ouyang et al., 2024; Zhang et al., 2022).In geology and mineral exploration, ML plays a crucial role in LD, enhancing the objectivity and accuracy of lithological classification (Shayeganpour et al., 2021a; Xie et al., 2018). However, challenges persist in recognizing certain lithology types, such as complex crystalline rocks. Multi-spectral bands of passive satellite images, such as ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) and Landsat-8 OLI (Operational Land Imager), integrated with ML algorithms, have been widely used for LD (Shirmard et al., 2022). Additionally, integrating these with the Digital Elevation Model (DEM) data can provide necessary bands for LD of crystalline rocks in arid mountainous regions (Baid et al., 2023; Ge et al., 2018b; Marzouki and Dridri, 2023).

Supervised ML classification (SMLC) models are widely used for LD, relying on training datasets containing labeled samples that provide essential information about the target variable (Kotsiantis, 2007; Witten et al., 2002). Notable SMLC methods used in LD include the maximum likelihood classifier (MLC)(Behnia et al., 2012; Ge et al., 2018b), K-nearest neighbors (K-NN) (Cracknell and Reading, 2014; Ge et al., 2018b), naïve Bayes (NB)(Cracknell and Reading, 2014; Horrocks et al., 2015), support vector machines (SVM) (Bachri et al., 2019; Shayeganpour et al., 2021b), Random Forest Classifier (RFC) (He et al., 2015; Kuhn et al., 2018), and Artificial Neural Network (ANN) (Cracknell and Reading, 2014; Latifovic et al., 2018). Although SVM, RFC, and ANN models have performed best based on prior research, the RFC algorithm is preferred for LD classification for several reasons: (a) it reduces overfitting, which usually occurs in SVM; (b) ANN is more suitable for hyperspectral satellite images (Liu et al., 2021); (c) it works effectively with high-dimensional data (multi-spectral RS data); (d) it performs well in arid crystalline rocks (Cracknell and Reading, 2014; Kuhn et al., 2018); (e) it can handle imbalanced data where ground truth points are scarce in uncommon and limited lithological units; (f) it exhibits notable resilience against outliers, can remove the "black-box" constraint of artificial neural networks, and provides an innovative approach to LD by assessing the relative importance of effective features and detecting the most significant ones (Palczewska et al., 2014); (g) it has limited bias and variance due to the vast number of trees.

Previous literature has shown that integrating state-of-the-art ML algorithms yields superior outcomes (Ge et al., 2018b; Temenos et al., 2023). Understanding and explaining the model's details can significantly aid in assessing its capabilities. Numerous AI-based systems, particularly those based on ML, face criticism for being perceived as "black boxes." These systems showcase impressive decision-making performance based on their learned models but lack transparency in elucidating the rationale or mechanism behind their decisions (Linardatos et al., 2021). Traditionally, much of the previous AI research on LD has utilized these tools without delving deeply into the factors influencing their algorithmic behavior. Therefore, model explainability and interpretability are crucial for trusted AI. Explainable AI (XAI) techniques have become increasingly important in ML, especially for applications in critical domains like healthcare and finance (Sierra-Botero et al., 2024). These techniques provide interpretability and transparency to complex models, allowing stakeholders to understand and trust the decision-making process. Among the commonly used explanatory techniques in ML are Local Interpretable Model-agnostic Explanations (LIME)(Garreau and Luxburg, 2020), Feature Importance Scores (Letoffe et al., 2024; Little et al., 2024), and SHapley Additive exPlanations (SHAP) (Arslan et al., 2022). LIME is a popular technique for explaining individual predictions by approximating the complex model locally with an interpretable model. It works by perturbing the input and observing how the predictions change, then fitting a simple model around the prediction to be explained (Garreau and Luxburg, 2020, 2020; Zafar and Khan, 2019; Z. Zhou et al., 2021). Feature importance scores, such as Gini importance or permutation importance, provide a global view of which features contribute most to a model's predictive accuracy (Little et al., 2024). On the other hand, SHAP values, based on Shapley values from cooperative game theory, have gained significant popularity in recent years. SHAP offers several advantages, including a solid theoretical foundation for fairly attributing predictions among features, the ability to provide both local and global explanations, consistency in ensuring that changes in feature importance align with changes in the model, and model-agnostic applicability(Arslan et al., 2022; Futagami et al., 2021). SHAP has been successfully applied in various domains, including healthcare for predicting post-stroke adverse mental outcomes, finance for enhancing interpretability in time-series forecasting models, and cybersecurity for explaining deep learning models in malware classification. While LIME, feature importance scores, and SHAP are all valuable explanatory techniques, SHAP is often preferred for several reasons (Arslan et al., 2022). However, the choice of technique often depends on the specific use case and requirements. LIME may be preferred for quick, intuitive explanations of individual predictions, while feature importance scores are useful for a simple, global understanding of feature relevance (Arslan et al., 2022). SHAP is ideal for comprehensive model interpretation, especially in high-stakes applications.

Therefore, from the above, it is evident that applying explanatory machine learning (XML) in remote sensing image classification would provide better explanations for the model's predictions rather than just focusing on achieving high classification accuracy. This is especially relevant in domains such as environmental monitoring, urban planning, and agriculture, where understanding the justification behind a categorization is as vital as the classification itself (Kakogeorgiou and Karantzalos, 2021; Temenos et al., 2023). Feature importance analysis is a crucial technique in explanatory ML for determining the value of various variables (such as spectral bands) in formulating predictions. SHAP values (Futagami et al., 2021; Mangalathu et al., 2020; Zeng et al., 2020) can be used to determine which features have the most influence in identifying different forms of land cover or land use (H. H. Chen et al., 2023; Hosseiny et al., 2022; Temenos et al., 2023). We believe that XAI techniques can help identify the most important features or spectral bands that contribute to improving AI models for LD. This information is valuable for geologists and helps in (a) enhancing the understanding of spectral signatures associated with different rock types, (b) guiding future data collection efforts, and (c) improving the design of RS instruments for geological applications. Consequently, by understanding how the model makes decisions, researchers can identify weaknesses or biases in the system and refine the algorithms or training data as needed (Antonini et al., 2024; Chen and Zhang, 2023; Kong et al., 2024; Yang et al., 2023). This iterative process can lead to more accurate and reliable lithological mapping over time. To our knowledge, this research is the first to apply multi-spectral RS-based explanatory ML in lithological classification. Our innovative approach combines cost-effective, multidisciplinary methodologies, including Geographic Information System (GIS) integration, minimal RS feature engineering, the RFC model, and advanced techniques such as feature importance analysis and SHAP for explaining LD. Feature extraction (FE) in ML is the process of transforming raw data into a set of features that can assist ML algorithms to perform well(Khalid et al., 2014; Suhaidi et al., 2021). There are two main approaches to feature extraction: Manual feature extraction, which involves human experts using their domain knowledge to identify and select relevant features from the raw data, and Automated feature extraction, which uses algorithms and computational techniques to identify and extract relevant features from the data. In our Study, the

PCA is used as the automated feature extraction approach. XML is used to understand and interpret raw data without domain experts manually designing and selecting features based on their knowledge and understanding of the problem. The objective is to identify and clarify the key spectral features critical for LD in arid crystalline terrains characterized by complex geological structures. The Hammash area in the southeastern desert of Egypt, known for its diverse crystalline rock formations and substantiated by reliable field studies, was selected as our primary research site. We also apply our approach to a comparable area in western Saudi Arabia, namely, Wadi Fatimah, to validate the model's effectiveness and adaptability to similar geological settings. These insights are expected to apply to other regions with similar geological conditions regionally and globally.

2. Study areas

This study focuses on two crystalline arid areas located in the Arabian-Nubian Shield (ANS) in Egypt and Saudi Arabia. The ANS comprises Precambrian crystalline basement rocks formed by the Pan-African Orogeny. The ANS was subdivided into the Nubian Shield and Arabian Shield due to the opening of the Red Sea in the Early Neogene. In the Nubian Shield, the Hammash gold mine area was selected for this study, while the Wadi Fatimah area was chosen for investigation in the Arabian Shield (Fig. 1).

The Hammash area covers about 41 km^2 and includes a gold mine located 60 km south of the Idfu-Marsa Alam Road in the southeastern desert of Egypt. It is characterized by the Red Sea Mountains with igneous and metamorphic rocks, a barren and scarce water supply, and large drought wadis trending in various directions. Based on previous work (El-Desoky et al., 2021; Gharib et al., 2021; Helmy and Kaindl, 1999; Hilmy and Osman, 1989), field observations, and published geological maps (Abdel-Rahman et al., 2023; Aboelkhair et al., 2021; El-Desoky et al., 2021; Kassem and Rahim, 2014; Sakran et al., 2009) (Fig. 1A), the rocks in the Hammash area are arranged from oldest to youngest as follows: ophiolitic serpentinites and related rocks (talc--carbonate and quartz-carbonate rocks), meta-gabbro, ophiolitic

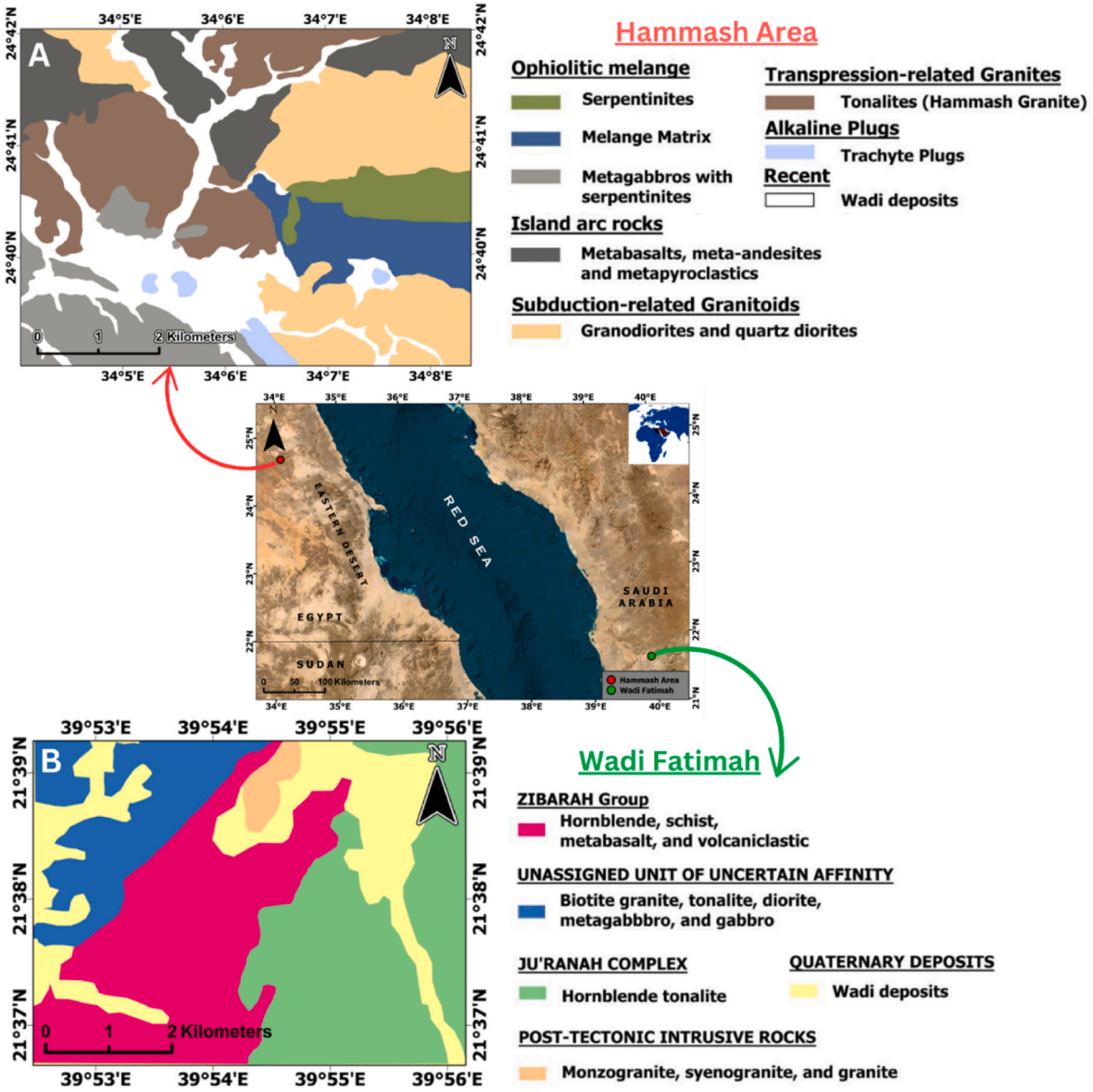

Fig. 1. High-resolution regional satellite image of southeastern Egypt and western Saudi Arabia, highlighting the geographic location of the study areas showing their geological maps; (a) Hammash goldmine area following (Abdel-Rahman et al., 2023; Said, 2010; Sakran et al., 2009); (b) a cropped area from the Wadi Fatimah region following (Alshehri and Abdelrahman, 2023; Moore and Al-Rehaili, 1989a,b).

mélange (a complex geological formation mainly composed of serpentinite and meta-gabbro blocks embedded in a quartz-feldspathic mélange matrix), island arc-related meta-volcanic rocks (meta-basalt, meta-pyroclastics, and meta-andesite), subduction-related granitoids (diorite and granodiorite), Hammash granite (mainly tonalites), dikes, trachyte plugs, and Wadi deposits (Quaternary deposit). A field survey was conducted by the co-author to study the petrology and mineralization in the Hammash mine area (Fig. 2) (Said, 2010).

The second study area encompasses 35 km^2 in the Wadi Fatimah area, near Makkah, western Saudi Arabia. It is very similar to the Hammash area in Egypt, as both areas are primarily composed of Precambrian crystalline rocks and are in arid regions. Based on previous studies (GRAINGER, 1992) and published geological maps (Alshehri and Abdelrahman, 2023; Moore, T.A. and Al-Rehaili, M.H.,1989) (Fig. 1b), this area has many complex geological formations, ordered from oldest to youngest as follows: Zibarah group (hornblende schist, meta-basalt, and volcaniclastics), an unassigned unit (a complex geological formation consisting of biotite granite, tonalite, diorite, and meta-gabbro), Ju'ranah complex group (mainly composed of hornblende tonalite), post-tectonic intrusive rocks (monzogranite, syenogranite, and granite), and Quaternary deposits (Wadi deposits).

3. Experimental design and layout

3.1. Data sources and software

Visible and near-infrared (VNIR) and shortwave-infrared (SWIR) radiometers of ASTER multi-spectral images were collected from Earth Science Data Systems (ESDS) Program (https://www.earthdata.nasa.gov/), while Landsat-8 OLI multi-spectral images and SRTM-DEM were acquired from USGS Earth Explorer (https://earthexplorer.usgs.gov/) (Table S1). The DEM is utilized in this study due to the mountainous terrain of the research area. The elevation range (360–524 m.a.s.l.) plays a crucial role in differentiating the wadi from the surrounding crystalline rock formations (Y. Y. Chen et al., 2023; Kumar et al., 2022; Yu et al., 2012). Additionally, the DEM can assist in distinguishing between different rock units, such as serpentinites and meta-gabbro, both mafic and ultramafic rocks in the study area, which are typically found

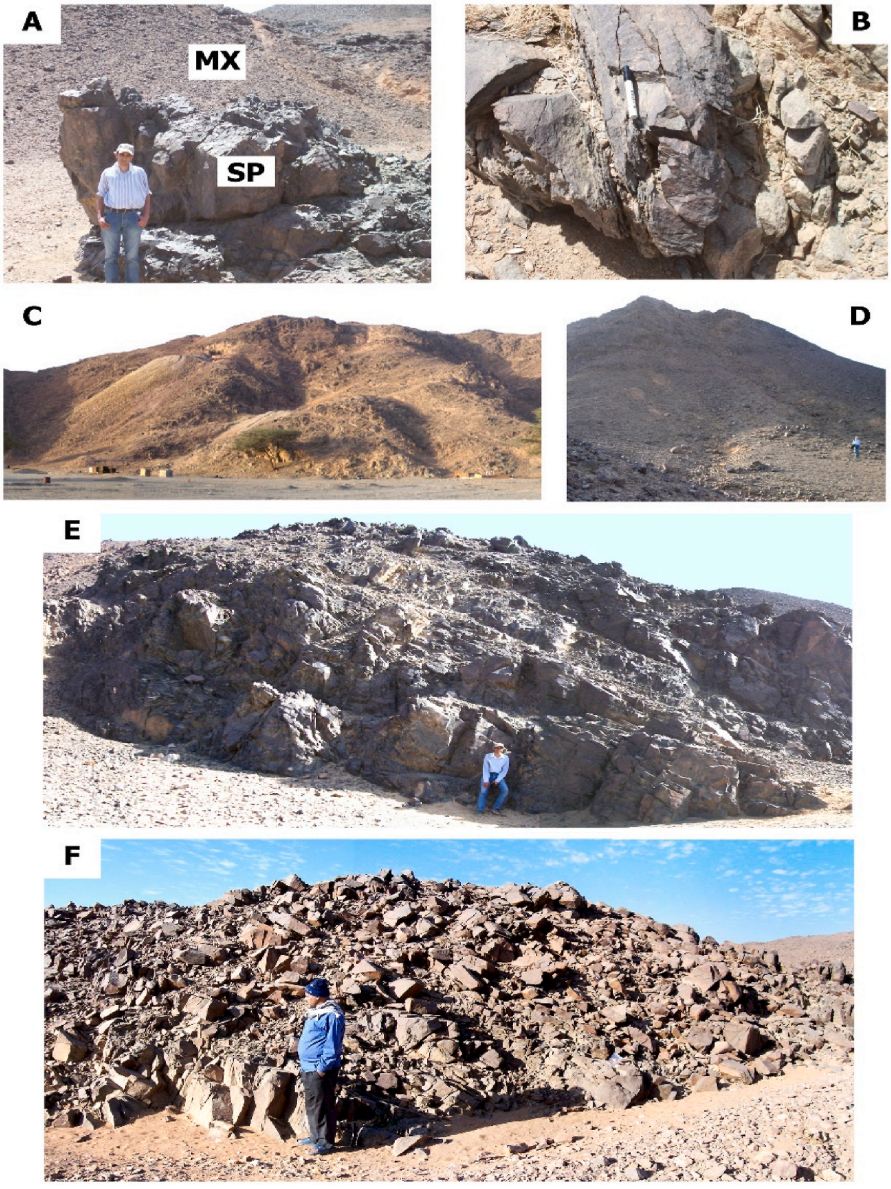

Fig. 2. Field images (Said, 2010) validating the main rock units in Hammash area including: (A) Serpentinite rocks (SP) embedded in mélange matrix (Mx), (B) Shearing in meta-gabbros, (C) Hammash granite, (D) Trachyte plug, (E) The arc meta-volcanics, (F) Highly jointed subduction-related granitoids.

at higher elevations compared to felsic rocks like granite.

The lithological map of the Hammash area used in this research was extracted from published geological maps (Abdel-Rahman et al., 2023; Aboelkhair et al., 2021; El-Desoky et al., 2021; Kassem and Rahim, 2014; Sakran et al., 2009), which were integrated with a detailed field study conducted by the third author, creating about 520 ground truth points (GTP). Approximately 300 GTP for the Fatimah area were extracted from previously published geological maps (Alshehri and Abdelrahman, 2023; Moore and Al-Rehaili, 1989a,b).

ArcGIS Pro 3.2.1 software was used for generating and handling the necessary features, integrating Python with ML libraries such as Scikit-Learn and geospatial libraries like ArcPy to execute RFC model and determine variable importance. SHAP was applied to provide both global and local insights into the model's predictions, utilizing multiple spectral bands within the dataset. The computational platform consisted of an Intel(R) Xeon(R) W-2245 CPU running at 3.90 GHz, with 64 GB of RAM and SSD storage, and Python 3.11.5 was the primary programming language used. Although no GPU was utilized, the CPU provided sufficient computational power. Additionally, Anaconda Navigator 4.4 served as a graphical user interface (GUI), providing access to Jupyter Notebook and a wide range of data science libraries, which, in combination with the SHAP Python library, allowed for a deeper interpretation of the model's outcomes.

4. Methodology

In this research, we used the "Machine Learning Life Cycle" to extract knowledge from data, including LD (Ashmore et al., 2021). Fig. 3 shows a flow chart of the complete ML cycle, which extracts the most useful multi-spectral bands for LD in crystalline arid regions with minimal independent feature engineering using RFC in the Hammash area and validates them in the Wadi Fatimah area, as follows: (a) labeling 520 GTP to cover every lithological unit using a published geological map and field verification in the Hammash area; (b) preparing VNIR, SWIR, and PCA for ASTER and OLI, and PCA for Landsat-8 and Shuttle Radar Topography Mission Digital Elevation Model (SRTM-DEM) as 21 independent raster features; (c) selecting independent features using multi-collinearity analysis combined with spectral and petrological analysis; (d) using 90% of randomly selected GTP on independent features to train different ML models and then optimizing the hyperparameters to achieve the best model performance; (e) evaluating the model using equations in Table S1; (f) determining and explaining the importance of independent variables affecting LD in the Hammash area using the SHAP library in Python 3 within Jupyter Notebook; (g) using the best ten independent features in the Wadi Fatimah area with the same model characteristics for validation.

5. Results and discussion

5.1. Dependent and independent features preparation

Dependent features must be labeled and classified before implementing supervised ML classification algorithms (Kotsiantis et al., 2006). To encompass the diverse lithological composition of the Hammash area, a total of 520 GTPs were established across the eight distinct rock units based on field survey and petrological analysis. Serpentine minerals comprise most serpentinites, along with a few residual pyroxenes, including secondary talc and chlorite, accessory opaques, and carbonates. Antigorite makes up the majority of the 90% of serpentine minerals (Figs. S1 and S2). Plagioclase and hornblende, along with the serpentine mineral antigorite, are the primary minerals found in meta-gabbros with serpentinites rocks (Figs. S3 and S4). Recordings of actinolitic hornblende (uralite) are also common. Ophiolitic mélange matrix rocks (Fig. S5) are very complicated rocks mainly composed of plagioclase and quartz, with sericite alteration of feldspar and subordinate number of accessory opaques. Rock shows a very well-pronounced

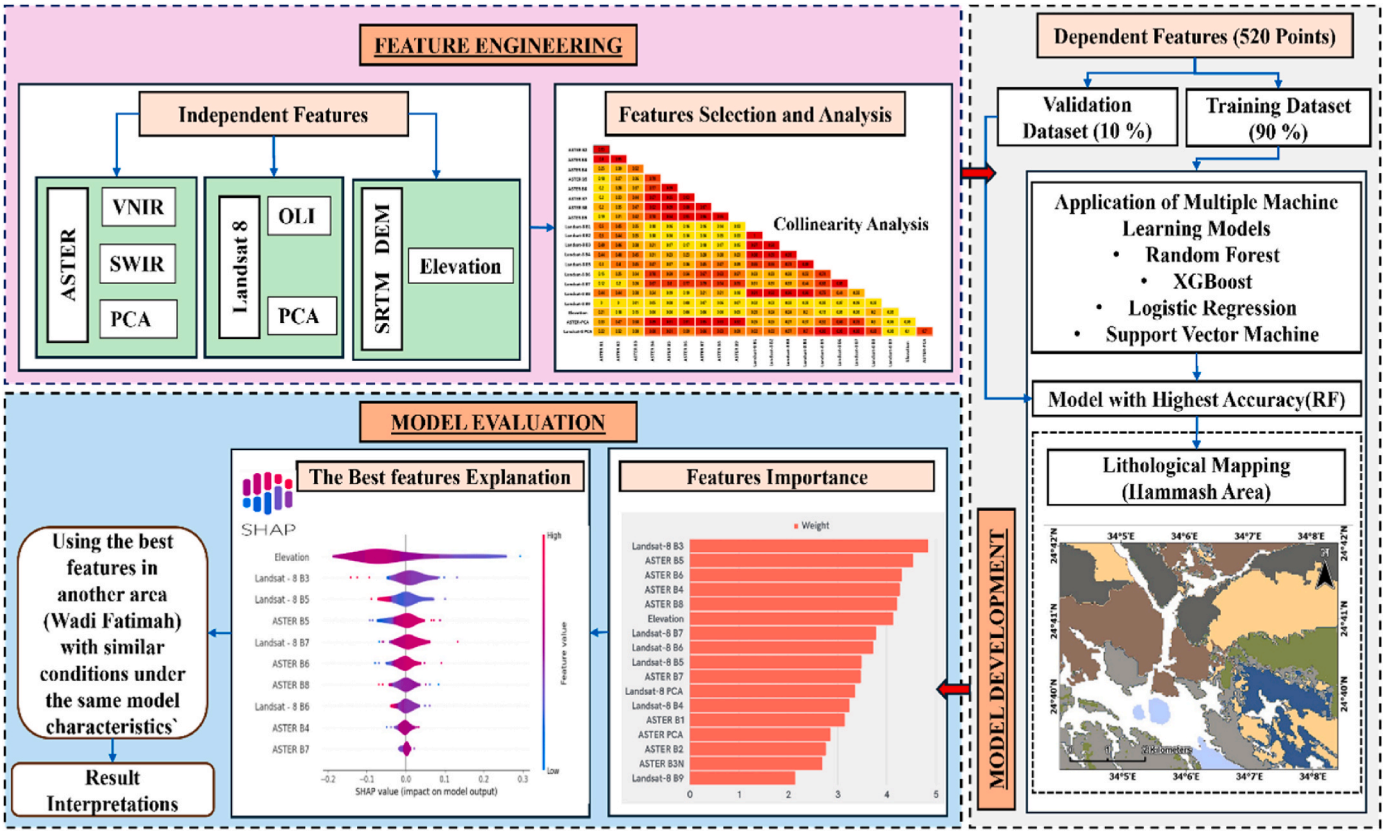

Fig. 3. Flowchart of a MLmodel developed for lithological discrimination, designed to identify the most important multispectral features for crystalline arid regions.

preferred orientation. Island arc rocks (Figs. S6, S7, S8) are composed of meta-basalts, meta-andesites and pyroclastics. The essential minerals encountered in these rocks are plagioclase, augite and olivine. Secondary sericite (paragonite), epidote, calcite, chlorite, serpentine, pumpellyite as well as accessory celadonite, prehnite, quartz, carbonates and opaques are observed. These rocks exhibit a porphyritic texture and locally flow texture. Some of the examined samples show amygdaloidal texture. The main minerals observed in subduction-related granitoid rocks are plagioclase, quartz and hornblende. Secondary chlorite and sericite are recorded with accessories represented by zircon, epidote and opaques. These rocks are mainly equigranular, showing a hypidiomorphic granular fabric. Transpression-related Granites (Hammash granites) (Fig. S9) are economically important because it includes gold mineralization (Figs. S10–S13). The main minerals encountered in these rocks are plagioclase and quartz with secondary sericite, kaolinite and chlorite. Further accessories are mainly microcline, muscovite and opaques. The main minerals found in Trachyte plugs (Fig. S14) are sanidine and plagioclase with secondary kaolinite and sericite as well as accessory opaques.

Careful consideration must be given to the uneven distribution of these units to avoid an imbalanced dataset. For example, the trachyte plugs, covering approximately 0.58 square kilometers, will be represented by 34 GTPs, whereas the much larger granodiorite and quartz diorite unit, spanning around 10 square kilometers, will be assigned 150 GTPs. minimal feature engineering was used in this study resulting in the usage of 21 independent raster features, generated from VNIR/SWIR

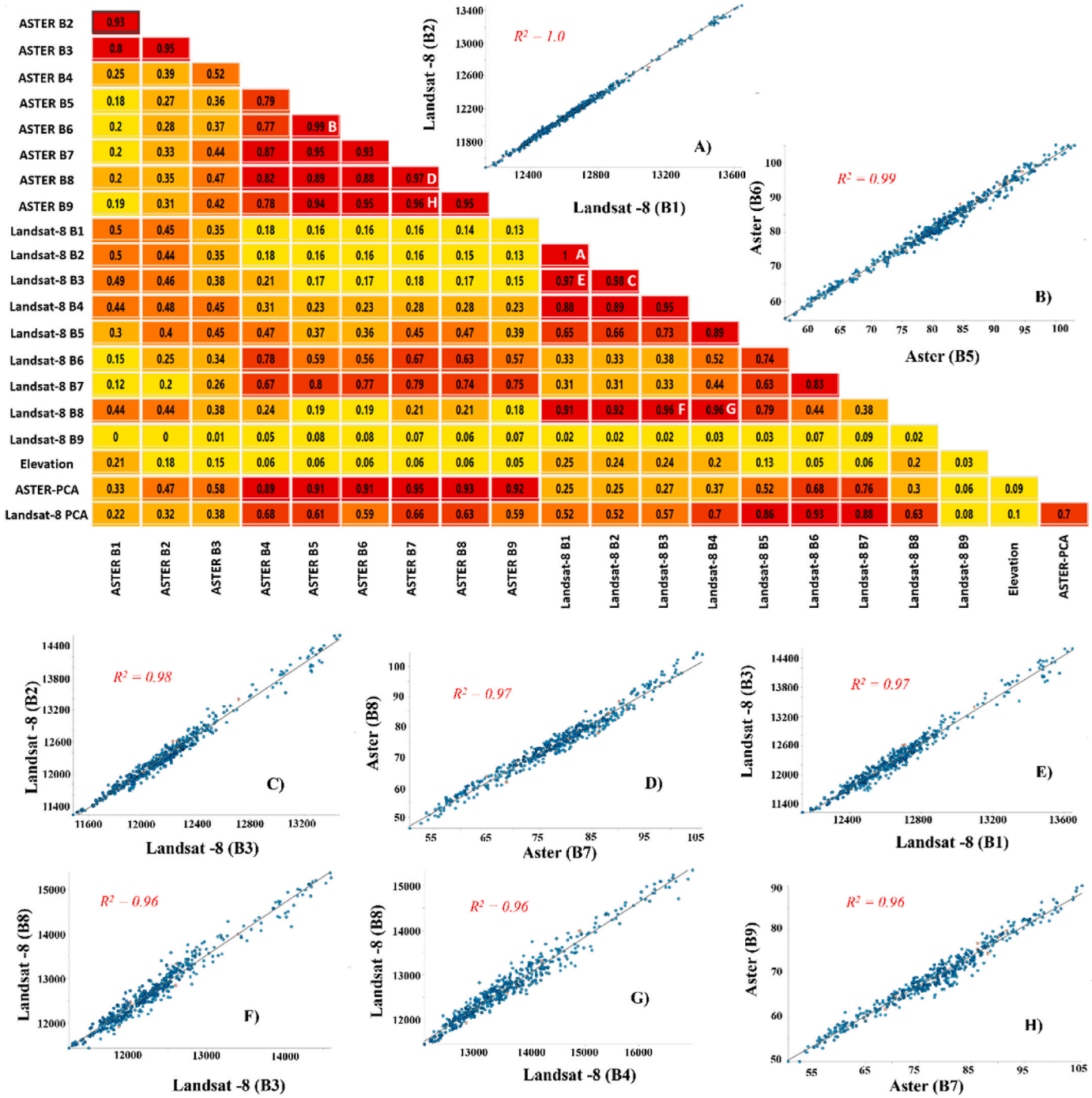

Fig. 4. Multicollinearity Analysis (MCA) for the independent features.

bands, and PCA utilization for ASTER, OLI, Landsat-8 and SRTM-DEM. PCA rasters were the only feature extraction method used and the thermal infrared bands in ASTER and Landsat-8 were excluded due to their lower spatial resolution.

6. Independent features selection and analysis

Multicollinearity Analysis (MCA), a statistical method used to evaluate the linear association between independent features, with R-squared being a commonly used metric is used here. MCA plays a crucial role in feature selection prior to training ML models (Chen et al., 2020; Ebrahimi-Khusfi et al., 2021; Morgan et al., 2023). An exceedingly high R-squared value (>0.95) poses a significant issue within the training dataset, reducing model performance and leading to inaccurate results (Daoud, 2018).

Fig. 4 shows the results of the MCA performed on the 21 independent images influencing the LD in the Hammash area. There are notably high R-squared values (>0.95) that can be handled by excluding some features in the pre-modeling stage: (a) B1, B2, and B8 in Landsat-8 were removed because they are highly correlated with B4 and B3; (b) B9 in ASTER was removed because it is highly correlated (R^2 = 0.96) with ASTER B7.

Despite the robust correlations between ASTER B5 and ASTER B6, as well as between ASTER B7 and ASTER B8, no bands were omitted based on petrological and spectral analysis due to the following reasons: (a) The spectra of serpentine minerals, such as antigorite, which constitute about 90% of serpentinite rocks (Fig. 5), has a unique absorption pattern at 2.3 μm (Fig. 6) resulting from the vibrational actions of the Mg-OH bond (Ge et al., 2018a; King and Clark, 1989; Rajendran et al., 2014), as illustrated in ASTER B8 based on the USGS spectral library Version 7 (https://crustal.usgs.gov/speclab/); (b) Mafic and ultramafic rocks, such as meta-basalt-related minerals, have an absorption peak between 2.1 and 2.3 μm based on previous studies(Ghrefat et al., 2021; Kamel et al., 2016), represented in B5, B6, B7, and B8 in ASTER and B7 in Landsat-8; (c) Gabbro in the Hammash area possesses special characteristics due to metamorphism, consisting mainly of 40% hornblende and 50% andesitic plagioclase (Fig. 5), that results in higher reflectance in B5, B6, B7, and B8 in ASTER (Fig. 6); (d) Subduction-related granitoids, the largest rock unit in the study area, consist of plagioclase and hornblende, which result in higher reflectance; (e) Andesitic plagioclase and quartz, constituting 95% of the mélange matrix, result in the highest reflectance in B8 in ASTER.

7. Utilization and optimization of RFC model

Based on the highest accuracy and F1 score, as shown in Table S3, RF outperformed SVM, logistic regression, XGBoost, and KNN, making it the optimal model for lithological interpretation in the study area. In recent studies, RF has been widely used for its ability to handle high-dimensional datasets and non-linear interactions between features. For instance (Breiman, 2001), introduced RF as a method that enhances model stability and accuracy by aggregating the predictions of multiple decision trees, which helps improve its overall performance. Despite RF's limitation in terms of interpretability, its accuracy and ability to deal with diverse data sources, such as remote sensing imagery and geospatial data, made it the most suitable method for this study. This is

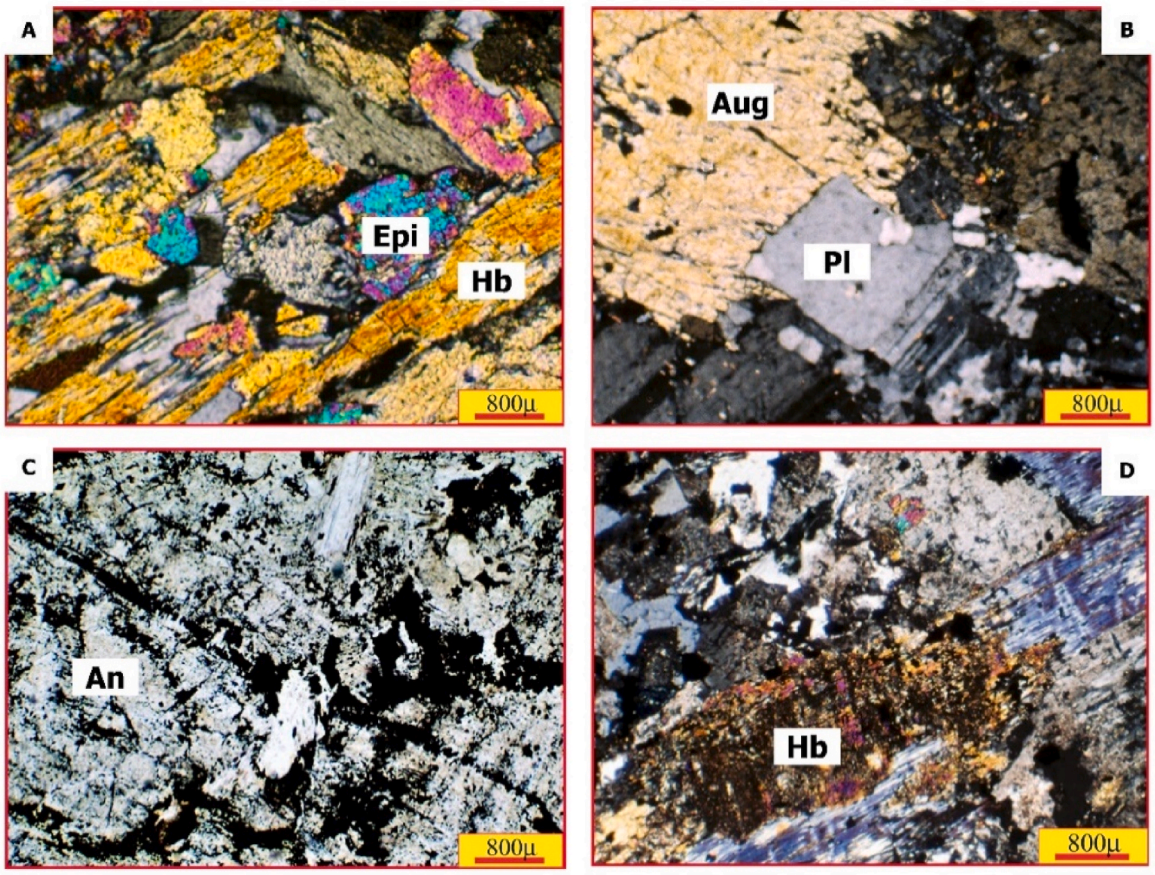

Fig. 5. (A) Epidote crystals (Epi) with hornblende (Hb) in ophiolitic meta-gabbro (XPL), (B) Relict sub-ophitic texture in ophiolitic meta-gabbro. Pl = Plagioclase, Aug = Augite (XPL), (C) Mesh texture exhibited by opaques in serpentinite. An = Antigorite (PPL), (D) Fresh hornblende (Hb) crystal in quartz diorite (XPL) (Said, 2010).

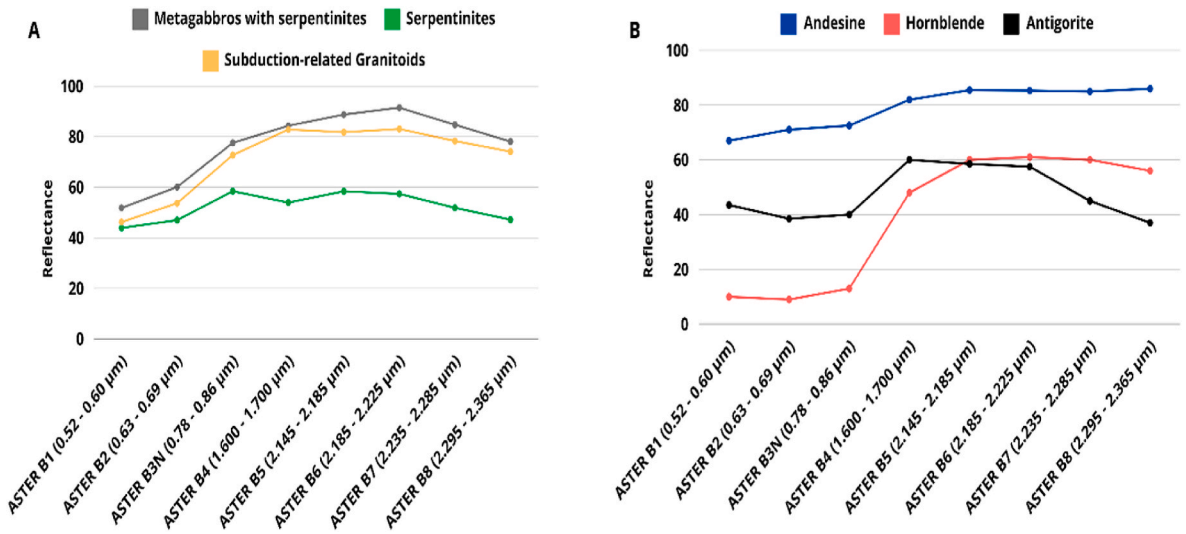

Fig. 6. Reflectance spectral curves for rock units and minerals in the Hammash area A) Rock units: Meta-gabbros with Serpentinites, Serpentinites, and Subduction-related Granitoids), (B) Minerals based on USGS spectral library Version 7: Andesine, Hornblende, and Antigorite.

particularly important for classifying remote sensing data in arid crystalline environments, where the landscape complexity and sparse vegetation present challenges that RF effectively addresses(Belgiu and Drăguţ, 2016; Pal, 2005; Rodriguez-Galiano et al., 2012; H. Zhou et al., 2021).

To address this (Lundberg and Lee, 2017), developed SHAP, which provides a unified framework for interpreting the predictions of any ML model, including RF. SHAP's ability to generate consistent and fair feature importance scores has led to its growing adoption in various fields, such as environmental modeling and healthcare (Lundberg et al., 2020). Compared to other approaches, such as feature importance scores from RF (Strobl et al., 2007), SHAP provides a more granular and theoretically sound explanation by assigning Shapley values. In our analysis here and for classification, the RFC model uses a resampling technique that randomly modifies the predictive features to maximize diversity within each tree (Jin et al., 2018) (Fig. S15A). This adds a layer of insight into our model's predictions, making it more transparent and interpretable than conventional modeling methods.

This method uses several decision trees to illustrate the geographical relationship between 17 independent LD features and dependent features. Each decision tree is built using a bootstrap sample of raw data, enabling accurate failure evaluation with the residual validating set, known as the out-of-bag (OOB) collection. The overall mean square error (MSE_{OOB}) of all 500 trees we built was about 24, computed using Eq. (1)in Table S1. The number of trees is an essential hyperparameter for RFC performance optimization. This is clear since the MSE_{OOB}

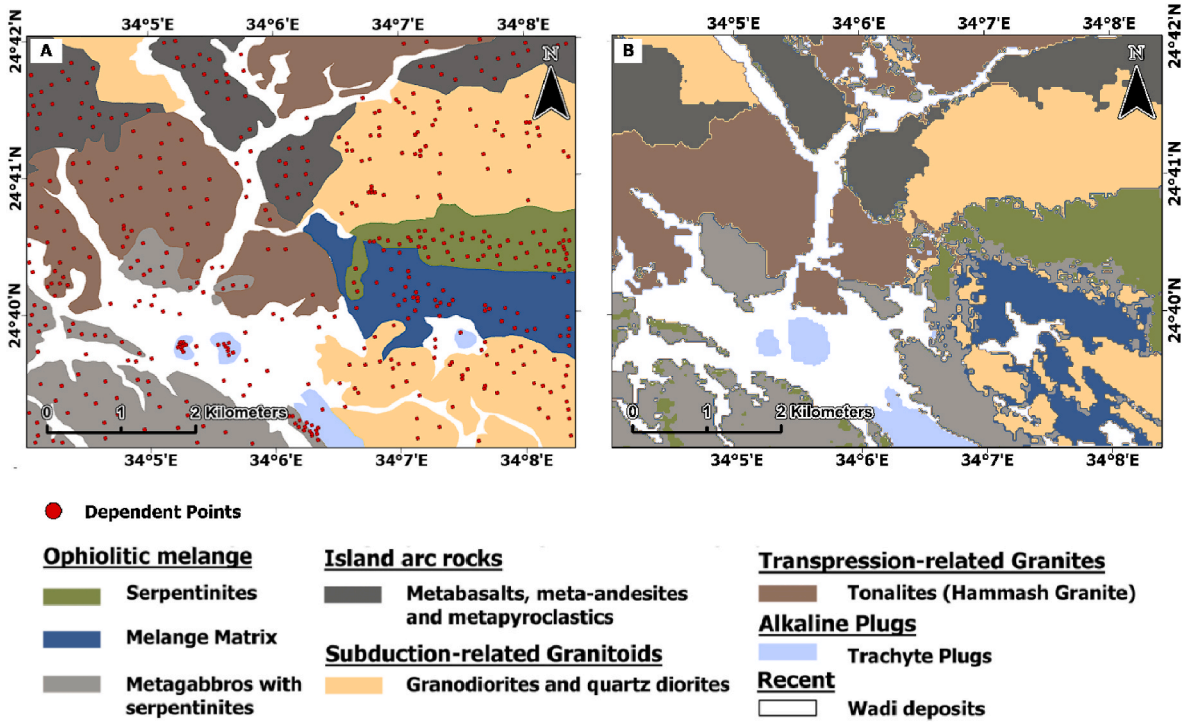

Fig. 7. (A) Geological map of the Hammash area associated with 520 GTP as dependent points that were examined during field surveys to determine the rock types; (B) The result of the RFC model in the Hammash area.

declined from around 24 to 23 as the number of trees increased from 500 to 10,000 (Fig. S15B).

Fig. 7A shows the position of dependent points on the Hammash geological map, alongside the predicted LD in Fig. 7B, based on the trained RFC model with improvements, using the model diagnostics in Table S2.

The model demonstrated a high level of accuracy in segmenting and classifying the principal rock units within the study area, including island arc rocks, the Hammash granite, and the Wadi deposits. These results indicate the robustness of the RFC model in distinguishing between these distinct lithological units. The ophiolitic mélange, which comprises serpentinite, mélange matrix, and meta-gabbro with serpentinite, presents a significant challenge due to its inherent complexity and the difficulty in differentiating these components in the field. Despite this, the model effectively classified these units, showcasing its capability to handle complex geological formations that are often problematic in manual field mapping. The predicted LD outperformed the traditional geological map in segmenting the intrusions between the mélange matrix and the subduction-related granitoid in the lower right corner of the study area. This improved segmentation suggests that the model offers enhanced resolution in identifying and delineating geological boundaries that may be less apparent in conventional mapping techniques. The model accurately identified and localized serpentinites within the meta-gabbro with the serpentinite group, further validating its precision in distinguishing closely associated lithological units. This level of detail is crucial for understanding the spatial distribution of these units and their geological significance (see Table 1).

Due to the unbalanced categorization data set, accuracy alone cannot be used to evaluate the model's performance. Based on the confusion matrix illustrated in the supplementary material (Fig. S16), the following computations can be performed using Eqs. (2)–(6) (Chicco and Jurman, 2020; Sokolova and Lapalme, 2009) in Table S1. The findings of various strategies, as reported in Table 2, were used to evaluate the RFC model and show it is an excellent match with overall accuracy (94%) and precision (81%) during validation. The overall sensitivity decreased to 76% due to the imbalanced data set, with training sample scarcity in trachyte plugs and the complexity of ophiolite mélange (mélange matrix and meta-gabbro with serpentinite), which mix in many places.

8. Independent features importance and explanation for LD

The RFC model's "variable importance" is an XAI tool that shows the proportional importance of the 17 LD independent factors (Fig. S17).

In this scenario, the Landsat-8 B3 band, coupled with elevation data, emerged as critical for distinguishing Wadi deposits from neighboring rock formations. This distinction is primarily because Quaternary deposits reflect more in the B3 band and are typically present at lower elevations according to the DEM. The variable importance was reinforced, with ASTER bands B5, B6, B7, and B8 initially expected to be excluded during the MCA, turning out to be crucial in differentiating ophiolitic mélange units and felsic rocks. While the PCA derived from both ASTER and Landsat-8 data was the sole effective FE method in this research, its contribution to LD analysis was not as significant as initially expected.

When the RFC model was rerun and evaluated using only the top 10 features in the Hammash area, the Landsat-PCA feature was excluded as it ranked eleventh. Since the goal of the research is to optimize the RFC with FE, the results detailed in the supplementary material show an accuracy of 90.3%, with only the top 10 features proven to be effective. We used SHAP to explain the impact of each independent feature of the LD using equation (6) in Table S1. The weight of each feature used for lithological units' delineation is presented (Fig. 8A) and the Beeswarm plot (Fig. 8B). Using SHAP here provides us with a consistent approach to feature importance identification by ensuring that each feature contribution to the prediction is attributed. Unlike traditional feature importance identification methods in RF, which might simply rank features based on their contribution to the model's decision-making process, SHAP here fairly distributed the prediction among the features (Lundberg et al., 2019; Mangalathu et al., 2020; Nohara et al., 2022). Such distribution ensures that each feature's importance is calculated in a manner that considers all possible combinations of features, offering a more robust and accurate assessment. While traditional feature importance in RF provides a global view of which features are most important overall, SHAP extends this by offering both global and

Table 1
Data used for independent features.

Subsystem	Sensor	Band No.	Spectral Wavelength Range (μm)	Spatial Resolution (m)	Date
VNIR	ASTER	B1	0.52–0.60	15	May 31, 2000
	ASTER	B2	0.63–0.69	15	May 31, 2000
	ASTER	B3	0.78–0.86	15	May 31, 2000
SWIR	ASTER	B4	1.6–1.7	30	May 31, 2000
	ASTER	B5	2.145–2.185	30	May 31, 2000
	ASTER	B6	2.185–2.225	30	May 31, 2000
	ASTER	B7	2.235–2.285	30	May 31, 2000
	ASTER	B8	2.295–2.365	30	May 31, 2000
	ASTER	B9	2.36–2.43	30	May 31, 2000
OLI	Landsat-8	B1	0.43–0.45	30	May 1, 2013
	Landsat-8	B2	0.45–0.51	30	May 1, 2013
	Landsat-8	B3	0.53–0.59	30	May 1, 2013
	Landsat-8	B4	0.64–0.67	15	May 1, 2013
	Landsat-8	B5	0.85–0.88	30	May 1, 2013
	Landsat-8	B6	1.57–1.65	30	May 1, 2013
	Landsat-8	B7	2.11–2.29	30	May 1, 2013
	Landsat-8	B8	0.50–0.68	30	May 1, 2013
	Landsat-8	B9	1.36–1.38	30	May 1, 2013
	SRTM-DEM	Elevation		1 arc-second	February 11, 2000

Table 2
RFC model evaluation for the Hammash area.

Class	F1-Score	MCC[a]	Sensitivity	Precision	Accuracy
Training					
Overall	1	1	1	1	1
Validation					
Serpentinites	0.75	0.68	0.75	0.75	0.88
Meta-gabbros with serpentinites	0.5	0.44	0.43	0.6	0.88
Melange Matrix	0.67	0.7	0.5	1	0.98
Island arc rocks	0.83	0.81	0.83	0.83	0.96
Subduction-related Granitoids	0.93	0.91	1	0.87	0.96
Transpression-related Granites	0.83	0.83	1	0.71	0.96
Trachyte Plugs	0.67	0.64	0.6	0.75	0.94
Wadi deposits	1	1	1	1	1
Overall	0.77	0.75	0.76	0.81	0.94

[a] Matthews correlation coefficient.

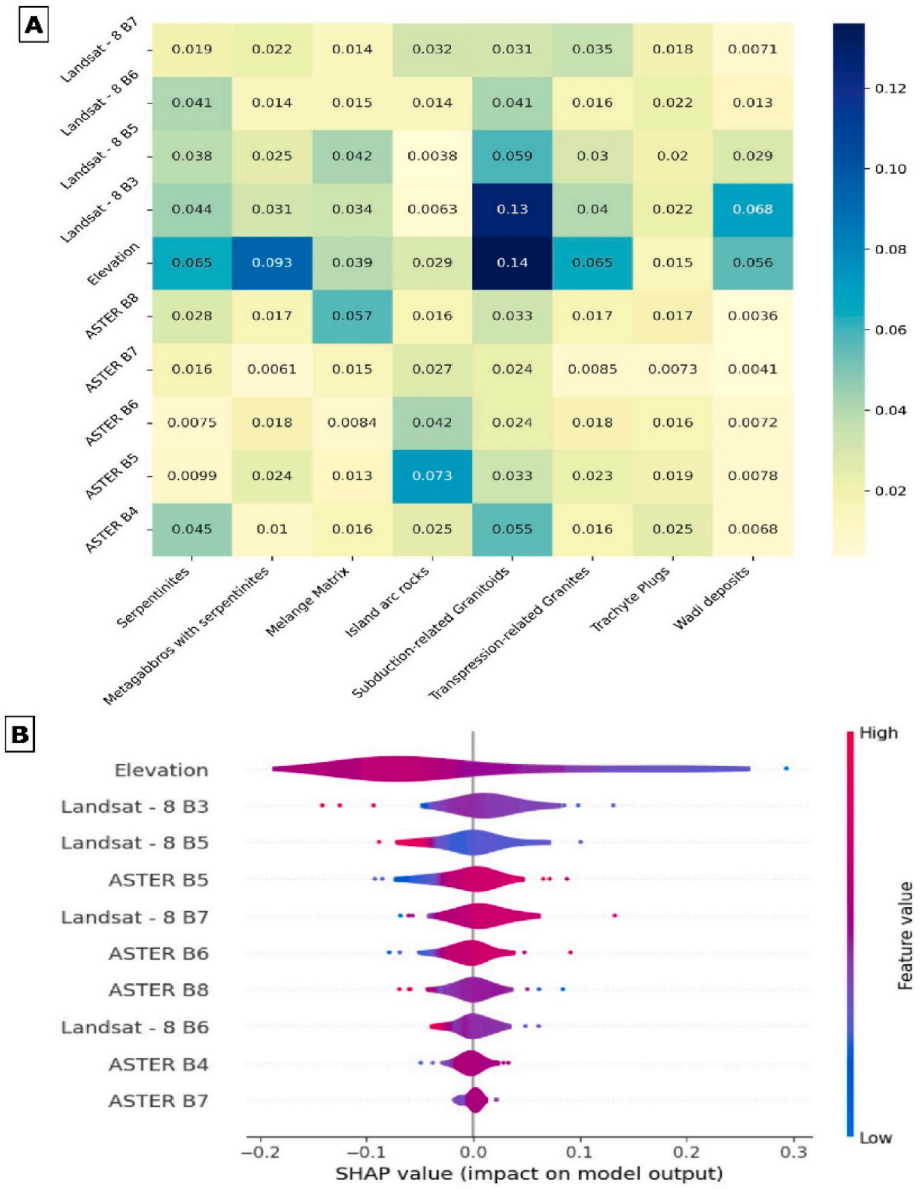

Fig. 8. (A) Heat map of the weight of each feature for delineating each lithological unit; (B) Beeswarm plot created using SHAP shows a detailed overview of the ways that top features in a dataset affect the model output.

local interpretability. This means that SHAP can explain the overall feature importance for the entire dataset (global) as well as for individual predictions (local) (Ekanayake et al., 2022). This is particularly important in lithological mapping, where local variations and specific geological formations might have significant implications.

SHAP values for each feature along with each data point value are plotted, with the x location of each dot representing how high or low the SHAP value is. The dots pile up over each feature row, y-location, illustrating its density, which is represented by how wide or narrow the plot is over different SHAP values. Colors indicate the feature's original value that allows us to gain insights about the patterns driving the model's segmentation process and provide a comprehensive view of the model's behavior at a global level. Fig. 8A highlights each feature's importance for different lithological units, while Fig. 8B identifies the order of that feature in relation to its actual value. Every point here indicates a Shapley value for an independent feature. The Shapley value determines the position on the x-axis, whereas the independent feature determines the location on the y-axis. The color denotes the feature's value, which ranges from low to high. Sometimes features may overlap features in the y-axis direction, giving us an idea of the spread of Shapley values per feature, which is not the case here. It is also noteworthy that features are ranked based on their significance in our lithological identification analysis. The feature rankings derived from random forest and SHAP were similar in some features (like ASTER B7 and Landsat-8 B6), and slight differences (like Landsat-8 B3, ASTER-B5, and ASTER-B8) and strong differences in the ranking across the others due to the distinct calculation methods. Random forest feature importance is typically calculated using heuristic approaches, such as the decrease in Gini impurity, which assesses how much each feature improves the model's accuracy. While effective for ranking features overall, this method does not provide explanations at the individual prediction level. In contrast, SHAP values, grounded in cooperative game theory, compute the marginal contribution of each feature by evaluating all possible combinations of features. This allows SHAP to capture interactions between features, providing more detailed and interpretable insights into the model's decision-making process.

For instance, the wide spread of SHAP values for elevation feature shows its significance for LD for this mountainous study area (360–524

m.a.s.l) (Fig. 8B). We saw that elevation feature is influential in the segmentation of most lithological units, especially subduction-related granitoids, meta-gabbros with serpentinites, serpentinites, Wadi deposits, and mélange matrix (Fig. 8A). Secondly, Landsat-8 B3 proves to be one of the most important bands for delineating subduction-related granitoids and Wadi deposits while still important in other rock unit segmentation (Fig. 8A) Similarly, we found that ASTER B5, ASTER B6, and Landsat-8 B7are significant for classifying island arc rocks. Although the ophiolitic mélange matrix is one of the most complex rock units, it was well classified mainly based on ASTER B8, elevation, Landsat-8 B3, ASTER B7, and Landsat-8 B5. While elevation plays a significant role in identifying serpentinite, it is just one of several key factors essential for distinguishing between serpentine in pure serpentinite and serpentine within meta-gabbro with serpentinite. Other critical contributors include Landsat-8 bands B3, B6, and B5, as well as ASTER bands B8 and B4. So, SHAP confirms the mineral spectroscopy analysis using ASTER, where ASTER B4 is important for Serpentinites (Antigorite) and Subduction-related granitoids. B5 and B6 in ASTER significantly differentiate island arc rocks containing meta-basalt. ASTER B8 is important for the Mélange Matrix, which contains fresh hornblende (Hb) crystals in quartz diorite.

9. Implementation and evaluation of the best features for a similar area

To validate our previously outlined approach, its implementation while utilizing the previously indicated best ten independent features, based on variable importance from the RFC model, over a different area was attempted. We used the same model characteristics (Table S3) in another area with similar conditions, namely Wadi Fatimah to verify our findings over Hammash area. Fig. 9A illustrates the published geological map of Wadi Fatimah, associated with 300 GTP used for model training and validation. Despite the complexity and mixing of geological units, the RFC model succeeded in discriminating between lithological units using the best ten variables approach implemented previously (Fig. 9B).

Table 3 summarizes the model validation in the Wadi Fatimah area based on the confusion matrix in the supplementary material (Fig. S18). During validation, the model evaluation showed excellent overall accuracy (93%) and precision (92%). The overall sensitivity and MCC slightly decreased to 91% and 86%, respectively, due to the complexity of the Ju'ranah complex group, an unassigned unit of the uncertain affinity group, and the Zibarah group, which interconnect in various areas.

10. Conclusion

This study pioneers the use of SHAP and advances LD in arid crystalline environments by integrating XAI and multi-spectral remote sensing data. By combining GIS, ML, and geological insights, this research demonstrates the effectiveness of RFC and SHAP analysis in identifying key features for distinguishing rock types in different geological settings, like the Hammash region in Egypt and the Wadi

Table 3
RFC model evaluation for the Wadi Fatimah area using the best ten features.

Class	F1-Score	MCC[a]	Sensitivity	Precision	Accuracy
Training					
Overall	1	1	1	1	1
Validation					
UNASSIGNED UNIT OF UNCERTAIN AFFINITY	0.80	0.67	0.91	0.71	0.83
ZIBARAH Group	1.00	1.00	1.00	1.00	1.00
JU'RANAH COMPLEX	0.76	0.65	0.67	0.89	0.83
POST-TECTONIC INTRUSIVE ROCKS	1.00	1.00	1.00	1.00	1.00
QUATERNARY DEPOSITS	1.00	1.00	1.00	1.00	1.00
Overall	0.91	0.86	0.91	0.92	0.93

[a] Matthews correlation coefficient.

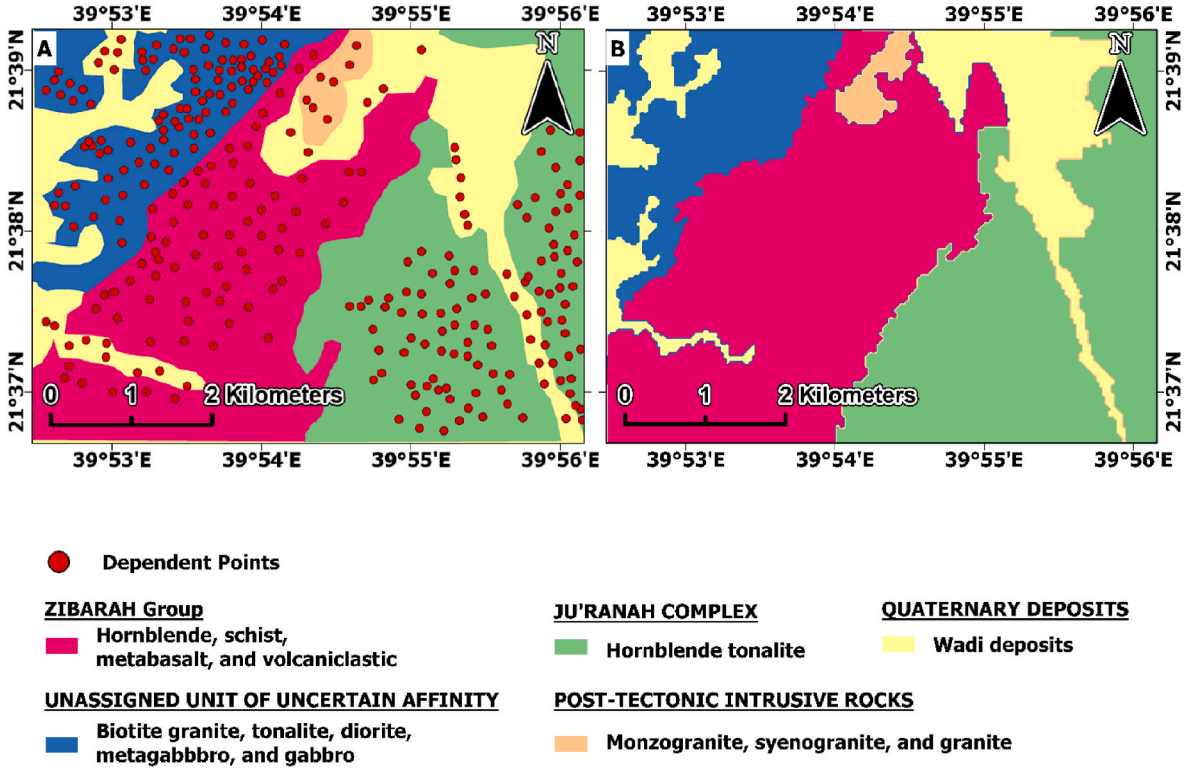

Fig. 9. (A) Previously published geological map incorporating 300 GTP-dependent points; (B) RFC model results over Wadi Fatimah area.

Fatimah area in the Kingdom of Saudi Arabia. The model, supported by 520 ground truth data points collected in the Hammash area, achieved a high overall accuracy of 94% in Hammash and 93% in Wadi Fatimah, confirming its robustness in challenging environments.

Based on the initial analysis as well as the validation step here, we found that B3, B5, B6, and B7 in Landsat-8, B4, B5, B6, B7, and B8 in ASTER, and SRTM-DEM are the most critical and effective multi-spectral features for LD in crystalline arid areas. Variable importance is an efficient XAI tool for evaluating RFC features and makes it trusted. Over time, RFC has proven its strength compared to other ML models like SVM (Bachri et al., 2019; Shebl et al., 2022).

Thermal infrared bands from ASTER and Landsat-8 were excluded due to their lower spatial resolution, leaving ASTER PCA and Landsat-8 PCA as the most effective methods for feature extraction. Before modeling, MCA was conducted to remove highly correlated bands, resulting in 17 independent features. The RFC model, optimized by fine-tuning the number of trees as a hyperparameter, processed these features, using 90% of the ground truth points for training and 10% for validation. Despite the complexity posed by the ophiolitic mélange, the model achieved an impressive 94% overall accuracy and 81% precision in rock unit segmentation.

The model's use of SHAP to interpret key features ensures transparency in decision-making, allowing for a deeper understanding of the most critical spectral bands and elevation data contributing to LD. Achieving high accuracy demonstrates its potential for application in geological mapping, resource exploration, and environmental monitoring, especially in remote and geologically complex areas. Furthermore, the ability to generalize the model across different arid regions makes it a powerful tool for global geological studies, with applications ranging from mineral exploration to hazard assessment.

Due to the fundamental nature of the datasets, which used multi-spectral bands with tree-based ML, this study highlights some limitations for other rock types since our effective features may only be suitable for arid mountainous crystalline rocks. Although this study proved its strength in determining the bands suitable for discriminating a wide variety of major crystalline rocks and complex crystalline formations, it showed weakness in classifying minor and limitedly distributed crystalline rocks such as trachyte plugs. Finally, our analysis displayed strength in delineating ore-bearing rocks but not the ore grading within the rocks, such as Hammash granite, which is well-defined.

Looking ahead, integrating high spectral resolution hyperspectral satellite imagery, such as that from PRISMA (Hyperspectral Precursor of the Application Mission) and Earth Surface Mineral Dust Source Investigation (EMIT), with convolutional neural networks (CNN) models, which excel in internal feature extraction, along with XAI techniques, could enhance LD. This approach has the potential to more effectively identify relevant bands for lithological mapping, benefiting not only crystalline arid regions but also sedimentary rock areas. By leveraging the advanced capabilities of hyperspectral imagery and CNN models, future research could significantly improve the accuracy and applicability of lithological discrimination in a wide range of geological settings.

CRediT authorship contribution statement

Hesham Morgan: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ali Elgendy:** Validation, Resources, Methodology. **Amir Said:** Writing – review & editing, Validation, Resources, Investigation. **Mostafa Hashem:** Formal analysis. **Wenzhao Li:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Investigation, Formal analysis. **Surendra Maharjan:** Visualization, Validation, Formal analysis. **Hesham El-Askary:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition.

Code availability section

Name of the code/library: Final_Model
Contact: maharjan@chapman.edu.
Hardware requirements: Intel(R) Xeon(R) W-2245 CPU @ 3.90 GHz 3.91 GHz Program language: Python 3.11.5.
Software required: Python.
Program size: 569 KB.
The source codes are available for downloading at the link: https://github.com/EssDs-lab/Lithology_Paper.git.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:
Hesham El Askary reports financial support was provided by US Department of Education. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to acknowledge the support by the US Department of Education with the award number P116Z230273, "Promoting the Integration of Earth Observations for Sustainable Development Goals". The authors would also like to extend their thanks and appreciation for Prof. M. A. El Sharkawi, Prof. Sh. Sakran and Dr. A. Zakariya (Cairo University) and Dr. Z. El Alfy (EMRA) for their help in the field campaign in the study regions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cageo.2024.105738.

Data availability

Data will be made available on request.

References

Abdel-Rahman, A.M., El-Desoky, H.M., Shebl, A., El-Awny, H., Amer, Y.Z., Csámer, Á., 2023. The geochemistry, origin, and hydrothermal alteration mapping associated with the gold-bearing quartz veins at Hamash district, South Eastern Desert, Egypt. Sci. Rep. 13, 1–27. https://doi.org/10.1038/s41598-023-42313-9.

Aboelkhair, H., Ibraheem, M., El-Magd, I.A., 2021. Integration of airborne geophysical and ASTER remotely sensed data for delineation and mapping the potential mineralization zones in Hamash area, South Eastern Desert, Egypt. Arabian J. Geosci. 14. https://doi.org/10.1007/s12517-021-07471-y.

Alifu, H., Tateishi, R., Johnson, B., 2015. A new band ratio technique for mapping debris-covered glaciers using Landsat imagery and a digital elevation model. Int. J. Rem. Sens. 36, 2063–2075. https://doi.org/10.1080/2150704X.2015.1034886.

Alshehri, F., Abdelrahman, K., 2023. In: Groundwater Potentiality of Wadi Fatimah, Western Saudi Arabia: Geophysical and Remote Sensing Integrated Approach, vol. 15. https://doi.org/10.3390/w15101828. Water (Switzerland).

Amer, R., Kusky, T., Ghulam, A., 2010. Lithological mapping in the central eastern desert of Egypt using ASTER data. J. Afr. Earth Sci. 56, 75–82. https://doi.org/10.1016/J.JAFREARSCI.2009.06.004.

Antonini, A.S., Tanzola, J., Asiain, L., Ferracutti, G.R., Castro, S.M., Bjerg, E.A., Ganuza, M.L., 2024. Machine learning model interpretability using SHAP values: application to igneous rock classification task. Applied Computing and Geosciences 23, 100178. https://doi.org/10.1016/j.acags.2024.100178.

Arslan, Y., Lebichot, B., Allix, K., Veiber, L., Lefebvre, C., Boytsov, A., Goujon, A., Bissyandé, T.F., Klein, J., 2022. Towards refined classifications driven by SHAP explanations. In: Holzinger, A., Kieseberg, P., Tjoa, A.M., Weippl, E. (Eds.), Machine Learning and Knowledge Extraction. Springer International Publishing, Cham, pp. 68–81. https://doi.org/10.1007/978-3-031-14463-9_5.

Ashmore, R., Calinescu, R., Paterson, C., 2021. Assuring the Machine Learning Lifecycle, vol. 54. ACM Computing Surveys (CSUR). https://doi.org/10.1145/3453444.

Bachri, I., Hakdaoui, M., Raji, M., Teodoro, A.C., Benbouziane, A., 2019. Machine learning algorithms for automatic lithological mapping using remote sensing data: a case study from souk arbaa sahel, sidi ifni inlier, western anti-atlas, Morocco. ISPRS

Int. J. Geo-Inf. 8, 248. https://doi.org/10.3390/IJGI8060248, 2019, Vol. 8, Page 248.
Baid, S., Tabit, A., Algouti, Ahmed, Algouti, Abdellah, Nafouri, I., Souddi, S., Aboulfaraj, A., Ezzahzi, S., Elghouat, A., 2023. Lithological discrimination and mineralogical mapping using Landsat-8 OLI and ASTER remote sensing data: igoudrane region, jbel saghro, Anti Atlas, Morocco. Heliyon 9, e17363. https://doi.org/10.1016/j.heliyon.2023.e17363.
Behnia, P., Harris, J.R., Rainbird, R.H., Williamson, M.C., Sheshpari, M., 2012. Remote predictive mapping of bedrock geology using image classification of Landsat and SPOT data, western Minto Inlier, Victoria Island, Northwest Territories, Canada. Int. J. Rem. Sens. 33, 6876–6903. https://doi.org/10.1080/01431161.2012.693219.
Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogrammetry Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324/METRICS.
Chen, D., Zhang, L., 2023. Monotonicity for AI ethics and society: an empirical study of the monotonic neural additive model in criminology, education, health care, and finance. https://doi.org/10.48550/arXiv.2301.07060.
Chen, H., Yang, L., Wu, Q., 2023. Enhancing land cover mapping and monitoring: an interactive and explainable machine learning approach using google Earth engine. Rem. Sens. 15, 4585. https://doi.org/10.3390/RS15184585, 2023, Vol. 15, Page 4585.
Chen, W., Chen, Y., Tsangaratos, P., Ilia, I., Wang, X., 2020. Combining evolutionary algorithms and machine learning models in landslide susceptibility assessments. Rem. Sens. 12, 1–26. https://doi.org/10.3390/rs12233854.
Chen, Y., Dong, Y., Wang, Y., Zhang, F., Liu, G., Sun, P., 2023. Machine learning algorithms for lithological mapping using Sentinel-2 and SRTM DEM in highly vegetated areas. Front. Ecol. Evol. 11. https://doi.org/10.3389/fevo.2023.1250971.
Chicco, D., Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genom. 21, 1–13. https://doi.org/10.1186/S12864-019-6413-7/TABLES/5.
Clark, R.N., Roush, T.L., 1984. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. J. Geophys. Res. Solid Earth 89, 6329–6340. https://doi.org/10.1029/JB089IB07P06329.
Cloutis, E.A., 1996. Review Article Hyperspectral geological remote sensing: evaluation of analytical techniques. Int. J. Rem. Sens. 17, 2215–2242. https://doi.org/10.1080/01431169608948770.
Cracknell, M.J., Reading, A.M., 2014. Geological mapping using remote sensing data: a comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. Comput. Geosci. 63, 22–33. https://doi.org/10.1016/J.CAGEO.2013.10.008.
Daoud, J.I., 2018. Multicollinearity and regression analysis. J. Phys. Conf. 949. https://doi.org/10.1088/1742-6596/949/1/012009.
Ebrahimi-Khusfi, Z., Nafarzadegan, A.R., Dargahian, F., 2021. Predicting the number of dusty days around the desert wetlands in southeastern Iran using feature selection and machine learning techniques. Ecol. Indicat. 125, 107499. https://doi.org/10.1016/j.ecolind.2021.107499.
Ekanayake, I.U., Meddage, D.P.P., Rathnayake, U., 2022. A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). Case Stud. Constr. Mater. 16, e01059. https://doi.org/10.1016/j.cscm.2022.e01059.
El-Desoky, H.M., Shahin, T.M., Amer, Y.Z., 2021. Characteristic of gold mineralization associated with granites at Hamash old gold mine, South Eastern Desert, Egypt. Arabian J. Geosci. 14, 1–14. https://doi.org/10.1007/S12517-021-06923-9/FIGURES/9.
Futagami, K., Fukazawa, Y., Kapoor, N., Kito, T., 2021. Pairwise acquisition prediction with SHAP value interpretation. Journal of Finance and Data Science 7, 22–44. https://doi.org/10.1016/j.jfds.2021.02.001.
Garreau, D., Luxburg, U., 2020. Explaining the explainer: a first theoretical analysis of LIME. In: Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics. Presented at the International Conference on Artificial Intelligence and Statistics. PMLR, pp. 1287–1296.
Ge, W., Cheng, Q., Jing, L., Armenakis, C., Ding, H., 2018a. Lithological discrimination using ASTER and Sentinel-2A in the Shibanjing ophiolite complex of Beishan orogenic in Inner Mongolia, China. Adv. Space Res. 62, 1702–1716. https://doi.org/10.1016/J.ASR.2018.06.036.
Ge, W., Cheng, Q., Tang, Y., Jing, L., Gao, C., 2018b. Lithological classification using sentinel-2A data in the shibanjing ophiolite complex in inner Mongolia, China. Rem. Sens. 10, 638. https://doi.org/10.3390/RS10040638, 2018, Vol. 10, Page 638.
Gharib, M.E., Maurice, A.E., Ahmed, A.H., Asimow, P.D., Bakhit, B.R., Selim, H.A., 2021. Evolution of a Neoproterozoic island arc in the northern Arabian-Nubian Shield: volcanic rocks and their plutonic equivalents in the Hamash area, south Eastern Desert, Egypt. Precambrian Res. 358, 106145. https://doi.org/10.1016/J.PRECAMRES.2021.106145.
Ghrefat, H., Kahal, A.Y., Abdelrahman, K., Alfaifi, H.J., Qaysi, S., 2021. Utilization of multispectral landsat-8 remote sensing data for lithological mapping of southwestern Saudi Arabia. J. King Saud Univ. Sci. 33, 101414. https://doi.org/10.1016/J.JKSUS.2021.101414.
Girija, R.R., Mayappan, S., 2019. Mapping of mineral resources and lithological units: a review of remote sensing techniques. International Journal of Image and Data Fusion 10, 79–106. https://doi.org/10.1080/19479832.2019.1589585.
GRAINGER, D.J., 1992. A geological excursion in the Wadi Fatima area, near Jiddah, Saudi Arabia. Geol. Today 8, 215–219. https://doi.org/10.1111/J.1365-2451.1992.TB00458.X.
Gupta, R.P., 2017. Remote sensing geology. In: Remote Sensing Geology, third ed. https://doi.org/10.1007/978-3-662-55876-8/COVER Third Edition 1–428.
He, J., Harris, J.R., Sawada, M., Behnia, P., 2015. A comparison of classification algorithms using Landsat-7 and Landsat-8 data for mapping lithology in Canada's Arctic. Int. J. Rem. Sens. 36, 2252–2276. https://doi.org/10.1080/01431161.2015.1035410.
Helmy, H.M., Kaindl, R., 1999. Mineralogy and fluid inclusion studies of the Au-Cu quartz veins in the Hamash area, South-Eastern Desert, Egypt. Mineral. Petrol. 65, 69–86. https://doi.org/10.1007/BF01161577.
Hilmy, M.E., Osman, A., 1989. Remobilization of gold from a chalcopyrite-pyrite mineralization Hamash gold mine, Southeastern Desert, Egypt. Miner. Deposita 24, 244–249. https://doi.org/10.1007/BF00206386.
Horrocks, T., Holden, E.J., Wedge, D., 2015. Evaluation of automated lithology classification architectures using highly-sampled wireline logs for coal exploration. Comput. Geosci. 83, 209–218. https://doi.org/10.1016/J.CAGEO.2015.07.013.
Hosseiny, B., Abdi, A.M., Jamali, S., 2022. Urban land use and land cover classification with interpretable machine learning – a case study using Sentinel-2 and auxiliary data. Remote Sens. Appl.: Society and Environment 28, 100843. https://doi.org/10.1016/J.RSASE.2022.100843.
Jin, Y., Liu, X., Chen, Y., Liang, X., 2018. Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: a case study of central Shandong. Int. J. Rem. Sens. 39, 8703–8723. https://doi.org/10.1080/01431161.2018.1490976.
Kakogeorgiou, I., Karantzalos, K., 2021. Evaluating explainable artificial intelligence methods for multi-label deep learning classification tasks in remote sensing. Int. J. Appl. Earth Obs. Geoinf. 103, 102520. https://doi.org/10.1016/J.JAG.2021.102520.
Kamel, M., Youssef, M., Hassan, M., Bagash, F., 2016. Utilization of ETM+ landsat data in geologic mapping of wadi ghadir-gabal zabara area, central eastern desert, Egypt. The Egyptian Journal of Remote Sensing and Space Science 19, 343–360. https://doi.org/10.1016/J.EJRS.2016.06.003.
Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H.A., Kumar, V., 2019. Machine learning for the geosciences: challenges and opportunities. IEEE Trans. Knowl. Data Eng. 31, 1544–1554. https://doi.org/10.1109/TKDE.2018.2861006.
Kassem, O.M.K., Rahim, S.H.A.E., 2014. Finite strain analysis of metavolcanics and metapyroclastics in gold-bearing shear zone of the Dungash area, Central Eastern Desert, Egypt. Geotectonics 48, 483–495. https://doi.org/10.1134/S0016852114060041.
Khalid, S., Khalil, T., Nasreen, S., 2014. A survey of feature selection and feature extraction techniques in machine learning. In: 2014 Science and Information Conference. Presented at the 2014 Science and Information Conference, pp. 372–378. https://doi.org/10.1109/SAI.2014.6918213.
Khalifa, A., Bashir, B., Çakir, Z., Kaya, Ş., Alsalman, A., Henaish, A., 2021. Paradigm of geological mapping of the adıyaman fault zone of eastern Turkey using landsat 8 remotely sensed data coupled with pca, ica, and mnfa techniques. ISPRS Int. J. Geo-Inf. 10. https://doi.org/10.3390/ijgi10060368.
King, T.V.V., Clark, R.N., 1989. Spectral characteristics of chlorites and Mg-serpentines using high-resolution reflectance spectroscopy. J. Geophys. Res. Solid Earth 94, 13997–14008. https://doi.org/10.1029/JB094IB10P13997.
Kong, W., Chen, J., Zhu, P., 2024. Machine learning-based uranium prospectivity mapping and model explainability research. Minerals 14, 128. https://doi.org/10.3390/min14020128.
Kotsiantis, S.B., 2007. Supervised machine learning: a review of classification techniques. Informatica 31, 249–268.
Kotsiantis, S.B., Zaharakis, I.D., Pintelas, P.E., 2006. Machine learning: a review of classification and combining techniques. Artif. Intell. Rev. 26, 159–190. https://doi.org/10.1007/S10462-007-9052-3/METRICS.
Kuhn, S., Cracknell, M.J., Reading, A.M., 2018. Lithologic mapping using Random Forests applied to geophysical and remote-sensing data: a demonstration study from the Eastern Goldfields of Australia. Geophysics 83, B183–B193. https://doi.org/10.1190/GEO2017-0590.1/ASSET/IMAGES/LARGE/FIGURE8.JPEG.
Kumar, C., Chatterjee, S., Oommen, T., Guha, A., Mukherjee, A., 2022. Multi-sensor datasets-based optimal integration of spectral, textural, and morphological characteristics of rocks for lithological classification using machine learning models. Geocarto Int. 37, 6004–6032. https://doi.org/10.1080/10106049.2021.1920632.
Latifovic, R., Pouliot, D., Campbell, J., 2018. Assessment of convolution neural networks for surficial geology mapping in the South Rae geological region, Northwest Territories, Canada. Rem. Sens. 10. https://doi.org/10.3390/rs10020307.
Letoffe, O., Huang, X., Asher, N., Marques-Silva, J., 2024. From SHAP scores to feature importance scores. https://doi.org/10.48550/arXiv.2405.11766.
Linardatos, P., Papastefanopoulos, V., Kotsiantis, S., 2021. Explainable ai: a review of machine learning interpretability methods. Entropy 23, 1–45. https://doi.org/10.3390/e23010018.
Little, C.O., Lina, D.H., Allen, G.I., 2024. Fair feature importance scores for interpreting decision trees. Transactions on Machine Learning Research.
Liu, H., Wu, K., Xu, H., Xu, Y., 2021. Lithology classification using TASI thermal infrared hyperspectral data with convolutional neural networks. Rem. Sens. 13. https://doi.org/10.3390/rs13163117.
Lundberg, S., Lee, S.-I., 2017. A unified approach to interpreting model predictions [WWW Document]. arXiv.org. URL. https://arxiv.org/abs/1705.07874v2, 9.7.24.
Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2, 56–67. https://doi.org/10.1038/s42256-019-0138-9.
Lundberg, S.M., Erion, G.G., Lee, S.-I., 2019. Consistent individualized feature attribution for tree ensembles. https://doi.org/10.48550/arXiv.1802.03888.

Mangalathu, S., Hwang, S.H., Jeon, J.S., 2020. Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. Eng. Struct. 219, 110927. https://doi.org/10.1016/j.engstruct.2020.110927.

Marzouki, A., Dridri, A., 2023. Lithological discrimination and structural lineaments extraction using Landsat 8 and ASTER data: a case study of Tiwit (Anti-Atlas, Morocco). Environ. Earth Sci. 82, 1–17. https://doi.org/10.1007/S12665-023-10831-4/METRICS.

Mehr, S.G., Ahadnejad, V., Abbaspour, R.A., Hamzeh, M., 2013. Using the mixture-tuned matched filtering method for lithological mapping with Landsat TM5 images. Int. J. Rem. Sens. 34, 8803–8816. https://doi.org/10.1080/01431161.2013.853144.

Moore, T.A., Al-Rehaili, M.H., 1989a. Geologic map of Makkah quadrangle. Geoscience Map.

Moore, T.A., Al-Rehaili, M.H., 1989b. Geologic Map of Makkah Quadrangle, Sheet 21D, Kingdom of Saudi Arabia DGMR. Geoscience Map GM-107C, 1(250,000). References - Scientific Research Publishing (n.d).

Morgan, H., Madani, A., Hussien, H.M., Nassar, T., 2023. Using an ensemble machine learning model to delineate groundwater potential zones in desert fringes of East Esna-Idfu area, Nile valley, Upper Egypt. Geoscience Letters 10, 1–19. https://doi.org/10.1186/S40562-023-00261-2/FIGURES/11.

Nohara, Y., Matsumoto, K., Soejima, H., Nakashima, N., 2022. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. Comput. Methods Progr. Biomed. 214, 106584. https://doi.org/10.1016/j.cmpb.2021.106584.

Ouyang, S., Chen, W., Qin, X., Yang, J., 2024. Geological background prototype learning-enhanced network for remote-sensing-based engineering geological lithology interpretation in highly vegetated areas. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 17, 8794–8809. https://doi.org/10.1109/JSTARS.2024.3385541.

Pal, M., 2005. Random forest classifier for remote sensing classification. Int. J. Rem. Sens. 26, 217–222. https://doi.org/10.1080/01431160412331269698.

Pal, M., Rasmussen, T., Porwal, A., 2020. Optimized lithological mapping from multispectral and hyperspectral remote sensing images using fused multi-classifiers. Rem. Sens. 12. https://doi.org/10.3390/RS12010177.

Palczewska, A., Palczewski, J., Robinson, R.M., Neagu, D., 2014. Interpreting random forest classification models using a feature contribution method. Adv. Intell. Syst. Comput. 263, 193–218. https://doi.org/10.1007/978-3-319-04717-1_9/FIGURES/12.

Rajendran, S., Nasir, S., Kusky, T.M., al-Khirbash, S., 2014. Remote sensing based approach for mapping of CO2 sequestered regions in Samail ophiolite massifs of the Sultanate of Oman. Earth Sci. Rev. 135, 122–140. https://doi.org/10.1016/j.earscirev.2014.04.004.

Rodriguez-Galiano, V.F., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson, P.M., Jeganathan, C., 2012. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. Remote Sensing of Environment 121, 93–107. https://doi.org/10.1016/j.rse.2011.12.003.

Rowan, L.C., Mars, J.C., 2003. Lithologic mapping in the mountain pass, California area using advanced Spaceborne thermal emission and reflection radiometer (ASTER) data. Remote Sensing of Environment 84, 350–366. https://doi.org/10.1016/S0034-4257(02)00127-X.

Said, A., 2010. In: Petrology, Structure and Gold Mineralization of Hammash - Umm Hagalig Area, South Eastern Desert, Egypt. Unpublished Ph. D. 149.

Sakran, S.M., Said, A., Alfy, Z.E., Sharkawi, M.A.E., 2009. Hammash releasing bend and its control of gold mineralization, Hammash gold mine area, South Eastern Desert, Egypt. Egypt. J. Geol. 53, 87–99.

Shayeganpour, S., Tangestani, M.H., Gorsevski, P.V., 2021a. Machine learning and multi-sensor data fusion for mapping lithology: a case study of Kowli-kosh area, SW Iran. Adv. Space Res. 68, 3992–4015. https://doi.org/10.1016/J.ASR.2021.08.003.

Shayeganpour, S., Tangestani, M.H., Homayouni, S., Vincent, R.K., 2021b. Evaluating pixel-based vs. object-based image analysis approaches for lithological discrimination using VNIR data of WorldView-3. Front. Earth Sci. 15, 38–53. https://doi.org/10.1007/s11707-020-0848-7.

Shebl, A., Kusky, T., Csámer, Á., 2022. Advanced land imager superiority in lithological classification utilizing machine learning algorithms. Arabian J. Geosci. 15, 1–13. https://doi.org/10.1007/s12517-022-09948-w.

Shirmard, H., Farahbakhsh, E., Müller, R.D., Chandra, R., 2022. A review of machine learning in processing remote sensing data for mineral exploration. Remote Sensing of Environment 268, 112750. https://doi.org/10.1016/J.RSE.2021.112750.

Sierra-Botero, D., Molina-Taborda, A., Valdés-Tresanco, M.S., Hernández-Arango, A., Espinosa-Leal, L., Karpenko, A., Lopez-Acevedo, O., 2024. Selecting interpretability techniques for healthcare machine learning models [WWW Document]. arXiv.org. URL. https://arxiv.org/abs/2406.10213v1, 9.10.24.

Sokolova, M., Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. Inf. Process. Manag. 45, 427–437. https://doi.org/10.1016/J.IPM.2009.03.002.

Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinf. 8, 25. https://doi.org/10.1186/1471-2105-8-25.

Suhaidi, M., Kadir, R.A., Tiun, S., 2021. A review of feature extraction methods on machine learning. J. Inf. Technol. Manag. 6, 51–59.

Sultan, M., Arvidson, R.E., Sturchio, N.C., 1986. Mapping of serpentinites in the eastern desert of Egypt by using landsat thematic mapper data. Geology 14, 995–999. https://doi.org/10.1130/0091-7613(1986)14<995:MOSITE>2.0.CO;2.

Temenos, A., Temenos, N., Kaselimi, M., Doulamis, A., Doulamis, N., 2023. Interpretable deep learning framework for land use and land cover classification in remote sensing using SHAP. Geosci. Rem. Sens. Lett. IEEE 20. https://doi.org/10.1109/LGRS.2023.3251652.

Validabadi Bozcheloei, K., Tangestani, M.H., 2019. Prospecting for clinoptilolite-type zeolite in a volcano-sedimentary terrain using ASTER data: a case study from alborz Mountains, northern Iran. Nat Resour Res 28, 1317–1327. https://doi.org/10.1007/s11053-019-09452-1.

Witten, I.H., Frank, E., Geller, J., 2002. Data mining: practical machine learning tools and techniques with java implementations. SIGMOD Record 31, 76–77. https://doi.org/10.1145/507338.507355.

Xie, Y., Zhu, C., Zhou, W., Li, Z., Liu, X., Tu, M., 2018. Evaluation of machine learning methods for formation lithology identification: a comparison of tuning processes and model performances. J. Petrol. Sci. Eng. 160, 182–193. https://doi.org/10.1016/J.PETROL.2017.10.028.

Yang, W., Wei, Y., Wei, H., Chen, Y., Huang, G., Li, X., Li, R., Yao, N., Wang, X., Gu, X., Amin, M.B., Kang, B., 2023. Survey on explainable AI: from approaches, limitations and applications aspects. Hum-Cent Intell Syst 3, 161–188. https://doi.org/10.1007/s44230-023-00038-y.

Yu, L., Porwal, A., Holden, E.-J., Dentith, M.C., 2012. Towards automatic lithological classification from remote sensing data using support vector machines. Comput. Geosci. 45, 229–239. https://doi.org/10.1016/j.cageo.2011.11.019.

Zafar, M.R., Khan, N.M., 2019. DLIME: a deterministic local interpretable model-agnostic explanations approach for computer-aided diagnosis systems [WWW Document]. arXiv.org. URL. https://arxiv.org/abs/1906.10263v1, 9.10.24.

Zeng, W., Davoodi, A., Topaloglu, R.O., 2020. Explainable DRC hotspot prediction with random forest and SHAP tree explainer. Proceedings of the 2020 Design, Automation and Test in Europe Conference and Exhibition, DATE 2020, 1151–1156. https://doi.org/10.23919/DATE48585.2020.9116488.

Zhang, W., Gu, X., Tang, L., Yin, Y., Liu, D., Zhang, Y., 2022. Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: comprehensive review and future challenge. Gondwana Res. 109, 1–17. https://doi.org/10.1016/J.GR.2022.03.015.

Zhang, X., Li, P., 2014. Lithological mapping from hyperspectral data by improved use of spectral angle mapper. Int. J. Appl. Earth Obs. Geoinf. 31, 95–109. https://doi.org/10.1016/J.JAG.2014.03.007.

Zhou, H., Fu, L., Sharma, R.P., Lei, Y., Guo, J., 2021. A hybrid approach of combining random forest with texture analysis and VDVI for desert vegetation mapping based on UAV RGB data. Rem. Sens. 13, 1891. https://doi.org/10.3390/rs13101891.

Zhou, Z., Hooker, G., Wang, F., 2021. S-LIME: stabilized-LIME for model explanation. In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, KDD '21. Association for Computing Machinery, New York, NY, USA, pp. 2429–2438. https://doi.org/10.1145/3447548.3467274.