



OPEN ACCESS

EDITED BY

Puhong Duan,
Hunan University, China

REVIEWED BY

Xiaohui Wei,
Hunan Normal University, China
Weitao Chen,
China University of Geosciences Wuhan, China

*CORRESPONDENCE

Jane Southworth,
✉ jsouthwo@ufl.edu

RECEIVED 22 January 2024

ACCEPTED 08 April 2024

PUBLISHED 23 May 2024

CITATION

Southworth J, Smith AC, Safaei M, Rahaman M, Alruzuq A, Tefera BB, Muir CS and Herrero HV (2024), Machine learning *versus* deep learning in land system science: a decision-making framework for effective land classification. *Front. Remote Sens.* 5:1374862. doi: 10.3389/frsen.2024.1374862

COPYRIGHT

© 2024 Southworth, Smith, Safaei, Rahaman, Alruzuq, Tefera, Muir and Herrero. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Machine learning *versus* deep learning in land system science: a decision-making framework for effective land classification

Jane Southworth^{1*}, Audrey C. Smith¹, Mohammad Safaei¹, Mashoukur Rahaman¹, Ali Alruzuq¹, Bewuket B. Tefera¹, Carly S. Muir² and Hannah V. Herrero³

¹Department of Geography, University of Florida, Gainesville, FL, United States, ²Global Research Institute, William and Mary, Williamsburg, VA, United States, ³Department of Geography and Sustainability, University of Tennessee, Knoxville, TN, United States

This review explores the comparative utility of machine learning (ML) and deep learning (DL) in land system science (LSS) classification tasks. Through a comprehensive assessment, the study reveals that while DL techniques have emerged with transformative potential, their application in LSS often faces challenges related to data availability, computational demands, model interpretability, and overfitting. In many instances, traditional ML models currently present more effective solutions, as illustrated in our decision-making framework. Integrative opportunities for enhancing classification accuracy include data integration from diverse sources, the development of advanced DL architectures, leveraging unsupervised learning, and infusing domain-specific knowledge. The research also emphasizes the need for regular model evaluation, the creation of diversified training datasets, and fostering interdisciplinary collaborations. Furthermore, while the promise of DL for future advancements in LSS is undeniable, present considerations often tip the balance in favor of ML models for many classification schemes. This review serves as a guide for researchers, emphasizing the importance of choosing the right computational tools in the evolving landscape of LSS, to achieve reliable and nuanced land-use change data.

KEYWORDS

machine learning, deep learning, land system science, land cover classification, decision making framework, interdisciplinary research

Highlights

- Introduced a decision framework for ML vs. DL in Land System Science (LSS).
- Evaluated the strengths and limitations of ML and DL in LSS classification.
- Highlighted the predominance of ML in current LSS classification schemes.
- Identified future potential and challenges for DL in LSS research.
- Emphasized the role of interdisciplinary collaboration for LSS advancements.

1 Introduction

Positioned at the interface of natural and human systems, the interdisciplinary field of land system science (LSS) has emerged as a pivotal platform for the integration of different dimensions of global environmental change and sustainability research (Verburg et al., 2013). Central to LSS is the examination of how human activities redefine terrestrial systems' structural and functional dynamics in the current geologic epoch, unofficially termed the Anthropocene (Turner et al., 2007; Verburg et al., 2016). Inextricably linked with the Anthropocene concept is the notion that human activity is the dominant driver of global change, and newly focused attention on the bidirectional relationship between human and natural systems (Verburg et al., 2015). Human-induced alterations to the Earth's terrestrial surface have significant implications for the structure and function of ecosystems within the Earth system, as well as profound consequences for human wellbeing (Turner et al., 2007). Thus, land system changes and their implications for global environmental change and sustainability are a major research challenge for the human-environmental sciences. While this research is undertaken by various communities, the field of Land System Science (LSS), has emerged as central to global environmental change and sustainability studies across different scales and disciplinary perspectives, creating a key platform to bring different disciplines together for a common research agenda (Turner et al., 2007; Verburg et al., 2015). Engaging scientists from across the social, environmental, and natural sciences, the domain of LSS seeks to understand the complex interactions between human activities and the environment within the context of land use and land cover change dynamics (Rindfuss et al., 2004; Verburg et al., 2015).

The discipline of LSS encapsulates an in-depth analysis of landscape evolutions, including transformations in vegetation and croplands, alongside the repercussions of human-induced actions, such as intensive livestock grazing, on regional ecosystems (Turner et al., 2007). Research approaches within the field of LSS, treat land—or landscape—systems as complex social-ecological systems (SES), composed of the social and natural subsystems, where human-environment interactions extend across space and time scales (Lambin et al., 2001; Geist et al., 2006; Turner et al., 2016; Verburg et al., 2016). Thus, a comprehensive exploration of land system modifications necessitates a dual focus: one that considers the earth's biophysical attributes, termed as “land cover,” and another that investigates human-driven land alterations or “land-use” (Lambin et al., 2001; Verburg et al., 2013; Verburg et al., 2016). An important factor in further understanding global change and the role of both human drivers and human interaction of natural systems is demonstrated via land use and land cover change (LULCC). Based on this recognition, and in view of the role of land in providing goods and environmental services, attention to LULCC has sharply risen (Hibbard et al., 2010). Thereby, LSS strives to decode the temporal shifts in land-use and land cover within the framework of a human-environment nexus, employing an array of methodologies including, but not limited to, remote sensing and supplementary data acquisition techniques (Turner et al., 2021). The advent of sophisticated sensor technology and the proliferation of “big data” have propelled AI's predominance in this realm, with a significant tilt towards remote sensing. This inclination stems from remote sensing's robust capabilities in monitoring terrestrial

attributes and discerning pertinent alterations (Rindfuss et al., 2004; Turner et al., 2007; Turner et al., 2021).

In recent years, LSS has faced increasingly complex challenges due to factors such as rapid environmental changes, urban sprawl, deforestation, and the varying impacts of human activities on different landscapes. These complexities necessitate a more intricate understanding and a nuanced approach to monitoring and managing land resources. Concurrently, the advent of “big data” in environmental sciences provides researchers with an unprecedented volume of data, coming from diverse sources such as satellite imagery, aerial photographs, and ground-based sensors (Gorelick et al., 2017; Yang et al., 2017; Tamimnia et al., 2020; Hermosilla et al., 2022). The big data of remote sensing is characterized by diverse spatial, spectral, and temporal resolutions, representing a dual-edged sword in the study of land system dynamics. On one hand, it provides unprecedented opportunities for gaining comprehensive insights into these dynamics. Yet on the other hand, this abundance and complexity of remote sensing big data poses significant challenges, necessitating the use of more advanced tools and approaches for effective analysis. Traditional methods, which may have sufficed in simpler data contexts, are now insufficient in handling the complexity of multi-dimensional data, requiring more intricate processing and interpretation (Liu, 2015). Where traditional classification methods fall short, data-driven AI methods can provide computational platforms and tools for analyzing voluminous and complex remote sensing datasets. However, the effective application of AI technologies for land classification tasks demands a thorough understanding and careful consideration of several underlying factors, underscoring the importance of this discourse in contemporary LSS research (Lambin et al., 2001; Rindfuss et al., 2004; Turner et al., 2007).

The process of land cover classification, crucial in the realm of remote sensing and LSS involves allocating specific land cover classes to individual pixels captured by Earth Observation technologies. This procedure can be executed using an array of techniques, each with unique methodological frameworks. Notably, remote sensing data analysis transcends mere spatial or geographic considerations, incorporating a crucial temporal element. This integration enables researchers to meticulously track and quantify changes in land cover over sequential time periods, enhancing the monitoring capabilities central to the research objectives of LSS. Classification techniques in this context diverge significantly, often categorized based on their training regimen—supervised or unsupervised—or the underlying theoretical framework—parametric or non-parametric. The landscape of classification algorithms is diverse, encompassing options like the Maximum Likelihood Classifier (MLC), an exemplar of supervised parametric algorithms, and k-means nearest neighbor (KNN) clustering, representative of unsupervised non-parametric approaches. Among these, unsupervised classification is prevalent, primarily due to its algorithmic ability to assimilate pixels into groups based on spectral signatures without reliance on preliminary training data. Conversely, supervised classification necessitates predefined training data, allowing the algorithm to discern various cover types dispersed throughout the landscape. The specificity of supervised methods varies considerably, with each tailored to different data types and research imperatives. These

methods are not without their respective advantages and limitations. The optimal method for a given study hinges on the intricacies of the data in question, including spatial and temporal resolutions, and the overarching research goals. In essence, the nuanced process of classification analysis is indispensable for the comprehensive understanding and effective management of land resources. Its relevance and application within LSS are not only of paramount importance at present but will persist as a critical component of future endeavors in the field. This continuity underscores the necessity for ongoing advancements and refinements in classification methodologies.

In the era of remote sensing (RS) big data, the multi-source and multi-scale characteristics of remotely sensed data has posed significant challenges in managing, processing, and interpreting this unprecedented proliferation of data, and it is often difficult to employ traditional processing algorithms and models (Liu, 2015). Fortunately, the field of classification has seen the introduction of more advanced techniques. The ease of access to large satellite imagery and spatial datasets, coupled with better computational resources and the complexity of high-dimensional RS data has popularized the use of Artificial Intelligence (AI) in land-use and land-cover (LULC) classification. This trend is evident with the growing adoption of machine learning (ML) and deep learning (DL) methods for classifying remotely sensed imagery, opening new research avenues, and allowing for the utilization of large, diverse datasets (Chen et al., 2017; Zhang et al., 2019; Jin and Mountrakis, 2022). Given that LSS is an interdisciplinary field, it integrates knowledge from various disciplines such as geography, ecology, remote sensing, geology, economics, and sociology to analyze and model the processes shaping land systems at different scales. One of the key challenges in LSS is dealing with the heterogeneity and complexity of land cover features, especially in areas where irregular and sparse spatial distribution of features exists. ML and DL techniques can aid in the classification of complex land cover features where traditional methods may struggle to accurately capture the spatial distribution of such land cover classes, such as the novel DL technique (EG-UNet) used by Zhou et al. (2023b) to enhance the classification of open-pit mining areas. In their analysis their model outperformed ten different DL models, performing especially well in classifying classes with few training samples. Agricultural land suitability classification and crop suggestion represent another critical aspect of LSS, especially in semi-arid ecosystems where agricultural productivity is highly dependent on environmental factors. Previous research demonstrates the application of ML and spatial multicriteria decision analysis to assess the suitability of agricultural land and provide recommendations for crop selection, thereby optimizing land use practices and enhancing agricultural productivity (Agrawal et al., 2024). Furthermore, LSS often involves the classification and characterization of lithological units based on geological knowledge and satellite imagery analysis. A study by Zhou et al. (2023a) proposes a DL framework guided by geological knowledge to classify lithological units from optical stereo mapping satellite imagery, facilitating improved geological mapping and resource exploration. ML-based classification algorithms play a crucial role in monitoring land use and land cover practices, particularly in challenging terrains such as hilly areas where traditional methods may be limited in their effectiveness. The utilization of ML

algorithms for land use and land cover monitoring in hilly terrains is explored in various studies, including a recent study by Parashar et al. (2024), which highlights the importance of these techniques in accurately detecting and classifying land cover changes in dynamic landscapes. The intersection of various disciplines and the application of advanced techniques such as DL and ML contribute to the effectiveness of LSS in addressing complex environmental and societal challenges.

However, the integration of AI as a “third approach” in remote sensing, sitting between traditional observation and simulation methods, requires thorough evaluation. As AI techniques have become pervasive in almost all RS applications, it is crucial to assess whether a given method improves accuracy, aligns with research goals, and for what applications the use of a particular AI technique is most appropriate. The key is to identify the specific conditions for which the application of AI methods in remote sensing contribute effectively and enhance methodology robustness, and in which cases such methods may detract from the research goals and diminish methodological robustness. It is important to recognize that while AI has potential in remote sensing applications, it is not a panacea. There are situations in which AI methods are ineffective or even detrimental, yielding weaker results compared to more conventional methods. Similarly, there are research scenarios for which more advanced DL methods may yield less accurate results than traditional ML techniques (Liu et al., 2017; Abdi, 2020; Jamali et al., 2021; Zaabar et al., 2022). Identifying these scenarios is essential for ensuring that AI is integrated wisely within the existing methodological framework in LSS. A more nuanced understanding of the conditions under which an AI technique is appropriate and advantageous enables researchers to better incorporate these new technologies with existing methods for optimal outcomes. Such an understanding is critical for guiding the responsible use of AI in remote sensing, ensuring it enhances rather than hinders research quality, and advances LSS's aim to better understand human-driven land system change.

This study aims to critically evaluate the performance and effectiveness of ML and DL techniques in executing scientific classification tasks in the domain of LSS. Our focus is to dissect the capabilities and constraints of both ML and DL in precisely classifying land cover features, discerning land use patterns, and forecasting dynamics within land systems. The scope encompasses an examination of a variety of scientific classification tasks such as land cover and land use classification, along with the detection of changes in land systems. We present examples from diverse types of remotely sensed data, including satellite imagery, LiDAR, and aerial photographs, supplemented by ancillary data like geological and soil maps, to enhance classification accuracy. The exploration will cover conventional ML algorithms like Random Forest (RF), Support Vector Machines (SVM), and k-Nearest Neighbors, alongside DL architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), U-Net, and Long Short-Term Memory (LSTM) networks. To assess the models' performance, evaluation metrics including accuracy, precision, recall, F1-score, and confusion matrices will be discussed. The review will feature case studies from various land system contexts—urban, agricultural, forested, and natural reserves—to test the models' generalizability and robustness. A thorough

comparative analysis will delve into the strengths and weaknesses of ML and DL techniques, considering factors like classification accuracy, computational efficiency, model complexity, and interpretability. The conclusion will synthesize these findings to offer insights and recommendations on the most suitable ML and DL approaches for land system classification tasks, aiming to advance land cover classification methodologies and guide researchers and practitioners in their methodological choices for land cover change and system analysis.

In the context of land system science, ML and DL serve as advanced techniques for classifying land cover and monitoring changes using remote sensing data. ML encompasses a variety of algorithms, such as decision trees (DT), RF, and SVM, that can identify patterns within large-scale land cover data (Kotaridis and Lazaridou, 2023). In LSS, ML methods have been instrumental in classifying land cover, predicting land use changes, and assessing environmental impacts. Key ML methods include Sparse Coding, Low-Rank Learning, Manifold learning, and the use of Hybrids. Sparse Coding focuses on finding a sparse representation of input data among an overcomplete basis set. Sparse coding is particularly useful in LSS for compressing large datasets and identifying significant features for land cover classification. Low-Rank Learning aims to discover a low-dimensional structure in high-dimensional data, and as such, this method is beneficial for processing and analyzing satellite imagery, where it can help in denoising and completing missing information in temporal sequences. Manifold Learning seeks to determine a lower-dimensional space in which the data points lie. It is adept at capturing the intrinsic geometry of spatial data, making it valuable for analyzing geographical patterns and trends in LSS. Hybrids result from the combining of two or more ML techniques, to leverage the strengths of each approach, and have been used in LSS to improve prediction accuracy and robustness over single-method approaches. In comparison are DL methods, which are a more complex subset of ML, involving neural networks with many layers. These methods have significantly advanced the field of LSS by enabling more sophisticated image classification, object detection, and semantic segmentation tasks. DL's ability to automatically learn hierarchical features from data makes it especially powerful for analyzing complex land system dynamics. DL is commonly used in fine-scale mapping and object-based image analysis. This approach includes techniques like CNN's, fully convolutional networks (FCN), and deep neural networks (DNN) (Xie and Niculescu, 2021; Zaabar et al., 2022). Both ML and DL operate on the principle of training models with labeled data to recognize patterns corresponding to various land covers. However, they differ significantly in complexity and application. Traditional ML algorithms, which are generally "shallower", work with fewer layers and parameters, making them simpler but sometimes less versatile in handling intricate data patterns. DL models, in contrast, generally feature many layers, allowing them to capture more complex, high-level abstractions in data, often achieving greater accuracy, particularly in image recognition tasks (Solórzano et al., 2021; Taye, 2023).

ML stands out for its automatic feature extraction, flexibility, and scalability, though it requires substantial training data and is prone to overfitting. While supervised ML offers more accuracy, unsupervised methods are useful when labeled data is scarce.

Conversely, DL's strength lies in its ability to learn directly from data, eliminating the need for manual feature engineering and efficiently handling high-dimensional data. However, DL models typically demand more computational power, larger datasets for training, and are more susceptible to overfitting, which can limit their practicality in certain scenarios (Sarker, 2021; Cheng et al., 2023). In practical terms, ML algorithms like RF are often employed for tasks such as large-scale land cover mapping due to their effectiveness and lesser computational demands. DL methods, with CNNs as a prime example, are preferable for detailed image classification tasks due to their proficiency in learning local patterns within data, making them ideal for intricate image-based land cover classification (Fujiyoshi et al., 2019). RNNs, another DL methodology, are suited for analyzing time series data, useful in monitoring temporal changes in land cover (Campos-Taberner et al., 2020; 2023; Masolele et al., 2021). However, their training process can be more complex, and they may need smaller datasets compared to CNNs. The choice between using ML or DL should factor in the specific project requirements, data availability, and computational resources, ensuring the selected method aligns with the research objectives.

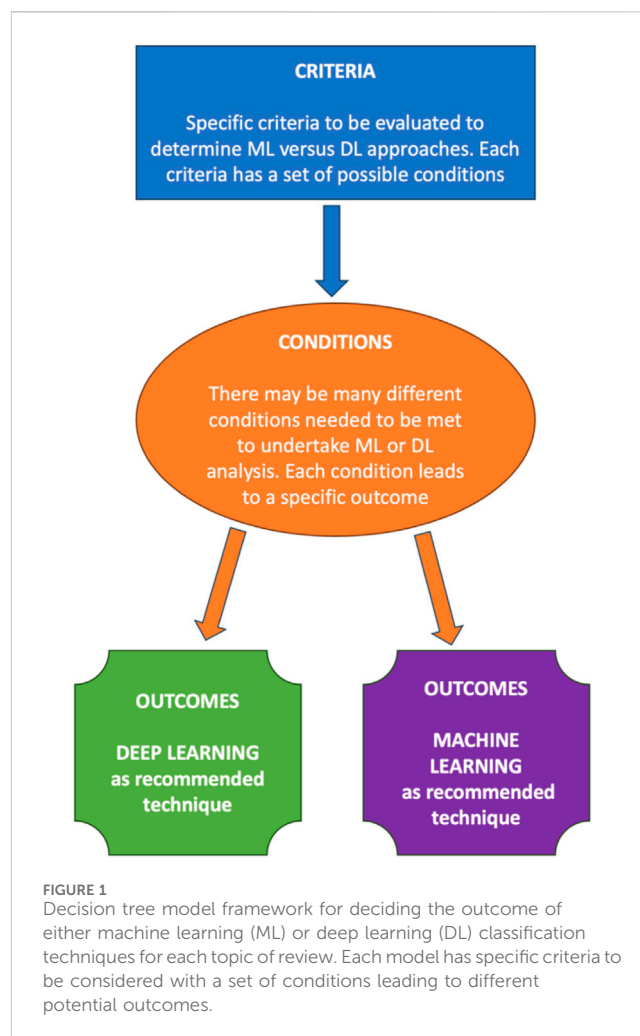
In this paper, we review the current status of ML and DL approaches for LSS research, with a specific focus on the application of AI techniques in land cover classification and change analyses. As providing a detailed description of every method is beyond the scope of this paper, we focus on representative characteristics of the ML and DL models most commonly used. We highlight these common or typical approaches with the use of significant research examples and featuring specific studies. As such, most of the limitations, advantages, disadvantages, and other considerations of each of the ML and DL approaches discussed refer to those models currently in common use. We then develop a decision tree approach to help guide researchers in selecting the more suitable of the two methods (ML or DL) for distinct research objectives and questions; datasets of varied spatial and temporal resolution; and different skill levels and computing power available to a researcher or research team. Given the somewhat recent inundation of remote sensing analyses with new or more accessible AI methods and the inclination of many researchers to employ ML/DL approaches, the need to pause and evaluate the usefulness of these advanced techniques is opportune. Although AI conjures an impression of "new and improved" or a higher degree of learning and sophistication, the assumption of new methods irrefutably being better equates to the age-old logical fallacy called the "*ad novitalem argument*." The excitement that often accompanies innovation and the emergence of new technology often obscures the practical application of such new and exciting tools. The logical approach in developing research methods may be obscured as researchers are focused on the latest available technologies. Indeed, trends in recent publications evidence that many RS researchers appear to be pushing to utilize ML and DL techniques in a race to be on the cusp of the latest technique while discarding more traditional- and possibly in some scenarios more appropriate-methods (Liu et al., 2017; Abdi, 2020; Jamali, 2021; Zaabar et al., 2022). Thus, an evaluation of the usefulness of AI methods in RS and considerations of when the different approaches should be utilized, is very timely. Given that the determination of which

technique to use is very much a function of the questions being asked, this paper examines the use of ML *versus* DL approaches within the field of LSS only. As many of the factors being considered in this review are already known, it is possible to develop a decision tree model to highlight which of the most commonly used ML/DL classification approaches is most suitable for a given study. In addition, in the discussion section, we will also revisit newer methods and approaches now being developed, to overcome some of these discussed ML and DL limitations or decision points, in order to highlight some of the possible future directions, where models are currently being developed to overcome some of the discussed limitations or constraints.

The primary objective of this paper is to articulate and provide a structured decision-making framework, specifically a decision tree, to guide researchers in selecting an appropriate computational method for land cover classification analysis within the field of LSS. The decision framework is presented as a decision tree, which serves as a pragmatic tool to aid researchers in navigating the complexities of method selection, given the distinct nature of frequently employed ML and DL techniques. The need for such a tool has become increasingly evident with the diversification of data sources and the exponential growth in data volume, coupled with the sophisticated demands of contemporary human-environment research. This decision tool is fundamental and must consider several factors, each of which may significantly influence the integrity of the research outcome. These considerations are as follows: 1) the researcher's familiarity with AI techniques and computing expertise, 2) the computing requirements and available resources, 3) what type of research question within LSS the researcher is addressing, 4) the types of image data available, 5) the time period the study is covering which also links back to image data availability, 6) what training data is available for the classification analysis, and 7) the required accuracy level. Determining whether ML or DL is more appropriate for a land classification task is contingent on these multifaceted considerations. Researchers must reflect on their objectives, resources, and constraints before selecting a methodology, ensuring it not only addresses their central question effectively but also aligns with the logistical parameters of their study. This balanced decision-making process paves the way for more robust, efficient, and suitable research outcomes in LSS.

2 Evaluating ML and DL for LSS applications

In the realm of LSS, understanding and modeling LULC transformations are paramount, significantly benefiting from advancements in remote sensing bolstered by ML and DL techniques. These advanced algorithms have revolutionized LULC tasks, extending from classification to intricate predictive modeling, primarily when employed with satellite or aerial imagery. The research landscape is replete with instances where ML and DL have been instrumental in enhancing LULC classifications (Singh et al., 2021; C. Zhang et al., 2020a; X. Zhang et al., 2020b), characterizations (Ahmed and Lin, 2021; Akar and Tunc Gormus, 2022), and comprehensive mapping (Alhassan et al., 2020). Notably,



Zhang et al. (2019) leveraged CNNs in their LULC study to bridge the gaps between predictive outputs and expected results. In a similar vein, Alhassan et al. (2020) utilized deep neural networks for nuanced LULC mapping, categorizing individual pixels of multispectral satellite imagery into diverse LULC types. These advancements, however, underscore the importance of selecting the appropriate technical approach with research objectives and resources, as this choice can significantly influence a project's trajectory and outcome. To address this, our review delineates the competencies essential for employing popular ML and DL methodologies in LSS research. A pivotal component of this analysis is a decision-tree-style flowchart (Figure 1), synthesized from comprehensive reviews and direct research insights. This guide serves as a navigational tool, underscoring the skills, data prerequisites, and specific criteria that dictate the choice between commonly employed ML and DL pathways. Our objective extends beyond mere comparison, constructing a fully developed decision model as the culmination of this review. This model encapsulates the critical facets pertinent to LSS, offering researchers a clear direction based on their unique constraints, resources, and expertise levels. Consequently, it will enable informed decision-making regarding the adoption of common ML or DL techniques in individual LSS endeavors.

2.1 Data type and system studied

2.1.1 Research question within LSS

LULC classification has emerged as an essential application within LSS, particularly in the interpretation of remote sensing imagery. The era of ‘big data’ has ushered in a deluge of Earth Observation data, presenting researchers with new avenues for groundbreaking methodologies and approaches (Karpatne et al., 2016). The intricacy lies in aligning the AI technique with the specific LSS research question at hand, thereby selecting the most appropriate approach within the vast spectrum of AI methodologies.

ML, employing a bottom-up approach, excels in deciphering patterns within massive data sets—patterns often imperceptible to human analysis. Conversely, traditional earth science modeling, characterized by a top-down approach, leverages established understanding of physical laws, often at the expense of computational simplicity and the agility needed to uncover unexpected insights. A synergistic integration of these methodologies has been the focus of contemporary research, striving to embed physical laws within the flexible architectures of ML and DL models. This fusion facilitates more interpretable models, capable of addressing the multifaceted challenges faced in the earth sciences (Karpatne et al., 2016). In LSS, research objectives vary widely, influencing the data requirements, study area parameters, and even the pixel characteristics within remote sensing images. Consequently, the research design and methodologies in LULCC studies must be meticulously customized. Notably, DL techniques, especially CNNs, have enhanced capabilities for object detection, segmentation, and traditional classification, empowering detailed image analyses (Barbierto et al., 2020; Du et al., 2020; Guo et al., 2021; Pešek et al., 2022). Furthermore, DL autonomously learns data features, a marked advancement over traditional ML, eliminating the need for manual, domain-specific feature engineering (Liu et al., 2017; Shih et al., 2019; Magalhães et al., 2022).

With the advent of commercial satellite imagery and advancements in computer vision, DL techniques have revolutionized the monitoring of land cover and land-use changes. Harnessing DL, researchers can now distill invaluable insights from decades-long satellite data repositories, providing a dynamic view of environmental changes and aiding in crucial decision-making processes. One noteworthy development is ML’s role in monitoring temporal processes, such as land cover transitions, with significant implications for policy decisions. For instance, insights derived from ML analyses can guide investments in agricultural water management or initiatives to rehabilitate deteriorated landscapes (Yomo et al., 2023).

In LSS, multi-spectral and hyperspectral imagery stand out as the predominant data types in ML and DL image classification. Multi-spectral imagery, encompassing specific wavelength bands, is instrumental in distinguishing various land cover types (Chen et al., 2017). In contrast, hyperspectral imagery offers a higher spectral resolution, enhancing the identification precision for specific cover types—essential for detailed applications like mineralogy, vegetation studies, and pollution monitoring (Ghamisi et al., 2016; Yu et al., 2017; Li et al., 2019; Lei et al., 2021). Additionally, panchromatic imagery, with its higher spatial resolution, finds its niche in

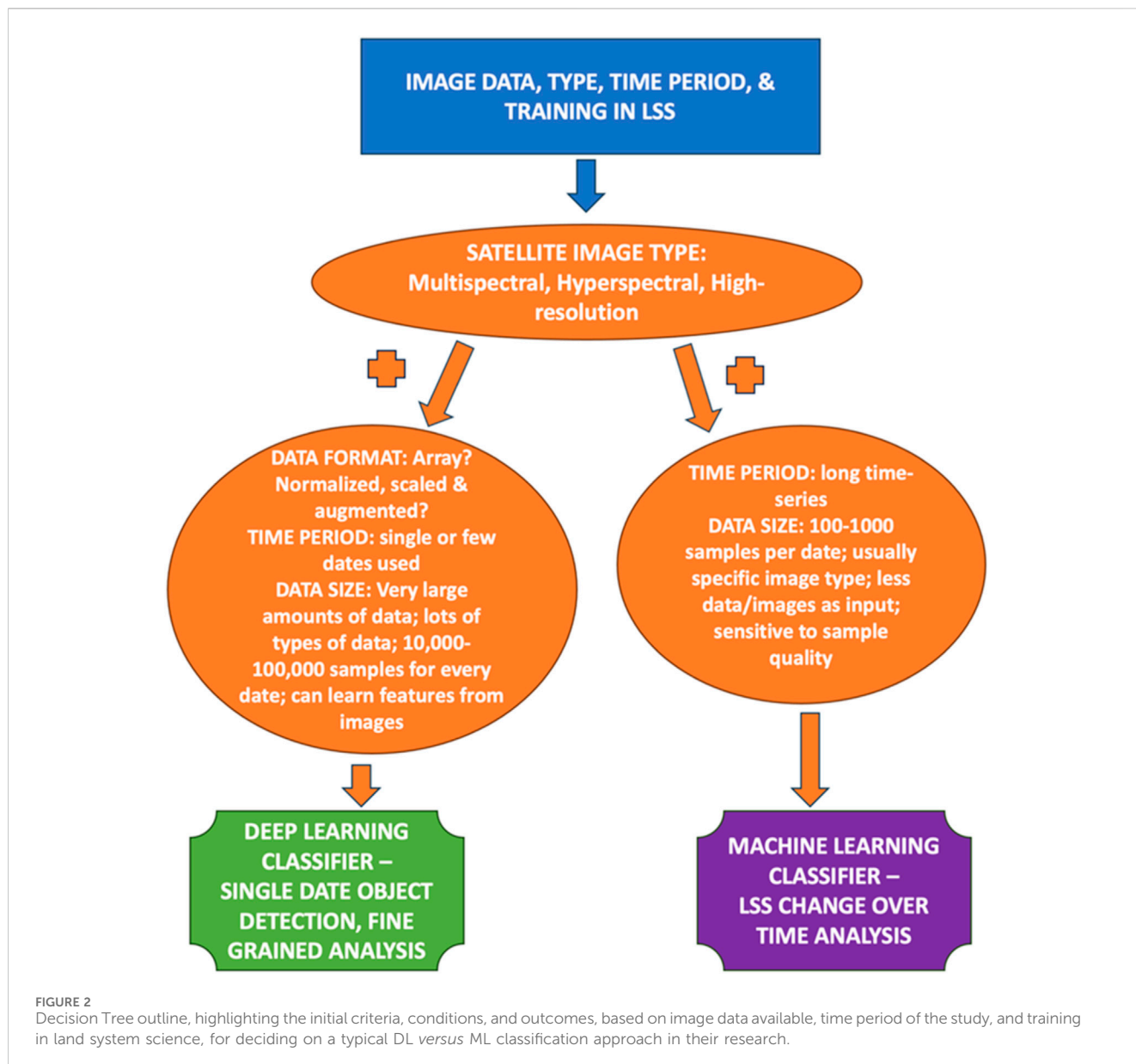
analyzing feature textures and shapes, critical for detailed object recognition tasks (Zhang X. et al., 2020; Sertel et al., 2022; Zhang et al., 2022). Similar to ML, DL models in LSS commonly utilize multi-spectral, hyperspectral, and panchromatic imagery, benefiting from their capacity to learn directly from data. These models, particularly adept with CNNs, have proven instrumental for a plethora of tasks, including land cover mapping, crop identification, and change detection, capitalizing on the rich information provided by these advanced imaging methods (Barbierato et al., 2020; Du et al., 2020; Guo et al., 2021; Wang J. et al., 2022; Wang L. et al., 2022; Pešek et al., 2022).

The judicious application of AI in LSS requires a nuanced understanding of the specific research questions, the characteristics of the available data, and the computational tools at one’s disposal. The convergence of ML and DL with traditional scientific methods is paving the way for innovative solutions in LSS, transcending the limitations of each approach when used in isolation. As we venture further into this era of big data and AI, the strategic integration of these technologies in LSS will continue to play a pivotal role in shaping our understanding and stewardship of the Earth’s land systems.

2.1.2 Image availability

Both ML and DL approaches used in LULC classification fundamentally depend on high-quality image data, but they differ significantly in their specific data requirements and preprocessing steps. ML algorithms require preprocessed, accurately labeled image data, where each pixel or segment is clearly categorized into different land cover types. These algorithms can efficiently process various data formats, such as raster (e.g., GeoTIFF) and vector formats (e.g., shapefiles), but their effectiveness heavily relies on the quality and representativeness of the training data, emphasizing the importance of careful data curation (Hermosilla et al., 2022).

In contrast, DL algorithms, though fundamentally reliant on similar types of data as ML, are more forgiving with respect to data quality and representativeness. They demand more extensive preprocessing, which often involves transforming image data into array formats suitable for feeding into neural networks. This difference in data handling underscores DL’s adaptability in managing a wider range of data qualities and volumes for LULC classification. Rigorous data transformations, including normalization, scaling, and sometimes augmentation, are indispensable for optimizing model performance. The advent of RNNs equipped with Attention Mechanisms (AM) marks a significant evolution in the DL landscape. As Campos-Taberner et al. (2020) illustrate, these advanced networks have gained traction beyond traditional applications, demonstrating proficiency in enhancing classifications derived from high-resolution imagery, notably in time series predictions and text classifications. A distinguishing characteristic of DL methodologies, particularly in the context of land cover classification, is their voracious demand for extensively labeled datasets. Given the intricate architectures of DL models, the training phase mandates substantial data volumes to effectively ‘learn’ and generalize. This requirement often compels researchers to employ strategies like data augmentation, expanding the effective size of available datasets. These techniques are crucial, especially in scenarios where access to vast repositories of labeled images is limited.



While both ML and DL paradigms mandate access to labeled image data, the DL algorithms distinguish themselves through their demand for larger datasets and specific data formatting prerequisites. Moreover, they exhibit a resilience against fluctuations in data quality and representativeness, unlike their ML counterparts. This relative robustness, alongside the advanced capabilities of models like RNNs with AM, highlights the compelling advantages of DL in tackling the complexities of land cover classification within the burgeoning domain of LSS.

2.1.3 Time period of study

The evolution of Earth Observation (EO) technologies has precipitated a paradigm shift in the temporal resolution of remotely sensed images utilized in computational modeling, enhancing the depth and frequency of time-series data for land cover studies (Southworth and Muir, 2021). This progression diverges markedly from historical methodologies that relied on

sparse two or three-date change detection analyses. Contemporary approaches enable researchers to dissect land cover transformations across extended periods, encompassing diverse seasons and years. Such temporal breadth enhances the comprehension of terrestrial dynamics integral to discerning land cover traits. However, the wealth of time-related data also brings its own set of challenges. The inundation of data, while contributing to accuracy, introduces ancillary complications like extraneous noise in predictive modeling and computational burdens (Davenport et al., 2012; Chen et al., 2017; Cheng et al., 2020; Arruda et al., 2021). These challenges underscore the necessity for advanced analytical tools and robust computing infrastructures.

Recent innovations, such as the integration of red-edge spectral bands in Sentinel-2 sensors, have revolutionized accuracy levels in land cover classification, independent of denser time-series data (Delegido et al., 2013; 2011a; Xie et al., 2018; Abdi, 2020). These technological leaps signify that enhanced sensor capabilities, even in

TABLE 1 Published research analyses that utilized machine learning and deep learning classification approaches for land cover, with specific details regarding the various methodological approaches, characteristics of the input imagery, and study conclusions and limitations.

Citation	Study purpose	Classification approach: Machine learning (ML) or deep learning (DL)	Imagery type	Time period	Conclusions and limitations [when available]
Barbierato et al. (2020)	Classification of urban forests from proximate sensing techniques	DL- CNN (Semantic Segmentation Network)	MS, LiDAR UltraCam XP (four band red, green, and blue + NIR); Google Street view images	Single date, Oct. 2013	- NDVI and LiDAR data were combined and used.
					- With respect to the classification of vegetated areas, the results of the study confirm the efficacy of the combined use of NDVI and LiDAR and semantic segmentation using DL for Google Street View (GSV)images
Du et al. (2020)	Map forested wetland inundation	DL- CNN U-net Semantic Segmentation	Worldview- 3	2015	- The model accurately determined inundation status with 95% accuracy (Kappa = 0.90) and high overlap (IoU = 70%) with field data and lidar intensity-derived labels.
Campos-Taberner et al. (2023)	Decision making on the European Common Agricultural Policy	DL- RNN	Sentinel-2	May 2019 and April 2020.	- Ground truth data captured the spatial variability of land use classes.
			bands + kNDVI + EkNDVI		
Shih et al. (2019)	Evaluate classifiers in various software	ML- RF, DT, SVM	Landsat TM and ETM+	November 1, 28 May 2010, 2011, and	- Model selection is a key factor for classification. For instance, the accuracy of LULC maps derived from ML classification is lower for the FNNR dataset due to limited training data and challenging terrain and illumination conditions.
		DL- ANN		16 August 2011.	- The limitation might be the use of manually digitized training data for six LULC classes from 2000 to 2003 Google VHR images and the original Landsat image, which overestimated the overall accuracy of trained classifiers.
Maxwell et al. (2018)	Evaluate model selection, input data and parameters, computational costs	ML- SVM, DT, RF, boosted DT, k-NN.	AVIRIS (hyperspectral), GEOBIA	Single date images from multiple years (1992, 2014)	- ML techniques are more robust than parametric methods for remote sensing analysis, but ideal classifiers, parameters and sample size may differ based on the analyst's goal.
		DL- ANN			
Abdi (2020)	Comparing classification performance of four nonparametric algorithms	ML- SVM, RF, Xgboost	Sentinel-2 multispectral	2017 and 2018	- DL underperformed, with lowest accuracy for open land and artificial surfaces (PA = 0.44 and 0.36). The use of tanh activation likely caused the underperformance.
					- Support vector machines produced the highest classification accuracy due to their small number of complex decision boundaries.
					- The limitation could be the training data, which might not be representative over large areas.

(Continued on following page)

TABLE 1 (Continued) Published research analyses that utilized machine learning and deep learning classification approaches for land cover, with specific details regarding the various methodological approaches, characteristics of the input imagery, and study conclusions and limitations.

Citation	Study purpose	Classification approach: Machine learning (ML) or deep learning (DL)	Imagery type	Time period	Conclusions and limitations [when available]
Pešek et al. (2022)	Cloud detection in desert landscapes	DL- CNN: U-Net, SegNet	VEN μ S	Nine tiles obtained between Apr. 2018 - Aug. 2019	<ul style="list-style-type: none"> - The CNN architectures outperform current cloud-masking algorithms by improving the false positive detection ratio.
		DeepLab			
Guo et al. (2021)	Mapping mangrove land cover from multi-scale contextual information	DL- CNN, U-net	Landsat time series		<ul style="list-style-type: none"> - Developing Capsule-Unet which introduce the capsule concept to U-net to extract mangroves with high accuracy via learning the spatial links among the objects in the remote sensing imagery.
Masolele et al. (2021)	Mapping land-use change driven by deforestation	DL- 2D and 3D CNN, RNN LSTM, Hybrid models	Landsat 5 & 7	Median composites for two annual values/year (2000–2005)	<ul style="list-style-type: none"> - In the pan-tropical setting, Hybrid CNN-LSTM, ConvLSTM, and CNN-MHSA models outperformed 2D-CNN, LSTM, and 3D-CNN in terms of classification accuracy.
Campos-Taberner et al. (2020)	Land use classification using RNN	DL- Recurrent Neural Network	Sentinel-2 time series	Sep. 2017–Aug.2018	<ul style="list-style-type: none"> - Sentinel-2 bands convey the most useful information.
					<ul style="list-style-type: none"> - Sentinel-2's near-infrared and red bands are particularly valuable for vegetation categorization due to their ability to detect differences in leaf area index and pigmentation
					<ul style="list-style-type: none"> - LSTM networks overcome this with a mechanism that allows for memory, deciding what information to retain or discard.
					<ul style="list-style-type: none"> - The 2-BiLSTM network achieved an overall accuracy of 98.7%, surpassing other classification algorithms.
					<ul style="list-style-type: none"> - Accuracy rates were consistently high, reaching up to 99.9% for rice crops.
Alshari et al. (2023)	Implementing the novel technique for land use classification	ML- Artificial Neural Network	Sentinel-2A, Landsat-8, and Normalized digital elevation	2016	<ul style="list-style-type: none"> - They did geometric and radiometric corrections, which improved categorization and corrected degraded images.
					<ul style="list-style-type: none"> - The combinations of classifiers showed good results.
X. Zhang et al. (2020b)	Comparing state-of-the-art DL for high-res image classification and object detection to traditional methods, using land cover and wind turbine detection case studies	DL - SegNet, U-Net, PSPNet	RapidEye: 5-m resolution imagery in five-bands: blue, green, red, red edge, and near infrared.	2012	<ul style="list-style-type: none"> - Images are too large to pass through a DCNN (deep convolutional neural network) under current GPU memory limitations.
					<ul style="list-style-type: none"> - Used sliding window to split the training data.
					<ul style="list-style-type: none"> - Label-preserving transformations were used for data augmentation to rotate, flip, and mirror.

(Continued on following page)

TABLE 1 (Continued) Published research analyses that utilized machine learning and deep learning classification approaches for land cover, with specific details regarding the various methodological approaches, characteristics of the input imagery, and study conclusions and limitations.

Citation	Study purpose	Classification approach: Machine learning (ML) or deep learning (DL)	Imagery type	Time period	Conclusions and limitations [when available]
					<ul style="list-style-type: none"> - DL models' convolution layers handle spatial-spectral information with a convolution kernel, boosting classification accuracy, especially in hyperspectral image classification.
Arruda et al. (2021)	Mapping spatial fire dynamics for ecosystem management	DL- DNN (Multi-Layer Perceptron Network Structure)	Landsat bands 2–7 MCD64A1, INPE	Individual model runs for May & December of 2017	<ul style="list-style-type: none"> - 97% accuracy; 36% agreement with other datasets (due to cloud cover)
Bhargava et al. (2021)	Estimate erosion and progradation of mangroves (case study)	ML -K means clustering algorithm	Global Mangrove Watch (GMW) Global Surface Water Explorer the GSWE dataset contains the spatial and temporal distribution of surface water and provides the extent of and change in water surfaces from	1984–2018	<ul style="list-style-type: none"> - The accuracy evaluation was greater than 90%. - The overall kappa coefficient was highest. - The GEE platform utilized cloud computing and big data methods, providing cost and time efficiency as the data and platform usage were free of charge. Additionally, it was fast, efficient, and capable of scaling to meet increasing demands.

isolation, can precipitate substantial gains in classification precision. In the context of employing DL and ML for image classification, the selection of appropriate satellite imagery is contingent on the specificity of the research objectives and the quantity of labeled data available. Typically, high-resolution imagery is coveted for its intricate detail, conducive to a broader spectrum of analytical tasks. Nonetheless, the robust nature of DL models permits the assimilation of expansive data volumes, rendering them compatible with lower-resolution images when high-definition resources are scarce. DL models demonstrate proficiency in extracting features directly from diverse image datasets, spanning multispectral, panchromatic, and hyperspectral imagery. This versatility, combined with the models' capacity to manage extensive data from vast geographical expanses, positions DL as a potent tool for satellite imagery analysis. Conversely, traditional ML models, while potent, often necessitate tailored feature engineering and may exhibit restrictions in data handling capabilities, particularly for extensive datasets typical in satellite observations. The crux of effective DL application, however, resides in the availability of substantial labeled datasets, a prerequisite that might be elusive in certain remote sensing ventures. Herein, conventional ML models may hold an advantage, albeit with the trade-offs of requiring feature engineering and potential limitations in data processing breadth (Figure 2).

Expanding the temporal domain of DL image classification introduces additional challenges. A key issue is the lack of labeled training data when using older images. Over time, changes in sensor technology, differences in image resolution, and varying image quality make it harder to analyze images

consistently over a long period. To deal with these problems, it is important to use a standardized method for preparing data and choose images from similar types of sensors or those with comparable features. Also, to use old data effectively, we need more labeled datasets that accurately show how land cover has changed over time. So, while DL is promising for long-term studies of land cover, it is crucial to carefully consider the limitations related to data labeling, changes in sensor technology, and differences in image quality over time (Alhassan et al., 2020; Xu et al., 2021; Wang J. et al., 2022). The authors have compiled an extensive list of recently published articles that utilize ML and DL for LSS and include some of the considerations detailed above (Table 1). This table serves as an integral data analysis component for this research, highlighting published research analyses that utilized ML and DL classification approaches for land cover, with specific details regarding the various methodological approaches, characteristics of the input imagery, and basic information regarding the study conclusions and limitations. Additionally, it may also serve as a resource for those wanting to expand their ML/DL analyses for LSS research (Table 1).

2.1.4 Available training data

Training data is the cornerstone for both ML and DL models used in image classification. However, the requirements differ significantly between the two, primarily in terms of data volume, quality, and representation. For ML, the emphasis is on the precise labeling and quality of data, even though the quantity required is relatively modest (Wang J. et al., 2022; Wang L. et al., 2022). The datasets must be meticulously curated to avoid misleading the

models, necessitating perhaps thousands of samples, each clearly annotated to represent different classes accurately. This process demands considerable time and expertise, especially in cleaning the data of noise and errors that can skew results. Conversely, DL models demand extensive datasets, often requiring massive training data. However, DL models are inherently designed to sift through data to detect patterns and correlations, allowing for some leniency regarding inconsistencies or anomalies within the datasets. Yet, the sheer volume of training data needed poses a challenge, as insufficient data leads to underdeveloped models. The representativeness of the data is also crucial, ensuring all possible variations of each class are included, preventing the model's performance from degrading from lack of diversity in the training phase. The stark requirement for large datasets in DL necessitates resources often beyond reach for many projects, leading to reliance on data augmentation methods or synthetic data to bolster training sets. This approach has its pitfalls, as synthetic data may not capture the complexity of real-world variations (Han et al., 2023).

Furthermore, the practical challenges of acquiring and processing different data types—as this forms the backbone of the training process—must also be addressed. Optical data from platforms like Landsat is commonplace, but often requires significant preprocessing to be usable in training models. Radar data, such as that from the Sentinel-1 mission, offers all-weather capabilities but brings complexities in interpreting and utilizing data for non-expert users. Moreover, while airborne and ground-based data collection methods provide higher resolution and detailed data, they come with logistical, financial, and technical challenges, restricting their use in broader studies. These methods, though highly detailed, are not always scalable or feasible for large-area studies due to cost, accessibility, and data processing constraints (Shirmard et al., 2022). Data fusion is another underexplored area that deserves attention. Integrating different data types (optical, radar, etc.) can significantly enhance model robustness, allowing more reliable performance across varying conditions and terrains (Zhu et al., 2017). However, this requires sophisticated processing techniques and expertise to ensure compatibility and meaningful synthesis of diverse data sources. The choice between ML and DL is not merely a technical decision but a strategic one, hinging on available resources, expertise, and project objectives (Figure 2). While DL holds promise with its ability to process and learn from vast datasets, its resource-intensive nature and need for substantial, diverse training data can limit its applicability. Conversely, ML can operate on leaner datasets but requires more rigorous data preparation and may not capture complex patterns within data as effectively as DL models.

2.1.5 Level of accuracy required

In LULCC research, particularly studies focusing on specifics like urban expansion, deforestation, or cropland conversion, DL techniques often surpass traditional ML methods in accuracy. These studies typically analyze vast quantities of remote sensing imagery, requiring precise detection of often subtle changes in land cover or use. DL methods, especially CNNs, excel in identifying these changes due to their superior object and pattern recognition capabilities (Kussul et al., 2017; Yu et al., 2017; Zhang et al.,

2018). ML, despite requiring fewer training data, serves well in applications like data mining and complex software applications (Lary et al., 2016). In contrast, DL, with its multilayered NNs, deals with high-level data abstractions. Modern CNNs, with their elaborate architecture, are particularly adept at handling diverse data arrays, such as those in image data (Fu and Menzies, 2017). However, these networks require extensive training datasets, significantly larger than those needed for ML, making data augmentation techniques a necessity, albeit not a panacea, especially for more complex or larger models.

The emphasis in using CNNs for remote sensing imagery is often on achieving high classification accuracy, sometimes at the expense of training time efficiency. Advanced computing resources and methodologies, including GPUs, can expedite model training and testing. However, pitfalls like over-fitting due to insufficient training data remain, potentially harming the model's performance on test data. Techniques like dropout can mitigate over-fitting by randomly disabling certain nodes, enhancing overall network performance (Gupta and Nanda, 2022). DL's data-hungry nature aligns it with big data, demanding not only vast amounts of varied data but also counter examples for training (Raj et al., 2019). Data is processed through sophisticated, pretrained CNNs, producing high-level features for subsequent categorization—a beneficial approach in RS, where acquiring labeled data is costly and labor-intensive (Marmanis et al., 2016). Despite their potential, DL models face limitations due to the scarcity of training data, with the large-scale nature of the datasets introducing new analytical challenges. Discriminating between classes becomes difficult with large spatial coverages, as internal variability within classes increases. This issue is compounded when using moderate or coarse-resolution images, where field data cannot comprehensively represent a pixel's entire area, leading to potential scaling problems (Elmes et al., 2020).

The extent of accuracy improvement DL offers over ML varies by the task and dataset. However, DL, particularly through CNNs, generally achieves superior results (Table 2). For instance, CNNs outperformed traditional ML in land cover classification tasks, reaching 94.6% accuracy compared to 88.7% with ML (RF) in satellite imagery (Kussul et al., 2017), and 85.24% versus 80.95% (SVM) in aerial imagery (Yu et al., 2017). Yet, higher accuracy with DL is not guaranteed. From 29 studies comparing ML and DL in LULC mapping, DL offered less than 10% improvement in 17 cases, and in 10 instances, the increase was under 4%. Surprisingly, ML even surpassed DL in four studies (Liu et al., 2017; Abdi, 2020; Jamali, 2021; Zaabar et al., 2022). Thus, the choice between ML and DL extends beyond just accuracy, encompassing considerations like computational resources and model interpretability. While DL aims to enhance LULC classification accuracy, various factors determine the choice of approach, and accuracy is just one of them. Table 2 compiles the model accuracy results from studies employing both ML and DL for LULC mapping. Results indicate that while DL can lead to accuracy improvements, these are not always substantial (Ienco et al., 2019; Interdonato et al., 2019; Xie and Niculescu, 2021; Magalhães et al., 2022). Performance varied widely, with accuracy ranging from 51.56% to 99.59% for ML and 61.45%–99.79% for DL, highlighting the impact of factors beyond the classification algorithm, such as training data and imagery type.

TABLE 2 Reported accuracies and potential accuracy improvement by the use of deep learning (DL) classification approaches from 29 published studies where both traditional machine learning (ML) methods and DL methods were used for land cover classification.

Citation/Title	Accuracy			Modeling method		Imagery type
	ML	DL	^a Change in DL over ML (%)	ML	DL	
Magalhães et al. (2022) Comparing machine and DL methods for the phenology-based classification of land cover types in the amazon biome using Sentinel-1 time series	92.16	93.49	1.44	SVM	Bi-GRU	Sentinel-1
Wang et al. (2022b) Evaluation of a deep-learning model for multispectral remote Sensing of land use and crop classification	73.47	97.28	32.40	RF	UNet++	Sentinel-2A
	79.04	97.04	22.77	RF	UNet++	Sentinel-2A
	80.23	97.72	21.79	RF	UNet++	Sentinel-2A
	72.69	96.86	33.25	RF	UNet++	Sentinel-2A
	70.96	75.34	6.17	RF	UNet++	Gaofen PMS
	74.72	97.72	30.78	RF	UNet++	Gaofen PMS
	80.75	87.66	8.55	RF	UNet++	Gaofen PMS
	70.65	91.17	29.04	RF	UNet++	Gaofen PMS
Zaabar et al. (2022) Application of convolutional neural networks with object-based image analysis for land cover and land use mapping in coastal areas: a case study in Ain Témouchent, Algeria	91.8	93.5	1.85	RF-OBIA	OB-CNN	Pléiades VHSR
	91.0	83.4	-8.35	RF-OBIA	OB-CNN	Sentinel-2A
Jamali (2021) Land use land cover modeling using optimized ML classifiers: a case study of Shiraz, Iran	93.55	91.02	-2.70	SVM	DNN	Landsat-5
	98.37	95.29	-3.13	SVM	DNN	Landsat-8
Rousset et al. (2021) Assessment of DL techniques for land use land cover classification in Southern New Caledonia	77.55	81.41	4.97	XGboost	DeepLab	SPOT-6
	51.56	61.45	19.18	XGboost	DeepLab	SPOT-6
Solórzano et al. (2021) Land use land cover classification with U-Net: advantages of combining Sentinel-1 and Sentinel-2 imagery	0.53	0.76	43.39	RF	U-Net	Sentinel-1 and Sentinel-2
Xie and Niculescu (2021) Mapping and monitoring of land cover/land use (LCLU) changes in the Crozon Peninsula (Brittany, France) from 2007 to 2018 by ML algorithms (support vector machine, random forest, and convolutional neural network) and by post-classification comparison (PCC)	80.23	83.11	3.58	RF	CNN	SPOT-5
	78.14	79.85	2.18	SVM	CNN	Sentinel-2B
Abdi (2020) Land cover and land use classification performance of ML algorithms in a boreal landscape using Sentinel-2 data	0.75	0.73	-2.66	SVM	multilayered feed-forward deep neural network	Sentinel-2
Campos-Taberner et al. (2020) Understanding DL in land use classification based on Sentinel-2 time series	94.9	98.7	4.00	RF	2-BiLSTM	Sentinel-2
Ienco et al. (2019) Combining Sentinel-1 and Sentinel-2 satellite image time series for land cover mapping via a multi-source DL architecture	88.27	89.88	1.82	RF	TWINNS	Sentinel-1 and Sentinel-2
	86.00	87.50	1.74	RF	TWINNS	Sentinel-1 and Sentinel-2
Interdonato et al. (2019) DuPLO: A DUAL view Point DL architecture for time series classification	82.99	83.72	0.87	RF	DuPLO	Sentinel-2
	96.04	96.36	0.33	RF	DuPLO	Sentinel-2
Li et al. (2019) DL for hyperspectral image classification: an overview	76.88	84.56	9.98	SVM	DFFN	Airborne sensor
	90.78	98.57	8.58	SVM	DFFN	ROSIS-03
	90.92	99.71	9.66	SVM	DFFN	AVIRIS
Wan et al. (2019) Multiscale dynamic graph convolutional network for hyperspectral image classification	88.34	93.47	5.80	JSDF	MDGCN	AVIRIS

(Continued on following page)

TABLE 2 (Continued) Reported accuracies and potential accuracy improvement by the use of deep learning (DL) classification approaches from 29 published studies where both traditional machine learning (ML) methods and DL methods were used for land cover classification.

Citation/Title	Accuracy			Modeling method		Imagery type
	ML	DL	^a Change in DL over ML (%)	ML	DL	
	90.82	95.68	5.35	JSDF	MDGCN	RODIS
	95.69	99.79	4.28	JSDF	MDGCN	AVIRIS
Zhang et al. (2019) Joint DL for land cover and land use classification	82.38	89.64	8.81	SVM	Joint DL	Vexcel UltraCam Xp
	80.26	87.58	9.12	OBIA-SVM	Joint DL	Vexcel UltraCam Xp
	82.26	90.72	10.28	SVM	Joint DL	Vexcel UltraCam Xp
	80.42	88.26	9.74	OBIA-SVM	Joint DL	
Mahdianpari et al. (2018) Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery	76.08	96.17	26.40	RF	InceptionResNetV2	RapidEye
Ndikumana et al. (2018) Deep recurrent neural network for agricultural classification using multitemporal SAR Sentinel-1 for Camargue, France	87.1	89.6	2.87	SVM	GRU	Sentinel-1
Zhang et al. (2018) An object-based convolutional neural network (OCNN) for urban land use classification	79.54	89.52	12.54	OBIA-SVM	OCNN	Vexcel UltraCam Xp
	80.37	90.87	13.06	OBIA-SVM	OCNN	
Guidici and Clark (2017) One-dimensional convolutional neural network land-cover classification of multi-seasonal hyperspectral imagery in the San Francisco Bay Area, California	87.5	88.0	0.57	SVM	CNN	AVIRIS
	89.5	89.9	0.44	SVM	CNN	AVIRIS
Ienco et al. (2017) Land cover classification via multitemporal spatial data by deep recurrent neural networks	74.28	75.18	1.21	RF	LSTM	Pléiades VHSR
	81.59	86.23	5.68	SVM	LSTM	Landsat-8
Kussul et al. (2017) DL classification of land cover and crop types using remote sensing data	88.7	94.6	6.65	RF	CNN	Landsat-8 and Sentinel-1A
Liu et al. (2017) SVM or DL? A comparative study on remote sensing image classification	92.38	91.12	-1.36	SVM-RBF	SAE	RODIS-03
	96.19	94.92	-1.32	SVM-RBF	SAE	RODIS-03
	99.59	97.79	-1.80	SVM-RBF	SAE	AVIRIS
	93.59	70.89	-24.25	SVM-RBF	SAE	AVIRIS
	91.38	86.58	-5.25	SVM-RBF	SAE	AVIRIS
Sharma et al. (2017) A patch-based convolutional neural network for remote sensing image classification	75.22	85.60	13.79	RF	a deep patch-based CNN system	Landsat-8
Wu and Prasad (2017) Convolutional recurrent neural networks for hyperspectral data classification	95.43	98.61	3.33	SVM-RBF	CRNN-LOP	ITRES-CASI
	95.13	98.08	3.10	SVM-RBF	CRNN-LOP	ProSpecTIRinstrument
Yu et al. (2017) Convolutional neural networks for hyperspectral image classification	58.31	64.19	10.08	SVM	CNN	AVIRIS
	80.95	85.24	5.29	SVM	CNN	AVIRIS
	61.24	67.85	10.79	SVM	CNN	RODIS

(Continued on following page)

TABLE 2 (Continued) Reported accuracies and potential accuracy improvement by the use of deep learning (DL) classification approaches from 29 published studies where both traditional machine learning (ML) methods and DL methods were used for land cover classification.

Citation/Title	Accuracy			Modeling method		Imagery type
	ML	DL	^a Change in DL over ML (%)	ML	DL	
Ghamisi et al. (2016) A self-improving convolution neural network for the classification of hyperspectral data	78.20	81.66	4.42	SVM	SICNN	AVIRIS
	78.21	82.67	5.70	SVM	SICNN	ROISIS-03
Li et al. (2016) Stacked autoencoder-based DL for remote-sensing image classification: a case study of African land-cover mapping	77.86	78.99	1.45	ANN	SAE	Landsat
Geng et al. (2015) High-resolution SAR image classification via deep convolutional autoencoders	76.92	88.11	14.54	SVM	DCAE	TerraSAR-X
Hu et al. (2015) Deep convolutional neural networks for hyperspectral image classification	87.60	90.16	2.92	RBF-SVM	CNN	AVIRIS
	91.66	92.60	1.02	RBF-SVM	CNN	AVIRIS
	90.52	92.56	2.25	RBF-SVM	CNN	ROISIS
Ly et al. (2015) Urban land use and land cover classification using remotely sensed SAR data through deep belief networks	0.76	0.81	6.57	SVM	DBN	RADARSAT-2 SAR

^aWhere percentage of change, is calculated as: Percentage change = [(New Value - Old Value)/Old Value] * 100.

2.2 Familiarity with AI techniques and computing expertise

Understanding the nuanced requirements of both ML and DL applications in LSS, specifically for LULC research, necessitates a comprehensive grasp of various skill sets and computational proficiencies. Diving into ML and DL methodologies for analyzing LULC through remote sensing requires not only theoretical knowledge but also practical competencies in software and algorithmic applications (Zhang et al., 2018; Ienco et al., 2019; Xie and Niculescu, 2021). Researchers venturing into this domain need a solid foundation in several critical areas.

1. Proficiency in Programming: Essential for crafting and manipulating algorithms, with languages like Python, R, or MATLAB being the cornerstones in both ML and DL realms.
2. Mastery of ML Algorithms: This encompasses understanding simpler, more accessible algorithms for those venturing into ML, including but not limited to DT, RF, and SVM.
3. Expertise in DL Algorithms: DL requires a deeper dive into more complex model structures, notably CNNs and their ilk, demanding a higher level of understanding and resources.
4. Data Management Skills: These involve preprocessing, feature extraction from image data, and understanding the sensitivity of DL to input data, which often demands a more intricate handling process.
5. Visualization Competencies: Both ML and DL necessitate skills in result interpretation and the visual representation of classified data, employing tools like confusion matrices and Receiver Operating Characteristic (ROC) curves.

6. Domain-Specific Knowledge: A comprehensive understanding of LULC changes, remote sensing parameters, and the specific environmental context of the study is paramount.
7. Software and Framework Literacy: Familiarity with specialized tools, libraries, and frameworks—such as TensorFlow, Keras, and PyTorch—is vital, especially in the DL context, for effective image classification research.
8. Cloud Computing and Hardware Understanding: Given the intensive computational demands, especially of DL, researchers need to be adept at utilizing cloud platforms (Amazon Web Service (AWS), Google Cloud Platform (GCP), Microsoft Azure) and understanding the hardware specifics like graphics processing units (GPUs) and tensor processing units (TPUs) (Ferreira et al., 2020; St. Onge et al., 2020).

The intersection of these skills defines the researcher's preparedness for engaging with ML or DL. While there is substantial overlap, DL generally demands a deeper engagement with aspects like specific libraries/frameworks, data sensitivity, and computational resources. Furthermore, the underlying statistical and mathematical models forming the bedrock of ML algorithms necessitate proficiency in statistics. These statistical models not only form the basis of the algorithms but also serve as evaluation metrics (F-score, accuracy, precision, etc.) ensuring the reliability of the models (Bhosle and Musande, 2019; C. Zhang et al., 2020a). Skills in Natural Language Processing (NLP) and data structures also become relevant in this regard, with a spectrum of algorithm types (Supervised, Unsupervised, and Reinforcement) each having unique applications and requirements (Rousset et al., 2021; Mou et al., 2022). Additionally, LULC tasks demand particular domain knowledge, including classification strategies, image interpretation,

and modeling methodologies (X. Zhang et al., 2020b). Effective LULC classification leverages various image features, necessitating a nuanced understanding of these elements for accurate land cover categorization. More recently, DL has emerged as a robust technique, setting new standards in LULC classification tasks and broader remote sensing analyses (Pashaei et al., 2020). While the foundational skills for ML and DL in LULC studies are closely related, the depth and specificity of knowledge, particularly when delving into DL, require a more robust, comprehensive skill set. Whether a researcher or a collaborative effort, the venture into LULC via ML or DL demands a blend of technical, analytical, and domain-specific expertise.

Finally, in addition to these critical aspects, there are other more fluid considerations a researcher should consider. The discussion above heavily implies the need for various skills and theoretical understanding, but the importance of practical, hands-on experience is also key (Barbierato et al., 2020; Campos-Taberner et al., 2023). Real-world problem-solving enhances theoretical knowledge, and this aspect is often as critical as the academic background itself. The integration of ML and DL into LSS and specifically LULC studies does not occur in isolation. As such, it is important to highlight the value of interdisciplinary knowledge and collaboration between data scientists, environmental scientists, and domain experts in geography, remote sensing, etc., to ensure the models' applicability and efficacy. There is also a growing need for researchers to understand the ethical implications of these technologies. Researchers must be aware of and know how to mitigate biases in their models and understand the ethical considerations of utilizing satellite or aerial data, privacy concerns, and potential impacts on policies or communities (Holloway and Merngersen, 2018; Elmes et al., 2020; Campos-Taberner et al., 2023). Finally, the field of AI and ML is rapidly evolving, and new techniques, algorithms, and best practices are continually emerging, and as such, researchers must commit to ongoing learning and adaptability with emerging technologies (Sarker, 2021; Southworth and Muir, 2021).

Enhancing classification accuracy in LSS significantly depends on the effective integration of diverse data sources. This integration enriches ML and DL models, providing a more comprehensive dataset for analysis. Key to this process is the utilization of public databases and application programming interfaces (APIs) from organizations such as the US Geological Survey (USGS), NASA (EarthData), and the European Space Agency (ESA). These platforms offer a rich array of socioeconomic, environmental, and satellite imaging data, crucial for LSS analysis. Additionally, citizen science initiatives and crowdsourcing platforms like OpenStreetMap and Global Forest Watch introduce unique datasets that blend satellite data with ground-level observations, providing a multifaceted view of land use patterns. To effectively combine these varied data sources, robust preprocessing and data fusion techniques are employed. These methods, including image registration, normalization, and transformation, address the challenges of inconsistent data coverage, formats, and resolutions, preparing the data for comprehensive analysis. Data standardization and the adoption of interoperable formats, such as GeoTIFF for spatial data, are essential for facilitating seamless integration across different datasets. Standards set by the Open Geospatial Consortium

(OGC) play a pivotal role in ensuring data compatibility, enhancing the ease of data manipulation and analysis.

Advanced ML techniques further support data integration by enabling efficient feature extraction and dimensionality reduction. Methods such as principal component analysis (PCA) and autoencoders help in transforming diverse data into an analytically suitable format. The application of unsupervised learning techniques also aids in identifying latent patterns within the integrated data, enriching the analysis. The synthesis of data from various sources not only expands the feature space for ML and DL models but also improves classification outcomes. Techniques like domain adaptation and transfer learning demonstrate the potential for enhanced model generalizability and accuracy across different contexts. Rigorous model evaluation and validation, through AUC-ROC curves, confusion matrices, and cross-validation, are critical in assessing the impact of data integration on classification accuracy. The integration of heterogeneous data sources presents a formidable avenue for boosting classification accuracy in LSS. As such, the strategies for data access, standardization, and analytical processing, as mentioned here briefly, highlights the significant advancements in ML and DL models' analytical capabilities, underscored by comprehensive data integration.

2.3 Machine and deep learning software and libraries for land use land cover

Implementing ML and DL in LULC classifications requires a nuanced selection of software and libraries, each offering distinct advantages. The burgeoning field of LSS necessitates an in-depth understanding of these tools for optimal outcomes in classification tasks.

1. PyTorch: An open-source machine learning framework known for its flexibility, dynamic computational graph, and effective GPU support. It is widely used in LULC for its ability to process large-scale visual data (Song et al., 2021; Kim et al., 2022; Liu et al., 2022).
2. TensorFlow: this comprehensive platform is suitable for both machine learning and deep learning. It is known for handling high-dimensional datasets and is versatile in deployment environments (Baylor et al., 2017; Silaparasetty, 2020), and Cherif et al. (2022), highlight its proficiency in intricate LULC tasks.
3. Keras: A user-friendly neural network library that works on top of TensorFlow. It is known for its ease of use in developing and testing ML models, making it a popular choice for LULC studies (Nhu et al., 2020; Cherif et al., 2022).
4. H2O.ai: Offers distributed computing for machine learning and deep learning. It is recognized for its efficient algorithms and user-friendly interface (Balakrishnan, 2019; GitHub, 2022).
5. Neural Designer: Focuses on neural network-based analysis with an intuitive interface. It is suitable for various data analytics applications in LULC (Yilmaz et al., 2020; Wang et al., 2021).
6. Microsoft Cognitive Toolkit (CNTK): Known for scalability and speed, this toolkit is suitable for developing deep learning

- models and is compatible with multiple programming languages (Chrisbasoglu, 2022).
7. ConvNetJS: A JavaScript library for browser-based computation, allowing for on-the-fly training of neural networks. It is practical for real-time LSS research scenarios (Andrej et al., 2022).
 8. Geospatial Data Abstraction Library (GDAL): GDAL is a widely used open-source translator library for raster and vector geospatial data formats (GDAL/OGR contributors, 2024). GDAL was originally developed to meet the requirements for efficient and flexible file format translation, with the capability to adapt to new formats. It plays a crucial role in machine learning (ML) and deep learning (DL) workflows by enabling the preprocessing of remote sensing data. This includes tasks such as representing and transforming coordinate reference systems, format conversion, and resampling, making data compatible for ML/DL models (Warmerdam, 2008).

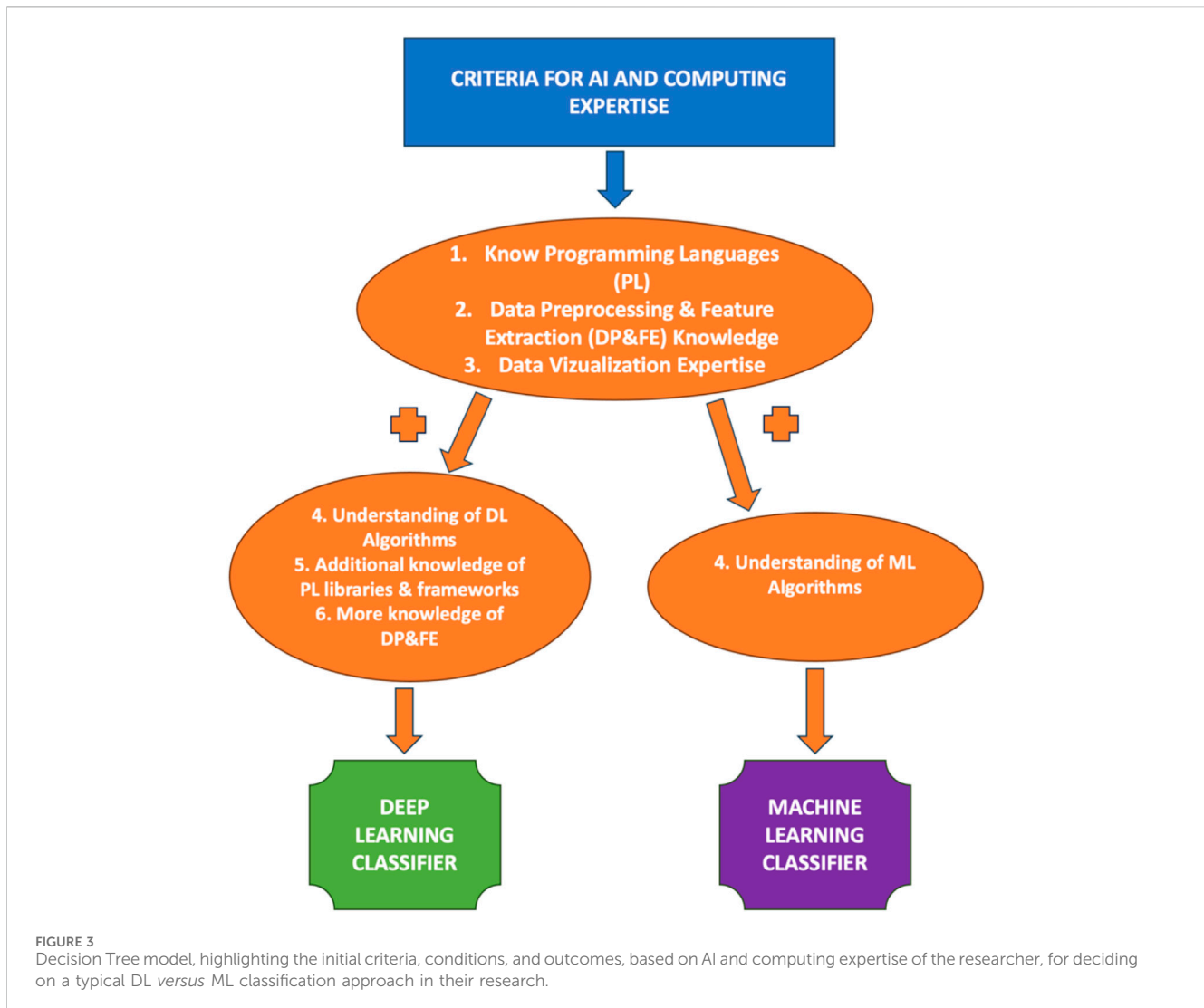
There are several other notable tools and libraries essential in the realm of LULC and LSS research. Among them is Scikit-learn, a Python package that is essential for LULC classifications and is praised for its simplicity and effectiveness in machine learning (Feizizadeh et al., 2023). With support for a broad range of raster and vector data formats necessary for LULC preprocessing, GDAL (Geospatial Data Abstraction Library) stands out for its key role in processing geospatial data (Hütt et al., 2020; Sawant et al., 2023). One of the major players in remote sensing is ERDAS IMAGINE, which is well-known for its sophisticated processing of satellite imagery and aerial photography and is frequently used in LULC analysis (Vivekananda et al., 2021). With its vast satellite imaging resources and analytical tools, Google Earth Engine emerges as a cutting-edge cloud-based platform for environmental data analysis. This makes it an invaluable tool for LULC and environmental monitoring (Tassi and Vizzari, 2020; Feizizadeh et al., 2023). Using high-level components for quick and effective LULC applications, Fast. ai, created on PyTorch, democratizes deep learning (Vivekananda et al., 2021; Neupane et al., 2023). The open-source deep learning framework MXNet is widely used in LULC due to its huge dataset management capabilities and scalability across GPUs and machines (Neupane et al., 2021). A crucial piece of software for remote sensing is ENVI (Environment for Visualizing Images), which provides powerful tools for image processing and analysis and is especially pertinent for LULC research (Saharan et al., 2018; Olorunfemi et al., 2020). Finally, the R programming language is a powerhouse in LULC research because it provides extensive statistical and analytical tools, particularly with packages like “caret,” “random forest,” and “raster” (Amini et al., 2022; Aryal et al., 2023).

While there is insufficient space in this review to discuss the strengths and weaknesses of each option, the main tools used (Keras, PyTorch, and TensorFlow) may be compared briefly. When considering PyTorch and TensorFlow, the choice is often contingent upon the user's preference between flexibility and a rich feature set. PyTorch offers more control for real-time changes and is rapidly gaining ground for research purposes. In contrast, TensorFlow's broad adoption in the industry is due to its comprehensive infrastructure and scalability, essential for deploying large-scale applications. On the other hand, Keras is not directly

comparable with PyTorch and TensorFlow, as it is a high-level API capable of running on top of TensorFlow. It abstracts many of the complexities of building a DL model, making it accessible and reducing the need for detailed technical knowledge. However, for intricate and customized model building, researchers might prefer the control offered by PyTorch or the depth of resources available in TensorFlow. The selection of a tool or library hinges on the specific needs of the project, the team's expertise, and the long-term goals associated with the application's development and deployment. Directly comparing these tools based on these parameters allows teams to align their choice with their strategic objectives.

These tools are renowned for their seamless compatibility with diverse computing environments, including integration with cloud platforms, third-party APIs, and various data systems. For example, TensorFlow's robustness is exemplified by its deep integration with Google Cloud, while PyTorch enjoys widespread support due to its alliances with multiple cloud service platforms. Similarly, H2O.ai's platform-independent nature makes it a versatile choice for various operating systems, enhancing its reach and applicability. Moreover, the open-source character of these libraries—particularly PyTorch, TensorFlow, Keras, and H2O.ai—democratizes access, allowing a wide spectrum of users, from researchers to enterprise-level users, to leverage advanced ML capabilities without hefty licensing fees. The Neural Designer, while not entirely open source, offers comprehensive solutions in NN applications, often justifying its cost. In contrast, Microsoft Cognitive Toolkit provides an enterprise-grade solution, reflecting its pedigree and focus on scalability and performance. ConvNetJS stands out by enabling ML directly in the browser, eliminating certain logistical hurdles and inviting a broader audience of developers into the ML sphere. This accessibility underscores the inclusive philosophy driving innovation in these tools. While these software and libraries are generally accessible due to low or no initial acquisition costs, it is prudent for users to consider ancillary expenses. These might include costs related to the computational resources, especially in DL applications, and potential investments into cloud services for data storage, processing, or model deployment, particularly in large-scale projects (Zhu et al., 2017; Cheng et al., 2020; Wang L. et al., 2022).

Ultimately, the choice of a specific tool or library will hinge not only on its standalone capabilities but also on its synergistic fit within the user's existing technological infrastructure, budget constraints, and the specific requirements of their projects. This holistic assessment, considering both the direct and indirect costs, ensures that users can harness the full potential of these advanced technologies in their ML and DL endeavors. In conclusion, the landscape of ML and DL software and libraries is rich and varied, each contributing uniquely to LULC classifications within LSS. While traditional ML algorithms like RF, SVM, and KNN excel with smaller, less complex datasets, DL models (e.g., CNN, RNN) are poised to handle larger, more intricate datasets. The choice between these paradigms hinges on the specific requirements and constraints of the research project, underscoring the need for proficiency in different tools and a strategic approach to their application (Figure 3). Both ML and DL demand a skill set that spans programming, data handling, and visualization, critical for navigating the complexities and maximizing the potential of LSS studies.

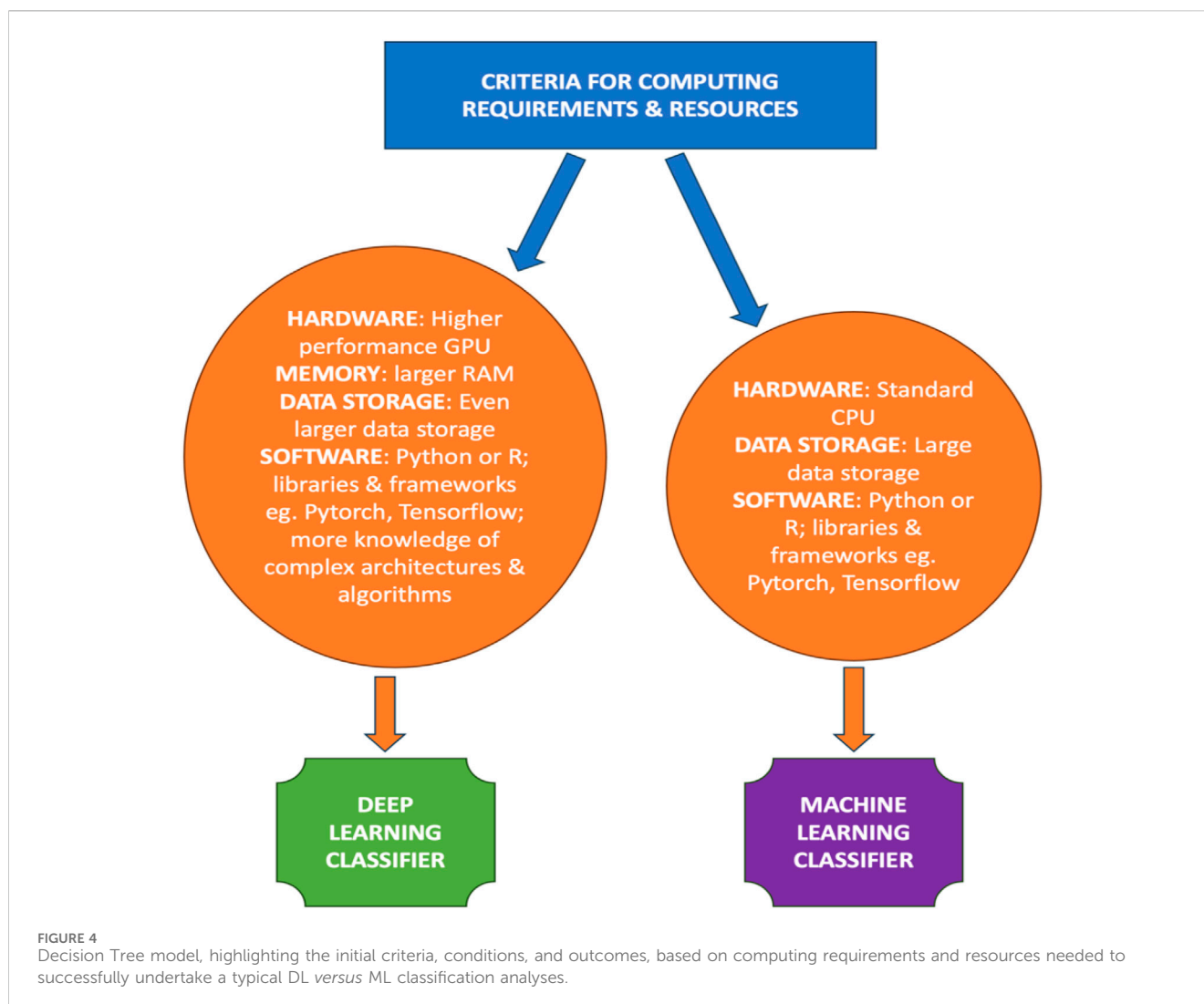


2.4 Computing requirements for analysis and available resources

The nuanced computing requirements for image classification in ML versus DL pivot significantly around the hardware utilized and the associated resources. Analyzing these requirements necessitates a deep dive into the processors involved, namely, central processing units (CPUs) and graphics processing units (GPUs), as well as the burgeoning field of application-specific integrated circuits (ASICs). CPUs, adept at sequential complex calculations, have historically sufficed for traditional ML models. In contrast, the advent of DL, particularly in tasks like image processing, has necessitated GPUs, recognized for their parallel processing capabilities. GPUs expedite the efficiency of learning and inference algorithms, making them indispensable for DL. This need arises from the GPU's architecture, a single instruction, multiple data (SIMD) device, allowing extensive parallelization and enhanced performance due to its numerous simpler cores, compared to the more complex CPU cores (Ota et al., 2017). The demand for specialized hardware escalates further with ASICs, such as Google's Tensor Processing Unit (TPU) and innovations like ShiDianNao, a custom ASIC engineered to optimize

memory access, thus slashing latency and energy costs (Chen and Ran, 2017). Beyond the hardware, the requirements extend to local memory for data and code storage, with sophisticated algorithms, especially in deep NNs, demanding rapid, extensive memory. Simpler models, like DTs, are less taxing on resources.

Recent leaps in cloud computing technologies have been revolutionary, with platforms like Google Earth Engine (GEE) and Amazon Web Services (AWS) offering substantial computing power, often complimentary, and direct access to extensive satellite data repositories from National Aeronautics and Space Administration (NASA), United States Geological Survey (USGS), and European Space Agency (ESA) (Sagan et al., 2020). These advancements are crucial, especially in remote sensing, circumventing the traditional high processing costs that impede research progress. The landscape of computing in DL extends to cluster and grid computing, with the latter now embracing extreme heterogeneity in high-performance computing systems, thanks to technological evolution over time. Grid computing, especially, has surfaced as a pivotal technology in managing remote sensing data across dispersed computing environments. Cloud computing further augments this by integrating disparate specialized nodes,



presenting a resilient, cost-effective methodology for handling voluminous remote sensing data (Duffy et al., 2023). However, the computational needs in an overparameterized setting spiral, often quadratically, relative to the data points, as training DL models correlate with the number of parameters and data points. This underscores the necessity for a more-than-linear surge in training data to achieve linear performance enhancements (Yao et al., 2022.). A revelation in this domain is the potential for sustaining performance upgrades without escalating energy consumption, courtesy of algorithmic advancements and the burgeoning parameters. This trend mirrors the auto industry's evolution, where increased efficiency has not negated the overall rise in vehicles, thus energy usage, underscoring the challenge of balancing technological progress with sustainable practices (Holloway and Mengersen, 2018; Cheng et al., 2020; Sarker, 2021).

Contrasting the computing requisites for ML and DL, it is evident that DL, especially in image classification, mandates more formidable hardware, given the extensive NNs' training. The memory requirements are also steeper, owing to the networks' complexity and parameter abundance. While both fields require substantial data storage and specific software (e.g., Python, R, TensorFlow, Keras, scikit-learn, PyTorch), DL necessitates greater

familiarity with these tools due to its intricate algorithms. Furthermore, though cloud computing can host both ML and DL tasks, the latter's resource intensity typically translates to higher costs. DL's demands in image classification outstrip those of ML, encompassing hardware, memory, data storage, and specialized software knowledge. Though cloud platforms can accommodate both, DL's complexity and resource intensity often imply a costlier process. This comparative analysis elucidates the distinct considerations pivotal in choosing between ML and DL for typical image classification endeavors (Figure 4).

3 Discussion: guiding LSS researchers through ML or DL method selection with a decision tree model

In the continuously evolving field of AI, the allure of novel tools and techniques is undeniable. However, it is imperative for researchers to exercise caution and undertake a meticulous suitability analysis before employing these advancements within diverse research landscapes. In this study, we have explored the application of ML and DL techniques in the scientific classification

task of land systems. Our findings reveal both strengths and weaknesses inherent in these methodologies, as well as illuminate potential directions for future development in this domain. This discussion synthesizes the distinct demands, advantages, and constraints of employing ML *versus* DL techniques in image classification tasks, aiming to provide a navigational aid for researchers. This examination is especially pertinent given the expanding prevalence of these tools within academic and practical applications. Our comprehensive analysis culminates in a decision tree model, delineating critical factors warranting consideration when oscillating between ML and DL methodologies for image classification in LSS studies.

1. **Assessment of Expertise & Resources:** Researchers must introspectively evaluate their proficiency and the resources at their disposal. DL demands a deeper understanding, encompassing familiarity with specific programming frameworks and the need for robust computational backbones. Conversely, ML is more accessible, demanding less esoteric knowledge and offering broader software compatibility.
2. **Computational Demands:** The availability of computational power is pivotal. DL's reliance on substantial computational support, particularly potent GPUs and extensive memory for model training, contrasts with ML's modest resource requirements, compatible with standard computing setups. The computational demands of machine learning models, including deep learning models, can differ greatly depending on parameters including model structure, size, complexity, hardware for training and inference, and optimization methods used. Training deep neural networks usually requires repetitive calculations on extensive datasets, which can use up significant energy, particularly when utilizing powerful processing hardware such as GPUs or TPUs. Simple machine learning models such as decision trees or linear models may consume less energy during both training and inference stages when compared to more complicated deep learning models. Deep learning models may demand substantial energy for training and inference, but energy consumption is influenced by multiple factors. It is inaccurate to generalize that they consistently consume more energy than other machine learning models.
4. **Research Questions & Information Complexity:** The nature of the research query and the intricacy of the requisite information significantly influence the methodological choice. DL thrives on high-dimensional data scenarios where factors like varying light conditions, scaling, and orientation diversities come into play. ML, however, suffices for inquiries that demand less exactitude or engage less convoluted data.
5. **Image Data Availability:** The decision hinges on the availability and quality of accessible image data. DL's hunger for voluminous training datasets contrasts starkly with ML's resilience to data imperfections and lower data appetite.
6. **Study Timeline Constraints:** Temporal aspects of the study also dictate methodology suitability. The quick deployment of ML models offers advantages over the more prolonged, intricate training demands of DL systems. Particularly for longitudinal

studies, the dearth of historical data fitting DL's appetite often results in the preference towards ML methodologies.

7. **Training Data Sufficiency:** The volume and integrity of data available for training directly influences the decision. Where DL necessitates a substantial repository of labeled imagery, ML operates efficiently with far less.
8. **Accuracy Imperatives:** Researchers must balance their accuracy requisites with pragmatic considerations like resource allocation and model interpretability. While DL, with the right conditions, typically surpasses ML in accuracy for image classification tasks, this edge must be substantial enough to justify the increased workload and resource commitment (as the improvement often hovers below 10%, and sometimes less than 4%, as per [Table 2](#)).

The field of artificial intelligence is continually evolving, offering researchers a myriad of sophisticated tools and techniques. However, the allure of novelty must be measured against the genuine utility these innovations offer research domains. In this discussion, we have sought to distill the essential factors which might dictate the selection of either ML or DL for image classification within the realm of LSS. While both ML and DL have carved out their own respective niches, it is imperative to remember that each has its strengths and drawbacks ([Lary et al., 2016](#); [Maxwell et al., 2018](#); [Cheng et al., 2020](#); [Gupta et al., 2021](#); [Sarker, 2021](#)). Our exploration serves as a roadmap, guiding researchers through the labyrinth of considerations pivotal to choosing one method over the other. The fact that DL tools have gained such a dominant presence in recent remote sensing literature underscores the timeliness and relevance of our analysis ([Campos-Taberner et al., 2023](#); [Barbierato et al., 2020](#); [Du et al., 2020](#); [Table 1](#)).

The decision tree structure depicted in [Figure 5](#) reveals the primary factors for consideration. Firstly, the expertise requisite to wield these tools is non-trivial. DL demands a deeper dive into specialized territories, necessitating familiarity with specific programming frameworks and an arsenal of robust computational assets. In contrast, ML is more forgiving in this realm, offering a wider array of user-friendly software packages that do not demand as deep a domain expertise. Next, the computational constraints must be addressed. DL, with its intricate NNs, devours computational power and memory—often compelling researchers to turn to specialized GPUs. Conversely, ML's modest computational appetite means it can typically operate comfortably on standard systems. The nature of the research question is paramount. High-dimensional challenges, those dealing with vast image datasets, and questions demanding an acute level of precision find a formidable ally in DL. This technique's prowess lies in its ability to decipher nuances in lighting, scale, and rotation. On the other hand, ML emerges as the method of choice for less intricate problems, where a slightly compromised accuracy is acceptable. The data pool from which these models draw is also vital. DL requires substantial volumes of data for training. However, its precision can falter if the data quality is compromised. ML, in this context, displays more resilience, operating effectively even with smaller data sets. Longitudinal studies, especially those delving deep into historical archives where data for earlier timelines might be unavailable, might find ML to be a more pragmatic choice.

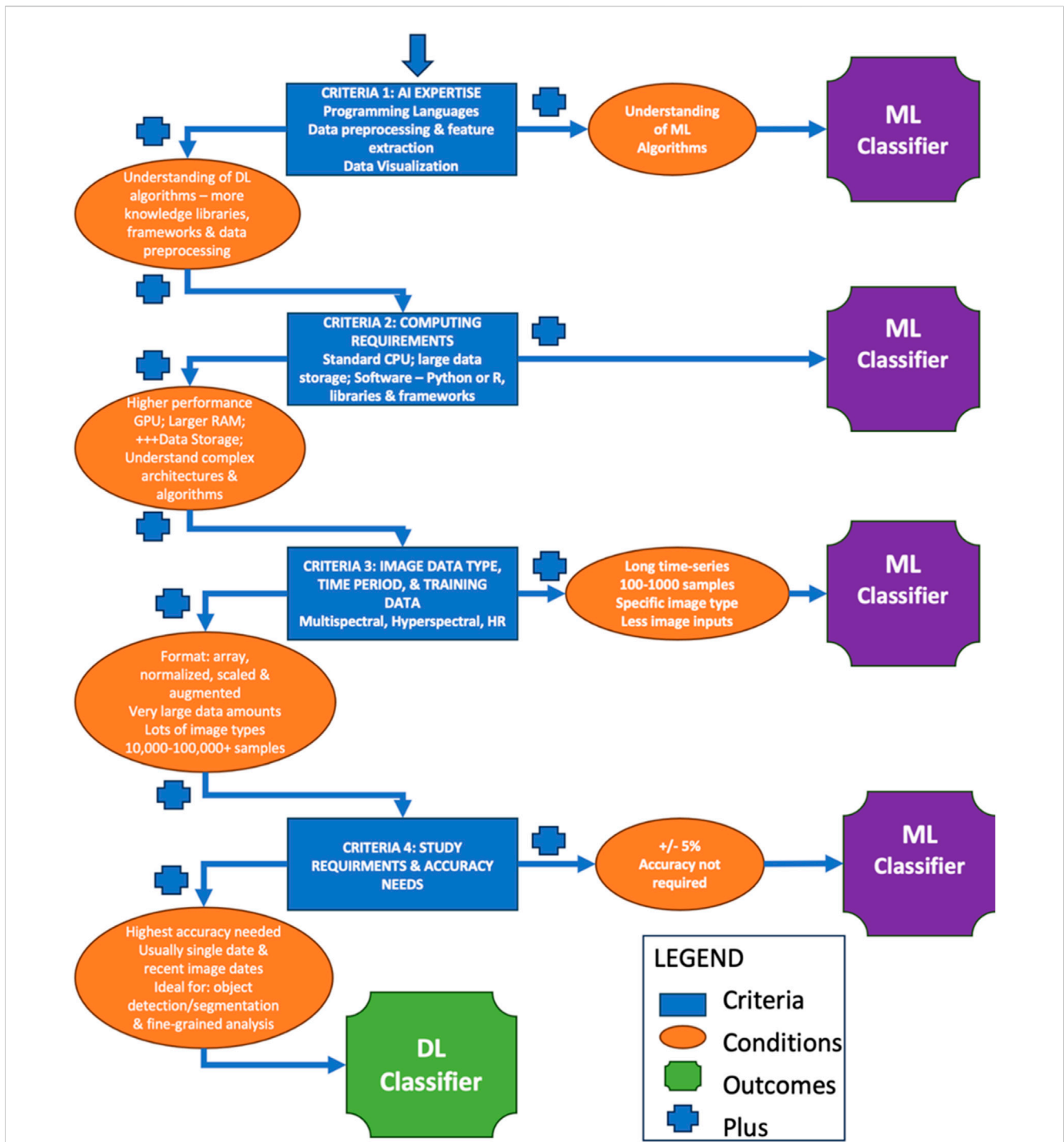


FIGURE 5
Final Decision Tree highlighting the four sets of criteria explored to determine the ideal classification approach related to typical machine learning (ML) or deep learning (DL) classification techniques in Land System Science: 1) AI Expertise, 2) Computing Requirements, 3) Image Data Type, Time Period, and Training Data and 4) Study Requirements and Accuracy needs.

Within the field of LSS, the choice between ML and DL for image classification hinges on a mosaic of factors. DL may be the most effective choice when precision is paramount, and data is abundant. Its capacity to discern complex patterns in vast datasets renders it invaluable for intricate projects (Sharma et al., 2017; Li et al., 2019; Wan et al., 2019). However, when data is scarce, or when the research objectives do not demand the peak of accuracy, ML, with its versatility and user-friendly nature, may be the most

effective choice (Abdi, 2020; Jamali, 2021; Wang et al., 2021; Wang L. et al., 2022; Magalhães et al., 2022; Zaabar et al., 2022). This decision tree, therefore, serves as a compass, enabling researchers to navigate the intricate terrain of AI-driven image classification with clarity and confidence. The findings presented, along with Tables 1, 2, coupled with the insights from Figure 5, underscore a critical observation: in the present landscape of LSS, the constraints associated with deploying DL models are often more

significant than their advantages. While there is great potential for DL to transform the field, Figure 5 unambiguously indicates that, in most scenarios, ML models currently deliver optimum outcomes for most classification tasks. It is a reminder that while we aspire for future innovations, it is essential to recognize and harness the proven capabilities of existing tools. While the field of LSS is not fully utilizing DL yet, the future does indeed look bright. The evolving DL tools and their advanced pattern recognition capabilities are expected to greatly benefit the field and allow for addressing much more complex environmental questions. As LSS continues to progress, the potential of DL to revolutionize various aspects of image classification and analysis is unmistakable. This optimism, paired with the current effectiveness of ML, suggests a future where DL not only complements but significantly enhances our understanding and exploration in LSS. In addition, the model developers themselves are acutely aware of many of the current limitations to ML and DL models, as have been described here for typical model studies. As such, new methods are being developed that do address some of these limitations. For example, new approaches being developed by Hong et al. (2021) are pushing the boundaries of ML approaches, with data sizes exceeding 100,000 (Figure 2, this is a listed limitation of ML approaches). Similarly, DL methods are being developed that do not have the current requirements for large memory (Figure 4, listed as a limitation of DL methods), being termed 'lightweight' DL methods (Choudhary et al., 2022). However, while these methods are in development, they are by no means mainstream, and the typical or current approaches discussed here still dominate in LSS research.

3.1 Future directions of development

We have identified five areas on which future research should focus: 1) the integration of multisource data; 2) semi-supervised and unsupervised learning; 3) enhanced interpretability; 4) domain adaptation and transfer learning; and 5) integration with decision support system. Research in the field of LULCC and LSS should continue to expand and enhance the integration of multisource data—including satellite imagery, ground-based observations, and socioeconomic data—in ML and DL applications to improve the accuracy and robustness of land system classification models. Future research on the efficacy and appropriateness of ML and DL techniques in LSS should explore the evolving nature of these techniques and their relative applications in the context of multisource data integration. Further, the development of semi-supervised and unsupervised learning approaches that can mitigate the reliance on labeled data, allowing for more efficient utilization of available datasets will prove invaluable to advancing the field of LSS and its transdisciplinary research objectives. Issues of trust in and interpretability of AI models is an area of critical importance that necessitates the attention of researchers. As AI methods become increasingly employed in LULCC analyses, research that seeks to elevate the interpretability of ML and DL models through techniques such as attention mechanisms, model distillation, and *post hoc* explanations will be crucial for increasing trust and acceptance of automated classification results. Future work should also focus on developing techniques for domain adaptation and transfer learning that can facilitate the

transferability of classification models across different geographical regions and environmental conditions, reducing the need for extensive retraining. Lastly, the future applicability of AI techniques in the classification and change analysis of remote sensing data will rely on the integration of ML and DL classification models with decision support systems that can enable real-time monitoring and management of land systems, supporting sustainable land use planning and natural resource management efforts. While ML and DL techniques offer significant promise for land system classification, addressing their inherent limitations and exploring new avenues for development will be essential for realizing their full potential in scientific research and environmental management.

4 Conclusion

The role of ML and DL in LSS image classification is evident (Tables 1, 2), especially when looking at Figure 5. This decision-making framework is not just an academic tool; it has wider implications beyond research. Integrating various data sources, like satellites, drones, and ground observations, can improve land change detection. Advanced DL architectures, especially CNNs, are leading this shift. Techniques like unsupervised learning can efficiently identify patterns in large sets of satellite imagery. However, challenges remain. Acquiring quality data for certain regions is difficult. The computational demands of DL are high, requiring significant resources. The complexity of some DL models makes them difficult to interpret, and overfitting remains a risk. Additionally, high-resolution imagery, while ideal, can be hard to obtain or costly. To address these challenges, a multifaceted approach is needed. Using diverse training datasets can lead to more resilient models. Regular evaluations keep models accurate, and combining different models can improve results. Collaborations between domain experts and ML/DL specialists can produce more relevant solutions. Making LULC datasets publicly available can stimulate new research, and it is vital to consider the ethical aspects of this technology. As the ML and DL fields evolve, staying updated is essential.

The benefits of this framework are wide-ranging. Better data quality can lead to more informed environmental policies, aiding in resource conservation, biodiversity, and carbon management. This improved data can also have societal impacts, potentially influencing migration, urban development, and socio-economic trends. For stakeholders in fields like agriculture and urban planning, this data is invaluable for making informed decisions. There are also potential economic benefits from improved land management. At its core, this approach can advance academic research in LSS, promoting interdisciplinary collaboration and a deeper understanding of land systems. In conclusion, as LSS evolves, the combined use of ML and DL, supported by a solid decision-making framework, promises significant advancements in research, policy, and practice.

Author contributions

JS: Conceptualization, Data curation, Formal Analysis, Methodology, Resources, Supervision, Visualization, Writing—original draft, Writing—review and editing. AS: Data curation, Formal Analysis, Methodology, Resources,

Writing—original draft, Writing—review and editing. MS: Data curation, Formal Analysis, Methodology, Resources, Writing—original draft, Writing—review and editing. MR: Data curation, Formal Analysis, Methodology, Resources, Writing—original draft, Writing—review and editing. AA: Data curation, Formal Analysis, Methodology, Resources, Writing—original draft, Writing—review and editing. BT: Data curation, Formal Analysis, Methodology, Resources, Writing—original draft, Writing—review and editing. CM: Data curation, Formal Analysis, Methodology, Resources, Writing—original draft, Writing—review and editing. HH: Writing—original draft, Writing—review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

References

- Abdi, A. M. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience Remote Sens.* 57 (1), 1–20. doi:10.1080/15481603.2019.1650447
- Agrawal, N., Govil, H., and Kumar, T. (2024). Agricultural land suitability classification and crop suggestion using machine learning and spatial multicriteria decision analysis in semi-arid ecosystem. *Environ. Dev. Sustain.*, 1–38. doi:10.1007/s10668-023-04440-1
- Ahmed, M. H., and Lin, L.-S. (2021). Dissolved oxygen concentration predictions for running waters with different land use land cover using a quantile regression forest machine learning technique. *J. Hydrology* 597, 126213. doi:10.1016/j.jhydrol.2021.126213
- Akar, O., and Tunc Gormus, E. (2022). Land use/land cover mapping from airborne hyperspectral images with machine learning algorithms and contextual information. *Geocarto Int.* 37 (14), 3963–3990. doi:10.1080/10106049.2021.1945149
- Alhassan, V., Henry, C., Ramanna, S., and Storie, C. (2020). A deep learning framework for land-use/land-cover mapping and analysis using multispectral satellite imagery. *Neural Comput. Appl.* 32 (12), 8529–8544. doi:10.1007/s00521-019-04349-9
- Alshari, E. A., Abdulkareem, M. B., and Gawali, B. W. (2023). Classification of land use/land cover using artificial intelligence (ANN-RF). *Front. Artif. Intell.* 5, 964279. doi:10.3389/frai.2022.964279
- Amini, S., Saber, M., Rabiei-Dastjerdi, H., and Homayouni, S. (2022). Urban land use and land cover change analysis using random forest classification of Landsat time series. *Remote Sens.* 14 (11), 2654. Article 11. doi:10.3390/rs14112654
- Andrej, A. (2022). ConvNetJS [JavaScript]. Available at: <https://github.com/karpathy/convnetjs>.
- Arruda, V. L., Piontekowski, V. J., Alencar, A., Pereira, R. S., and Matricardi, E. A. (2021). An alternative approach for mapping burn scars using Landsat imagery, Google Earth Engine, and Deep Learning in the Brazilian Savanna. *Remote Sens. Appl. Soc. Environ.* 22, 100472. doi:10.1016/j.rsase.2021.100472
- Aryal, J., Sitaula, C., and Frery, A. C. (2023). Land use and land cover (LULC) performance modeling using machine learning algorithms: a case study of the city of Melbourne, Australia. *Sci. Rep.* 13 (1), 13510. Article 1. doi:10.1038/s41598-023-40564-0
- Balakrishnan, D. S. (2019). An overview of H2O: an open source AI platform open source for you. Available at: <https://www.opensourceforu.com/2019/07/an-overview-of-h2o-an-open-source-ai-platform/>.
- Barbierato, E., Bernetti, I., Capecchi, I., and Saragosa, C. (2020). Integrating remote sensing and street view images to quantify urban forest ecosystem services. *Remote Sens.* 12 (2), 329. doi:10.3390/rs12020329
- Baylor, D., Breck, E., Cheng, H.-T., Fiedel, N., Foo, C. Y., Haque, Z., et al. (2017). “TFX: a TensorFlow-based production-scale machine learning platform,” in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, August 13–17, 2017, 1387–1395.
- Bhargava, R., Sarkar, D., and Friess, D. A. (2021). A cloud computing-based approach to mapping mangrove erosion and progradation: case studies from the Sundarbans and French Guiana. *Estuar. Coast. Shelf Sci.* 248, 106798. doi:10.1016/j.ecss.2020.106798
- Bhosle, K., and Musande, V. (2019). Evaluation of deep learning CNN model for land use land cover classification and crop identification using hyperspectral remote sensing images. *J. Indian Soc. Remote Sens.* 47 (11), 1949–1958. doi:10.1007/s12524-019-01041-2
- Campos-Taberner, M., García-Haro, F. J., Martínez, B., Izquierdo-Verdiguier, E., Atzberger, C., Camps-Valls, G., et al. (2020). Understanding deep learning in land use classification based on Sentinel-2 time series. *Sci. Rep.* 10 (1), 17188. doi:10.1038/s41598-020-74215-5
- Campos-Taberner, M., García-Haro, F. J., Martínez, B., Sánchez-Ruiz, S., Moreno-Martínez, Á., Camps-Valls, G., et al. (2023). Land use classification over smallholding areas in the European Common Agricultural Policy framework. *ISPRS J. Photogrammetry Remote Sens.* 197, 320–334. doi:10.1016/j.isprsjprs.2023.02.005
- Chen, B., Huang, B., and Xu, B. (2017). Multi-source remotely sensed data fusion for improving land cover classification. *ISPRS J. Photogrammetry Remote Sens.* 124, 27–39. doi:10.1016/j.isprsjprs.2016.12.008
- Chen, J., and Ran, X. (2019). Deep learning with edge computing: a review. *Proc. IEEE* 107 (8), 1655–1674. doi:10.1109/jproc.2019.2921977
- Cheng, G., Xie, X., Han, J., Guo, L., and Xia, G. S. (2020). Remote sensing image scene classification meets deep learning: challenges, methods, benchmarks, and opportunities. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sens.* 13, 3735–3756. doi:10.1109/jstars.2020.3005403
- Cheng, X., Sun, Y., Zhang, W., Wang, Y., Cao, X., and Wang, Y. (2023). Application of deep learning in multitemporal remote sensing image classification. *Remote Sens.* 15 (15), 3859. Article 15. doi:10.3390/rs15153859
- Cherif, E., Hell, M., and Brandmeier, M. (2022). DeepForest: novel deep learning models for land use and land cover classification using multi-temporal and -modal sentinel data of the Amazon basin. *Remote Sens.* 14 (19), 5000. doi:10.3390/rs14195000
- Choudhary, K., DeCost, B., Chen, C., Jain, A., Tavazza, F., Cohn, R., et al. (2022). Recent advances and applications of deep learning methods in materials science. *npj Comput. Mater.* 8, 59. doi:10.1038/s41524-022-00734-6
- Chrisbasoglu (2022) *The Microsoft cognitive toolkit—cognitive toolkit—CNTK*. Available at: <https://learn.microsoft.com/en-us/cognitive-toolkit/>.
- Davenport, T. H., Barth, P., and Bean, R. (2012). How ‘big data’ is different. *MIT Sloan Manag. Rev.* 54 (1).
- Delegido, J., Vergara, C., Verrelst, J., Gandia, S., and Moreno, J. (2011b). Remote estimation of crop chlorophyll content by means of high-spectral-resolution reflectance techniques. *Agron. J.* 103 (6), 1834–1842. doi:10.2134/agronj2011.0101
- Delegido, J., Verrelst, J., Alonso, L., and Moreno, J. (2011a). Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. *Sensors* 11 (7), 7063–7081. doi:10.3390/s110707063
- Delegido, J., Verrelst, J., Meza, C. M., Rivera, J. P., Alonso, L., and Moreno, J. (2013). A red-edge spectral index for remote sensing estimation of green LAI over agroecosystems. *Eur. J. Agron.* 46, 42–52. doi:10.1016/j.eja.2012.12.001
- Du, L., McCarty, G. W., Zhang, X., Lang, M. W., Vanderhoof, M. K., Li, X., et al. (2020). Mapping forested wetland inundation in the Delmarva Peninsula, USA using deep convolutional neural networks. *Remote Sens.* 12 (4), 644. doi:10.3390/rs12040644
- Duffy, M. L., Medeiros, B., Gettelman, A., and Eidhammer, T. (2023). Perturbing parameters to understand cloud contributions to climate change. *J. Clim.* 37 (1), 213–227. doi:10.1175/JCLI-D-23-0250.1

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Elmes, A., Alemohammad, H., Avery, R., Caylor, K., Eastman, J. R., Fishgold, L., et al. (2020). Accounting for training data error in machine learning applied to earth observations. *Remote Sens.* 12 (6), 1034. doi:10.3390/rs12061034
- Feizizadeh, B., Omarzadeh, D., Kazemi Garajeh, M., Lakes, T., and Blaschke, T. (2023). Machine learning data-driven approaches for land use/cover mapping and trend analysis using Google Earth Engine. *J. Environ. Plan. Manag.* 66 (3), 665–697. doi:10.1080/09640568.2021.2001317
- Ferreira, K. R., Queiroz, G. R., Camara, G., Souza, R. C. M. D., Vinhas, L., Marujo, R. D. F. B., et al. (2020). “Using remote sensing images and cloud services on AWS to improve land use and cover monitoring,” in 2020 IEEE Latin American GRSS and ISPRS Remote Sensing Conference (LAGIRS), Santiago, Chile, 22–26 March 2020 (IEEE), 558–562.
- Fu, W., and Menzies, T. (2017). “Easy over hard: a case study on deep learning,” in Proceedings of the ACM SIGSOFT Symposium on the Foundations of Software Engineering, Part F130154, 49–60.
- Fujiyoshi, H., Hirakawa, T., and Yamashita, T. (2019). Deep learning-based image recognition for autonomous driving. *IATSS Res.* 43 (4), 244–252. doi:10.1016/j.iatssr.2019.11.008
- GDAL/OGR contributors (2024). GDAL/OGR geospatial data abstraction software library open source geospatial foundation. Available at: <https://gdal.org/index.html>.
- Geist, H., McConnell, W., Lambin, E. F., Moran, E., Alves, D., and Rudel, T. (2006). “Causes and trajectories of land-use/cover change,” in *Land-use and land-cover change: local processes and global impacts* (Berlin, Heidelberg: Springer Berlin Heidelberg), 41–70.
- Geng, J., Fan, J., Wang, H., Ma, X., Li, B., and Chen, F. (2015). High-resolution SAR image classification via deep convolutional autoencoders. *IEEE Geoscience Remote Sens. Lett.* 12 (11), 2351–2355. doi:10.1109/lgrs.2015.2478256
- Ghamisi, P., Chen, Y., and Zhu, X. X. (2016). A self-improving convolution neural network for the classification of hyperspectral data. *IEEE Geoscience Remote Sens. Lett.* 13 (10), 1537–1541. doi:10.1109/lgrs.2016.2595108
- Github, G. (2022) *H2O.ai*. GitHub. Available at: <https://github.com/h2oai>.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. doi:10.1016/j.rse.2017.06.031
- Guidici, D., and Clark, M. L. (2017). One-Dimensional convolutional neural network land-cover classification of multi-seasonal hyperspectral imagery in the San Francisco Bay Area, California. *Remote Sens.* 9 (6), 629. doi:10.3390/rs9060629
- Guo, Y., Liao, J., and Shen, G. (2021). Mapping large-scale mangroves along the maritime silk road from 1990 to 2015 using a novel deep learning model and Landsat data. *Remote Sens.* 13 (2), 245. doi:10.3390/rs13020245
- Gupta, R., and Nanda, S. J. (2022). Cloud detection in satellite images with classical and deep neural network approach: a review. *Multimedia Tools Appl.* 81 (22), 31847–31880. doi:10.1007/s11042-022-12078-w
- Gupta, S., Dwivedi, R. K., Kumar, V., Jain, R., Jain, S., and Singh, M. (2021). “Remote sensing image classification using deep learning,” in Proceedings of the 2021 10th International Conference on System Modeling and Advancement in Research Trends, SMART 2021, Moradabad, India, 10 – 11 December 2021, 274–279.
- Han, W., Zhang, X., Wang, Y., Wang, L., Huang, X., Li, J., et al. (2023). A survey of machine learning and deep learning in remote sensing of geological environment: challenges, advances, and opportunities. *ISPRS J. Photogrammetry Remote Sens.* 202, 87–113. doi:10.1016/j.isprs.2023.05.032
- Hermosilla, T., Wulder, M. A., White, J. C., and Coops, N. C. (2022). Land cover classification in an era of big and open data: optimizing localized implementation and training data selection to improve mapping outcomes. *Remote Sens. Environ.* 268, 112780. doi:10.1016/j.rse.2021.112780
- Hibbard, K., Janetos, A., van Vuuren, D. P., Pongratz, J., Rose, S. K., Betts, R., et al. (2010). Research priorities in land use and land-cover change for the Earth system and integrated assessment modelling. *Int. J. Climatol.* 30 (13), 2118–2128. doi:10.1002/joc.2150
- Holloway, J., and Mengersen, K. (2018). Statistical machine learning methods and remote sensing for sustainable development goals: a review. *Remote Sens.* 10 (9), 1365. doi:10.3390/rs10091365
- Hong, D., Hu, J., Yao, J., Chanussot, J., and Zhu, X. X. (2021). Multimodal remote sensing benchmark datasets for land cover classification with a shared and specific feature learning model. *ISPRS J. Photogrammetry Remote Sens.* 178, 68–80. doi:10.1016/j.isprs.2021.05.011
- Hu, W., Huang, Y., Wei, L., Zhang, F., and Li, H. (2015). Deep convolutional neural networks for hyperspectral image classification. *J. Sensors* 2015, 1–12. Article ID 258619. doi:10.1155/2015/258619
- Hütt, C., Waldhoff, G., and Bareth, G. (2020). Fusion of sentinel-1 with official topographic and cadastral geodata for crop-type enriched LULC mapping using FOSS and open data. *ISPRS Int. J. Geo-Information* 9 (2), 120. Article 2. doi:10.3390/ijgi9020120
- Ienco, D., Gaetano, R., Dupaquier, C., and Maurel, P. (2017). Land cover classification via multitemporal spatial data by deep recurrent neural networks. *IEEE Geoscience Remote Sens. Lett.* 14 (10), 1685–1689. doi:10.1109/lgrs.2017.2728698
- Ienco, D., Interdonato, R., Gaetano, R., and Minh, D. H. T. (2019). Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. *ISPRS J. Photogrammetry Remote Sens.* 158, 11–22. doi:10.1016/j.isprs.2019.09.016
- Interdonato, R., Ienco, D., Gaetano, R., and Ose, K. (2019). DuPLO: a DUal view Point deep Learning architecture for time series classification. *ISPRS J. photogrammetry remote Sens.* 149, 91–104. doi:10.1016/j.isprs.2019.01.011
- Jamali, A. (2021). Land use land cover modeling using optimized machine learning classifiers: a case study of Shiraz, Iran. *Model. Earth Syst. Environ.* 7 (3), 1539–1550. doi:10.1007/s40808-020-00859-x
- Jin, H., and Mountrakis, G. (2022). Fusion of optical, radar and waveform LiDAR observations for land cover classification. *ISPRS J. Photogrammetry Remote Sens.* 187, 171–190. doi:10.1016/j.isprs.2022.03.010
- Karpatne, A., Jiang, Z., Vatsavai, R. R., Shekhar, S., and Kumar, V. (2016). Monitoring land-cover changes: a machine-learning perspective. *IEEE Geoscience Remote Sens. Mag.* 4 (2), 8–21. doi:10.1109/MGRS.2016.2528038
- Kim, S., Wimmer, H., and Kim, J. (2022). “Analysis of DL libraries: Keras, PyTorch, and MXnet,” in 2022 IEEE/ACIS 20th International Conference on Software Engineering Research, Management and Applications (SERA), Las Vegas, USA, May 22–25, 2022, 54–62.
- Kotiridis, I., and Lazaridou, M. (2023). Cnns in land cover mapping with remote sensing imagery: a review and meta-analysis. *Int. J. Remote Sens.* 44 (19), 5896–5935. doi:10.1080/01431161.2023.2255354
- Kussul, N., Lavreniuk, M., Skakun, S., and Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience Remote Sens. Lett.* 14 (5), 778–782. doi:10.1109/lgrs.2017.2681128
- Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., et al. (2001). The causes of land-use and land-cover change: moving beyond the myths. *Glob. Environ. change* 11 (4), 261–269. doi:10.1016/s0959-3780(01)00007-3
- Lary, D. J., Alavi, A. H., Gandomi, A. H., and Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geosci. Front.* 7 (1), 3–10. doi:10.1016/j.gsf.2015.07.003
- Lei, J., Li, M., Xie, W., Li, Y., and Jia, X. (2021). Spectral Mapping with adversarial learning for unsupervised hyperspectral change detection. *Neurocomputing* 465, 71–83. doi:10.1016/j.neucom.2021.08.130
- Li, S., Song, W., Fang, L., Chen, Y., Ghamisi, P., and Benediktsson, J. A. (2019). Deep learning for hyperspectral image classification: an overview. *IEEE Trans. Geoscience Remote Sens.* 57 (9), 6690–6709. doi:10.1109/tgrs.2019.2907932
- Li, W., Fu, H., Yu, L., Gong, P., Feng, D., Li, C., et al. (2016). Stacked Autoencoder-based deep learning for remote-sensing image classification: a case study of African land-cover mapping. *Int. J. remote Sens.* 37 (23), 5632–5646. doi:10.1080/01431161.2016.1246775
- Liu, P. (2015). A survey of remote-sensing big data. *Front. Environ. Sci.* 3, 45. doi:10.3389/fenvs.2015.00045
- Liu, P., Choo, K. K. R., Wang, L., and Huang, F. (2017). SVM or DL? A comparative study on remote sensing image classification. *Soft Comput.* 21, 7053–7065. doi:10.1007/s00500-016-2247-2
- Liu, Z.-Q., Tang, P., Zhang, W., and Zhang, Z. (2022). CNN-enhanced heterogeneous graph convolutional network: inferring land use from land cover with a case study of park segmentation. *Remote Sens.* 14 (19), 5027. doi:10.3390/rs14195027
- Lv, Q., Dou, Y., Niu, X., Xu, J., Xu, J., and Xia, F. (2015). Urban land use and land cover classification using remotely sensed SAR data through deep belief networks. *J. Sensors* 2015, 1–10. doi:10.1155/2015/538063
- Magalhães, I. A. L., de Carvalho Júnior, O. A., de Carvalho, O. L. F., de Albuquerque, A. O., Hermuche, P. M., Merino, É. R., et al. (2022). Comparing machine and deep learning methods for the phenology-based classification of land cover types in the Amazon biome using sentinel-1 time series. *Remote Sens.* 14 (19), 4858. doi:10.3390/rs14194858
- Mahdianpari, M., Salehi, B., Rezaee, M., Mohammadimanes, F., and Zhang, Y. (2018). Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sens.* 10 (7), 1119. doi:10.3390/rs10071119
- Marmanis, D., Datcu, M., Esch, T., and Stilla, U. (2016). Deep learning earth observation classification using ImageNet pretrained networks. *IEEE Geoscience Remote Sens. Lett.* 13 (1), 105–109. doi:10.1109/LGRS.2015.2499239
- Masolele, R. N., Sy, V. D., Herold, M., Marcos, D., Verbesselt, J., Gieseke, F., et al. (2021). Spatial and temporal deep learning methods for deriving land-use following deforestation: a pan-tropical case study using Landsat time series. *Remote Sens. Environ.* 264, 112600. doi:10.1016/j.rse.2021.112600

- Maxwell, A. E., Warner, T. A., and Fang, F. (2018). Implementation of machine-learning classification in remote Sensing: an applied review. *Int. J. Remote Sens.* 39 (9), 2784–2817. doi:10.1080/01431161.2018.1433343
- Mou, L., Saha, S., Hua, Y., Bovolo, F., Bruzzone, L., and Zhu, X. X. (2022). Deep reinforcement learning for band selection in hyperspectral image classification. *IEEE Trans. Geoscience Remote Sens.* 60, 1–14. doi:10.1109/TGRS.2021.3067096
- Ndikumana, E., Ho Tong Minh, D., Baghdadi, N., Courault, D., and Hossard, L. (2018). Deep recurrent neural network for agricultural classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sens.* 10 (8), 1217. doi:10.3390/rs10081217
- Neupane, B., Aryal, J., and Rajabifard, A. (2023). *Cnns for remote extraction of urban features: A Survey-driven benchmarking (SSRN scholarly paper 4537529)*. doi:10.2139/ssrn.4537529
- Neupane, B., Horanont, T., and Aryal, J. (2021). Deep learning-based semantic segmentation of urban features in satellite images: a review and meta-analysis. *Remote Sens.* 13 (4), 808. doi:10.3390/rs13040808
- Nhu, V.-H., Hoang, N.-D., Nguyen, H., Ngo, P. T. T., Thanh Bui, T., Hoa, P. V., et al. (2020). Effectiveness assessment of Keras based deep learning with different robust optimization algorithms for shallow landslide susceptibility mapping at tropical area. *CATENA* 188, 104458. doi:10.1016/j.catena.2020.104458
- Olunfemi, I. E., Fasinmirin, J. T., Olufayo, A. A., and Komolafe, A. A. (2020). GIS and remote sensing-based analysis of the impacts of land use/land cover change (LULCC) on the environmental sustainability of Ekiti State, southwestern Nigeria. *Environ. Dev. Sustain.* 22 (2), 661–692. doi:10.1007/s10668-018-0214-z
- Ota, K., Dao, M. S., Mezaris, V., and Natale, F. G. D. (2017). Deep learning for mobile multimedia: a survey. *ACM Transactions on Multimedia Computing, Commun. Appl. (TOMM)* 13 (3s), 1–22. doi:10.1145/3092831
- Parashar, D., Kumar, A., Palni, S., Pandey, A., Singh, A., and Singh, A. P. (2024). Use of machine learning-based classification algorithms in the monitoring of Land Use and Land Cover practices in a hilly terrain. *Environ. Monit. Assess.* 196 (1), 8. doi:10.1007/s10661-023-12131-7
- Pashaie, M., Kamangir, H., Starek, M. J., and Tissot, P. (2020). Review and evaluation of deep learning architectures for efficient land cover mapping with uas hyper-spatial imagery: a case study over a wetland. *Remote Sens.* 12 (6), 959. Article 6. doi:10.3390/rs12060959
- Pešek, O., Segal-Rozenhaimer, M., and Karnieli, A. (2022). Using convolutional neural networks for cloud detection on VENUS images over multiple land-cover types. *Remote Sens.* 14 (20), 5210. doi:10.3390/rs14205210
- Raj, R., Luostarinen, T., Pursiainen, E., Posti, J. P., Takala, R. S., Bendel, S., et al. (2019). Machine learning-based dynamic mortality prediction after traumatic brain injury. *Sci. Rep.* 9 (1), 17672. doi:10.1038/s41598-019-53889-6
- Rindfuss, R. R., Walsh, S. J., Turner, B. L., Fox, J., and Mishra, V. (2004). Developing a science of land change: challenges and methodological issues. *Proc. Natl. Acad. Sci.* 101 (39), 13976–13981. doi:10.1073/pnas.0401545101
- Rousset, G., Despinoy, M., Schindler, K., and Mangeas, M. (2021). Assessment of deep learning techniques for land use land cover classification in southern New Caledonia. *Remote Sens.* 13 (12), 2257. doi:10.3390/rs13122257
- Sagan, V., Peterson, K. T., Maimaitijiang, M., Sidike, P., Sloan, J., Greeing, B. A., et al. (2020). Monitoring inland water quality using remote sensing: potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth-Science Rev.* 205, 103187. doi:10.1016/j.earscirev.2020.103187
- Saharan, M. A., Vyas, N., Borana, S. L., and Yadav, S. K. (2018). CLASSIFICATION AND ASSESSMENT OF THE LAND USE – LAND COVER CHANGES IN JODHPUR CITY USING REMOTE SENSING TECHNOLOGIES. *Int. Archives Photogrammetry, Remote Sens. Spatial Inf. Sci.* XLII-5, 767–771. doi:10.5194/isprs-archives-XLII-5-767-2018
- Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *Sn Comput. Sci.* 2 (6), 420. doi:10.1007/s42979-021-00815-1
- Sawant, S., Garg, R. D., Meshram, V., and Mistry, S. (2023). Sen-2 LULC: land use land cover dataset for deep learning approaches. *Data Brief* 51, 109724. doi:10.1016/j.dib.2023.109724
- Sertel, E., Ekim, B., Etehad Osgouei, P., and Kabadayi, M. E. (2022). Land use and land cover mapping using deep learning based segmentation approaches and VHR worldview-3 images. *Remote Sens.* 14 (18), 4558. doi:10.3390/rs14184558
- Sharma, A., Liu, X., Yang, X., and Shi, D. (2017). A patch-based convolutional neural network for remote sensing image classification. *Neural Netw.* 95, 19–28. doi:10.1016/j.neunet.2017.07.017
- Shih, H., Stow, D. A., and Tsai, Y. H. (2019). Guidance on and comparison of machine learning classifiers for Landsat-based land cover and land use mapping. *Int. J. Remote Sens.* 40 (4), 1248–1274. doi:10.1080/01431161.2018.1524179
- Shirmard, H., Farahbakhsh, E., Heidari, E., Beiranvand Pour, A., Pradhan, B., Müller, D., et al. (2022). A comparative study of convolutional neural networks and conventional machine learning models for lithological mapping using remote sensing data. *Remote Sens.* 14 (4), 819. doi:10.3390/rs14040819
- Silaparasetty, N. (2020) *Machine learning concepts with Python and the jupyter notebook environment: using tensorflow 2.0*. Berkeley, CA: Apress. doi:10.1007/978-1-4842-5967-2
- Singh, R. K., Singh, P., Drews, M., Kumar, P., Singh, H., Gupta, A. K., et al. (2021). A machine learning-based classification of LANDSAT images to map land use and land cover of India. *Remote Sens. Appl. Soc. Environ.* 24, 100624. doi:10.1016/j.rsase.2021.100624
- Solórzano, J. V., Mas, J. F., Gao, Y., and Gallardo-Cruz, J. A. (2021). Land use land cover classification with U-net: advantages of combining sentinel-1 and sentinel-2 imagery. *Remote Sens.* 13 (18), 3600. doi:10.3390/rs13183600
- Song, Q., Yang, F., Yang, L., Liu, C., Hu, M., and Xia, L. (2021). Learning point-guided localization for detection in remote sensing images. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sens.* 14, 1084–1094. doi:10.1109/JSTARS.2020.3036685
- Southworth, J., and Muir, C. (2021). Specialty grand challenge: remote sensing time series analysis. *Front. Remote Sens.* 2. doi:10.3389/frsen.2021.770431
- St-Onge, C., Kara, N., Wahab, O. A., Edstrom, C., and Lemieux, Y. (2020). Detection of time series patterns and periodicity of cloud computing workloads. *Future Gener. Comput. Syst.* 109, 249–261. doi:10.1016/j.future.2020.03.059
- Tamimnia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S., and Brisco, B. (2020). Google Earth Engine for geo-big data applications: a meta-analysis and systematic review. *ISPRS J. photogrammetry remote Sens.* 164, 152–170. doi:10.1016/j.isprsjprs.2020.04.001
- Tassi, A., and Vizzari, M. (2020). Object-oriented LULC classification in Google earth engine combining SNIC, GLCM, and machine learning algorithms. *Remote Sens.* 12 (22), 3776. doi:10.3390/rs12223776
- Taye, M. M. (2023). Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers* 12 (5), 91. doi:10.3390/computers12050091
- Terra, J. (2020). Pytorch vs tensorflow vs Keras: here are the difference you should know. Available at: <https://www.simplilearn.com/keras-vs-tensorflow-vs-pytorch-article>.
- Thanga Manickam, M., Karthik Rao, M., Barath, K., Shree Vijay, S., and Karthi, R. (2022). “Convolutional neural network for land cover classification and mapping using Landsat images,” in *Innovations in computer science and engineering*, 385 Editors H. S. Saini, R. Sayal, A. Govardhan, and R. Buyya (Singapore: Springer), 221–232. doi:10.1007/978-981-16-8987-1_24
- Truong, T. X., Nhu, V.-H., Phuong, D. T. N., Nghi, L. T., Hung, N. N., Hoa, P. V., et al. (2023). A new approach based on TensorFlow deep neural networks with ADAM optimizer and GIS for spatial prediction of forest fire danger in tropical areas. *Remote Sens.* 15 (14), 3458. doi:10.3390/rs15143458
- Turner, B. L., Lambin, E. F., and Reenberg, A. (2007). The emergence of land change science for global environmental change and sustainability. *Proc. Natl. Acad. Sci.* 104 (52), 20666–20671. doi:10.1073/pnas.0704119104
- Turner, B. L., Lambin, E. F., and Verburg, P. H. (2021). From land-use/land-cover to land system science: this article belongs to Ambio’s 50th Anniversary Collection. Theme: agricultural land use. *Ambio* 50, 1291–1294. doi:10.1007/s13280-021-01510-4
- Turner II, B. L., Geoghegan, J., Lawrence, D., Radel, C., Schmook, B., Vance, C., et al. (2016). Land system science and the social–environmental system: the case of Southern Yucatán Peninsular Region (SYPR) project. *Curr. Opin. Environ. Sustain.* 19, 18–29. doi:10.1016/j.cosust.2015.08.014
- Verburg, P. H., Crossman, N., Ellis, E. C., Heinimann, A., Hostert, P., Mertz, O., et al. (2015). Land system science and sustainable development of the earth system: a global land project perspective. *Anthropocene* 12, 29–41. doi:10.1016/j.ancene.2015.09.004
- Verburg, P. H., Dearing, J. A., Dyke, J. G., Van Der Leeuw, S., Seitzinger, S., Steffen, W., et al. (2016). Methods and approaches to modelling the Anthropocene. *Glob. Environ. Change* 39, 328–340. doi:10.1016/j.gloenvcha.2015.08.007
- Verburg, P. H., Erb, K. H., Mertz, O., and Espindola, G. (2013). Land System Science: between global challenges and local realities. *Curr. Opin. Environ. Sustain.* 5 (5), 433–437. doi:10.1016/j.cosust.2013.08.001
- Vivekananda, G., Swathi, R., and Sujith, A. (2021). Multi-temporal image analysis for LULC classification and change detection. *Eur. J. Remote Sens.* 54 (Suppl. 2), 189–199. doi:10.1080/22797254.2020.1771215
- Wan, S., Gong, C., Zhong, P., Du, B., Zhang, L., and Yang, J. (2019). Multiscale dynamic graph convolutional network for hyperspectral image classification. *IEEE Trans. Geoscience Remote Sens.* 58 (5), 3162–3177. doi:10.1109/tgrs.2019.2949180
- Wang, J., Bretz, M., Dewan, M. A. A., and Delavar, M. A. (2022a). Machine learning in modelling land-use and land cover-change (LULCC): current status, challenges and prospects. *Sci. Total Environ.* 822, 153559. doi:10.1016/j.scitotenv.2022.153559
- Wang, J., Zhong, Y., Zheng, Z., Ma, A., and Zhang, L. (2021). RSNNet: the search for remote sensing deep neural networks in recognition tasks. *IEEE Trans. Geoscience Remote Sens.* 59 (3), 2520–2534. doi:10.1109/TGRS.2020.3001401
- Wang, L., Wang, J., Liu, Z., Zhu, J., and Qin, F. (2022b). Evaluation of a deep-learning model for multispectral remote Sensing of land use and crop classification. *Crop J.* 10 (5), 1435–1451. doi:10.1016/j.cj.2022.01.009

- Warmerdam, F. (2008). "The geospatial data abstraction library," in *Open source approaches in spatial data handling* (Berlin, Heidelberg: Springer Berlin Heidelberg), 87–104.
- Wu, H., and Prasad, S. (2017). Convolutional recurrent neural networks for hyperspectral data classification. *Remote Sens.* 9 (3), 298. doi:10.3390/rs9030298
- Xie, G., and Niculescu, S. (2021). Mapping and monitoring of land cover/land use (LCU) changes in the crozon peninsula (Brittany, France) from 2007 to 2018 by machine learning algorithms (support vector machine, random forest, and convolutional neural network) and by post-classification comparison (PCC). *Remote Sens.* 13 (19), 3899. doi:10.3390/rs13193899
- Xie, Q., Dash, J., Huang, W., Peng, D., Qin, Q., Mortimer, H., et al. (2018). Vegetation indices combining the red and red-edge spectral information for leaf area index retrieval. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sens.* 11, 1482–1493. doi:10.1109/jstars.2018.2813281
- Xu, X., Chen, Y., Zhang, J., Chen, Y., Anandhan, P., and Manickam, A. (2021). A novel approach for scene classification from remote sensing images using deep learning methods. *Eur. J. Remote Sens.* 54 (Suppl. 2), 383–395. doi:10.1080/22797254.2020.1790995
- Yang, D., Fu, C. S., Smith, A. C., and Yu, Q. (2017). Open land-use map: a regional land-use mapping strategy for incorporating OpenStreetMap with earth observations. *Geo-spatial Inf. Sci.* 20 (3), 269–281. doi:10.1080/10095020.2017.1371385
- Yao, S., Chen, H., Thompson, E. J., and Cifelli, R. (2022). An improved deep learning model for high-impact weather nowcasting. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sens.* 15, 7400–7413. doi:10.1109/jstars.2022.3203398
- Yilmaz, T., Olgun, R., and Ak, M. K. (2020). NDVI indicated changes in vegetation and their relations to climatic comfort factors in Demre-Akçay Sub basin, Turkey. *J. Environ. Biol.* 41 (2), 344–350. doi:10.22438/jeb/41/2(SI)/JEB-10
- Yomo, M., Yalo, E. N., Gnazou, M. D. T., Silliman, S., Larbi, I., and Mourad, K. A. (2023). Forecasting land use and land cover dynamics using combined remote sensing, machine learning algorithm and local perception in the Agoènyivé Plateau, Togo. *Remote Sens. Appl. Soc. Environ.* 30, 100928. doi:10.1016/j.rsase.2023.100928
- Yu, S., Jia, S., and Xu, C. (2017). Convolutional neural networks for hyperspectral image classification. *Neurocomputing* 219, 88–98. doi:10.1016/j.neucom.2016.09.010
- Zaabar, N., Niculescu, S., and Kamel, M. M. (2022). Application of convolutional neural networks with object-based image analysis for land cover and land use mapping in coastal areas: a case study in Ain Témouchent, Algeria. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sens.* 15, 5177–5189. doi:10.1109/jstars.2022.3185185
- Zhang, C., Harrison, P. A., Pan, X., Li, H., Sargent, I., and Atkinson, P. M. (2020a). Scale Sequence Joint Deep Learning (SS-JDL) for land use and land cover classification. *Remote Sens. Environ.* 237, 111593. doi:10.1016/j.rse.2019.111593
- Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., et al. (2018). An object-based convolutional neural network (OCNN) for urban land use classification. *Remote Sens. Environ.* 216, 57–70. doi:10.1016/j.rse.2018.06.034
- Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., et al. (2019). Joint Deep Learning for land cover and land use classification. *Remote Sens. Environ.* 221, 173–187. doi:10.1016/j.rse.2018.11.014
- Zhang, X., Han, L., Han, L., and Zhu, L. (2020b). How well do DL-based methods for land cover classification and object detection perform on high resolution remote sensing imagery? *Remote Sens.* 12 (3), 417. doi:10.3390/rs12030417
- Zhang, Y., Chen, G., Myint, S. W., Zhou, Y., Hay, G. J., Vukomanovic, J., et al. (2022). UrbanWatch: a 1-meter resolution land cover and land use database for 22 major cities in the United States. *Remote Sens. Environ.* 278, 113106. doi:10.1016/j.rse.2022.113106
- Zhou, G., Chen, W., Qin, X., Li, J., and Wang, L. (2023a). Lithological unit classification based on geological knowledge-guided deep learning framework for optical stereo mapping satellite imagery. *IEEE Trans. Geoscience Remote Sens.* 61, 1–16. doi:10.1109/tgrs.2023.3327774
- Zhou, G., Xu, J., Chen, W., Li, X., Li, J., and Wang, L. (2023b). Deep feature enhancement method for land cover with irregular and sparse spatial distribution features: a case study on open-pit mining. *IEEE Trans. Geoscience Remote Sens.* 61, 1–20. doi:10.1109/tgrs.2023.3241331
- Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., et al. (2017). Deep learning in remote sensing: a comprehensive review and list of resources. *IEEE geoscience remote Sens. Mag.* 5 (4), 8–36. doi:10.1109/mgrs.2017.2762307