Remote Sensing Applications: Society and Environment 24 (2021) 100643

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

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

Stacked vector multi-source lithologic classification utilizing Machine Learning Algorithms: Data potentiality and dimensionality monitoring

Ali Shebl [a,b,*], Árpád Csámer [a]

[a] Department of Mineralogy and Geology, University of Debrecen, Hungary
[b] Department of Geology, Tanta University, Egypt

ARTICLE INFO

Keywords:
Lithologic classification
Support vector machine
Artificial neural network
Maximum likelihood classifier
Sentinel 2
Sentinel 1
ASTER

ABSTRACT

Machine Learning Algorithms (MLAs) have recently introduced considerable lithologic mapping. Thus, this study scrutinizes the efficacy of Artificial Neural Network (ANN), Maximum Likelihood Classifier (MLC) and Support Vector Machine (SVM) over hybrid datasets including optical (Sentinel 2, ASTER, Landsat OLI and Earth-observing 1 Advanced Land Imager (ALI)), radar (Sentinel 1 and ALOS PALSAR), DEMs and their derivatives (Slope, and Aspect). The study aims to (1) monitor the effect of data dimensionality in enhancing categorization accuracy. (2) disclose the most efficient MLA and most powerful dataset in labeling rock units accurately. (3) highlight the impact of embedding topographical and radar data in lithologic classification. (4) outline the best relation between the number of training pixels and number of utilized bands, in delivering reliable allocation. To achieve these aims, we selected training and testing pixels meticulously, in concordance with a recently published geological map of the study area. We adopted a stacked vector approach for handling the implemented multi-sensor data. Results show that diversifying information sources raised the classification accuracy by approximately 10% for each classifier. SVM and MLC are much better than ANN. Slope is better than aspect and both are less qualified when compared to DEM. Sentinel 1 (C-band) and ALOS PALSAR (L-band) effects are not so different whatever the implemented polarization. Landsat OLI is less qualified in lithologic classification when compared to Sentinel 2, ASTER and ALI. The utilized training pixels should be at least 30N for (N) channels submitted to the classifiers.

1. Introduction

Accurate geological mapping is the key for economic minerals localization and geo-hazards mitigations. Enormous improvements in lithologic mapping have been introduced by the availability of a wide range of remote sensing datasets (Grebby et al., 2011). The digital format of the latter exceedingly prompted Machine Learning Algorithms to be applied in land cover and geologic mapping; using multispectral and/or hyperspectral data (Foody and Mathur, 2004; Ham et al., 2005; Lee et al., 2012; Liesenberg and Gloaguen, 2012; Pal and Mather, 2005). Efficient lithologic classification was a matter of interest for several researchers over the last decade. The classifier performance may vary depending on the implemented datasets, thus a wide range of datasets are implemented for mapping rock units using several classifiers. Most of the results highlighted SVM (Bachri et al., 2019; Bentahar and Raji, 2021; Kumar et al., 2021; Manap and San, 2018), MLC (Fatima et al., 2013; Ge et al., 2018; Hadigheh and Ranjbar, 2013; Jellouli et al., 2016; Mehr et al., 2013), and ANN (Grebby et al., 2011; He et al., 2015). Thus, we decided to intensively investigate the performance of these classifiers over a wide range of the utilized data sets. On the other hand and for datasets, previous studies highly recommend using ASTER in lithologic mapping (Jellouli et al., 2016; Othman and Gloaguen, 2014, 2017; Yu et al., 2012), others (Bentahar and Raji, 2021; Fatima et al., 2013; He et al., 2015) utilized Landsat data effectively. With the advent of the recently launched Sentinel 2 (S2) data, several studies highlighted the advantages of S2 in categorization rock units due to its higher spectral resolutions (Bentahar and Raji, 2021; Ge et al., 2018). However (Pour and Hashim, 2014), stated that ALI is an efficient source for lithological and mineralogical mapping due to its excellent VNIR spectral separability. Thus and for the first time, this study implied the four widely used sensors in lithologic mapping.

With the proven role of geomorphological and topographical data in enhancing lithological allocation (Bachri et al., 2019; Grebby et al., 2011; Othman and Gloaguen, 2017; Yu et al., 2012), the current study embedded not only DEM but also its derivatives (slope and aspect maps) to inspect their role in identifying rock units. Recently (Kumar et al.,

* Corresponding author. Department of Mineralogy and Geology, University of Debrecen, Hungary.
E-mail address: ali.shebl@science.tanta.edu.eg (A. Shebl).

https://doi.org/10.1016/j.rsase.2021.100643
Received 14 June 2021; Received in revised form 22 August 2021; Accepted 29 September 2021
Available online 2 October 2021

2021), utilized sentinel 1 data in geological mapping, thus Sentinel 1 radar data is also considered in this study to enhance the mapping process. Authors of (Kumar et al., 2021) highly recommend the integration of several sensors for superb geological mapping especially with MLAs. Consequently, and for the first time, This study integrated Sentinel 2, ASTER, Landsat OLI, and Earth-observing 1 Advanced Land Imager, Sentinel 1, ALOS PALSAR-1, ALOS PALSAR DEM, Slope, and Aspect to map 9 lithological units. Moreover, and due to a large number of the included bands the study closely monitors the effect of data dimensionality on the classifier performance and overall accuracy. Furthermore, a detailed examination of the effect of topographic and radar data when added to each optical sensor is inspected to highlight the reliable data combinations for delivering superior outputs, through performing more than 100 classification process. The investigated area was selected for this work as it is famous for gold mining and have various rock units, as well as, it has a recently published geological map which is useful for comparing the results and verification.

2. Materials and methods

2.1. Study area description

Um Salatit –Mueilha is located at the extreme southern part of the Central Eastern Desert (CED), Egypt as shown in Fig. 1. It extends between latitudes 24° 40″ to 25° 20″ N and longitudes 33° 45″ to 34° 05″ E, and covered by an intricate complex of Precambrian rugged igneous and metamorphic rocks, in addition to, Phanerozoic sedimentary Nubian sandstone. Authors of (Zoheir et al., 2019a) stated that the area is covered mainly by a widely distributed stretch of Neoproterozoic ophiolitic mélange constituted mainly of allocthonous ophiolitic fragments embedded in a sheared matrix, as well as, other different mappable units (Fig. 1). The vastly disseminated unit is the ophiolitic mélange. The other mappable units in the area include metavolcanics, metagabbro-diorites and granitic rocks.

2.2. Data characteristics and preprocessing

The widely used Landsat-8 (L8) was launched on February 11, 2013,

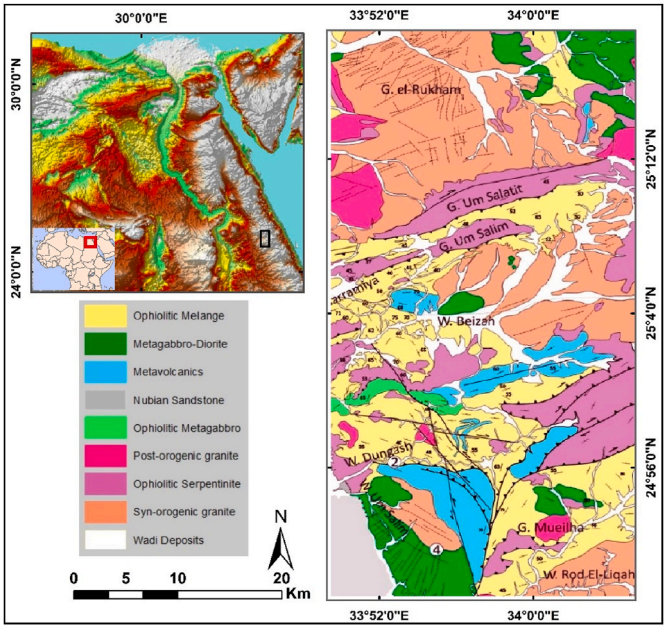

Fig. 1. Location of the study area (small black rectangle over the Egyptian Eastern Desert); and geological map of the study area showing its lithologic units, after (Zoheir et al., 2019a).

toting two sensors namely OLI (Operational Land Imager) and TIRS (Thermal Infrared Sensor) to acquire spectral data in the visible and Near-Infrared (VNIR), Short-Wave Infrared (SWIR) and Thermal Infrared (TIR) regions. OLI data are recorded in nine spectral bands, while TIRS data give information only in two bands (Roy et al., 2014), as shown in Table 1. Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) are commonly used in lithologic discrimination (Pour and Hashim, 2014), and detect radiance in fourteen bands covering spectral areas from VNIR, SWIR, and TIR regions as shown in Table 1. Advanced Land Imager (ALI) sensor of Earth Observing-1 satellite, recorded data in ten spectral bands (Czapla-Myers et al., 2016), as shown in Table 1. ALI operates in a pushbroom fashion, with a spatial resolution of 30 m for the multispectral bands, and 10 m for the panchromatic band. The standard scene width is 37 km, while the standard scene length is 42 km. Sentinel 2 (S2) was developed by the European Space Agency (ESA) and launched (S2A) on June 23, 2015, to provide spectral data in 13 bands in various spectral regions and variable spatial resolutions (10, 20, or 60m) (Drusch et al., 2012), as shown in Table 1.

Radar data includes Sentinel-1, the premier satellite of the Copernicus Programme satellite constellation conducted by ESA. The first satellite, Sentinel-1A, was launched on April 3, 2014, and Sentinel-1B was launched on April 25, 2016. They carry a C-band ($\lambda = 5.6$ cm, $f = 5.4$ GHz) synthetic-aperture radar (SAR) instrument, which provides a collection of data in all weather conditions, day or night. Sentinel-1A of Interferometric Wide (IW) mode (VH, VV polarizations), Ground Range Detected High Resolution (GRDH) is obtained from ESA to fulfill the aim of this study. Phased Array L-type band Synthetic Aperture Radar (PALSAR) device was on the board of Advanced Land Observing Satellite (ALOS), which was launched on January 24, 2005. PALSAR sensor is an active microwave sensor. Fine Beam Double polarization (HH, HV), Geocoded product, 1.5 processing level is utilized in this study. L-band (1.27 GHz) synthetic aperture radar aids at achieving high-resolution DEM products. Thus, ALOS-PALSAR FBS (Fine Beam Single polarization mode, HH) RT1 (Radiometric Terrain Corrected with Pixel spacing is 12.5 m) DEM is used to form topographical combinations with optical and radar data, as well as, derivation for slope and aspect of the study area.

All of these data are attained through the U.S. Geological Survey (https://earthexplorer.usgs.gov/) and the European Space Agency (ESA) with the following IDs: LC81740432019298LGN00 (Landsat 8), AST_L1A_00303062007083043 (ASTER), S2A_MSIL1C_20200505T081611_N0209_R121_T36RWN_20200505T095132 (Sentinel 2), EO1A1740422003070110PZ (EO1 ALI), S1A_IW_GRDH_1SDV_20200926T154707_20200926T154736_034532_0404DC_2A0B (Sentinel 1), ALOS-FBD_GEC_1P-ORBIT_00022767_D20100503-T201857038 (ALOS-PALSAR 1), and AP_10689_FBS_F0480_RT1 (ALOS-PALSAR DEM). The cloud-free multispectral data of Landsat-8 level 1T, ASTER, and EO-1 ALI are introduced to Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) atmospheric correction, and data resizing to the borders of the study area. All of these operations were carried out at the Environment for Visualizing Images (ENVI) software version 5.6. Sentinel-2 Dataset was reprojected to UTM WGS 84 zone 36N then, radiometrically-corrected using sen2cor processor in Sentinel Application Platform (SNAP). Sentinel-1 data was preprocessed by applying the precise orbit of acquisition, removing thermal and image border noise, radiometric calibration, and range doppler and terrain correction using Sentinel-1 toolbox. ALOS PALSAR1 data are also subjected to reprojection, radiometric calibration and speckle reduction using SNAP.

2.3. Training and testing samples

Nine main information classes cover the study area. Thus, representative, and well-distributed nine classes (for training and testing samples), are delineated from Landsat 8 and ASTER data, based on the geological map and visual interpretation of the images. The full

Table 1
Characteristics of the utilized multispectral datasets.

Landsat 8			ASTER			Sentinel 2			EO1 ALI		
B.n	C.W. (μm)	S.R (m)	B.n	C.W. (μm)	S.R (m)	B.n	C.W. (μm)	S.R (m)	B.n	C.W. (μm)	S.R (m)
1	0.442	30	1	0.560	15	1	0.443	60	Pan.	0.585	10
2	0.483	30	2	0.660	15	2	0.490	10	1	0.443	30
3	0.561	30	3N	0.820	15	3	0.560	10	2	0.482	30
4	0.654	30	3B	0.820	15	4	0.665	10	3	0.565	30
5	0.864	30	4	1.650	30	5	0.704	20	4	0.660	30
6	1.609	30	5	2.165	30	6	0.740	20	5	0.790	30
7	2.203	30	6	2.205	30	7	0.782	20	6	0.867	30
8	0.598	15	7	2.260	30	8	0.842	10	7	1.250	30
9	1.373	30	8	2.330	30	8a	0.865	20	8	1.650	30
10	10.90	100	9	2.395	30	9	0.945	60	9	2.215	30
11	12.00	100				10	1.375	60			
						11	1.610	20			
						12	2.190	20			

Band number (B.n), Central wavelength (C.W), Spatial Resolution (S.R) and Pan for panchromatic.

description of training and ground truth pixels is given in Table 2, and their spectral signatures from Landsat OLI are shown in Fig. 2. Prototype pixels are used to train the classifiers to produce thematic maps for the classes, and ground truth samples are employed to assess the accuracy of these outputs.

2.4. Machine learning classifiers

2.4.1. Artificial Neural Network (ANN)

A frequently used non-parametric supervised classifier. Mimicking the human brain, ANN is chiefly formed from neurons (processing nodes) embedded into input, middle (maybe more than one), and output layers (Haykin, 2010). Binding and connecting these nodes during the process forms a complicated network depending on the number of the utilized layers and classes. Through this network, a decision surface could be managed, and various classes could be separated. One of the main ANN strengths is the ability to create non-linear decision surfaces for inherently not separable data with a simple linear decision surface (Richards and Jia, 1999). For this study, we used ENVI 5.6 software to perform multi-layer feed-forward ANN by the logistic activation function. The best results are revealed by assigning the training root mean square (RMS) exit criterion as 0.1, training threshold contribution value as 0.9, training rate as 0.2, and training momentum as 0.9.

2.4.2. Maximum Likelihood Classifier (MLC)

Statistical parametric supervised classifier, where the classes are separated depending on probability distributions. Each class signature is specified by its mean vector and covariance matrix. Unclassified pixel belongs to a certain class only when it possesses a higher probability for that class (Chen et al., 2007). Consequently and based on discriminant functions, decision surfaces could be handled and separating different classes, according to the following equations (Richards and Jia, 1999),

$$x \in \omega_i, \text{ if } p(\omega_i|x) > p(\omega_j|x) \quad \text{for all } j \neq I, \tag{1}$$

Where x represents a pixel vector (column of values), ωi spectral class. Bayes' theorem control the relation between the coveted p(ωi|x) and the obtainable (from training data) p(x|ωi) probabilities,

$$p(\omega i|x) = p(x|\omega i)\ p(\omega i)/p(x) \tag{2}$$

$$x \in \omega_i, \text{ if } p(x|\omega_i)p(\omega_i) > p(x|\omega_j)p(\omega_j) \quad \text{for all } j \neq I \tag{3}$$

Where p(ωi) is the probability that class ωi occurs in the image, and is called prior probability. Then last equation (3) is transformed to decision rule (utilized in MLC);

$$x \in \omega_i, \text{ if } g_i(x) > g_j(x) \quad \text{for all } j \neq I \tag{4}$$

Where gi(x) denotes discriminant functions. For the current study, lithologic generalization using MLC is carried out using ENVI 5.6 software.

Table 2
Areas, training and testing pixels, and abbreviations of the Lithologic classes.

	Lithologic Unit	Area (Km^2)	Training Samples	Testing Samples
1	Wadi Deposits (Wdp)	1.39	3546	879
2	Nubian Sandstone (Nss)	0.61	1599	338
3	Post-orogenic granite (Pog)	0.45	1190	292
4	Syn-orogenic granite (Sog)	0.86	2215	599
5	Metagabbro-Diorite (MgD)	0.62	1605	367
6	Metavolcanics (Mvs)	0.79	2080	286
7	Ophiolitic Melange (Ome)	0.94	2404	438
8	Ophiolitic Metagabbro (Omg)	0.54	1410	188
9	Ophiolitic Serpentinite (Osp)	0.96	2518	389
Total		7.16	18567	3776

2.4.3. Support vector machine (SVM)

SVM is a non-parametric, supervised classifier that proved its leverage in remote sensing classifications since introduced by (Bayliss et al., 1998). The unique character of SVM is that pixels in the proximity of the separating hyperplane, accurately detect, where it should pass. Thus, these pixels are called support vectors, as they delicately locate the optimal separator hyperplane, via maximizing the distance between support pixels at class margins, then constructing two marginal hyperplanes. Generally, for (N) number of bands, (w_i) set of coefficients, delineate the linear surface as

$$w_1x_1 + w_2x_2 + \ldots\ldots + w_Nx_N + w_{N+1} = 0$$

$$\text{or, } w^t_x + w_{N+1} = 0 \tag{5}$$

Where x is the co-ordinate vector and w is called the weight vector. The transpose(t) operation for transforming column vector to row vector. After training and specification for decision surface, the decision rule is

$$x \in \text{class 1, if } w^t_x + w_{N+1} > 0 \tag{6}$$

$$x \in \text{class 2, if } w^t_x + w_{N+1} < 0 \tag{7}$$

Kernel function (such as linear, polynomial, radial basis function and sigmoid (Wang et al., 2017) is required for SVM. In this study, radial basis function kernel is applied due to its considered interpolation capabilities (Zhu et al., 2011), the penalty parameter was set to 100, and the gamma parameter in the kernel function was the inverse of the band number (Othman and Gloaguen, 2014).

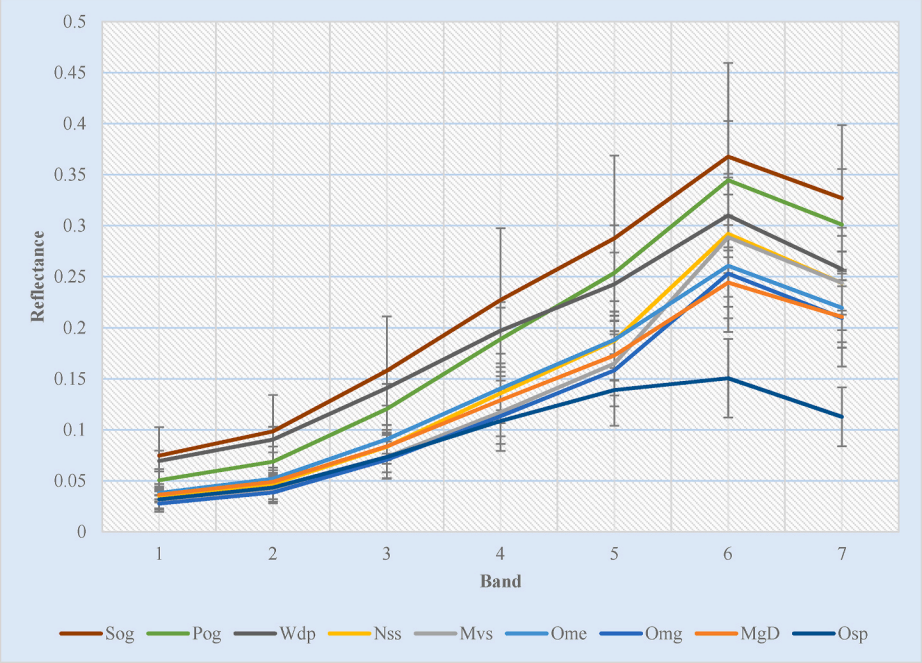

Fig. 2. Average of surface reflectance from training pixels selected from Landsat OLI data (7 bands) for nine information classes.

2.5. Stacked vector method

The utilized data are hybrid and have various spectral and spatial characteristics. Consequently, all the datasets are reprojected to the same coordinate system (WGS, 1984 Zone 36), then, cubic convolution resampling to a pixel size 20 m, was performed, to ensure the same size and location for each pixel in the dimensional data. Combinations are established as extended pixel vectors by aggregating spectral (i.e. optical sensors) and nonspectral (i.e. DEMs and Radar) datasets, according to the following form

$$X = \left[x_1^t,\ x_2^t,\ \ldots\ldots x_s^t\right]^t \quad (8)$$

Where $\mathbf{x}_1 \ldots \ldots \mathbf{x}_S$ are data vectors included in S number of individual layers (Richards and Jia, 1999). The superscript "t" indicates a vector transpose operation. X denotes the stacked vector introduced to the three classifiers. In this study, 37 stacked vectors are originated and tested for the classification process as described in Table 3.

3. Results

We find several results from fulfilling 111 classification processes (37 processes for each classifier). For all the outputs, we assessed the allocation accuracy via the widely used error matrix and kappa coefficient. Overall Accuracies (OAs) and kappa coefficient (K) for all the outputs are shown in Table 3. The latter clearly shows the notability of SVM and MLC over ANN, and obviously, confirm (just by comparing the numbers at the top and the bottom of the table) the effect of data dimensionality (by increasing the number of participated channels) in boosting classification accuracy. What is better displayed in Fig. 3a, where the OAs are gradually increased from around 75% to up to 91% (i.e. growing by 16%) by augmentation of the data input to the classifier. Fig. 3a, also demonstrated SVM superiority over MLC in lithologic classification, approximately at all the processes. Moreover, the increasing rate between (10–20) bands is higher than that between bands (20–30). An OA peak value is obtained when the participated bands are 31, beyond this number, no considered OA increase is evident whatever the participated number of bands. Up and down curve swinging for the increasing trend is attributed to the various characteristics of datasets introduced to the classifier. Moreover, data spectral characteristics can variously affect the generalization OAs, even though at the same dimensionality, as shown in Fig. 3b.

Among the utilized multispectral datasets (S2, L8, AST, and ALI), the highest labeling efficiency is recorded for ALI, the lowest for L8. ASTER and S2 give reasonable results. Combining two optical sensors results in variable accuracies depending on the spectral characteristics of the combined datasets, the number of bands, and the utilized classifier. SVM highest OA was 87.57% for AST + ALI, MLC gives OA of S2+ALI as 86.65%, and ANN best result is specified to L8+ALI by 79.34%. Jointing three different optical datasets takes the categorization OAs a little bit higher (89.67% by SVM and 89.35% by MLC) When the classifiers are fed by ALI + S2+ASTER data. Stacking the four optical datasets raised the accuracy of the resultant thematic maps to 89.40% and 90.01% when using MLC and SVM, respectively.

Embedding topographical data (especially DEM), to any collection of multispectral data is prodigious. DEM introduces a permanent but variable OA raise, depending on the other included data vectors and dimensions, as shown in Fig. 3c. For DEM derivatives, the slope is preferable than aspect in boosting the allocation output, but DEM is more advisable when compared to both. Immersing radar (S1 and ALOS PALSAR1) data in multispectral and topographical data can faintly improve the lithologic identification as shown in Fig. 3d.

However, tangible OA excess is achieved by incorporating all the datasets for the classification (Fig. 3d). It is worth mentioning to note that, the highest two OAs are delivered by ALI + S2+AST + DEM + S1+PALSAR1+slope (43 band, 91.41% by SVM) and ALI + S2+AST + DEM (31 band, 91.23% by SVM). This in turn, boosts the second combination to be applied in further classifications (only 0.18% difference in OAs despite the presence of 13 excessive bands in the first combination). By investigation of the confusion matrix of this combination, reasonable accuracies are observed. Average of producer and user accuracies for each class of ALI + S2+AST + DEM combination are shown in Fig. 3e. Consequently, the number of utilized training pixels and (31) channels are supposed to be well related in the current study. The lowest number of training pixels, employed in this study was 1190 for Pog (i.e. 38 N). Thus, we believed that classification OA reaches its peak for a (N) number of bands (depending on the data, classifiers, and classes) by implementing number training pixels ranging from 30 to 40N, and no

Table 3
Overall accuracies (OAs) and Kappa coefficients (K) for the utilized datasets and classifiers.

Data	N	ANN		MLC		SVM	
		OA	K	OA	K	OA	K
L8	7	71.05	0.664	74.81	0.710	76.43	0.726
AST	9	73.11	0.688	77.07	0.735	80.27	0.771
ALI	9	**77.15**	**0.735**	**81.36**	**0.784**	**83.87**	**0.812**
S2	12	74.29	0.701	78.71	0.753	79.42	0.761
Optical sensor + DEM							
L8+DEM	8	18.94	0.0565	81.78	0.7896	80.93	0.7791
AST + DEM	10	70.84	0.6614	82.47	0.797	86.84	0.8475
ALI + DEM	10	42.85	0.332	**84.75**	**0.8238**	**86.89**	**0.8481**
S2+DEM	13	**82.07**	**0.7927**	81.51	0.786	82.73	0.7998
2 Optical sensors together							
S2+AST	21	78.681	0.7525	86.626	0.8451	86.89	0.8481
S2+L8	19	76.85	0.7315	78.416	0.7509	80.985	0.7794
S2+ ALI	21	77.171	0.7345	**86.652**	**0.8453**	85.725	0.8344
AST + L8	16	73.384	0.6915	79.449	0.7633	84.666	0.8221
AST + ALI	18	74.973	0.71	84.401	0.8194	**87.579**	**0.8559**
L8+ ALI	16	**79.343**	**0.7595**	84.322	0.8187	85.301	0.8294
2 Optical sensors + DEM							
S2+AST + DEM	22	**82.256**	**0.7937**	87.314	0.8531	88.268	0.8641
AST + ALI + DEM	19	73.119	0.6873	**87.394**	**0.8543**	**89.91**	**0.8831**
3 Optical sensors together							
S2+AST + L8	28	77.092	0.7337	86.732	0.8465	87.791	0.858
AST + L8+ ALI	25	73.49	0.693	87.764	0.8585	88.135	0.8623
L8+S2+ ALI	28	**78.125**	**0.7459**	84.613	0.8213	86.99	0.849
ALI + S2+ AST	30	77.807	0.7422	**89.3538**	**0.8766**	**89.671**	**0.8802**
3 Optical sensors + DEM							
S2+AST + L8+DEM	29	81.197	0.7818	88.082	0.8621	89.036	0.873
AST + L8+ ALI + DEM	26	71.18	0.666	88.638	0.8687	89.48	0.878
L8+S2+ ALI + DEM	29	81.56	0.785	86.573	0.8444	86.99	0.849
ALI + S2+AST + DEM	31	**83.26**	**0.805**	**89.857**	**0.8826**	**91.23**	**0.898**
Synergetic use of 4 Optical sensors							
S2+AST + L8+ ALI (4Opt)	37	79.21	0.758	89.40	0.877	**90.01**	**0.884**
4 Optical sensors + topographic data							
S2+AST + L8+ ALI + DEM	38	**82.28**	**0.7944**	89.6	0.88	**91.18**	**0.897**
S2+AST + L8+ ALI + slope	38	78.76	0.75	89.72	0.880	90.78	0.893
S2+AST + L8+ ALI + aspect	38	75.45	0.71	**89.85**	**0.882**	89.88	0.882
4 Optical sensors + radar data							
S2+AST + L8+ ALI + S1VV	38	76.50	0.727	89.83	0.882	90.09	0.885
S2+AST + L8+ ALI + S1VH	38	76.24	0.72	89.93	0.883	90.30	0.887
S2+AST + L8+ ALI + S1VH + S1VV	39	76.69	0.72	89.93	0.883	90.09	0.885
S2+AST + L8+ ALI + Palsar (HH + HV)	39	76.48	0.72	90.06	0.885	89.80	0.881
4Opt + S1+ Palsar	41	**77.48**	**0.73**	**90.22**	**0.886**	**90.14**	**0.885**
4 Optical sensors + topographic and radar data							
4Opt + DEM+2 S1+2pal	42	**82.83**	**0.80**	**90.59**	**0.891**	90.51	0.890
4Opt + DEM+2 S1+2pal + slope	43	75.92	0.72	90.28	0.887	**91.41**	**0.90**
4Opt + DEM+2 S1+2pal + aspect	43	81.48	0.78	89.67	0.880	90.17	0.886
4Opt + DEM+2 S1+2pal + slope + aspect	44	82.28	0.793	89.98	0.884	90.83	0.893

The four optical sensors together (4 Opt), Sentinel 1 (S1), ALOS Palsar data (pal) and Number of bands (N).

considered increase is obtained beyond that limit unless a new training data could be included for the classification process. Consequently and for achieving potent categorization, this study assumed that the number of training pixels for a single region of interest (ROI) is preferable to be at least 30 N, rather than the wide range (10N–100N per class) assumed by the authors of (Davis et al., 1978). The resultant thematic maps of the previously mentioned stacked data collections are shown in Fig. 4.

4. Discussion

Various rock units could be reasonably discriminated by multispectral data depending on their spectral characteristics and the imaging sensor properties as displayed in Table 4, which shows continuous interest in lithologic mapping over the last 10 years. Table 4 compares results of 15 previous studies with the current study that implemented the frequently recommended MLAs over a wide range of integrated data sets.

Using MLAs and adding topographical and/or radar data could enhance the discrimination process, especially if the classes have closer spectral characteristics and are poorly recognized by optical sensors solely. This study revealed that discrimination ability is higher for DEM and slope when compared to radar data. For instance, serpentinites, post-orogenic granites and metavolcanics are easily identified by their higher elevations (from DEM) and steeper slopes as shown in Fig. 5a, b. On contrary, wadi deposits and syn-tectonic granites are distinguished, by their relative lower elevations. In between, mélange assemblages and the other rock units could be differentiated with the aid of their spectral signatures (from multispectral data). Thus, topographical data can set apart most of the lithologic classes. However, radar data can hardly disclose differences between rock units (due to approximately their similar interaction with active microwaves) and the obvious discrimination is evident only for fine wadi deposits and weathered syn-tectonic granites from other hard rocks, as shown in Fig. 5c.

Classifying and mapping Precambrian rocks (particularly ophiolitic) is a complicated process not only from remote sensing data (often mixed spectral signatures are encountered) but also during fieldwork. For example, several researchers introduced geological maps for parts of the study area, based on field studies, and many variations are detected between the produced maps (Aboelkhair et al., 2010; Helba et al., 2001; Zoheir et al., 2019b; Zoheir and Weihed, 2014). Consequently,

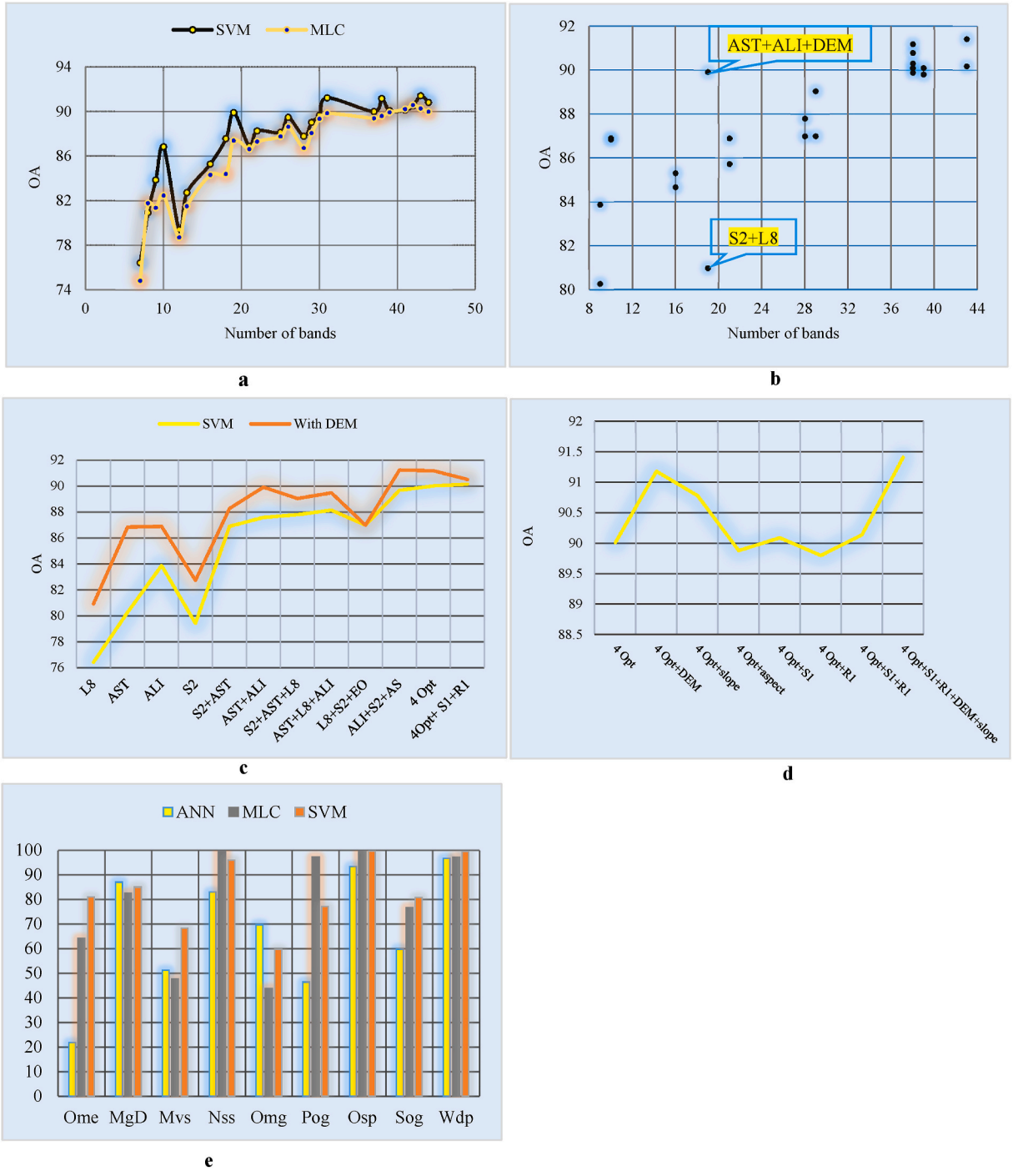

Fig. 3. (a) The effect of increasing data dimensionality on classification accuracy. Beyond 31 participated band, the Overall Accuracy (OA) excess is barely perceptible; (b) various OA results for different datasets having the same dimensions, differences could reach to about 9% (between S2+L8 and AST + ALI + DEM); (c) DEM influence in boosting the classification OA; (d) Topographical and radar data impact on categorization OA; (e) Average accuracies of the information classes, classified by ANN, MLC, and SVM, from ALI + S2+AST + DEM.

homogenous monotone training pixels are carefully selected, aiming to resolve these discrepancies. The resultant thematic maps are closely analogous to the reference geological map as shown in Fig. 4. Misclassifications are sometimes introduced among ophiolitic mélange, ophiolitic metagabbro, and metagabbro-diorite complex. This is attributed to closeness in their spectral signatures (as shown in Fig. 2), cognate elevations, and similar appearance in radar data. Thus, these ground cover classes are not exemplary resolved when compared to the reference map. ALI + S2+AST + DEM combination (31 bands) gather the advantages of ALI (six unique VNIR wavelength bands (Hubbard et al., 2003; Hubbard and Crowley, 2005) and improved Signal-to-Noise Ratio (Bryant et al., 2003; Lencioni et al., 1999; Lobell and Asner, 2003; Mendenhall et al., 2000), ASTER (high spectral separability of most functional groups and their minerals using eminent six SWIR bands (Pour and Hashim, 2014)), S2 (higher spectral characteristics for both VNIR and SWIR (Ge et al., 2018) and DEM (topographical characteristics discernment), thus delivered precise results. Furthermore, this combination is strongly recommended for lithologic classifications due to two main reasons: 1- adding any supporting bands can hardly enhance the allocation accuracy, as seen for ALI + S2+AST + DEM +

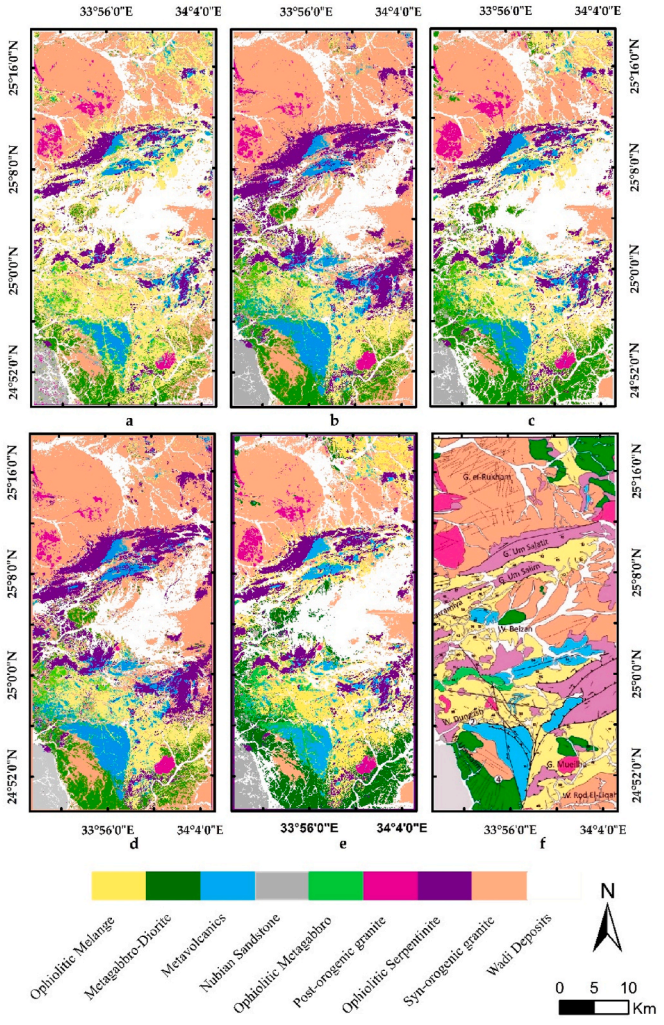

Fig. 4. Lithologic classification outputs utilizing ALI + S2+AST + DEM data from (a) ANN; (b) MLC and; (c) SVM. (d) and (e) are the results of generalization over ALI + S2+AST + DEM + S1+PALSAR1+slope data utilizing MLC and SVM, respectively. (f) Geological map for comparison.

S1+PALSAR1+slope combination (43 bands), only gives 0.18% improvement by adding 11 supporting layers. 2- Removing any constituent can significantly reduce the OA as shown in Table 3. Our result is harmonized with (Ge et al., 2018), who preferred S2+AST + DEM in lithologic classification, however, we strongly recommend adding ALI to the previous combination as it improves OAs by approximately 1%, 2.5%, and 3% for ANN, MLC, and SVM respectively. It should be emphasized that this study applied transfer learning as it is still a powerful method when multisource remote sensing images are implemented in the allocation process (Dong et al., 2021), which is the case in the current study. However, other studies efficiently combined deep and transfer learning to get efficient results (Dong et al., 2021; Shi et al., 2021; Xie et al., 2021; Yuan et al., 2020; Zhou et al., 2019).

MLC reasonable results are interpreted by representing training signatures for each class by both covariance matrices and mean vectors. However, sometimes and due to data dimensionality and heterogeneity, incompatible statistics and quadratic cost increase are encountered resulting in lower OAs compared to SVM results. For SVM, the class boundaries are determined by the collection of weights (equations (5)–(7)) that make the accurate labeling process is feasible, as well as a penalty for misclassifications is introduced to achieve better results compared to ANN. For data with N dimensions, 30–40N is assigned to be the appropriate numbers of training pixels required for fulfilling accurate classification (by attaining the exact location of the decision surface). If the used training pixels are insufficient (less than 30N), the location of the separating surface may be inadequate, resulting in various misclassifications depending on the utilized algorithm. Also,

Table 4
Comparison of the obtained results over the last decade with the current study.

	Study	Classifiers	Data used	Results
1	Grebby et al. (2011)	ANN (Self-Organizing Map), MLC	Airborne multispectral and LiDAR data	ANN results shows greater enhancement
2	Yu et al. (2012)	MLC, SVM	ASTER, DEM	SVM outputs is better than MLC results
3	Mehr et al. (2013)	MLC	Thematic Mapper 5 (TM5)	MLC enhances the separability of units in the image by 7.3%
4	Fatima et al. (2013)	MLC, PP, MDM, MD, SAM	Landsat ETM+ and ASTER	MLC has the highest correlation with the geological map
5	Hadigheh and Ranjbar (2013)	MLC, SAM, SID	ASTER, IRS	MLC is the better than SAM and SID
6	A. A. Othman and Gloaguen (2014)	SVM	ASTER	SVM is efficient tool in mineral and lithological mapping
7	He et al. (2015)	ANN, RF, SVM, MLC	Landsat-7 and Landsat-8	ANN, RF, and SVM outperformed MLC
8	Jellouli et al. (2016)	MLC, SVM	ASTER	MLC results is better than that of SVM
9	Othman and Gloaguen (2017)	MLC, SVM, RF	ASTER, DEM	RF and SVM are better than MLC
10	Ge et al. (2018)	ANN, k-NN, MLC, RF, SVM	Sentinel-2A	MLC method offered the highest overall accuracy
11	Manap and San (2018)	MLC, RF, SVM	ASTER	MLC and RF have approximately the same accuracy, and SVM is more accurate than them.
12	Bachri et al. (2019)	SVM	Landsat 8, PALSAR DEM	SVM performs effective classification with 85% accuracy
13	Karimzadeh and H. Tangestani (2021)	SVM	WorldView-3	SVM performs effective classification with 88.36% accuracy
14	Bentahar and Raji (2021)	MLC, SVM	Landsat OLI, ASTER, and Sentinel 2A	the effectiveness of Sentinel 2A data and SVM is better than MLC
15	Kumar et al. (2021)	SVM and RF	ASTER, PALSAR, Sentinel 1	SVM is better than RF
16	Current Study	ANN, SVM, MLC	Sentinel 2, ASTER, Landsat OLI Advanced Land Imager, Sentinel 1, ALOS PALSAR-1, ALOS PALSAR DEM, Slope, Aspect	SVM is better than MLC which is in turn better than ANN, DEM enhances the classification, S2, ASTER and ALI are preferred when compared to Landsat OLI, Integration of multi-sensors strongly boost the outputs.

Artificial neural network (ANN), Maximum likelihood (MLC), support vector machine (SVM), k-nearest neighbor (k-NN), maximum likelihood classification (MLC), random forest classifier (RF), and parallelepiped (PP), minimum distance to mean (MDM), mahalanobis distance (MD), Spectral Angle Mapper (SAM), Spectral Information Divergence (SID), Indian Remote Sensing Satellite (IRS).

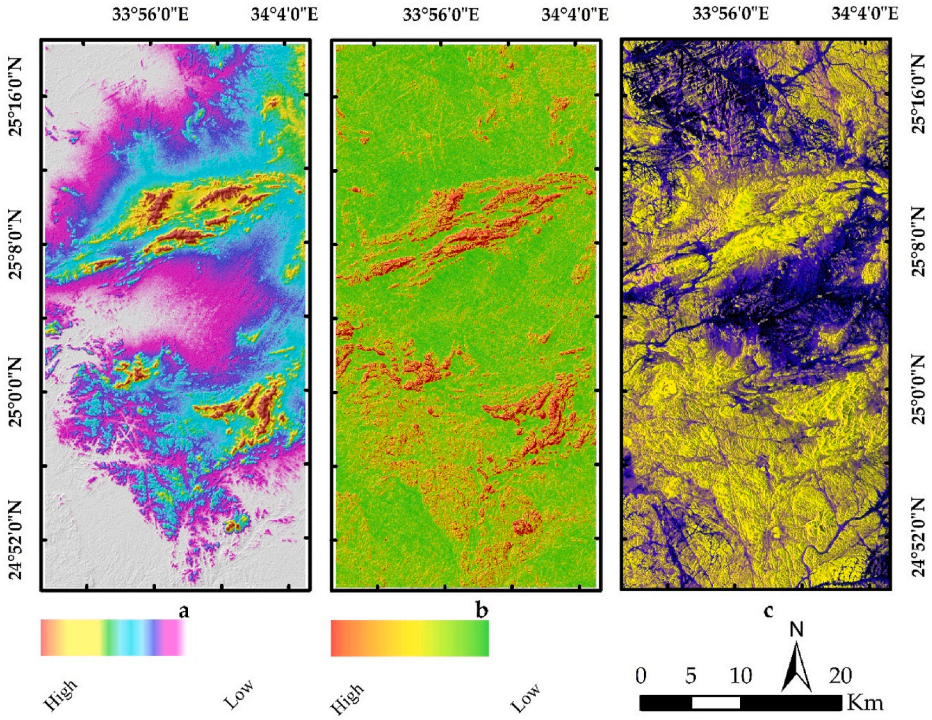

Fig. 5. Reasonable lithologic discrimination fro m (a) DEM (approximately brown Osp, green to cyan Mvs, white Wdp, and pink Sog) compared to; (b) Slope and; (c) RGB (HH, HV, HH/HV) ALOS PALSAR 1 combination.

excessively utilized training pixels (exceeding 40N) or exaggerated dimensions (for a constant number of training pixels) cannot significantly improve the classification because the separator hyperplane is already optimally located, and any support (either by training pixels or channels) cannot introduce significant revamping for the hyperplane placement between classes. we believe that the possible way to boost the allocation process is by increasing both training data and dimensions together, keeping the same percentage (30–40N).

5. Conclusions

In this study, three classifiers (ANN, MLC, and SVM) are fed by 37 data inputs including (Sentinel 2, ASTER, Landsat OLI, Earth-observing 1 Advanced Land Imager, Sentinel 1, ALOS PALSAR 1, ALOS PALSAR DEM, Slope, Aspect and their combinations) to enhance the lithologic mapping of the study area. Also, the dimensionality power in improving the classification output and the effect of these datasets are outlined. The study concludes the followings results:

- SVM, MLC, and ANN are eligible in lithologic classification. Results of SVM and MLC outperforms that of ANN.
- Data dimensionality increment can enhance the allocation process only when there is sufficient training data, compared to the number of utilized bands (N). The minimum number of the desired training pixels for executing accurate classification is 30N.
- DEM remarkably improves the allocation compared to slope, aspect, sentinel 1 and Alos palsar data.
- S2, ASTER and ALI are preferred when compared to Landsat OLI, in lithologic categorization. Finally,
- For further lithologic classifications, ALI+ S2+AST+DEM combination is robustly recommended. Moreover, this study can help researchers in detecting a suitable data input for their classification, as this study introduces 111 classification results for commonly used remote sensing datasets and their combinations.

Author statement

Ali Shebl: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing - original draft, Writing - review and editing. **Árpád Csámer**: Funding acquisition, Investigation, Supervision, Validation, Resources, Writing - review and editing.

Funding

This research received no external funding. This research is supported by University of Debrecen. Ali Shebl is funded by Stipendium Hungaricum scholarship under the joint executive program between Hungary and Egypt.

Data availability statement

The data that support the findings of this study are openly available in figshare at [https://figshare.com/s/fd6a261ce7e636307ddd], reference number[10.6084/m9.figshare.14701209].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are thankful to U.S. Geological Survey, and European Space Agency (ESA), for providing the data. Thanks to Prof. Mahmod Ashmawy, Prof. Mohamed Abdelwahed and Prof. Samir Kamh of Tanta University, for their support.

References

Aboelkhair, H., Ninomiya, Y., Watanabe, Y., Sato, I., 2010. Processing and interpretation of ASTER TIR data for mapping of rare-metal-enriched albite granitoids in the Central Eastern Desert of Egypt. J. Afr. Earth Sci. 58, 141–151. https://doi.org/10.1016/j.jafrearsci.2010.01.007.

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-Information 2019 8. https://doi.org/10.3390/IJGI8060248. Page 248 8, 248.

Bayliss, J.D., Gualtieri, J.A., Cromp, R.F., 1998. Analyzing hyperspectral data with independent component analysis. 26th AIPR Work. Exploit. New Image Sources Sensors 3240, 133–143. https://doi.org/10.1117/12.300050.

Bentahar, I., Raji, M., 2021. Comparison of Landsat OLI, ASTER, and sentinel 2A data in lithological mapping : a case study of rich area (central high Atlas, Morocco). Adv. Space Res. 67, 945–963. https://doi.org/10.1016/J.ASR.2020.10.037.

Bryant, R., Moran, S., McElroy, S.A., Holified, C., Thome, K.J., Miura, T., Biggar, S.F., 2003. Data continuity of earth observing 1 (EO-1) Advanced land imager (ALI) and Landsat TM and ETM+. IEEE Trans. Geosci. Rem. Sens. 41, 1204–1214. https://doi.org/10.1109/TGRS.2003.813213.

Chen, X., Warner, T.A., Campagna, D.J., 2007. Integrating visible, near-infrared and short-wave infrared hyperspectral and multispectral thermal imagery for geological mapping at Cuprite, Nevada. Remote Sens. Environ. 110, 344–356. https://doi.org/10.1016/j.rse.2007.03.015.

Czapla-Myers, J., Ong, L., Thome, K., McCorkel, J., 2016. Validation of EO-1 hyperion and Advanced land imager using the radiometric calibration test site at railroad valley, Nevada. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9, 816–826. https://doi.org/10.1109/JSTARS.2015.2463101.

Davis, S.M., Landgrebe, D.A., Phillips, T.L., Swain, P.H., Hoffer, R.M., Lindenlaub, J.C., Silva, L.F., 1978. Remote Sensing: the Quantitative Approach. mhi.

Dong, Y., Liang, T., Zhang, Y., Du, B., 2021. Spectral-spatial weighted Kernel Manifold embedded distribution alignment for remote sensing image classification. IEEE Trans. Cybern. 51, 3185–3197. https://doi.org/10.1109/TCYB.2020.3004263.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens. Environ. 120, 25–36.

Fatima, K., Khattak, U.K., Kausar, A.B., 2013. Selection of appropriate classification technique for lithological mapping of Gali Jagir area, Pakistan. Int. J. Earth Sci. Eng. 7, 964–971.

Foody, G.M., Mathur, A., 2004. A relative evaluation of multiclass image classification by support vector machines. IEEE Trans. Geosci. Rem. Sens. 42, 1335–1343. https://doi.org/10.1109/TGRS.2004.827257.

Ge, W., Cheng, Q., Jing, L., Armenakis, C., Ding, H., 2018. 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.

Grebby, S., Naden, J., Cunningham, D., Tansey, K., 2011. Integrating airborne multispectral imagery and airborne LiDAR data for enhanced lithological mapping in vegetated terrain. Remote Sens. Environ. 115, 214–226. https://doi.org/10.1016/j.rse.2010.08.019.

Hadigheh, S.M.H., Ranjbar, H., 2013. Lithological mapping in the eastern part of the central Iranian volcanic belt using combined ASTER and IRS data. J. Indian Soc. Remote Sens. 2013 414 41, 921–931. https://doi.org/10.1007/S12524-013-0284-1.

Ham, J.S., Chen, Y., Crawford, M.M., Ghosh, J., 2005. Investigation of the random forest framework for classification of hyperspectral data. In: IEEE Transactions on Geoscience and Remote Sensing, pp. 492–501. https://doi.org/10.1109/TGRS.2004.842481.

Haykin, S., 2010. Neural Networks: a Comprehensive Foundation. 1999. Mc Millan, New Jersey, pp. 1–24.

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, pp. 2252–2276. https://doi.org/10.1080/01431161.2015.1035410 36.

Helba, H.A., Khalil, K.I., Abou, N.M.F., 2001. Alteration patterns related to hydrothermal gold mineralizaition in meta-andesites at Dungash area, Eastern Desert, Egypt. Resour. Geol. 51, 19–30. https://doi.org/10.1111/j.1751-3928.2001.tb00078.x.

Hubbard, B.E., Crowley, J.K., 2005. Mineral mapping on the Chilean-Bolivian Altiplano using co-orbital ALI, ASTER and Hyperion imagery: data dimensionality issues and solutions. Remote Sens. Environ. 99, 173–186. https://doi.org/10.1016/j.rse.2005.04.027.

Hubbard, B.E., Crowley, J.K., Zimbelman, D.R., 2003. Comparative alteration mineral mapping using visible to shortwave infrared (0.4-2.4 μm) Hyperion, ALI, and ASTER imagery. IEEE Trans. Geosci. Rem. Sens. 41, 1401–1410. https://doi.org/10.1109/TGRS.2003.812906.

Jellouli, A., El Harti, A., Adiri, Z., El Ghmari, A., Bachaoui, E.M., Jellouli, A., El Harti, A., Adiri, Z., El Ghmari, A., Bachaoui, E.M., 2016. Lithological mapping using ASTER data in the Moroccan Anti Atlas belt. EGUGA 18, EPSC2016–13872.

Karimzadeh, S., Tangestani, M.H., 2021. Evaluating the VNIR-SWIR datasets of WorldView-3 for lithological mapping of a metamorphic-igneous terrain using support vector machine algorithm; a case study of Central Iran. Adv. Space Res. 68, 2421–2440. https://doi.org/10.1016/J.ASR.2021.05.002.

Kumar, C., Chatterjee, S., Oommen, T., Guha, A., Mukherjee, A., 2021. Multi-sensor datasets-based optimal integration of spectral, textural, and morphological characteristics of rocks for lithological classification using machine learning models. https://doi.org/10.1080/10106049.2021.1920632.

Lee, S., Song, K.Y., Oh, H.J., Choi, J., 2012. Detection of landslides using web-based aerial photographs and landslide susceptibility mapping using geospatial analysis. Int. J. Rem. Sens. 33, 4937–4966. https://doi.org/10.1080/01431161.2011.649862.

Lencioni, D.E., Digenis, C.J., Bicknell, W.E., Hearn, D.R., Mendenhall, J.A., 1999. Design and performance of the EO-1 Advanced land imager. Sensors, Syst. Next-Generation Satell. III 3870, 269–280. https://doi.org/10.1117/12.373195.

Liesenberg, V., Gloaguen, R., 2012. Evaluating SAR polarization modes at L-band for forest classification purposes in eastern Amazon, Brazil. Int. J. Appl. Earth Obs. Geoinf. 21, 122–135. https://doi.org/10.1016/j.jag.2012.08.016.

Lobell, D.B., Asner, G.P., 2003. Comparison of earth observing-1 ALI and Landsat ETM+ for crop identification and yield prediction in Mexico. IEEE Trans. Geosci. Rem. Sens. 41, 1277–1282. https://doi.org/10.1109/TGRS.2003.812909.

Manap, H.S., San, B.T., 2018. Lithological mapping using different classification algorithms in western antalya, Turkey. Int. Multidiscip. Sci. GeoConference Surv. Geol. Min. Ecol. Manag. SGEM 18, 551–556. https://doi.org/10.5593/SGEM2018/2.2/S08.069.

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. http://doi.org/10.1080/01431161.2013.853144 34, 8803-8816.

Mendenhall, J.A., Lencioni, D.E., Evans, J.B., 2000. Earth Observing-1 Advanced Land Imager: Radiometric Response Calibration.

Othman, A., Gloaguen, R., 2014. Improving lithological mapping by SVM classification of spectral and morphological features: the Discovery of a new chromite body in the mawat ophiolite complex (kurdistan, NE Iraq). Rem. Sens. 6, 6867–6896. https://doi.org/10.3390/rs6086867.

Othman, A.A., Gloaguen, R., 2017. Integration of spectral, spatial and morphometric data into lithological mapping: a comparison of different Machine Learning Algorithms in the Kurdistan Region, NE Iraq. J. Asian Earth Sci. 146, 90–102. https://doi.org/10.1016/J.JSEAES.2017.05.005.

Pal, M., Mather, P.M., 2005. Support vector machines for classification in remote sensing. Int. J. Rem. Sens. 26, 1007–1011. https://doi.org/10.1080/01431160512331314083.

Pour, A.B., Hashim, M., 2014. ASTER, ALI and Hyperion sensors data for lithological mapping and ore minerals exploration. SpringerPlus 3, 130.

Richards, J.A., Jia, X., 1999. Remote Sensing Digital Image Analysis.

Roy, D.P., Wulder, M.A., Loveland, T.R., Woodcock, C.E., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., 2014. Landsat-8: science and product vision for terrestrial global change research. Remote Sens. Environ. 145, 154–172.

Shi, Y., Ma, D., Lv, J., Li, J., 2021. ACTL: Asymmetric Convolutional Transfer Learning for tree species identification based on deep neural network. IEEE Access 9, 13643–13654. https://doi.org/10.1109/ACCESS.2021.3051015.

Wang, F., Zhen, Z., Wang, B., Mi, Z., 2017. Comparative study on KNN and SVM based weather classification models for Day Ahead short term solar PV power forecasting. Appl. Sci. 8, 28. https://doi.org/10.3390/app8010028.

Xie, F., Gao, Q., Jin, C., Zhao, F., 2021. Hyperspectral image classification based on superpixel pooling convolutional neural network with transfer learning. Remote Sens. 13, 930. https://doi.org/10.3390/rs13050930.

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.

Yuan, X., Ou, C., Wang, Y., Yang, C., Gui, W., 2020. Deep quality-related feature extraction for soft sensing modeling: A deep learning approach with hybrid VW-SAE. Neurocomputing 396, 375–382. https://doi.org/10.1016/J.NEUCOM.2018.11.107.

Zhou, P., Han, J., Cheng, G., Zhang, B., 2019. Learning compact and discriminative stacked autoencoder for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 57, 4823–4833. https://doi.org/10.1109/TGRS.2019.2893180.

Zhu, X., Zhang, S., Jin, Z., Zhang, Z., Xu, Z., 2011. Missing value estimation for mixed-attribute data sets. IEEE Trans. Knowl. Data Eng. 23, 110–121. https://doi.org/10.1109/TKDE.2010.99.

Zoheir, B., El-Wahed, M.A., Pour, A.B., Abdelnasser, A., 2019a. Orogenic gold in transpression and transtension zones: field and remote sensing studies of the barramiya–mueilha sector, Egypt. Rem. Sens. 11, 2122.

Zoheir, B., Steele-MacInnis, M., Garbe-Schönberg, D., 2019b. Orogenic gold formation in an evolving, decompressing hydrothermal system: genesis of the Samut gold deposit, Eastern Desert, Egypt. Ore Geol. Rev. 105, 236–257.

Zoheir, B., Weihed, P., 2014. Greenstone-hosted lode-gold mineralization at Dungash mine, Eastern Desert, Egypt. J. Afr. Earth Sci. 99, 165–187.