XXIII Conference of PhD Students and Young Scientists

Interdisiplinary topics in Mining, Geology and Geomatics

POLSKA AKADEMIA NAUK
ODDZIAŁ WE WROCŁAWIU
KOMISJA NAUK GÓRNICZYCH

POLSKA AKADEMIA NAUK
SEKCJA GÓRNICTWA PODZIEMNEGO
RUD MIEDZI KOMITETU GÓRNICTWA
POLSKIEJ AKADEMII NAUK

Unsupervised classification of lithological units using ASTER satellite data by machine learning algorithms, case study in a region in the Makran subduction zone in southern Iran

Peyman Namdarsehat

Faculty of Geoengineering, Mining and Geology, Department of Geodesy and Geoinformatics, Wroclaw University of Science and Technology, Wroclaw, Poland

Abstract

Over the past decade, Machine Learning Algorithms (MLAs) have achieved great success in accurately detecting and mapping geological targets compared to traditional methods. MLAs use artificial intelligence derived from input data by analyzing their features and patterns with minimal human intervention. In this study, MLAs were used to classify geological units in a selected area in southern Iran. After preprocessing the ASTER satellite data and performing the K-Means method, an unsupervised classification was performed to find out how efficient the result of the K-Means method is in the complicated geological zones with marine sediments up to 10 km thick compared to previous geological maps.

In this study, nine clusters were used for classification. The analysis and results of this comparative study demonstrate the potential of K-means clustering for developing classified geological units that are prominent in the study area, particularly when adequate surveys such as litho-stratigraphic and structural studies are not available. It can also be used for quality control of geological maps. This method can help us save costs before preparing a geological map.

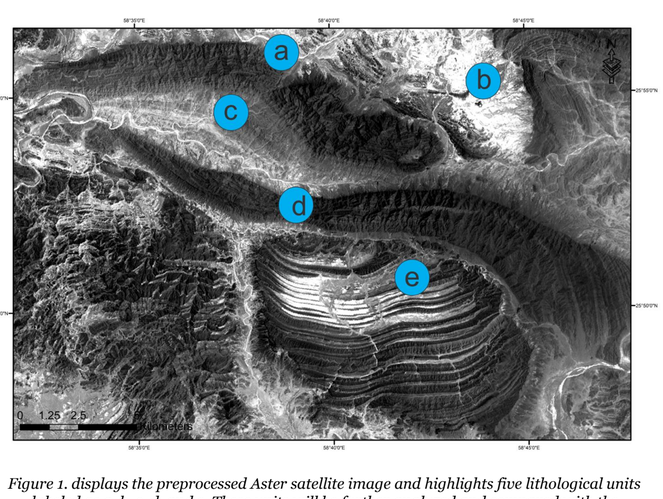
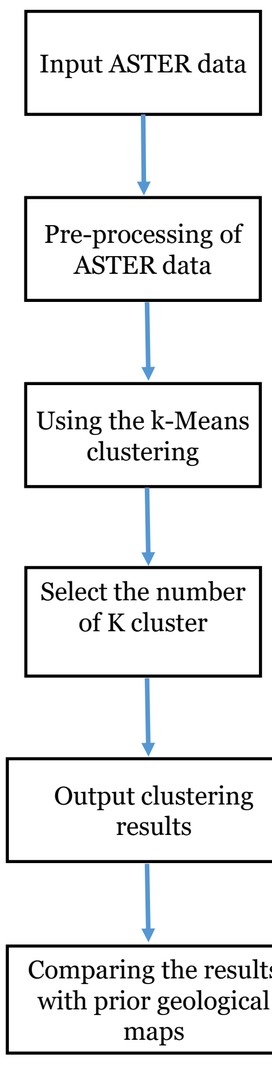


Figure 1. displays the preprocessed Aster satellite image and highlights five lithological units labeled as a, b, c, d, and e. These units will be further analyzed and compared with the results obtained from the K-Means algorithm and the geological map.

Introduction and Method

The introduction of machine learning algorithms (MLAs) has led to the development of semi-automatic tools for mapping targets in the geoscience field. The two commonly used learning methods are supervised and unsupervised approaches. The study was conducted using ASTER satellite data. The unique capabilities of ASTER, including its high-spatial-resolution multispectral bands, enable the discernment of various mineralogical distinctions. ASTER's data encompasses six bands in the SWIR (Short-Wave Infrared) range and five bands in the TIR (Thermal Infrared) range. These datasets provide ample information that can be effectively utilized through straightforward techniques like band ratios or more advanced analyses such as machine learning. In unsupervised classification, the pixels of a dataset are clustered based on statistics only, without defining any training classes. One popular unsupervised machine learning technique used in geology is K-Means clustering. The goal of K-Means clustering is to group similar data points based on their features, which can be helpful in identifying geological units based on their characteristics. The K-Means algorithm groups the dataset into K groups and minimizes the sum of the distances between each point and the nearest centroid. The method employed in this study offers several advantages over other unsupervised clustering techniques. It has low complexity and can perform fast computations, making it well-suited for handling large datasets.



Discussion and results

To begin the analysis, the algorithm places a predetermined number of centroids randomly in the data space. In this study, nine clusters were used to analyze ASTER satellite images following preprocessing. After integrating all the data, including the geological map, ASTER processing, and K-Means clustering, the comparison indicates that four units (b, c, d, and e) are consistent with the earlier geologic map that was produced independently of the classification. However, further field studies are needed to verify the classification of unit a, as it is not fully consistent with the previous map.

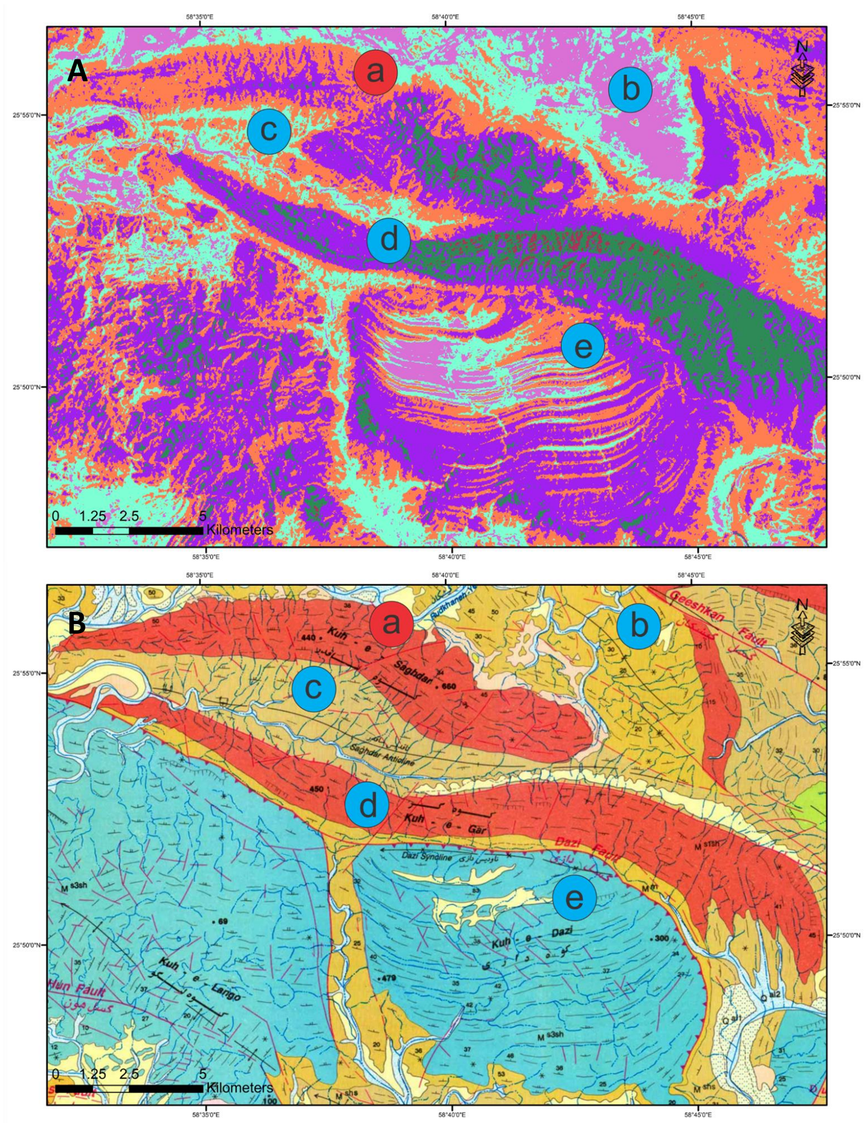


Figure 2. presents a comparison of the results obtained from the K-Means clustering (A) and the geological maps (B).

Conclusion

The processes performed on the satellite images in this study were aimed at improving the identification of geological and lithological units and determining their extent. These processes are of great value in the exploration phase of geological projects. Satellite classification plays a crucial role in this process and, when integrated with surface data, can provide very accurate results in a time-saving manner.

The results of this study highlight the effectiveness of unsupervised classification (K-Means algorithm) as a reliable method for producing accurate geological maps. These maps serve as a solid foundation for geological studies and can help reduce costs prior to project initiation. The use of unsupervised classification techniques can provide comprehensive and accurate geological information that facilitates informed decision making and minimizes the risks associated with exploration activities.

References

[1] Geological Survey of Iran, 1996. Geological quadrangle map of Iran, No. 7741, Pibeshk, Scale 1:100,000

[2] Govender, P. and Sivakumar, V., 2020. Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019). *Atmospheric pollution research*, *11*(1), pp.40-56.

[3]Michie, D., Spiegelhalter, D.J. and Taylor, C.C., 1994. Machine learning, neural and statistical classification.