Application of Machine Learning Techniques for Remote Sensing in Lithological Mapping: a Review

Tanushree Parmar and Romit S. Beed

February 11, 2025

Application of Machine Learning Techniques for Remote Sensing in Lithological Mapping: A Review

Tanushree Parmar[1], Romit S. Beed[2]

[1,2] Department of Computer Science, St. Xavier's College (Autonomous), Kolkata
[1] tanushree82.chauhan@gmail.com
[2] rbeed@yahoo.com

Abstract:

Identification of the geological features with the help of maps plays a crucial role in mineral exploration. The traditional way of geological data collection in remote areas is time-consuming and also challenging, thus the images extracted from the Remote Sensing technique are followed by image processing, which assists in classifying and improving geological mapping more accurately. Remote Sensing technology can now map different litho-units and associated structural features with greater precision and speed. In this paper, a comparative review is done based on remote sensing and image processing using different machine-learning methods. Images obtained from many satellites popularly used by geologists for geosciences have been explored and compared. Secondly, different machine-learning methods for processing these images are analyzed and their accuracy is compared. The comparative result concludes that the satellites that are first and foremost for such studies are of multispectral types (e.g. ASTER, Sentinel-2 and Landsat) due to the historical coverage. The survey concludes that when Landsat-8 images are used with the SVM give output accuracy of more than 80%. At the same time, the Random Forest is a technique, that uncovers the potential of remote sensing to address the emerging problems in Geographic Information Science. With only a sizable number of experiments performed through deep learning, provides results with more than 90% accuracy highlighting the supremacy of deep learning in geographical and remote sensing applications.

Keywords:

Lithological Mapping, Mineral exploration, Machine Learning, Remote Sensing, Spectral reflectance

Introduction:

The major constituents of the Earth's surface are water bodies such as oceans, rivers and lakes, only a small 29% of the Earth's surface is occupied by land in the form of mountains, hills, plains and islands. This small area includes a series of diverse elements such as forests, deserts, vegetation, urban areas, outcrops of different types of rocks and so on. Thus, the mapping of these rocks is a challenging task to conduct (Mani *et al*., 2021).

These maps contain structure details, lithological units, alteration types and many more features related to mineral identification. Traditionally, the expert did this work by extracting the information, collecting the data and exploring the area over time. The conventional geological mapping method now switches to the Remote sensing technique, to gather information about the Earth's surface. The data acquisition from the Remote Sensing technique is spectral, temporal and spatial; which helps the geologist to provide proper identification of minerals without accessing the remote area. It even helps in geological survey mapping, analysis and interpretation of the work (Wang *et al*., 2024). Further, it provides unique information related to the extraction of different geological features exposed on the earth's surface.

In remote sensing each pixel of the image corresponds to the spectral vector of reflectance value with a specific wavelength which is further compared with the spectral characteristics of the mineral object, this helps to detect the type of mineral with the same characteristics. In remote sensing sometimes it becomes difficult to capture the image of a mineral due to environmental disturbance. In such cases spectral absorption property of a mineral overlaps and even the light scattering effect leads to different spectral signatures. These limitations can be easily overcome with the help of Machine learning algorithms. It also helps in mineral exploration of even high-dimensional Data images. Unsupervised, supervised and semi-supervised are different types of machine learning algorithms that help in the proper exploration of the minerals from the image captured by the Remote Sensing technique. Integrating the remote sensing technique supervised by machine learning methods together generates low-cost and accurate solutions for lithological mapping. Geological maps provide detailed information regarding Earth's surface. The lithological map furnishes the characteristics of different types of rocks, their distribution and their properties (Omairi and Garouani, 2023)

1. Remote Sensing

It is a technique to acquire data about any object without being physically in contact with that object. It is mainly applied to detect and monitor the characteristics of any area by measuring its emitted or reflected radiation from a distance with the help of satellites or aircraft. (Blaschke *et al*., 2011, Mani *et al*., 2021). The electromagnetic spectrum technique used in remote sensing records the energy that Earth's surface reflects or emits. The properties of the object decide the reflected radiation from an object (Awange *et al*., 2013).

Remote sensing technology has enhanced the lithological mapping technique. It provides insight into details of even those areas that are unreachable or uncovered by the geologist using traditional methods. Lithological mapping uses various remote sensing techniques, for example, radar imaging (Radford *et al*., 2018), hyperspectral imaging (Peyghambari and Zhang, 2021) and multispectral imaging (Ghrefat *et al*., 2021). Multi-spectral imaging (MSI) differs from hyperspectral imaging (HIS) in just their band representation. MSI contains 3 to 10 bands while HIS has 100 to 1000 bands (Doneus *et al*., 2014). The mineral's spectral properties support the identification of different rocks or minerals by remote sensing technique (Ourhzif *et al*., 2019).

The multispectral imaging (MSI) technique has limits due to its low spectral resolution (Nicolis *et al*., 2021), making it quite hard to distinguish between surface objects with small variations. Nevertheless, these drawbacks can be removed through the use of machine learning techniques, which can stand for patterns as well as characteristics, thus increasing the speed and accuracy of multispectral images (Prost, 2014). The target of the study is to analyze the different types of satellites that are useful in mineral exploration and lithological mapping.

1.1 Satellite data

There are two types of Remote sensing processes classified according to the source of signal used to identify any object. Active Remote Sensing instruments explore the object by emitting their light while the passive Remote Sensing instruments depend on the sunlight emitted or reflected by the object. Each mineral has a unique reflectance value with the help of which it is been identified. For geological mapping, optical as well as radar data are used. Optical and data vary in terms of the range of wavelength used. Lee *et al*., (2020) in his paper reported that optical sensor works over the short range of the electromagnetic spectrum while Zhou and Guan, (2011) noticed that radio Remote Sensing Systems operate in the interval of 1 mm to 1m.

1.1.1 Optical data

ASTER deals with 14 spectral bands, which are imposed on the EOS Tera platform. The first three-band VNIR explores objects with a resolution of 15m while the next six-band SWIR has a resolution of 30m. The last five bands keep a record of thermal infrared radiation with a resolution of 90m. (Rowan and Mars 2003)

Table 2: ASTER Satellite Description

ASTER			
Swath Width= 60 km			
Year of Launch= 1999			
Subsystem →	**VNIR**	**SWIR**	**TIR**
Band number ↓	**Spectral Range**		
Band 1	0.520–0.600		
Band 2	0.630–0.690		
Band 3	0.780–0.860		
Band 4		1.600–1.700	
Band 5		2.145–2.185	
Band 6		2.185–2.225	
Band 7		2.235–2.285	
Band 8		2.295–2.365	
Band 9		2.360–2.430	
Band 10			8.125–8.475
Band 11			8.475–8.825
Band 12			8.925–9.275
Band 13			10.250–10.950
Band 14			10.950–11.650
Ground resolution (m)	15	30	90

Landsat 5, 7, 8 and 9 are satellites that work as optical data providers and are useful in the geological mapping process. For over four decades, they have been constantly tracking the Earth's surface to meet the needs of diverse information and data requests, to map various physical features of the Earth. (Wulder *et al*., 2008).

Landsat 5 was first active in 1984; it was equipped with both multispectral scanners and thematic mapper sensors. The Thermal Mapper band which is the thermal band collects the data over the visible Near Infrared (VNIR), shortwave infrared (SWIR) and thermal domains, it is also spatially resolved at 120 meters for the thermal band and at 30 meters for the rest of the bands (Banskota and Kumar, 2014).

On April 15, 1999, Landsat 7 was launched with an ETM + sensor. The major output that came out of the satellite was its work in visual and near-infrared bands, that is 8 spectral bands were used to collect the data which have a different special resolution. Out of 8 bands, the first four were of resolution of 30m, six bands were of thermal infrared with 60m resolution while the last two Bands having 15 M Resolution were panchromatic bands (Rajan *et al*., 2019).

Landsat 8 became functional on 11th February 2013 with a Thermal infrared sensor (TIRS) and OLI sensor. Landsat-8 contains 11 Spectral bands where the resolution of the image was the same as the ETM+ sensor that is, it contains VNIR and SWIR bands from range 1-7 while band 8 was a Panchromatic band. The special band 9 was mainly used to detect Cyrus cloud with a 30m resolution while the last two bands are thermal bands with a 100m resolution (Zang *et al*., 2016). The Landsat -9 is the latest satellite launched on September 27, 2021. It contains special features discriminating between natural changes and human changes, thus providing vital support in decision-making (Huanget *et al*., 2003).

Table 1: Landsat Satellites Description

LANDSAT 5		
Swath Width= 185 km		
Year of Launch= 1984		
Subsystem →	**TM**	**MSS**
Band number ↓	**Spectral Range**	
Band 1 Blue	0.45–0.52	
Band 2 Red	0.52–0.60	
Band 3 Green	0.63–0.69	
Band 4 NIR	0.76–0.90	
Band 5 SWIR 1	1.55–1.75	
Band 6 Thermal	10.40–12.50	
Band 7 SWIR 2	2.08–2.35	
Band 4 Green		0.50–0.60
Band 5 Red		0.60–0.70
Band 6 NIR 1		0.70–0.80
Band 7 NIR 2		0.80–1.10
Band 4 Green		0.50–0.60
Ground resolution (m)	30	57×79

LANDSAT 8		
Swath Width= 185 km		
Year of Launch= 2013		
Subsystem →	**OLI**	**TIRS**
Band number ↓	**Spectral Range**	
Band 1 Coastal Aerosol	0.43–0.45	
Band 2 Blue	0.45–0.51	
Band 3 Green	0.53–0.59	
Band 4 Red	0.64–0.67	
Band 5 NIR	0.85–0.88	
Band 6 SWIR 1	1.57–1.65	
Band 7 SWIR 2	2.11–2.29	
Band 8 Panchromatic	0.50–0.68	
Band 9 Cirrus	1.36–1.38	
Band 10 TIRS 1		10.60–11.19
Band 11 TIRS 2		11.50–12.51
Ground resolution (m)	30	100

LANDSAT 9		
Swath Width= 185 km		
Year of Launch= 2021		
Subsystem →	**OLI-2**	**TIRS-2**
Band number ↓	**Spectral Range**	
Band 1 Coastal Aerosol	0.43–0.45	
Band 2 Blue	0.45–0.51	
Band 3 Green	0.53–0.59	
Band 4 Red	0.64–0.67	
Band 5 NIR	0.85–0.88	
Band 6 SWIR 1	1.57–1.65	
Band 7 SWIR 2	2.11–2.29	
Band 8 Panchromatic	0.50–0.68	
Band 9 Cirrus	1.36–1.38	
Band 10 TIRS 1		10.60–11.19
Band 11 TIRS 2		11.50–12.51
Ground resolution (m)	30	100

**Source: https://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors

Sentinel-2 watches over Earth from a polar orbit. It snaps high-resolution, multi-colored pictures of the ground. Sentinel-2A took flight on June 23, 2015. Its buddy Sentinel-2B, joined the party on March 7, 2017. These space cameras help spot minerals on our planet's skin. VNIR and SWIR combined to form 13 bands. Four bands have a sharp 10 m view, the next six bands have a resolution of 20 m and the last three peek at 60 m (Drusch *et al*., 2012).

Table 3: Sentinel Satellite Description

SENTINEL			
Swath Width= 60 km			
Year of Launch= 1999			
Subsystem →	**VNIR**	**SWIR**	**TIR**
Band number ↓	**Spectral Range**		
Band 1	0.520–0.600		
Band 2	0.630–0.690		
Band 3	0.780–0.860		
Band 4		1.600–1.700	
Band 5		2.145–2.185	
Band 6		2.185–2.225	
Band 7		2.235–2.285	
Band 8		2.295–2.365	
Band 9		2.360–2.430	
Band 10			8.125–8.475
Band 11			8.475–8.825
Band 12			8.925–9.275
Band 13			10.250–10.950
Band 14			10.950–11.650
Ground resolution (m)	15	30	90

Digital Globe sells space pictures and map data. They also run civilian spy satellites like IKONOS, Quick Bird, GeoEye-1 and World View. Ball Aerospace and Technologies team was known to design these satellites. WorldView-1 was launched in 2007. It snaps black-and-white photos at 50 cm sharp. This camera has no color, but it takes detailed pics fast. It works great for making 3D maps (Good *et al*. 2018). WorldView-2 came next in 2009. It shoots 46 cm black-and-white images and 1.85m color images. It has color bands from 0.4 to 1.04 µm (Kruse and Perry, 2013). WorldView-3 blasted off in August 2014. It stands alone as the only commercial satellite with 16 bands of high-resolution Earth images. WorldView-3 sees eight bands at 3.7 m sharp from 1.2-2.33 µm. It also grabs eight bands at 1.2 m sharp from 0.42-1.04 µm. (Kruse *et al*., 2015).

2.1.2. Airborne data

In the case of Airborne Data, Specific Sensors are used, supported by drones or planes for data acquisition. Since 1997, the Geoscan Airborne Multispectral Scanner (AMSS) designed two sensors (MK I and MK II) that provide multispectral and hyperspectral airborne data. Lyon and Honey described the Geoscan AMSS Mk II scanner for the first time in 1990 and this tool was used for mineral exploration (Agar, 1994).

The first full spectral range imaging radiometer was AVIRIS (Advanced visual infrared imaging spectrometer). It was designed to scan objects with a spectral resolution of 10nm with the technique called whisk–broom scanning using 224 bands(Hamlin *et al*., 2011). The features of AVIRIS were improved and modified to design a new generation scanner called AVIRIS-NG, with an excellent signal-to-noise ratio. AVIRIS-NG is an HSI spectrometer with 380 nm to 2510 nm wavelength and 5 nm sampling (Tripathi and Gohil, 2019).

In 1999, Australia built and managed the hyperspectral mapper (HyMap). The sensor has 2 to 10 m of spatial resolution, covering wavelengths ranging from 0.45 to 2.48 µm with the help of 126 spectral bands (Ishidoshiro *et al*., 2016).

Hyperion Sensor initiated the Hyperspectral remote sensing technique. NASA's EO-1 Millennium Mission started in November 2000 when they launched Hyperion. It is one of the space-borne hyperspectral sensors with a 30m spatial resolution, 10nm spectral resolution of 10 nm, supported by 242 spectral bands (Pearlman *et al*., 2003).

Table 4: Airborne Satellite

Satellite/Sensor	**AVIRIS**	**AVIRIS-NG**	**Hy Map**	**Hyperion**
Year of Launch	1987	2012	1999	2000
Total Spectral Bands	224 spectral bands	427 spectral bands	126 spectral bands	242 spectral bands
Ground Resolution (m)	~20	~5	~5	30
Spectral Range (µm)	0.36–2.50	0.38–2.51	0.45–2.50	0.357–2.576
Swath Width (km)	~10.5	4–6	-	7.7×42

The extraction of datasets done with the help of different satellites or sensors was analyzed from the research papers and is visualized (Figure-1), where it is clear that the maximum multispectral satellites were used for mineral exploration with the help of remote sensing technology. It appears from Figure 1 that most of the study Landsat series satellites are used followed by ASTER and Sentinel-2.

The satellites used for the extraction of images as referred to in above mentioned papers are noted in Figure 1. It illustrates that most of the study used Landsat series satellites followed by ASTER and Sentinel-2. While the least used satellite is TOPSAR. Further, it appears that Landsat, being a widely used satellite because of its long historical archive from 1972 as well as its global coverage and consistency make it famous among geologists in mineral exploration.

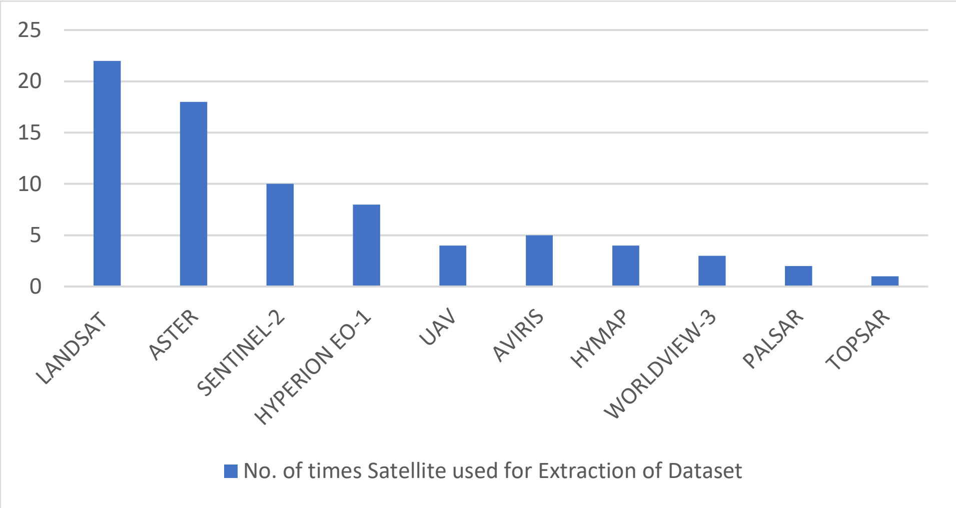

Figure 1: Bar diagram illustrating the use of satellites in mineral exploration from 2003 to 2024.

2. Machine Learning:

The important step in mineral exploration is accurately mapping different geological features to generate the lithological map. Machine learning plays an important role in developing different types of classified maps with the support of Remote Sensing data. It maps different lithological units, alteration units, structures and many more leading to an easy and inexpensive approach. Sun and Scanlon (2019) stated that Remote sensing techniques, combined with Machine learning methods, can be used to explore big data with high resolution.

ML methods are subdivided mainly into two parts- The First is Supervised Learning and another is Unsupervised Learning. ML Method which contains Supervised Learning, is implemented with labeled data to help in the prediction or classification of problems, which model input features and desired output (Kotsiantis, 2007). These input features generated through remote sensing techniques might be noisy or uncertain which is further improved using different ML techniques, thus making the model robust in nature concerning spectral analysis and ground truth measurement (Gewali *et al., 2018).*

In unsupervised Learning, target labels are not required to recognize the patterns. It supports data reduction techniques like PCA, ICA, MNF and clustering techniques like K-Means and ISO data. In the case of Image processing, clustering techniques are preferred (Xie *et al.*, 2020).

The Big data analysis follows the Dimension Reduction Method to make the data robust by removing noise and outliers, decreasing the complexity of a problem; as in smaller and proper datasets the accuracy to find the optimal solution increases (Caggiano *et al.*, 2018). ML makes strategic decisions based on data analysis and interpretation by processing high-dimensional data into low-dimensional data by predicting some trends or characteristics of data (Cracknell and Reading, 2014). Omairi and Garouani, (2023) stated that the integrated approach of

Remote Sensing with Machine Learning methods can enhance our knowledge related to insights features of geology.

2.1 Dimensionality Reduction Technique:

It aims to reduce the number of input features while retaining its maximum originality by compressing factors after analyzing its factors properly. PCA, ICA and MNF are the different techniques of reduction that transform input variables that are correlated to each other into independent components (Nielson, 2011, Shirmard *et al.,* 2020). For Gold exploration, Sheikhrahimi *et al.*, (2019), Traore *et al.,* (2020) and Abdolmaleki *et al.*, (2020) used PCA and MNF methods along with the images captured from satellites like ASTER, landsat8 and sentinel-2 respectively to recognize and map the area with maximum gold extract components. For image transformation, PCA was used to provide fast and cost-efficient means to delineate lithological and alteration units.

EI Atillah *et al*., (2019) used PCA and generated false-color composite (FCC) images for map lineament obtained from ASTER in the area of Morocco. This study resulted in 74% satisfying performance. Even different dimensionality reduction techniques were compared by Shirmard *et al*., (2020) and efficiently mapped different alteration types in geology. In the research papers, written by Zhang *et al*., (2020) and Wang *et al*., (2020), cracks and faults were identified respectively using the PCA concept.

Even some researchers used Aster and Landsat 8 OLI sensor data in lithological mapping. Wang *et al*., (2024) used an algorithm named as light gradient boosting machine (lightGBM) for geological mapping along with PCA and Random Forest. In this study, geochemical sample points were trained and a new model was designed which used the Softmax function. The function was iterated five times to identify the lithological unit. A new and accurate mapping technique was conducted in the Duolong mineral district, Tibet, China.

Table 5: Satellite and Dimensionality Reduction Techniques used in research paper

Dimensionality Reduction Techniques		
Research Paper	**Satellite / Sensor**	**Method**
Ghoneim *et al*., (2024)	Landsat 8/9	Principal Component Analysis
Wang *et al*., (2024)	ASTER, Landsat 8 OLI	
Zhang *et al.,* (2020)	UAV	
Wang *et al.,* (2020)	UAV	
Traore *et al.,* (2020)	Landsat 8	
Shirmard *et al*., (2020)	ASTER	
Abdolmaleki *et al.,* (2020)	Sentinel-2	
Sekandari *et al.,* (2020)	ASTER, Sentinel-2, Landsat 8	
Sheikhrahimi *et al.,* (2019)	ASTER	
El Atillah *et al.,* (2019)	ASTER	
Ghoneim *et al*., (2024)	Landsat 8/9	Minimum Noise Fraction
Lorenz *et al.,* (2021)	UAV	
Traore *et al.,* (2020)	Landsat 8	

2.2 Support Vector Machine (SVM):

It is a type of classification method. Wang and Zhang (2010) used SVM Classifier to extract the rock and mineral composition by mapping Hyperion images in the northwest region of China. Lithological and mineralogical knowledge procured from the Hyperion satellite is rather similar to the conventional geological map derived for gold extraction. (Shirmard *et al*., 2021). SVM also supports the design mineral prosperity map that justifies enough mineral extract in a particular area (Abdokmaleki *et al*., 2020). Lorenz *et al*. (2021) mentioned the use of drone technology along with SVM for mapping different mineralization zones.

SVM is used to identify the best hyperplane separating the input support vectors or training samples from different classes. It can be done by adjusting kernel functions like the Sigmoid function, polynomial function, or linear function and penalty parameters (Huang *et al*., 2002). SVM provides accurate results even when the sample size is small and has high dimensionality (Srivastava *et al*., 2012). SVM has revolutionized the concept known as spectral-based lithological Mapping. Various applications use SVM prominently (Othman and Gloaguen,2014). Yu *et al*., 2012 have reported an automated lithological categorization in their work in the northern part of India. While Bressan *et al*., (2020) compared SVM with Random neural network method to categorize geological features near offshores. Bachri *et al*., (2020) combined the spectral characteristics of Landsat8, DEM and Geomorphometric properties with the SVM technique to evaluate the effective lithological mapping.

Rezaei *et al*., (2020) identified different rock types with the support vector machine technique in mapping lithological units. In their study, 80% accuracy was calculated for using SVM. Band Ratio Spectral angle mapper (SAM) to classify primary rocks. In some papers, the object-based classification technique was compared with SVM. It was observed that object-based classification is more accurate than SVM, while using hyperspectral images. SVM scored a 76.30% accuracy rate and 0.719 kappa coefficient while the accuracy and kappa coefficient of Object-based classification were 81.30% and 0.779 (Petropoulos *et al*., 2012). Similarly, Ge *et al*., (2018) also used Sentinel -2 data, on which SVM and object-based classification method was applied in Mongolia regions. In this region, the lithological map was created to detect different sedimentary rocks and ultramafic rocks by using spatial and spectral features, an accuracy of 90.83% (kappa coefficient of 0.885) was achieved with object-based classification. Ourhzif *et al*., (2019) analyzed lithological groups using ASTER and Landsat 8 OLI data and calculated that the Landsat 8 OLI satellite gave 22.4% more accuracy than ASTER Data. The accuracy of ASTER data was 74.88% (kappa coefficient = 0.71) while the accuracy of Landsat 8 OLI data was 97.28% (kappa coefficient = 0.97).

Ghonein et al., (2024) in their study discriminated the various outcrop lithological rocks with the help of Landsat 8/9 satellite in the Eastern desert part of Egypt. They used the SVM classifier, PCA and MNF to design a geological map.

Table 6: Types of Satellite and SVM Techniques used in the research paper

Method = SVM	
Research Paper	**Satellite / Sensor**
Mahboob *et al*., (2024)	Landsat 8, Sentinal-2
Ghoneim *et al*., (2024)	Landsat 8/9
Shirmard *et al*., (2021)	Hyperion
Lorenz *et al*., (2021)	UAV
Abdolmaleki *et al*., (2020)	Sentinel-2
Cardoso-Fernandes *et al*., (2020b)	Sentinel-2
Bachri *et al*., (2020)	Landsat 8
Rezaei *et al*., (2020)	Landsat 8

Xu *et al.*, (2019)	ASTER
Ourhzif *et al.*, (2019)	ASTER, Lansat-8 OLI
Ge *et al.*, (2018)	Sentinel -2
Othman and Gloaguen (2014)	ASTER
Wang and Zhang (2010)	Hyperion

2.3 K-means clusters Method:

According to Lloyd (1982), K-means clusters (KMC) are a group of clusters made by separating N number of data points which have the smallest difference among its value or mean value. The K-means clustering technique can handle high-dimensional datasets (Tang *et al*.,2019). KMC is mainly used to recognize objects generated using hyperspectral images. Ren *et al*., (2019) in their new KMC algorithm used for mining purposes, classify minerals using hyperspectral images. In the study, he proved that the KMC algorithm shows a stronger clustering result as compared to the conventional algorithm method used for designing various mineral distribution maps. Even in many research papers, KMC is combined with the image extracted from Remote Sensing to generate accurate geological maps. The remote sensing data extracted from ASTER, Sentinel 2, Landsat-7 and Landsat-8 were combined with the KMC model to generate the lithological map (El Atillah *et al*., 2019). Later, the result was compared with the field to check the accuracy of the KMC model.

The iterative self-organizing data processing technique (ISODATA) is an unsupervised classification approach mainly applied to the multispectral and hyperspectral images generated from satellites (Karimi and Peng, 2004). The spectral reflectance of the band is also used to generate different clusters.

Table 7: Types of Satellite and K-Means Clustering Technique used in research paper

Method = K-means Clustering	
Research Paper	**Satellite / Sensor**
Ren et al. (2019)	AVIRIS
El Atillah et al. (2019)	ASTER, Sentinel-2, Landsat 8, Landsat 7
Karimi and Peng, (2004)	ASTER, Sentinel-2, Landsat 7

2.4 Random Forest Method:

Leo Breiman in 2001 introduced a new concept called Random Forest. It is the supervised learning method that is used to create multiple subtrees depending on the parameters (Ganuer *et al*., 2010). This classifier is non-parametric. It uses a voting procedure to ensemble trees for classification problems (Dragut,2016). Random Forest uses the concept called the Gini Index, which is used for splitting the input variable and finding the best threshold value as well while implementing this method, majority voting is also considered (Parmar *et al*., 2019). In this ensemble ML method is used, which is used to create each tree according to its subset while the overall prediction is based on all the trees in the ensemble therefore It is also called the bagging technique (Richman and Wuthrich, 2020).

Kuhn *et al*., (2018) reported that Random Forest acts as an essential tool in gold-rich mineral exploration when integrated with the Geo-physical data and Remote sensing data. Even Cardoso Fernandes *et al*., (2019) used SVM along with Random Forest to identify Li-bearing pegmatites using Sentinel-2 images. Remote sensing data provide textual, Geomorphic and spectral information. Bachri *et al*., (2020) Mapped lithological unit using Sentinel 2 and Palsar data along with random forest technique. Even sometimes to explore the efficiency of metals, multi-sensor data along with random forest are calculated (Wang *et al*., 2020). Random forest is also helpful in building a model for specifying the hyperspectral band for the classification of minerals like dolomite, talc and calcite (Chung *et al*., 2020).

Lu *et al*., (2022) used satellite Landsat 8 data, the data was collected based on parameters like moisture, greenness, reflectance, brightness and temperature on which the Kruskal-Wallis rank test was evaluated to find the significant difference among different Rock types on each parameter separately. When all the parameters were combined, 85.26% accuracy and 0. 77 Kappa coefficient was noted. It was a time series data so single data reflectance was checked. The combined accuracy indicates that the ground truth and procedure applied with random forest show the similarity.

Mahboob et al., (2024) use machine learning algorithms like SVM, Random Forest and CNN to find copper Cu deposits in the Northern Pakistan region. Landsat 8 and Sentinel-2 satellite data were used. A prospective map was designed which concluded that the Random Forest model was more accurate in detecting copper deposit areas. The prediction model was evaluated using a Confusion Matrix, statistical measures and Receiver Operating Characteristics (ROC) curve.

Algorithms like K nearest neighbor, Random Forest and Max likelihood were trained to find the best classification accuracy. Among these three, Random Forest proved itself more potential and gave 88.38% accuracy, which indicates high accuracy between ground truth examined and classified data (Bachri *et al*., 2022). Xi *et al*., (2022) also compared the random forest technique with the data extracted from Sentinal-2, Landsat 8 and ASTER; in which ASTER data accuracy came out to be highest at 81.80% followed by Sentinal-2 accuracy of 81.60%.

Table 8: Type of Satellite and Random Forest Technique used in the research paper

Method = Random Forest	
Research Paper	**Satellite / Sensor**
Wang *et al.*, (2024)	ASTER, Landsat 8 OLI
Mahboob *et al.*, (2024)	Landsat 8, Sentinal-2
Xi *et al*., (2022)	ASTER, Sentinel-2, Landsat 8
Bachri *et al.*, (2022)	ASTER
Lu *et al*., (2022)	Landsat-8
Wang *et al.*, (2020)	ASTER, Sentinel-2
Bachri *et al.*, (2020)	Sentinel-2, PALSAR
Chung *et al.*, *(*2020)	Sentinal-2
Cardoso-Fernandes *et al.*, (2019)	Sentinel-2
Kuhn *et al.*, (2018)	Landsat 5
Cracknell and Reading (2014)	Landsat 7

3. Deep Learning:

Guo *et al.*, (2016) in their study, show that the deep neural network method supports both unsupervised learning as well as supervised learning. The deep learning method even supports semi-supervised learning where some parts contain labeled examples while a large number of unlabeled examples are used to build a model. Large and complex datasets can be loosely grouped into recurrent neural networks and feed-forward architecture using deep learning methods (Shrestha and Mahmood, 2019).

3.1 Artificial neural network (ANN):

ANN working is based on neurons to control the task. This classifier is designed to illuminate how humans classify different patterns and how they learn to solve any tasks or problems. A machine learning method that deals with complex patterns keeps a record of the functionality of dependent variables and explanatory variables and is in non-linear form is termed an Artificial Neural Network (lek and Guegan,1999). A potential map of gold mineralization was designed along with the unknown potential area that can be further explored for gold mining. This study was conducted in the southeast part of Spain popularly known as the Rodalquilar gold mines, with the integration of Remote sensing data and ANN technology (Rigol-Sanchez *et al*.,2003).

Landsat5 and Hyperion data together can be used for lithological mapping. Thematic Mapper sensor data was used by Leverington, (2010) along with ANN to develop a model to discriminate different lithological groups. Even Wang *et al*., (2010) used a Neural network (Probabilistic type) to examine irregularity among multi-minerals caused by geo-field features. In the Henan region of China, the study was conducted to develop a potential map of Pb-Zn-Ag and molybdenum minerals. Especially Satellite imagery along with the regional geological maps were visualized using the ANN method and various lineaments were compared for mapping (Borisova *et al*.,2014)

Bouwafoud *et al*., (2021) compared the ANN model with the spectral distance index (SID) model with the help of Landsat 8 OLI data. In this study, ANN proved efficient with an accuracy of 92.56% while SID was 49. 61% accurate only.

Multilayer perception (MLP), is used to address problems related to different fields with geological mapping (Shirmard *et al*, 2022). In this ground truth data can be trained with the help of a multi-layer perception model to classify multispectral as well as hyperspectral satellite images which can provide highly accurate results. Various lithological maps were generated by the researchers using MLP for multispectral images using the Envy platform (Otele *et al*., 2021) MLP technique when used with multi-spectral images, leads to faster and more accurate results (Venkatesh and Raja, 2003). Shirmard *et al*., (2020) introduced the application of multi-layer perception and self-organizing maps in classifying multispectral images generated from Landsat satellite using PCA component as input. The deep neural network can be divided into a multi-layer perception (MLP), recurrent neural network (RNN), convolutional neural network (CNN) and graph neural network (GNN). Among them, CNN is useful in image processing purposes (Lecun and Yoshua, 1998).

Table 9: Showing Satellite and Artificial Neural Network Techniques used in the research paper

Method = ANN	
Research Paper	**Satellite / Sensor**
Bouwafoud *et al*., (2021)	Landsat -8 OLI
Shirmard *et al*., (2020)	Landsat-8
Borisova *et al*., (2014)	Landsat 7

Leverington (2010)	Hyperion, Landsat 5
Wang *et al.*, (2010)	Landsat 7
Rigol-Sanchez *et al.*, (2003)	Landsat 5
Venkatesh and Raja, (2003)	Landsat 5

3.2 Convolutional Neural Network (CNN)

Deep Neural Network is also called Convolutional Neural Network. CNN technique is mainly used in image processing and target identification by extracting features from different images (Li *et al.*, 2020). CNN uses pooling layers as well as convolution layers to extract spatial information from the images. Features maps are generated by convolution layers depending on the input given to the kernel. This kernel is in the form of a weight matrix (2-dimensional). Here the neighboring pixel plays an important role in the formation of different textures. Shirmard *et al.*, (2022) classify each pixel based on the digital numbers of the band. In this study, a chip (sets of neighboring pixels input in CNN) was created using the size of stride and window. Shuo and Kang, (2021) introduce an improved version of CNN, which deals with overcoming the problems related to gradient vanishing and overfitting. Newman *et al.*, (2017) focus on the high speed and performance ratio using the CNN model.

Tahmasebi *et al*., (2020) noted that many CNN architectures are openly available to solve the limitation of computation resources when dealing with large numbers of data sets. Elman (1990) stated that in case of the temporal or dynamic data, RNN proves itself more accurate in finding the result. Long short-term memory (LSTM) is another type of recurrent neural network that can be used for mineral explosion (Hochreiter and Schmidhuber, 1997). CNN when combined with LSTM, is useful to handle video processing data (Xia *et al.*, 2020). Even it can be useful for spatiotemporal datasets when used with remote sensing applications (Boulila *et al., 2021).* Wang and Zuo (2024) evaluated that CNN can be used for lithological mapping using hyperspectral images also. They concluded that minerals that closely resembled and are strongly correlated can be explored with the help of hyperspectral satellites. Gaofen-5 satellite images were examined and an accuracy of 97.40% was measured.

Table 10: Type of Satellite and Convolutional Neural Network Technique used in research paper

Method = CNN	
Research Paper	**Satellite / Sensor**
Mahboob *et al.*, (2024)	Landsat 8, Sentinal-2
Shirmard *et al.*, (2022)	Landsat-8, Sentinel-2, ASTER
Sang *et al.*, (2020)	UAV
Zhao *et al.*, (2020)	AVIRIS
Latifovic *et al.*, (2018)	Landsat 7, Landsat 5

4. Importance of Explainability

Explainability is vital in deploying machine learning models for lithological mapping, as it ensures trust and understanding among geologists and decision-makers. In fields like mineral exploration, where model outputs often guide expensive fieldwork or financial investments, being able to interpret and explain predictions is critical. Techniques like Random Forest and PCA offer relatively high explainability, providing insights into variable importance or the contribution of transformed components to the variance. On the other hand, deep learning models like CNNs and ANNs, while highly accurate, are often perceived as black-box models due to their complexity. This lack of interpretability can hinder their adoption in geoscience applications, where domain experts demand transparency to validate results against geological realities.

5. Practical Challenges of Deployment

Deploying machine learning models for lithological mapping faces several practical challenges. First, the availability of high-quality, labelled datasets is often a bottleneck, particularly in remote or unexplored areas. Computational resources pose another constraint, especially for advanced models like CNNs, which require significant processing power. Additionally, integrating model predictions with on-field geological data is complex, requiring iterative validation and fine-tuning to match real-world conditions. Scalability is also an issue; for instance, methods like SVM may perform well on smaller datasets but struggle with larger, high-dimensional data. Finally, deployment in the field demands robust, resource-efficient tools that can handle diverse environmental conditions and deliver actionable insights under constrained resources. Addressing these challenges is essential to harness the full potential of machine learning in geosciences.

Result

This article reviews the implementation of the different types of Machine-learning algorithms that can be used for mineral exploration. The remote sensing technique deals with the hyperspectral scanner as well as multispectral scanner, which extracts images from space-borne, airborne and ground-based sensors. The comparative study predicts that the satellites mainly used are multispectral e.g. ASTER, Sentinel-2 and Landsat, due to their historical coverage.

Table 10: Accuracy Rate measured while using Remote sensing and ML Techniques

Techniques	**Research Papers**	**Satellite**	**Accuracy**
SVM	Rezaei *et al., 2020*	Landsat-8	80%
	Petropoulos *et al., 2012*	Hyperion	76.30%
	Ourhzif *et al., 2019*	ASTER	74.88%
		Landsat-8	97.28%
Object-Based Classification	Petropoulos *et al., 2012*	Hyperion	81.30%
	Ge *et al., 2018*	Sentinal-2	90.83%
Random Forest	Lu *et al., 2022*	Landsat -8	85.26%
	Bachri *et al., 2022*	ASTER	88.38%
	Xi *et al., 2022*	ASTER	81.80%
		Sentinal-2	81.60%
ANN	Bouwafoud *et al., 2021*	Landsat-8	92.56%
Spectral distance Index (SID)			49.61%
CNN	Wang and Zuo, 2024	Goafen-5	97.40%

The SVM when used with Landsat-8 images provides more than 80% accuracy. Object-based classification works best when used with remote sensing techniques to explore the area while the Random Forest technique overall proved its excellence in exploring the remote sensing images. Only a few experiments are conducted using deep learning but the accuracy with which the results obtain are more than 90%. As Earlier only unsupervised learning methods were used along with HIS because it requires large numbers of labelled data to work upon. But in the last few years, CNN Deep learning techniques have been used to excess hyperspectral images for lithological mapping.

Annexure 1 contains the summarized table, including the evaluation metrics used in each model to check the accuracy rate. It even describes the data requirements, explainability, practical challenges and limitations while implementing the models.

Conclusion and Future Prospectives:

The key issue is the quality and availability of data, with limited ground truth data and the challenge of distinguishing between minerals with similar spectral signatures. The complexity of geological environments, where surface data may not reveal subsurface deposits, adds to the difficulty. While deep learning models like CNNs can handle complex data, they often lack interpretability, making it hard for geologists to trust their predictions. Overfitting and the integration of various data sources are additional hurdles. Looking forward, advancements in sensor technology, such as hyperspectral imaging and better integration of multisource data will enhance our ability to identify mineral deposits accurately.

In the flourishing land of mineral exploration, deep learning offers state-of-the-art methods that markedly boost our chances for mineral detection. Conventional Neural Networks (CNNs) are the technological marvels of spectral and spatial data, the two of which are the deciding factors for the minerals with similar signatures. Integrating the RNN and LSTM enables temporal snapshots of a mineral deposit while the Transformer models provide them with a tool that can perceive relationships in data. Generative Adversarial Networks (GANs) function as a bridging technique for data augmentation and discrepancy detection in instances where the available amount of labeled data is not much. Also, targeted designs, UNet and SegNet are specialists in image segmentation showing the path to extract mineral mapping. The increase in the use of self-supervised learning and transfer learning extends the horizons of deep learning by employing labeled and unlabeled data in the best possible way. Multimodal deep learning that considers different data sources combined with 3D CNNs analyzing volumetric data gives a more detailed picture of mineral subsurface deposits.

In summary, the techniques presented here mark new ways of finding lithological maps and monitoring areas such as mining exploration, geophysical surveys and environmental monitoring through state-of-the-art computer automation systems. The crucial factor is the right choice of methods concerning the final goals and the available data. Sentiment of the research, when joining hands with Earth observation and ML, will help among other things, in getting the understanding of the geophysics of the Earth and the influence of mining and smelting on the environment. This not only opens up new possibilities in various fields but also enriches our knowledge, driving innovation and progress through these cutting-edge approaches.

References

A. Parmar, R. Katariya, V. Patel, "A review on random forest: an ensemble classifier," in international conference on intelligent data communication technologies and internet of things (ICICI) 2018, in: J. Hemanth, X. Fernando, P. Lafata, Z. Baig (Eds.), Lecture Notes on Data Engineering and Communications Technologies, Springer International Publishing, Cham, (2019), 758–763, https://doi.org/10.1007/978-3-030-03146-6_86.

A. Rezaei, H. Hassani, P. Moarefvand, A. Golmohammadi, “Lithological mapping in Sangan region in Northeast Iran using ASTER satellite data and image processing methods”, Geol. Ecol. Landsc. 4 (1) (Jan. 2020) 59–70, https://doi.org/10.1080/24749508.2019.1585657

A. Shebl, T. Kusky, “Advanced land imager superiority in lithological classification utilizing machine learning algorithms”, Arabian J. Geosci. 15 (9) (May 2022) 923, https://doi.org/10.1007/s12517-022-09948-w.

Atangana Otele, C.G.; Onabid, M.A.; Assembe, P.S.; Nkenlifack, M. “Updated lithological map in the Forest zone of the Centre, South and East regions of Cameroon using multilayer perceptron neural network and Landsat images”. J. Geosci. Environ. Prot., (9) (2021) 120–134.

Bachri, I., Hakdaoui, M. Raji, M., Benbouziane, A., Mhamdi, H., “Identification of lithology using sentinel-2A through an ensemble of machine learning algorithms”, Int. J. Appl. Geospatial Res. (IJAGR) 13 (1) (Apr. 2022) 1–17, https://doi.org/10.4018/IJAGR.297524.

Bachri, I., Hakdaoui, M., Raji, M., Benbouziane, A., “Geological mapping using random forests applied to remote sensing data: A demonstration study from Msaidira-Souk Al Had, Sidi Ifni inlier (Western Anti-Atlas, Morocco)”, in: IEEE International conference of Moroccan Geomatics (Morgeo), IEEE. (2020) pp. 1–5. https://doi.org10.1109/Morgeo49228.2020.9121888.

Borisova, D., Jelev, G., Atanassov, V., Koprinkova-Hristova, P., Alexiev, K., “Algorithms for lineaments detection in the processing of multispectral images”, in: Michel, U., Schulz, K. (Eds.), Earth Resources and Environmental Remote Sensing/GIS Applications V, (2014) p. 92451L. https://doi.org/10.1117/12.2067245

Bouwafoud, M. Mouflih, A. Benbouziane, “Lithological mapping using Landsat 8 OLI in the meso-Cenozoic Tarfaya Laayoune basin (south of Morocco): comparison between ANN and SID classification”, Open J. Geol. 11 (12) (2021) 658–681, https://doi.org/10.4236/2021.1112032.

C. Huang, L.S. Davis, J.R.G. Townshend, “An assessment of support vector machines for land cover classification”, Int. J. Rem. Sens. 23 (4) (Jan. 2002) 725–749, https://doi.org/10.1080/01431160110040323

Cardoso-Fernandes, J., Teodoro, A.C.M., Lima, A., Roda-Robles, E., “Evaluating the performance of support vector machines (SVMs) and random forest (RF) in Li-pegmatite mapping: Preliminary results”, in Schulz, K., Nikolakopoulos, K.G., Michel, U. (Eds.), Earth Resources and Environmental Remote Sensing/GIS Applications X, (2019) p. 26. https://doi.org/10.1117/12.25325

Chung, B., Yu, J., Wang, L., Kim, N.H., Lee, B.H., Koh, S., Lee, S., “Detection of magnesite and associated gangue minerals using hyperspectral remote sensing—A laboratory approach.” Remote Sensing (12), (2020), 1325. https://doi.org/10. 3390/rs12081325.

Cracknell, M.J., Reading, A.M., “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”. Computers & Geosciences (63), (2014), 22–33. https://doi.org//10.1016/j.cageo.2013.10.008.

F. Masoumi, T. Eslamkish, A.A. Abkar, M. Honarmand, J.R. Harris, “Integration of spectral, thermal and textural features of ASTER data using Random Forests classification for lithological mapping”, J. Afr. Earth Sci. (129) (May 2017) 445–457, https://doi.org/10.1016/j.jafrearsci.2017.01.028.

Ghoneim, S. M., Hamimi, Z., Abdelrahman, K., Khalifa, M. A., Shabban, M., & Abdelmaksoud, A. S. “Machine learning and remote sensing-based lithological mapping of the Duwi Shear-Belt area, Central Eastern Desert, Egypt”. *Scientific Reports*, *14*(1), (2024) 17010.

G.P. Petropoulos, C. Kalaitzidis, K. Prasad Vadrevu, “Support vector machines and object-based classification for obtaining land-use/cover cartography from Hyperion hyperspectral imagery”, Comput. Geosci. 41 (Apr. 2012) 99–107, https://doi.org/10.1016/j.cageo.2011.08.019

Kuhn, S., Cracknell, M.J., Reading, A.M., “Lithologic mapping using random forests applied to geophysical and remote-sensing data: A demonstration study from the Eastern Goldfields of Australia”. GEOPHYSICS 83, (2018) B183–B193. https://doi.org/10.1190/geo2017-0590.1

Lek, S., Guegan, J., “Artificial neural networks as a tool in ecological modeling, an introduction”. Ecological Modelling 120 (2-3), (1999) 65–73.

Leverington, D.W., “Discrimination of sedimentary lithologies using Hyperion and Landsat Thematic Mapper data: A case study at Melville Island, Canadian High Arctic”. International Journal of Remote Sensing 31, (2010), 233– 260. https://doi.org/10.1080/01431160902882637.

Mahboob, M. A., Celik, T., & Genc, B. “Predictive Modelling of Mineral Prospectivity Using Satellite Remote Sensing and Machine Learning Algorithms”. *Remote Sensing Applications: Society and Environment*, (2024), 101316.

P.K. Srivastava, D. Han, M.A. Rico-Ramirez, M. Bray, T. Islam, “Selection of classification techniques for land use/land cover change investigation”, Advances in Space Research 50 (9) (Nov. 2012) 1250–1265, https://doi.org/10.1016/j.asr.2012.06.032.

R. Genuer, J.-M. Poggi, C. Tuleau-Malot, “Variable selection using random forests”, Pattern Recognition. Letters. 31 (14) (Oct. 2010) 2225–2236, https://doi.org/ 10.1016/j.patrec.2010.03.014.

R. Richman, M.V. Wuthrich, “Nagging Predictors,” Risks 8 (3) (Aug. 2020) 83, https://doi.org/10.3390/risks8030083.

Rigol-Sanchez, J., Chica-Olmo, M., Abarca-Hernandez, F., “Artificial neural networks as a tool for mineral potential mapping with GIS”. International Journal of Remote Sensing 24, (2003), 1151–1156. https://doi.org/10.1080/0143116021000031791.

Sergi, R.; Solaiman, B.; Mouchot, M.C.; Pasquariello, G.; Pósa, P. “Landsat-TM image classification using principal components analysis and neural networks”. In Proceedings of the International Geoscience and Remote Sensing Symposium: Quantitative Remote Sensing for Science and Applications, Firenze, Italy, Volume 3, (10–14 July 1995); pp. 1927–1929.

Shirmard, H.; Farahbakhsh, E.; Beiranvand Pour, A.; Muslim, A.M.; Müller, R.D.; Chandra, R. “Integration of selective dimensionality reduction techniques for mineral exploration using

ASTER satellite data". Remote Sensing. 12 (2020), 1261

Shirmard, H.; Farahbakhsh, E.; Heidari, E.; Beiranvand Pour, A.; Pradhan, B.; Müller, D.; Chandra, R. "A Comparative Study of Convolutional Neural Networks and Conventional Machine Learning Models for Lithological Mapping Using Remote Sensing Data". Remote Sensing. (14) (2022), 819. https://doi.org/10.3390/rs14040819

Venkatesh, Y.V.; Kumar Raja, S. "On the classification of multispectral satellite images using the multilayer perceptron". Pattern Recognition. (36) (2003), 2161–2175.

W. Ge, Q. Cheng, L. Jing, C. Armenakis, H. Ding, "Lithological discrimination using ASTER and Sentinel-2A in the Shibanjing ophiolite complex of Beishan orogenic in Inner Mongolia, China", Advances in Space Research 62 (7) (Oct. 2018) 1702–1716, https://doi.org/10.1016/j.asr.2018.06.036.

Wang, G., Yan, C., Zhang, S., Song, Y., "Probabilistic neural networks and fractal method applied to mineral potential mapping in Luan Chuan region, Henan Province, China", in: Sixth International Conference on Natural Computation, (2010), pp. 1003–1007. https://doi.org/10.1109/ICNC.2010.5582906.

Wang, Z., Zuo, R., Dong, Y., "Mapping of Himalaya leucogranites based on ASTER and Sentinel-2A datasets using a hybrid method of metric learning and random forest". IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13, (2020) 1925–1936. https://doi.org/10.1109/JSTARS. 2020.2989509.

Wang, W., Xue, C., Zhao, J., Yuan, C., & Tang, J. "Machine learning-based field geological mapping: A new exploration of geological survey data acquisition strategy". *Ore Geology Reviews*, (2024), 105959.

Y. Lu, C. Yang, R. He, "Towards lithology mapping in semi-arid areas using time-series Landsat-8 data", Ore Geol. Rev. 150 (Nov. 2022), 105163, https://doi.org/ 10.1016/j.oregeorev.2022.105163.

Y. Xi, A.M. Mohamed Taha, A. Hu, X. Liu, "Accuracy comparison of various remote sensing data in lithological classification based on random forest algorithm", Geocarto Int. 37 (26) (Dec. 2022) 14451–14479, https://doi.org/10.1080/10106049.2022.2088859

Ourhzif, Z., Algouti, A., & Hadach, F. "Lithological mapping using Landsat 8 OLI and aster multispectral data in imini-ounilla district south high atlas of Marrakech. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42* (2019), 1255-1262.

Annexure:1

Table 11: Summarized table describing the evaluation Metrics, Data requirements, practical challenges and limitations experienced while implementing the Machine Learning and Deep Learning models.

Method	**Accuracy**	**Evaluation Metrics**	**Data Requirements**	**Explainability**	**Practical Challenges**	**Remarks**	**Limitations**	**Best Use Case**
Support Vector Machine (SVM)	>80%	Kappa Coefficient: 0.72	Requires labelled data	Moderate (supports decision boundary visualization)	Struggles with scalability on large datasets	Effective for small labelled datasets	Struggles with large datasets	Spectral-based lithological mapping
Random Forest	88.38 %	ROC, Confusion Matrix	Moderate labelled data	High (variable importance can be derived)	Overfitting risk without proper parameter tuning	Works well with multisensory data	Can overfit without parameter tuning	Mineral deposit detection using multi-sensors
K-Means Clustering	Varies	Cluster Compactness	Minimal labelling required	Low (unsupervised; cluster interpretability can be subjective)	Requires pre-selection of cluster count	Handles high-dimensional datasets	Requires optimal cluster count selection	Quick clustering of lithological features
Convolutional Neural Network (CNN)	>90%	Precision, Recall, F1-Score	Requires large labelled datasets	Low (black-box nature; difficult to interpret results)	Computationally intensive; high resource demand	Superior for image processing tasks	Computationally intensive	Lithological mapping with hyperspectral imaging
Principal Component Analysis (PCA)	74%	Variance Explained	None	High (transforms data to explain maximum variance)	Limited to linear relationships	Efficient dimensionality reduction	Limited to linear relationships	Preprocessing for dimensionality reduction
Object-Based Classification	81.30 %	Kappa Coefficient: 0.779	Requires spatial data	Moderate (segmentation-based results are	Dependent on segmentation quality	More accurate than SVM for spatial	Dependent on segmentation quality	Mapping geological features in spatial

Method	**Accuracy**	**Evaluation Metrics**	**Data Requirements**	**Explainability**	**Practical Challenges**	**Remarks**	**Limitations**	**Best Use Case**
				interpretable)	and thresholds	datasets		data
Light Gradient Boosting Machine (LightGBM)	~85%	Softmax Function Iterations	Requires labelled geochemical points	Moderate (interpretability through feature importance)	Sensitive to hyperparameters; may require frequent tuning	High accuracy with minimal computation time	High sensitivity to parameter tuning	Geological mapping in complex terrains
Artificial Neural Network (ANN)	92.56 %	Accuracy, RMSE	Requires labelled training data	Low (black-box nature; limited interpretability)	Overfitting risk; requires large and clean datasets	Effective for complex non-linear relationships	May overfit with insufficient data	Potential mapping for resource exploration