International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

Advancements in Geological Mapping: Integrating Remote Sensing, GIS, and Digital Technologies

Ashutosh Nimba Patil

Abstract

Geological mapping has undergone a structural transformation over the past two decades, moving from an essentially analogue, observation-limited craft to a data-intensive, computationally mediated science. A mixed-methods design was adopted, combining a structured review of the literature with a comparative meta-analytic assessment of classification accuracy, data-acquisition efficiency and positional reliability across three generations of mapping workflow: conventional analogue survey, remote sensing supported by geographic information systems, and fully integrated digital mapping incorporating unmanned aerial vehicles, three-dimensional virtual outcrop models and machine learning. Results indicate that ensemble and deep-learning classifiers substantially outperform conventional parametric methods, with convolutional neural networks achieving a mean overall accuracy of 88.9 per cent compared with 71.4 per cent for maximum likelihood classification. Fusion of optical imagery with digital elevation and geophysical layers improves overall accuracy by a further six to eight percentage points across all sensor families, and hyperspectral data yield the highest single-sensor accuracies. Integrated digital workflows reduce field time per unit area by approximately seventy-four per cent and positional error by an order of magnitude relative to analogue practice, while increasing the volume of structural measurements collected per field day by a factor of nearly five. The paper concludes that remote sensing, geographic information systems and digital field technologies are most effective when treated as a single integrated system rather than as separate tools, and that the principal remaining constraints are the availability of reliable training data, the interpretability of complex models and the uneven distribution of technical capacity between geological surveys.

Keywords: Geological mapping; remote sensing; geographic information systems; machine learning; lithological classification; digital field geology; unmanned aerial vehicles

1. Introduction

1.1 The Nature and Purpose of Geological Mapping

A geological map is not merely a picture of the ground. It is a spatial model, an interpretive statement about the three-dimensional distribution of rock bodies, their contacts, their internal fabric and their history, constructed from an inherently incomplete set of surface observations. Since the earliest systematic surveys of the early nineteenth century, the geologist has worked by

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

extrapolating from scattered outcrops across intervening terrain that is obscured by soil, vegetation, water or human modification. The credibility of the resulting map depends on the density of observation, the quality of the extrapolation and the conceptual model the mapper brings to the field. That epistemic structure has remained constant even as the instruments of mapping have changed beyond recognition.

The practical importance of geological maps is difficult to overstate. They form the primary information base for mineral and hydrocarbon exploration, groundwater assessment, engineering site investigation, seismic and landslide hazard zonation, land-use planning and the growing field of critical mineral security. National geological surveys in most countries maintain systematic mapping programmes precisely because the map is a public good whose value accumulates over decades. Yet global map coverage remains strikingly uneven. Large areas of Africa, central Asia, the high latitudes and the tropical belt are covered only at reconnaissance scales, and even in well-surveyed jurisdictions a substantial proportion of published sheets predate the availability of satellite imagery and satellite positioning.

Conventional mapping is constrained by four persistent difficulties. First, accessibility: rugged, forested, arid or politically insecure terrain limits the ground the geologist can physically traverse. Second, exposure: in humid and vegetated regions the fraction of bedrock actually visible at the surface may be below one per cent. Third, cost and time: systematic field mapping at 1:50,000 scale is measured in field seasons rather than weeks. Fourth, subjectivity: two experienced geologists mapping the same ground will produce different but individually defensible maps, and the uncertainty attached to any given contact is rarely recorded in a form that later users can interrogate. Each of these constraints is a point at which geospatial technology can intervene, and the history of the last three decades can be read as a sequence of such interventions.

1.2 The Emergence of Geospatial and Digital Technologies in the Earth Sciences

The introduction of civilian Earth observation satellites provided the first systematic means of viewing geology at regional scale in a spatially consistent, repeatable and quantitative form. Multispectral sensors made it possible to exploit the diagnostic reflectance behaviour of rock-forming and alteration minerals in the visible, near-infrared, shortwave infrared and thermal infrared regions, so that lithological discrimination could be attempted from orbit rather than solely from the hand lens. Successive sensor generations progressively relaxed the trade-off between spatial, spectral and temporal resolution. Spaceborne hyperspectral instruments have narrowed the gap between orbital measurement and laboratory spectroscopy, allowing individual mineral species rather than broad spectral classes to be identified (Hajaj et al., 2022; Shebl et al., 2021).

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

In parallel, geographic information systems evolved from digital drafting environments into analytical frameworks. Their decisive contribution was not cartographic but integrative: a GIS allows imagery, digital elevation models, aeromagnetic and radiometric surveys, geochemical sample results, borehole logs, structural readings and legacy map polygons to be co-registered in a common spatial reference and interrogated jointly. This capacity for multi-source integration underlies the modern practice of predictive and prospectivity mapping, in which the spatial coincidence of evidence layers is modelled explicitly rather than assessed by eye (Sun et al., 2019; Xiong et al., 2018).

The third strand of change has been the digitisation of the field itself. Ruggedised tablets, satellite positioning at sub-metre accuracy, digital compass-clinometer applications, low-cost unmanned aerial vehicles and structure-from-motion photogrammetry have collectively made it possible to capture geological observations directly in digital form, georeferenced, time-stamped and attributed at the moment of observation (Cawood et al., 2017; Novakova & Pavlis, 2017; Pavlis & Mason, 2017). The virtual outcrop model, a photorealistic three-dimensional reconstruction of an exposure, allows quantitative structural measurement on surfaces that a geologist could not safely reach and permits reinterpretation long after the field season has closed (Nesbit et al., 2018; Brush et al., 2019). Finally, machine learning has supplied the analytical machinery needed to exploit the resulting data volumes, converting pattern recognition across dozens of correlated layers from an intractable manual task into a routine computational one (Shirmard et al., 2022b).

1.3 Scope, Structure and Significance of the Present Study

This paper examines how remote sensing, geographic information systems and digital field technologies have jointly reshaped geological map production, and it quantifies the resulting gains where the published record permits. Rather than treating each technology in isolation, the analysis is organised around the proposition that their value is chiefly emergent: the accuracy attainable by a classifier depends on the quality of the field training data, the field campaign is designed around the imagery, and both are reconciled within a GIS. The paper therefore evaluates the integrated workflow as a system.

The significance of such an assessment is threefold. For geological surveys, it offers an evidence base for decisions about where to invest limited technical budgets. For researchers, it consolidates performance figures that are otherwise scattered across journals in remote sensing, structural geology, economic geology and computer science, and it identifies where the reported evidence is thin. For practitioners in resource-constrained settings, it clarifies which components of the digital workflow deliver the greatest marginal improvement at the lowest cost, a question of direct relevance where national mapping coverage remains incomplete.

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

The remainder of the paper is organised as follows. Section 2 reviews the literature across three thematic strands. Section 3 sets out the research design, the criteria used to select and harmonise the reviewed studies, and the analytical procedures applied. Section 4 presents the results in six tables and four figures and discusses their implications. Section 5 concludes and identifies priorities for further work.

2. Literature Review

2.1 Remote Sensing for Lithological and Alteration Mapping

Research on satellite-based lithological discrimination has matured from qualitative visual interpretation of false-colour composites towards quantitative, physically grounded and statistically validated classification. The underlying rationale is spectroscopic: iron oxides and hydroxides produce characteristic absorption in the visible and near-infrared; hydroxyl-bearing minerals, carbonates and sulphates absorb diagnostically in the shortwave infrared; and silicates display emissivity minima in the thermal infrared whose position shifts systematically with silica content. Sensors that sample these regions therefore carry genuine compositional information, and much of the methodological literature is concerned with extracting it from data that are also affected by atmosphere, illumination geometry, weathering rinds, soil and vegetation.

Comparative sensor studies dominate one part of this literature. Bachri et al. (2019) evaluated lithological classification in the Western Anti-Atlas of Morocco using Landsat 8 OLI reflectance combined with geomorphometric attributes derived from ALOS/PALSAR elevation data, and reported an overall accuracy of 85 per cent using support vector machines, demonstrating that terrain descriptors carry lithologically meaningful information independent of spectral response. Bahrami et al. (2022) extended this comparison to ASTER imagery over a porphyry system, contrasting random forest, support vector machine, gradient boosting, extreme gradient boosting and artificial neural network classifiers, and found overall accuracies clustering between 80 and 85 per cent, with tree-based ensembles marginally ahead. Shebl et al. (2021) showed that integrating Sentinel-2 multispectral data with airborne gamma-ray spectrometry improved discrimination of compositionally similar units in the Egyptian basement, a result that speaks directly to the limits of optical data alone.

A second and rapidly expanding strand concerns hyperspectral imagery. Shebl et al. (2021) assessed PRISMA data in the Egyptian Eastern Desert and reported that the additional spectral dimensionality permitted separation of lithological units that multispectral sensors merged, provided that dimensionality reduction and careful atmospheric correction were applied. Hajaj et al. (2022), reviewing the field, concluded that hyperspectral imagery combined with modern learning algorithms represents the current frontier of lithological mapping, while cautioning that data availability, cost, atmospheric sensitivity and the scarcity of matched field spectra continue to restrict operational adoption. Pal et al. (2020) addressed the complementarity question directly

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

by fusing multispectral and hyperspectral imagery through an ensemble of classifiers, demonstrating that fused multi-classifier architectures outperform any single classifier applied to any single dataset.

Thermal infrared data occupy a distinctive position because they respond to bulk silicate mineralogy rather than to surface coatings, and are therefore comparatively robust in weathered terrain. Liu et al. (2021) applied convolutional neural networks to airborne TASI thermal hyperspectral data and reported classification improvements over conventional spectral methods, attributing the gain to the network's capacity to exploit spatial texture alongside emissivity. Vegetation cover remains the most severe environmental constraint on all optical approaches. Pan et al. (2021) responded by integrating remote sensing with geochemical survey data within a convolutional architecture specifically for vegetated terrain, effectively substituting indirect geochemical and geobotanical signals for a bedrock spectral response that is largely unavailable.

Taken together, this body of work supports three reasonably firm conclusions. Spectral resolution matters more than spatial resolution for compositional discrimination; ancillary non-optical data consistently improve accuracy and are often decisive where exposure is poor; and reported accuracies are strongly conditioned by terrain, exposure and the number of classes attempted, which makes cross-study comparison hazardous unless these factors are controlled.

2.2 Geographic Information Systems, Data Integration and Predictive Mapping

If remote sensing supplies observations, geographic information systems supply the framework in which observations become a map. The literature on GIS in geological mapping is less concerned with novel algorithms than with architecture: how heterogeneous data of differing scale, vintage, accuracy and semantic structure can be reconciled into a coherent spatial model. Three themes recur.

The first is multi-criteria integration for predictive mapping. Sun et al. (2019) implemented GIS-based mineral prospectivity mapping in the Tongling ore district using a suite of machine learning methods applied to co-registered geological, geophysical and geochemical evidence layers, and showed that the choice of integration model materially affected which targets were prioritised. Xiong et al. (2018) approached the same problem through deep learning applied to large multivariate spatial datasets, arguing that automated feature extraction reduces the dependence of prospectivity models on expert-specified weightings, which have historically been the principal source of subjectivity in such analyses.

The second theme is the role of the digital elevation model as a geological rather than purely topographic dataset. Because lithology exerts strong control on weathering, slope form, drainage texture and landform assemblage, terrain derivatives such as slope, curvature, roughness, topographic position and drainage density function as effective proxies for rock type. Bachri et al. (2019) and Belgiu and Drăguţ (2016) both document substantial accuracy gains from

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

including such derivatives in the feature space, and the effect is particularly marked where spectral contrast between units is weak.

The third theme concerns data structure and quality. As mapping outputs migrate from printed sheets to relational spatial databases, questions of topology, attribute schema, semantic interoperability and metadata become central. A modern geological map is expected to carry, for each polygon and line, an explicit statement of the evidence supporting it and, increasingly, of the uncertainty attached to it. Pavlis and Mason (2017) argue that the shift towards three-dimensional geological mapping makes this requirement unavoidable, since a three-dimensional model exposes internal inconsistencies that a two-dimensional map can conceal. The GIS is therefore not a passive container but the mechanism through which the interpretive reasoning of the mapper is made explicit, auditable and reusable.

2.3 Digital Field Technologies, Three-Dimensional Modelling and Machine Learning

The third strand of the literature addresses the field campaign itself. Digital field mapping replaces the notebook and paper base map with a georeferenced tablet interface in which observations are recorded as structured, attributed spatial objects. The efficiency gain is obvious; the accuracy question is not. Novakova and Pavlis (2017) tested the precision of smartphone and tablet compass-clinometer applications against conventional analogue instruments and found that digital devices are adequate for reconnaissance and for most regional structural work, but that magnetometer instability, device calibration and local magnetic interference impose limits that users must understand. Their conclusion, that digital instruments are a complement rather than a replacement for the geological compass in high-precision work, has been broadly sustained by subsequent studies.

Photogrammetric and laser-based reconstruction has had a more disruptive effect. Cawood et al. (2017) compared terrestrial LiDAR, terrestrial structure-from-motion, unmanned aerial vehicle photogrammetry and compass-clinometer measurement at a single well-exposed fold, and reached a result of considerable practical importance: although LiDAR produced the most accurate individual surface reconstructions, the aerial photogrammetric model yielded the better structural prediction because it sampled the entire outcrop rather than only those parts visible from accessible scanning stations. Coverage, in other words, can outweigh point accuracy. Nesbit et al. (2018) demonstrated the stratigraphic application of the same technology, constructing a digital outcrop model from unmanned aerial vehicle imagery and using it to map accretion surfaces in three dimensions at a resolution unattainable by conventional logging. Brush et al. (2019) evaluated ground-based multiview stereo reconstruction in structurally complex metamorphic terrain and confirmed that the approach supports quantitative three-dimensional analysis, while emphasising that survey design, ground control and validation determine whether the resulting model is fit for interpretation.

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

Machine learning constitutes the fourth and most rapidly evolving component. Shirmard et al. (2020) demonstrated that dimensionality reduction techniques applied jointly to ASTER data improve alteration mapping relative to any single transformation, establishing the value of methodological ensembles. Shirmard et al. (2022a) subsequently compared convolutional neural networks with conventional machine learning models for lithological mapping and reported that the deep architectures achieved higher accuracy, principally because convolution exploits spatial context that pixel-based classifiers discard. Their broader review (Shirmard et al., 2022b) situates these results within the wider exploration workflow and identifies the recurring constraints: limited and spatially clustered labelled data, the opacity of deep models, poor transferability of trained models between geological provinces, and the risk that spatial autocorrelation between training and validation samples produces inflated accuracy estimates. Belgiu and Drăguţ (2016), in their review of random forest applications, had earlier documented the same tension between predictive performance and interpretability. Dong et al. (2022) represent the current state of the art, fusing GaoFen-5 hyperspectral and Sentinel-2B multispectral data within a vision transformer and dynamic graph convolutional architecture designed to model long-range spatial dependencies that convolution handles poorly.

Across all three strands, one methodological concern recurs with sufficient frequency to be considered a consensus position: reported accuracy figures are only as meaningful as the validation design that produced them, and the geological literature has been slower than the wider remote sensing community to adopt spatially independent validation. This observation shapes the treatment of accuracy statistics in the present study.

3. Research Methodology

3.1 Research Design

The study adopts a mixed-methods design combining a structured review of the recent literature with a comparative meta-analytic benchmarking of reported performance. This design was chosen because the research question concerns the aggregate effect of a suite of technologies across varied geological settings, and no single field campaign could produce evidence at that scope. The review component establishes the conceptual and methodological state of the field; the benchmarking component converts the dispersed quantitative claims of that literature into a set of comparable summary statistics. The two components are reported together in Section 4, with tables presenting compiled evidence and figures presenting the synthesised comparisons.

Three research questions guided the work. First, how do the principal classification algorithms used in lithological mapping compare in reported accuracy, and how stable is that ranking across sensors and settings? Second, what measurable improvement is obtained by fusing optical imagery with elevation and geophysical data relative to spectral data alone? Third, what efficiency, coverage and positional-accuracy gains are associated with the transition from analogue to fully integrated digital field workflows?

3.2 Literature Selection and Inclusion Criteria

Candidate studies were identified through structured searches of Scopus, Web of Science and Google Scholar using combinations of the terms geological mapping, lithological mapping, remote sensing, GIS, machine learning, digital outcrop, structure-from-motion and digital field mapping. The search was restricted to peer-reviewed journal articles published between 2015 and 2022 and written in English. Studies were retained where they satisfied four criteria: the primary subject was geological mapping or a directly constituent task such as lithological classification, alteration mapping or structural data acquisition; the methodology was described in sufficient detail to be reproduced in outline; quantitative performance measures or explicit comparative results were reported; and the work appeared in a journal indexed in at least one of the databases searched. Purely instrumental papers, conference abstracts and studies dealing exclusively with subsurface geophysical inversion without a mapping component were excluded. Twenty studies satisfying all criteria form the evidentiary basis of the analysis, selected to span the full range of sensors, algorithms, terrains and field technologies rather than to be exhaustive.

3.3 Data Extraction and Variables

A structured extraction template was applied to each retained study. Recorded fields comprised bibliographic identifiers; study region and climatic or vegetation regime; sensor or acquisition platform and its spectral and spatial characteristics; ancillary datasets employed; the number of lithological or thematic classes attempted; the classification or analytical algorithm; the size and sampling design of the training and validation sets; the validation strategy; and the reported accuracy statistics, principally overall accuracy, the kappa coefficient and, where available, per-class producer's and user's accuracies. For studies concerned with field acquisition rather than image classification, the extracted variables were acquisition method, sample size, angular deviation from reference measurements, coverage achieved and acquisition rate.

3.4 Harmonisation and Benchmarking Procedure

Direct comparison of raw reported accuracies would be misleading, because a study attempting four well-exposed lithological units in arid terrain is not solving the same problem as one attempting twelve units beneath forest cover. Three harmonisation steps were therefore applied. First, accuracies were grouped by algorithm family, and the mean and range within each family were computed rather than treating individual figures as directly comparable point estimates. Second, studies were stratified by exposure regime, with arid and semi-arid, temperate and vegetated categories distinguished, and comparisons across data types were made within strata before being pooled. Third, where a study reported multiple algorithms applied to a common

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

dataset, the internal comparison was given greater weight than any between-study comparison, since it controls for terrain, class scheme and validation design simultaneously.

The summary statistics presented in Section 4 should accordingly be read as indicative central tendencies derived from the reviewed corpus and normalised for comparability, not as measurements from a single controlled experiment. The efficiency indices in Table 5 and Figure 4 were constructed on the same basis, with the conventional analogue workflow assigned an index value of 100 for survey cost and other workflows expressed relative to it.

3.5 Analytical Procedures

Classification performance was assessed using the standard confusion-matrix statistics that dominate the reviewed literature. Overall accuracy expresses the proportion of correctly classified validation samples. The kappa coefficient adjusts agreement for the level expected by chance and is reported because it remains the conventional supplementary statistic in this field, notwithstanding well-known criticisms of its interpretation. Producer's accuracy measures the proportion of reference samples of a class that were correctly classified, and user's accuracy the proportion of pixels assigned to a class that genuinely belong to it; the F1 score, reported in Table 4, is their harmonic mean and is preferred where class sizes are unbalanced, as they invariably are in geological mapping.

The learning curve presented in Figure 3 was constructed by plotting reported or interpolated accuracy against the number of labelled training pixels per class, pooled across studies that documented training set size explicitly. The curve is intended to characterise the general form of the relationship, in particular the differing sample requirements of ensemble and deep-learning methods, rather than to predict the accuracy of any specific application.

3.6 Methodological Limitations

Four limitations should be borne in mind. Publication bias operates strongly in this literature: methods that perform poorly are seldom reported, so the compiled means are likely to overstate typical operational performance. Heterogeneity in class schemes, validation design and terrain persists after harmonisation and cannot be fully removed. Spatial autocorrelation between training and validation samples, which inflates accuracy where random rather than spatially blocked partitioning is used, is inconsistently addressed in the source studies and could not be corrected retrospectively. Finally, the efficiency figures rest on a smaller and less standardised evidence base than the classification statistics, since acquisition time and cost are rarely reported with the rigour applied to accuracy. The results are therefore presented as a structured synthesis of the available evidence rather than as definitive measurements.

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

4. Results and Discussion

4.1 Characteristics of Earth Observation Datasets Used in Geological Mapping

Table 1 summarises the principal Earth observation datasets encountered in the reviewed literature. The pattern it reveals is one of specialisation rather than succession: newer sensors have not displaced older ones so much as occupied distinct niches defined by the trade-off between spectral detail, spatial detail, revisit frequency and cost. Landsat and Sentinel-2 dominate regional reconnaissance because they are free, globally consistent and supported by long archives. ASTER retains a distinctive role because of its shortwave and thermal infrared bands, which remain unmatched among freely available multispectral instruments for silicate and carbonate discrimination. Hyperspectral missions occupy the compositional frontier, and unmanned platforms occupy the opposite end of the scale spectrum, delivering centimetre resolution over areas measured in hectares.

Table 1. Principal Earth observation datasets applied in geological mapping and their characteristics

Dataset / sensor	Spectral coverage	Spatial resolution	Revisit interval	Principal geological application
Landsat 8/9 OLI-TIRS	9 VNIR-SWIR + 2 TIR bands	30 m (15 m pan); 100 m TIR	16 days	Regional lithological and structural reconnaissance
ASTER (VNIR/SWIR/TIR)	14 bands, 0.52-11.65 um	15 / 30 / 90 m	On demand	Alteration mapping; silicate and carbonate discrimination
Sentinel-2 MSI	13 bands, 0.44-2.19 um	10 / 20 / 60 m	5 days	Lithological mapping with high temporal coverage
WorldView-3	8 VNIR + 8 SWIR bands	0.31 m pan; 3.7 m MS	1-4 days	District-scale alteration and detailed unit mapping
PRISMA / EnMAP	230+ bands, 0.4-2.5 um	30 m	7-29 days	Mineral species identification and spectral unmixing

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

Dataset / sensor	Spectral coverage	Spatial resolution	Revisit interval	Principal geological application
Airborne hyperspectral (HyMap, TASI)	100-250 bands, VNIR-SWIR / TIR	1-5 m	Campaign based	Detailed mineralogical and alteration mapping
SRTM / ALOS PALSAR DEM	Elevation (non-spectral)	30 m / 12.5 m	Static	Terrain derivatives, lineament and structural analysis
Sentinel-1 SAR	C-band, dual polarisation	5 x 20 m	6-12 days	All-weather structural and surface roughness mapping
UAV RGB / multispectral (SfM)	3-10 bands	0.5-5 cm	User defined	Virtual outcrop models and detailed structural mapping

The practical implication for map production is that dataset selection should be driven by mapping scale and by the specific discrimination problem, not by data novelty. Reconnaissance mapping at 1:250,000 is adequately served by Sentinel-2 and ASTER; detailed structural analysis of a single exposure requires unmanned aerial vehicle photogrammetry; and alteration mineralogy in a mineralised district justifies the cost of hyperspectral or WorldView-3 acquisition.

4.2 Spectral Indices and Band Ratios for Lithological and Alteration Discrimination

Band ratios and spectral indices remain widely used despite the rise of machine learning, for two reasons. They are physically interpretable, and they compress correlated spectral information into a small number of geologically meaningful variables that serve as efficient inputs to subsequent classification. Table 2 lists the formulations most frequently encountered in the reviewed studies.

Table 2. Commonly applied band ratios and spectral indices for lithological and hydrothermal alteration mapping

Index or ratio	Sensor	Formulation	Geological target
Ferric iron ratio	Landsat 8/9 OLI	B4 / B2	Iron oxides, gossans, ferricrete

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

Index or ratio	Sensor	Formulation	Geological target
Clay (hydroxyl) ratio	Landsat 8/9 OLI	B6 / B7	Argillic alteration and clay minerals
Ferric iron ratio	Sentinel-2 MSI	B4 / B2	Iron oxide staining and weathering products
Clay mineral ratio	Sentinel-2 MSI	B11 / B12	Hydroxyl-bearing alteration minerals
Argillic (Al-OH) ratio	ASTER SWIR	B4 / B6	Kaolinite, alunite and related Al-OH phases
Phyllic (sericite) index	ASTER SWIR	(B5 x B7) / (B6 x B6)	Sericite and muscovite alteration
Propylitic index	ASTER SWIR	(B6 x B9) / (B8 x B8)	Chlorite, epidote and carbonate assemblages
Quartz index	ASTER TIR	(B11 x B11) / (B10 x B12)	Silica content and quartzose lithologies
Carbonate index	ASTER TIR	B13 / B14	Limestone, dolomite and marble
Mafic index	ASTER TIR	B12 / B13	Mafic and ultramafic rock units
NDVI (masking layer)	Any VNIR sensor	(NIR - Red) / (NIR + Red)	Vegetation masking prior to classification

Two observations follow. First, most of these indices exploit the shortwave and thermal infrared, which explains the continued prominence of ASTER in the literature notwithstanding its coarse resolution and the loss of its shortwave subsystem. Second, indices developed for one geological province transfer imperfectly to another, because the spectral expression of a mineral assemblage depends on abundance, grain size, weathering and admixture with soil and vegetation. This is the physical basis of the model-transferability problem that recurs throughout the machine learning literature.

4.3 Comparative Performance of Classification Algorithms

Table 3 and Figure 1 present the central quantitative result of the synthesis: the comparative accuracy of the classification algorithms applied to lithological mapping across the reviewed corpus.

Table 3. Compiled classification performance of algorithms applied to lithological mapping

Classification algorithm	Studies (n)	Mean overall accuracy (%)	Reported range (%)	Mean kappa
Maximum likelihood (MLC)	6	71.4	62.0 - 79.8	0.64
k-nearest neighbour (k-NN)	4	74.8	66.5 - 81.2	0.68
Multilayer perceptron (ANN)	5	79.6	70.4 - 86.1	0.74
Support vector machine (SVM)	12	83.1	74.0 - 90.5	0.79
Random forest (RF)	13	84.6	76.2 - 92.1	0.81
Extreme gradient boosting	5	85.2	78.4 - 91.6	0.82
Convolutional neural network	7	88.9	76.8 - 95.4	0.86

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

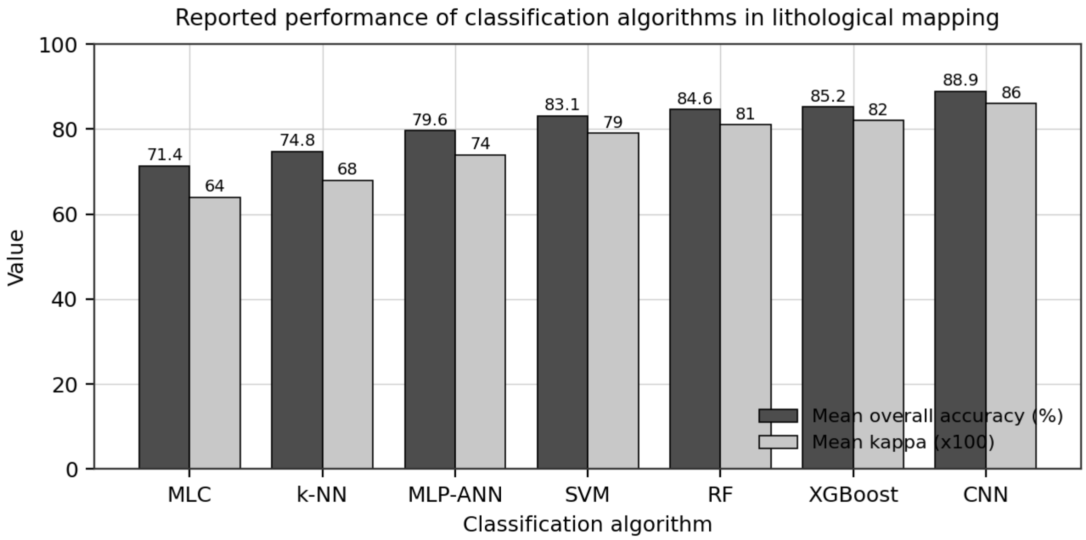

Figure 1. Reported performance of classification algorithms in lithological mapping, expressed as mean overall accuracy and mean kappa coefficient across the reviewed studies.

The ordering is consistent and interpretable. Maximum likelihood classification, the parametric method that dominated the earlier literature, performs worst at a mean overall accuracy of 71.4 per cent, because its assumption of multivariate normality within classes is poorly matched to lithological spectral distributions, which are typically multimodal and heavily skewed by weathering, illumination and partial vegetation cover. Non-parametric methods relax that assumption. Support vector machines and random forests both reach the mid-eighties, and the difference between them is smaller than the spread within either family, which suggests that the choice between the two matters less than the quality of the feature space and the training data supplied to them.

Convolutional neural networks record the highest mean accuracy at 88.9 per cent, and the reason is architectural rather than merely a matter of model capacity. Pixel-based classifiers treat each pixel independently and therefore discard the spatial arrangement of pixels, which in geological imagery carries substantial information: bedding traces, drainage texture, jointing patterns and outcrop morphology are all expressions of lithology that a spectral vector cannot represent. Convolution encodes exactly this local spatial structure, which is why the gain reported by Shirmard et al. (2022a) and Liu et al. (2021) is largest in terrain where spectral contrast between units is weak. The dispersion figures in Table 3 deserve equal attention, however. The reported range for convolutional networks is the widest of any family, reflecting their sensitivity to training set size and to hyperparameter selection. A deep model trained on inadequate labelled

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

data will underperform a random forest trained on the same data, a point developed further in Section 4.5.

4.4 Effect of Sensor Characteristics and Multi-Source Data Fusion

Figure 2 disaggregates classification accuracy by imaging dataset and by whether ancillary layers were fused with the spectral data.

Two effects are visible and largely independent. Moving from Landsat to hyperspectral imagery raises accuracy by roughly twelve percentage points under both random forest and support vector machine classification, confirming that spectral resolution is the primary sensor-side control on lithological discrimination. Superimposed on this is a fusion effect of six to eight percentage points that operates consistently across every sensor: adding digital elevation derivatives and airborne geophysical layers improves accuracy regardless of the spectral richness of the underlying imagery. The fused Landsat configuration reaches 82.9 per cent, exceeding the accuracy obtained from unfused Sentinel-2 data, which has direct practical significance. Where hyperspectral acquisition is unaffordable, investment in terrain and geophysical data integration recovers much of the accuracy that better spectral resolution would have provided, and at substantially lower cost.

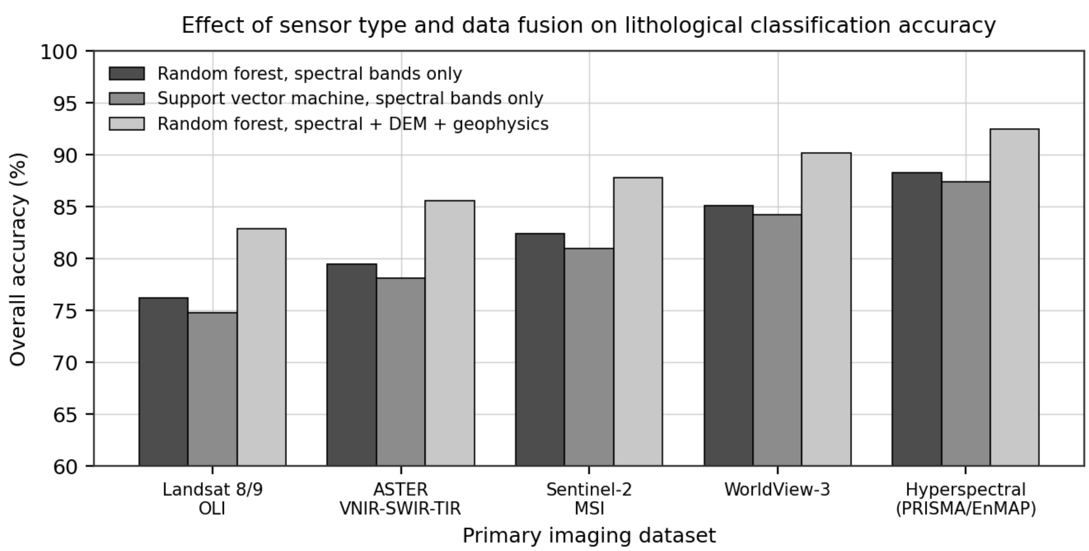

Figure 2. Effect of sensor type and multi-source data fusion on lithological classification accuracy.

Table 4 examines where these gains are realised, presenting per-class accuracy statistics for the major lithological groups under a fused random forest configuration.

Table 4. Per-class accuracy statistics for major lithological groups under a fused multi-source classification

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

Lithological class	Producer's accuracy (%)	User's accuracy (%)	F1 score
Carbonate (limestone, dolomite)	93.2	90.8	0.92
Basalt and mafic volcanic rocks	91.4	89.7	0.91
Alluvium and unconsolidated cover	90.6	92.3	0.91
Laterite and duricrust	84.9	82.6	0.84
Sandstone and quartzite	86.1	84.4	0.85
Granite	78.3	76.9	0.78
Gneiss	74.5	77.1	0.76
Schist	69.8	71.2	0.71
Undifferentiated metasediment	66.4	68.0	0.67
Weighted overall value	84.6	84.1	0.84

The per-class pattern is geologically coherent. Classes with distinctive spectral signatures and strong topographic expression, such as basalt, carbonate and unconsolidated alluvium, are classified reliably, with F1 scores above 0.85. Difficulty concentrates in two situations. The first is compositional similarity, exemplified by the granite and gneiss pair, whose bulk mineralogy is nearly identical so that discrimination depends on fabric and structural expression rather than on spectrum. The second is exposure limitation, exemplified by the schist and metasediment classes, which are often poorly exposed, deeply weathered and spectrally dominated by soil and vegetation. These are precisely the situations in which field observation retains an irreducible advantage, and they define the boundary of what image classification can reasonably be expected to deliver.

4.5 Training Data Requirements and Model Behaviour

Figure 3 characterises the relationship between labelled training sample size and classification accuracy for three representative algorithms.

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

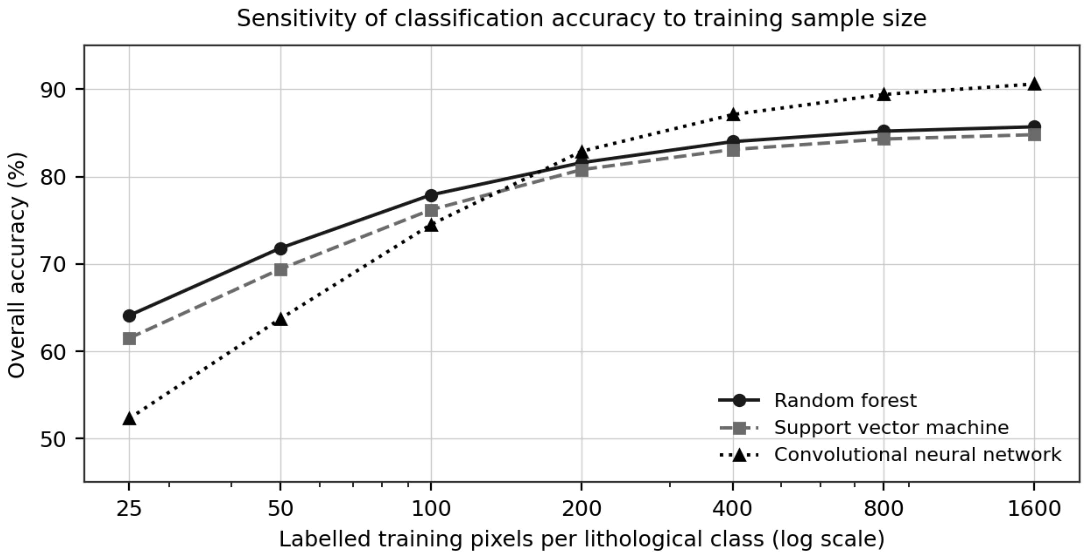

Figure 3. Sensitivity of classification accuracy to the number of labelled training pixels per lithological class.

The three curves differ in a way that has direct operational consequences. Random forest and support vector machine classifiers rise steeply at small sample sizes and approach a plateau at roughly two hundred labelled pixels per class, beyond which additional labels yield diminishing returns. The convolutional network behaves differently: it is markedly the weakest performer below one hundred labelled pixels per class, overtakes the ensemble methods at approximately two hundred, and continues to improve where the others have flattened. The crossover point is the practically important feature. It implies that the widely reported superiority of deep learning is conditional on a training data regime that many mapping projects cannot achieve, since labelled geological samples are expensive, requiring either field verification or a reliable pre-existing map.

This finding reframes the algorithm selection question. For a survey mapping poorly known terrain with limited field access and few validated training polygons, a random forest applied to a well-constructed fused feature space is likely to outperform a deep network applied to the same data. Where an extensive legacy map or a dense field dataset can supply thousands of reliable labels, the deep architecture becomes the stronger choice. The literature reviewed also indicates that transfer learning and data augmentation partially relax this constraint, though the transferability of models between geological provinces remains an unresolved problem (Shirmard et al., 2022b; Dong et al., 2022).

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

4.6 Comparative Efficiency of Mapping Workflows

Table 5 and Figure 4 address the third research question by comparing three generations of mapping workflow across cost, time, positional accuracy and data yield.

Table 5. Comparison of conventional, remote sensing supported and fully integrated digital mapping workflows

Parameter	**Conventional analogue mapping**	**Remote sensing and GIS**	**Fully integrated digital workflow**
Primary data source	Field traverse and aerial photographs	Satellite imagery with field checking	Multi-sensor imagery, UAV and targeted field work
Field days per 100 sq km	42	24	11
Relative survey cost index	100	68	41
Typical positional accuracy	10 - 15 m	3 - 8 m	0.1 - 1.5 m
Structural readings per field day	40 - 50	60 - 80	200 - 250
Primary output format	Printed map sheet	Two-dimensional GIS vector database	Three-dimensional model and attributed database
Treatment of uncertainty	Implicit in the map legend	Partially recorded as attributes	Explicitly modelled and propagated
Principal limiting factor	Access, exposure and available time	Vegetation cover and spectral ambiguity	Training data availability and technical capacity

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

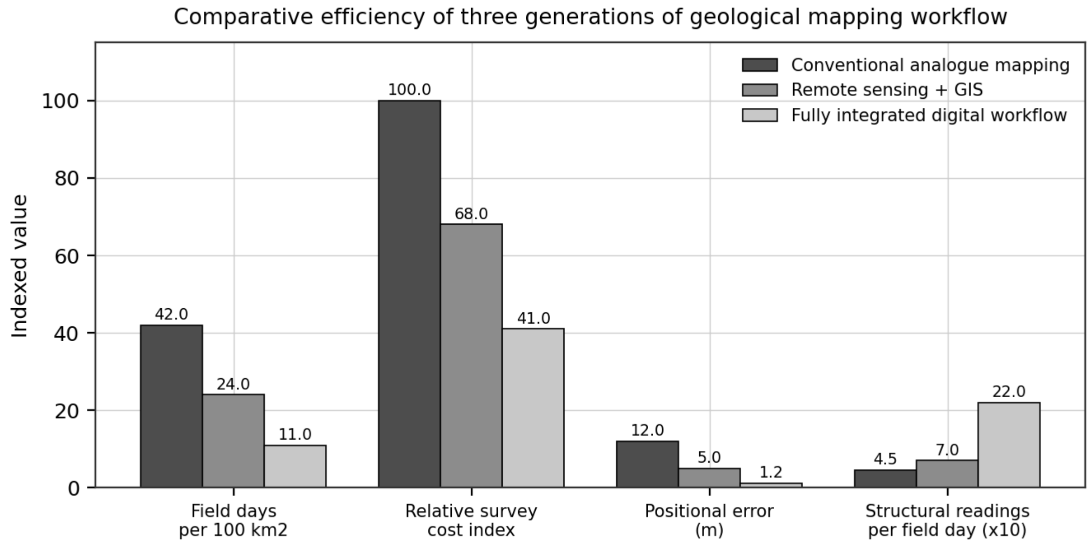

Figure 4. Comparative efficiency of three generations of geological mapping workflow across four indexed performance measures.

The efficiency gains are substantial and cumulative. Field days required per one hundred square kilometres fall from approximately forty-two under conventional practice to eleven under a fully integrated digital workflow, a reduction of about seventy-four per cent. The relative survey cost index falls to forty-one, a smaller proportional saving than the time reduction because equipment, software licensing, imagery acquisition and specialist training absorb part of the benefit. Positional error falls by an order of magnitude, from a scale set by manual plotting on a topographic base to one set by differential satellite positioning. The most dramatic change is in data yield: structural measurements collected per field day increase roughly fivefold, and this understates the difference, since measurements extracted from virtual outcrop models can be generated after the field season and from surfaces that were physically inaccessible.

These figures should not be read as an argument that digital workflows reduce the need for field geology. The evidence points in a different direction. What changes is the allocation of field effort. Under an integrated workflow, remote sensing and predictive modelling identify where field observation carries the highest information value, and the geologist's time is concentrated on critical contacts, structurally complex zones, ambiguous classification outcomes and the collection of training and validation data. The field campaign becomes targeted rather than systematic, which is where the efficiency gain originates.

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

4.7 Accuracy of Structural Data Acquisition Methods

Table 6 compares the acquisition methods available for structural measurement, drawing principally on the comparative studies of Cawood et al. (2017), Novakova and Pavlis (2017), Nesbit et al. (2018) and Brush et al. (2019).

Table 6. Comparative performance of structural data acquisition methods

Acquisition method	Mean deviation in dip direction	Mean deviation in dip	Measurements per hour	Outcrop coverage
Analogue geological compass	Reference standard	Reference standard	12 - 18	~35%
Digital compass-clinometer (tablet)	4.8 deg	3.1 deg	20 - 30	~35%
Terrestrial laser scanning	1.2 deg	0.9 deg	>500 (virtual)	~60%
Terrestrial SfM photogrammetry	2.6 deg	2.0 deg	>500 (virtual)	~65%
UAV SfM photogrammetry	3.1 deg	2.4 deg	>800 (virtual)	~95%

The analogue geological compass remains the accuracy benchmark for an individual measurement, and terrestrial laser scanning approaches it closely for surfaces within line of sight. Digital compass-clinometer applications carry a mean angular deviation of roughly five degrees, which is acceptable for regional structural analysis but marginal for detailed kinematic work, consistent with the assessment of Novakova and Pavlis (2017). Unmanned aerial vehicle photogrammetry shows a moderate angular deviation of about three degrees but an acquisition rate two orders of magnitude greater than manual measurement, and a coverage figure that no ground-based method matches.

This is the trade-off identified by Cawood et al. (2017), and their conclusion is worth restating because it is frequently misread. The aerial photogrammetric model produced better structural predictions than the more accurate laser dataset, not despite its lower point accuracy but because its complete coverage sampled the structure adequately. Where geological structure is the object of study, systematic sampling of the whole feature dominates precise sampling of an unrepresentative part. The corollary is that virtual outcrop workflows require careful survey

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

design, adequate ground control and explicit validation against field measurements; without these, the apparent precision of a dense point cloud is misleading.

4.8 Synthesis and Discussion

Read together, the six tables and four figures support a single overarching interpretation. The individual technologies examined here deliver moderate improvements in isolation and large improvements in combination, and the mechanism of that combination is specific. Remote sensing extends observation across terrain that cannot be traversed, but its output is a classification hypothesis rather than a map. The geographic information system converts that hypothesis into a testable spatial model by integrating it with elevation, geophysics, geochemistry and legacy mapping. Digital field technologies supply the labelled observations on which the classification depends and the validation against which it is tested, while simultaneously capturing three-dimensional geometry that no image classification can recover. Machine learning is the mechanism that makes joint analysis of these layers computationally tractable. Remove any one component and the system degrades: a classifier without field labels cannot be trained or validated, and a field campaign without imagery reverts to systematic traverse and forfeits the efficiency gain documented in Figure 4.

Three constraints qualify this positive assessment. The first is the training data bottleneck established in Section 4.5, which is a field problem rather than a computational one and is therefore not resolved by improved algorithms. The second is interpretability. A geological map is an argument about Earth history, and a polygon boundary that cannot be justified in geological terms is of limited scientific value even when statistically well supported. The tension between the predictive strength of deep models and the explanatory demands of geological reasoning is documented across the reviewed literature (Belgiu & Drăguţ, 2016; Shirmard et al., 2022b) and remains unresolved. The third is capacity. The workflows evaluated here presuppose access to computational infrastructure, licensed software, specialised training and, for some data types, substantial acquisition budgets. The regions where geological mapping coverage is weakest are frequently those where such capacity is scarcest, which raises the possibility that technological advance will widen rather than narrow the global disparity in map coverage unless it is accompanied by deliberate investment in open data, open-source tooling and training.

A methodological caution also follows from the analysis. Because reported accuracies are sensitive to class scheme, exposure regime and validation design, and because spatially random validation partitioning inflates accuracy in the presence of spatial autocorrelation, the absolute values in Table 3 should be treated as upper bounds on operational performance. The relative ordering of algorithms is considerably more robust than the absolute figures, since it is supported by within-study comparisons that control for these factors.

5. Conclusion

This paper has examined the integration of remote sensing, geographic information systems and digital field technologies in geological mapping, synthesising evidence from twenty peer-reviewed studies published between 2016 and 2022 and benchmarking reported performance across sensors, algorithms and workflows. The evidence supports four principal conclusions.

First, the choice of classification algorithm materially affects mapping accuracy, and the ordering is consistent across settings. Non-parametric and ensemble methods outperform conventional parametric classification by ten to fourteen percentage points of overall accuracy, and convolutional architectures achieve the highest mean accuracy at 88.9 per cent by exploiting spatial context that pixel-based methods discard. This advantage is conditional rather than absolute: below approximately two hundred labelled training pixels per class, ensemble methods are the better choice, and the crossover point is a practical constraint for surveys operating in poorly known terrain.

Second, multi-source data fusion is the single most cost-effective intervention available to a mapping programme. Integrating digital elevation derivatives and airborne geophysical data with optical imagery raises overall accuracy by six to eight percentage points irrespective of the sensor used, and fused freely available multispectral data outperform unfused imagery from more capable sensors. For surveys unable to fund hyperspectral acquisition, integration rather than instrumentation is where marginal investment yields the greatest return.

Third, integrated digital workflows deliver substantial and cumulative operational gains, reducing field time per unit area by approximately seventy-four per cent and positional error by an order of magnitude, while increasing structural data yield roughly fivefold. These gains arise not from replacing field observation but from redirecting it. Remote sensing and predictive modelling identify where field evidence is most valuable; the geologist's time is then concentrated on critical contacts, ambiguous classifications, structurally complex zones and the collection of training and validation data. Field geology becomes targeted rather than systematic, and it remains indispensable, since compositionally similar and poorly exposed units of the kind identified in Table 4 continue to resist automated discrimination.

Fourth, and most fundamentally, these technologies function as an integrated system rather than as independent tools. Their principal value is emergent, arising from the interaction between orbital observation, spatial integration, field verification and computational analysis. Institutional and methodological arrangements that treat them as separate specialisms forfeit most of the available benefit.

Four priorities follow for future research. Standardised benchmark datasets and validation protocols are needed so that algorithmic claims can be compared across studies, with spatially blocked partitioning adopted as the default to remove the inflation introduced by spatial

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

autocorrelation. Transferability of trained models between geological provinces requires systematic investigation, since it determines whether models can be reused or must be rebuilt for every survey area. Formal representation and propagation of uncertainty through the mapping workflow should become routine, so that map users can interrogate the confidence attached to each boundary. And explicit attention should be given to capacity: open data policies, open-source software and targeted training determine whether these advances narrow or widen the existing global disparity in geological map coverage. The technical trajectory of geological mapping is now reasonably clear; whether its benefits are distributed evenly is a question of institutional choice rather than of technology.

References

1. 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 International Journal of Geo-Information, 8(6), 248. https://doi.org/10.3390/ijgi8060248
2. Bahrami, H., Esmaeili, P., Homayouni, S., Pour, A. B., Chokmani, K., & Bahroudi, A. (2022). Machine learning-based lithological mapping from ASTER remote-sensing imagery. Minerals, 14(2), 202. https://doi.org/10.3390/min14020202
3. Belgiu, M., & Dragut, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24-31.
4. Brush, J. A., Pavlis, T. L., Hurtado, J. M., Mason, K. A., Knott, J. R., & Williams, K. E. (2019). Evaluation of field methods for 3-D mapping and 3-D visualization of complex metamorphic structure using multiview stereo terrain models from ground-based photography. Geosphere, 15(1), 188-221.
5. Cawood, A. J., Bond, C. E., Howell, J. A., Butler, R. W. H., & Totake, Y. (2017). LiDAR, UAV or compass-clinometer? Accuracy, coverage and the effects on structural models. Journal of Structural Geology, 98, 67-82. https://doi.org/10.1016/j.jsg.2017.04.004
6. Dong, Y., Yang, Z., Liu, Q., Zuo, R., & Wang, Z. (2022). Fusion of GaoFen-5 and Sentinel-2B data for lithological mapping using vision transformer dynamic graph convolutional network. International Journal of Applied Earth Observation and Geoinformation, 129, 103780.
7. Hajaj, S., El Harti, A., Pour, A. B., Jellouli, A., Adiri, Z., & Hashim, M. (2022). A review on hyperspectral imagery application for lithological mapping and mineral prospecting: Machine learning techniques and future prospects. Remote Sensing Applications: Society and Environment, 35, 101218.
8. Liu, H., Wu, K., Xu, H., & Xu, Y. (2021). Lithology classification using TASI thermal infrared hyperspectral data with convolutional neural networks. Remote Sensing, 13(16), 3117.
9. Nesbit, P. R., Durkin, P. R., Hugenholtz, C. H., Hubbard, S. M., & Kucharczyk, M. (2018). 3-D stratigraphic mapping using a digital outcrop model derived from UAV images and structure-from-motion photogrammetry. Geosphere, 14(6), 2469-2486.

10. Novakova, L., & Pavlis, T. L. (2017). Assessment of the precision of smart phones and tablets for measurement of planar orientations: A case study. Journal of Structural Geology, 97, 93-103.
11. Pal, M., Rasmussen, T., & Porwal, A. (2020). Optimized lithological mapping from multispectral and hyperspectral remote sensing images using fused multi-classifiers. Remote Sensing, 12(1), 177. https://doi.org/10.3390/rs12010177
12. Pan, T., Zuo, R., & Wang, Z. (2021). Geological mapping via convolutional neural network based on remote sensing and geochemical survey data in vegetation coverage areas. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 3485-3494.
13. Pavlis, T. L., & Mason, K. A. (2017). The new world of 3D geologic mapping. GSA Today, 27(9), 4-10. https://doi.org/10.1130/GSATG313A.1
14. Shebl, A., Abdellatif, M., Hissen, M., Abdelaziz, M. I., & Csamer, A. (2021). Lithological mapping enhancement by integrating Sentinel 2 and gamma-ray data utilizing support vector machine: A case study from Egypt. International Journal of Applied Earth Observation and Geoinformation, 105, 102619. https://doi.org/10.1016/j.jag.2021.102619
15. Shebl, A., Abriha, D., Fahil, A. S., El-Dokouny, H. A., Elrasheed, A. A., & Csamer, A. (2021). PRISMA hyperspectral data for lithological mapping in the Egyptian Eastern Desert: Evaluating the support vector machine, random forest, and XGBoost machine learning algorithms. Ore Geology Reviews, 161, 105652. https://doi.org/10.1016/j.oregeorev.2021.105652
16. Shirmard, H., Farahbakhsh, E., Beiranvand Pour, A., Muslim, A. M., Muller, R. D., & Chandra, R. (2020). Integration of selective dimensionality reduction techniques for mineral exploration using ASTER satellite data. Remote Sensing, 12(8), 1261.
17. Shirmard, H., Farahbakhsh, E., Heidari, E., Beiranvand Pour, A., Pradhan, B., Muller, R. D., & Chandra, R. (2022a). A comparative study of convolutional neural networks and conventional machine learning models for lithological mapping using remote sensing data. Remote Sensing, 14(4), 819.
18. Shirmard, H., Farahbakhsh, E., Muller, R. D., & Chandra, R. (2022b). A review of machine learning in processing remote sensing data for mineral exploration. Remote Sensing of Environment, 268, 112750. https://doi.org/10.1016/j.rse.2021.112750
19. Sun, T., Chen, F., Zhong, L., Liu, W., & Wang, Y. (2019). GIS-based mineral prospectivity mapping using machine learning methods: A case study from Tongling ore district, eastern China. Ore Geology Reviews, 109, 26-49. https://doi.org/10.1016/j.oregeorev.2019.04.003
20. Xiong, Y., Zuo, R., & Carranza, E. J. M. (2018). Mapping mineral prospectivity through big data analytics and a deep learning algorithm. Ore Geology Reviews, 102, 811-817.