IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 May 2026, accepted 12 June 2026, date of publication 22 June 2026, date of current version 29 June 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3705849

SURVEY

A Decade of Deep Learning in Remote Sensing: A Systematic Review of Reviews

ANASTASIIA SAFONOVA[1,2], RONNY HÄNSCH[3], (Senior Member, IEEE), RIBANA ROSCHER[4,5], GILBERTO CAMARA[6], BEGÜM DEMIR[7,8], (Senior Member, IEEE), JOAN SERRA-SAGRISTÀ[2], (Senior Member, IEEE), AND MASAHIRO RYO[1,9]

[1]Leibniz Centre for Agricultural Landscape Research (ZALF), 15374 Müncheberg, Germany
[2]Department of Information and Communications Engineering, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Spain
[3]German Aerospace Center (DLR), 82234 Weßling, Germany
[4]Institute of Geodesy and Geoinformation, University of Bonn, 53115 Bonn, Germany
[5]Institute of Bio- and Geosciences (IBG-2), Forschungszentrum Jülich GmbH, 52428 Jülich, Germany
[6]National Institute for Space Research (INPE), São José dos Campos, São Paulo 12227-010, Brazil
[7]Berlin Institute for the Foundations of Learning and Data (BIFOLD), 10587 Berlin, Germany
[8]Faculty of Electrical Engineering and Computer Science, Technische Universität Berlin, 10587 Berlin, Germany
[9]Environment and Natural Sciences Department, Brandenburg University of Technology Cottbus–Senftenberg, 03046 Cottbus, Germany

Corresponding author: Anastasiia Safonova (safonova.nastya1@gmail.com)

This work was supported in part by the Federal Ministry for Research, Technology, and Space, Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) under Project 16DKWN089; in part by the Marie Skłodowska-Curie Fellowship through the Horizon Europe Program, RAMON LLULL, under Grant 101126667; and in part by the Spanish Ministry of Science, Innovation, and Universities (MICIU) and the European Regional Development Fund (FEDER), funded by MICIU/Agencia Estatal de Investigación (AEI)/10.13039/501100011033/FEDER, Unión Europea (UE), under Grant PID2021-125258OB-I00 and Grant PID2024-156292OB-I00.

ABSTRACT Since 2014, Deep Learning (DL) has emerged as a powerful technology for analyzing Remote Sensing (RS) data, enabling advancements in land cover classification, object detection, change detection, and other RS tasks. With the rapid growth of research in this area, nearly ten thousand articles have been published with hundreds of review articles. This review paper aims to map the evolution of DL applications in the RS research domain, chronicling the major achievements from their widespread adoption from 2016 to March 2026. Instead of targeting a specific topic, we performed a systematic review of peer-reviewed "review" articles (namely, "*review of reviews*"). Following the scoping review framework, we identified and summarized 448 review articles across four major scientific databases: Web of Science, Scopus, IEEE Xplore, and arXiv. The included reviews cover convolutional neural networks, vision transformers, generative adversarial networks and hybrid architectures, as well as transfer learning, self-supervised learning and Foundation Models (FMs). These are applied to a range of applications, including agriculture, urban analysis, disaster monitoring, climate science, marine environments and hyperspectral imaging. A bibliometric analysis reveals that the annual growth rate for review publications is 74.9%, with agriculture and environmental monitoring being the most prevalent application domains. The review shows that, while the field has progressed from convolutional network-dominated pipelines towards transformer-based and FM architectures, persistent challenges in data scarcity, domain generalization, benchmark fragmentation and computational efficiency continue to constrain operational deployment. This article provides a comprehensive overview of achievements in DL applications for the RS domain, drawing on over a decade of research to offer evidence-based recommendations for future research.

INDEX TERMS Applications, deep learning, remote sensing, review of reviews.

The associate editor coordinating the review of this manuscript and approving it for publication was Gerardo Di Martino.

I. INTRODUCTION

Deep Learning (DL) has become a transformative technology in Remote Sensing (RS), enhancing the analysis and

© 2026 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

interpretation of remotely sensed data through automatic feature learning from large datasets. This has led to notable gains in accuracy and efficiency. The rapidly expanding Artificial intelligence (AI) in RS literature spans diverse applications, including agriculture [1], [2], environmental monitoring [3], urban planning [4], disaster management [5], and natural resource exploration [6], [7].

Since 2014, RS researchers have increasingly applied DL algorithms to image analysis tasks such as scene classification and object detection [8], [9]. DL, based on Artificial Neural Networks (ANNs) with many hidden layers, enables complex feature extraction and representation learning. Various architectures address specific tasks: Convolutional Neural Networks (CNNs), including AlexNet [10], were among the first adapted for RS; Recurrent Neural Networks (RNNs) address sequential data [11]; autoencoders support unsupervised feature learning [12]; Generative Adversarial Networks (GANs) enable image synthesis and augmentation [13]; and Long Short-Term Memory networks (LSTMs) model long-range temporal dependencies [14]. The AI in RS literature has grown rapidly, reaching nearly ten thousand publications by March 2026 (Figure 1).

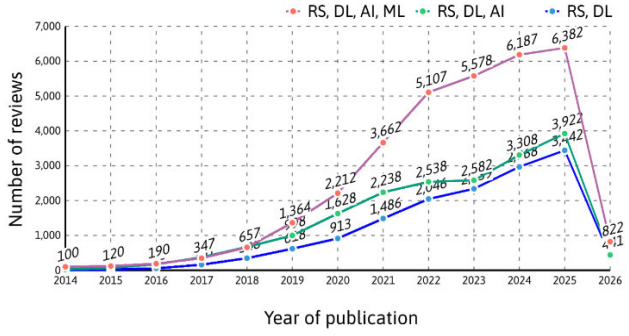

FIGURE 1. The number of papers since 2014 with the implementation of AI algorithms in RS domains based on the Web of Science database.

With the rapid expansion of the field, the number of review papers has grown steadily, now numbering in the hundreds, each contributing to consolidating knowledge within specific themes. A systematic review is normally conducted on a specific topic to deepen understanding within that area. However, this approach does not broaden generic understanding across topics. Offering a comprehensive overview of this rapidly expanding field is essential for researchers entering the domain and for understanding how AI in RS has evolved over the first decade of widespread adoption. No article providing such a cross-topic synthesis had been identified at the time of this study, and while several textbooks exist, they cannot guarantee coverage of all major themes and may reflect individual biases. There is therefore a need to systematically analyze peer-reviewed review articles to understand the full extent of the field.

The goal of this study is to comprehend the first decade of progress (2016–2026) in DL applied to RS by mapping a large number of reviewed themes through a review of review articles. While not providing deep understanding of any single topic, this approach efficiently maps the major themes, achievements, and gaps identified across the literature. The review identifies the principal application domains that have received attention over the decade, evaluates the performance of DL algorithms and architectures through synthesis of the reviewed literature, critically assesses methods including data preprocessing, network architectures, and optimization strategies, and identifies emerging trends, challenges, and future research directions that cannot be discerned by analyzing any single specific topic.

The scope of the review encompasses image classification, object detection, change detection, land cover mapping, time series analysis, and multimodal data fusion. The integration of multi-source data, Transfer Learning (TL), Self-Supervised Learning (SSL), and Explainable Artificial Intelligence (XAI) techniques are explored to understand advances in interpretability, generalizability, and decision support. Particular attention is given to the 2024–2026 literature, which documents the emergence of Foundation Models (FMs), diffusion-based generative architectures, state space models, and large-scale geospatial pretraining strategies that are reshaping the field's methodological trajectory.

The review is guided by the following research questions (RQs):

RQ1: What are the main application domains and thematic clusters of DL in RS covered in existing review articles?

RQ2: Which DL architectures and learning paradigms have been most frequently reported in RS reviews, and how has their relative emphasis evolved over the period 2016–2026?

RQ3: What methodological challenges, limitations, and future research directions are consistently identified across these reviews?

This review is organized as follows. Section II describes the scoping review methodology. Section III presents the results, covering general characteristics of included reviews, dominant DL architectures, learning paradigms, and explainability. Section IV discusses the principal challenges facing the field. Section V outlines future research directions. Section VI concludes the paper.

II. METHODS

This review follows the scoping review methodology proposed by Arksey and O'Malley [15], refined by Levac et al. [16] and Peters et al. [17]. This framework synthesizes knowledge by outlining prevailing trends and key issues within a research domain and is well suited to mapping the breadth of DL applications in RS. The methodology involves five steps [18]: (i) identifying relevant studies, (ii) selecting studies, (iii) charting data, (iv) summarizing and reporting results, and (v) identifying challenges and future directions.

A. DATA SOURCE AND SEARCH CRITERIA

A comprehensive literature search was conducted across the Web of Science Core Collection, Scopus, IEEE Xplore, and arXiv. The final search was performed on March 20, 2026, using refined domain-specific keywords: ''deep learning'' AND ''remote sensing'' AND (''review'' OR ''survey'' OR ''overview''). The search was not restricted by publication year or study type, except for limiting results to peer-reviewed review articles.

B. ELIGIBILITY CRITERIA

The eligibility criteria for inclusion were: (i) peer-reviewed review articles; (ii) full text accessible in English; and (iii) explicit inclusion of both ''deep learning'' and ''remote sensing'' as core topics.

Because many DL review papers first appear on arXiv prior to peer review, we included arXiv in the search stage to ensure comprehensive coverage. In total, we identified 60 review papers available on arXiv. For 29 of these, a peer-reviewed journal or conference version existed, and we used the peer-reviewed version in our dataset. The remaining 31 papers were available only on arXiv at the time of data collection. These were retained only when they presented a clear review methodology and strong relevance to deep learning in remote sensing. These arXiv-only reviews were included based on qualitative judgment of methodological clarity and topical relevance, and no formal quality scoring rubric or inter-rater reliability statistics were applied, which we acknowledge as a limitation of this review.

C. SELECTION OF REVIEWS

The selection of relevant review articles followed three sequential steps: (i) screening titles and abstracts to exclude papers not meeting the inclusion criteria; (ii) assessing full texts to remove non-relevant studies; and (iii) obtaining the full texts of all remaining articles for detailed analysis and synthesis.

To ensure consistency during the screening process, the initial title and abstract screening were performed by one author, while all ambiguous or borderline cases were jointly discussed among the co-authors until consensus was reached. We did not compute Cohen's Kappa or a similar inter-rater agreement statistic on a random subsample. Therefore, screening consistency cannot be quantified and should be considered a methodological limitation.

D. DATA SUMMARY AND SYNTHESIS

Articles were compiled into a single electronic table for screening. All filtered records retrieved from the four databases were exported in standardized bibliographic formats and merged into a harmonized dataset to support screening and subsequent synthesis. Duplicate entries were identified through normalized title matching.

III. RESULTS

A. GENERAL CHARACTERISTICS OF INCLUDED REVIEWS

1) SEARCH AND SELECTION PROCESS

The final literature search across four major databases retrieved 1,433 records before deduplication. Merging and deduplication resulted in 448 unique peer-reviewed review articles that met the eligibility criteria and were included in the final synthesis. The deduplication step revealed substantial overlap between databases, particularly between Scopus and Web of Science. Among the included reviews, 276 appeared in only one database and 172 were indexed in two or more databases. This overlap explains why per-database counts in the PRISMA diagram (Figure 2) cannot be summed [19], [20]. Scopus contributed the largest share of included records, followed by Web of Science, IEEE Xplore, and arXiv. A structured summary table and the full list of included reviews are publicly available (DOI: https://doi.org/10.5281/zenodo.19470643).

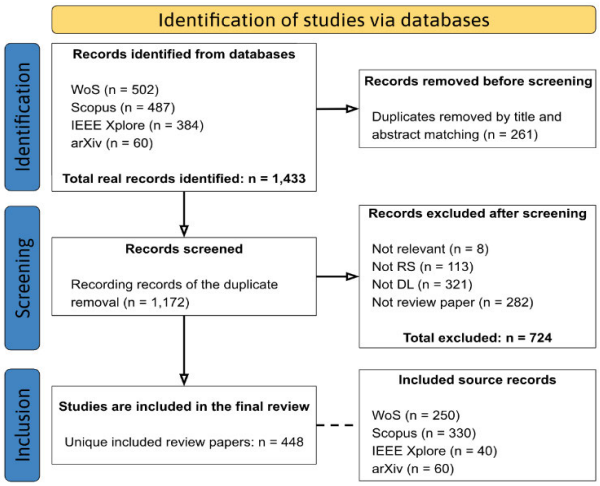

FIGURE 2. PRISMA flow diagram showing the identification, screening and inclusion of review articles from four databases: Web of Science, Scopus, IEEE Xplore and arXiv.

2) PUBLICATION TRENDS IN THE PAST TEN YEARS

The annual growth rate of review articles on DL in RS is 74.9%, reflecting a strong long-term upward trend despite the expected decline in the incomplete final year (Figure 3A). Most publications focus on general RS itself (48.4%, 166/343), followed by agriculture (18.8%, 84/448), climate and environmental studies (9.8%, 44/448), Hyperspectral Imaging (HSI) and Multispectral Imaging (MSI) (9.8%, 44/448), and urban and infrastructure applications (8.7%, 39/448) (Figure 3B).

3) MAJOR APPLICATION DOMAINS

DL has been applied across a wide spectrum of RS domains, with substantial growth in both methodological sophistication and application diversity. Analysis of 448 review papers indicates nine major application categories: agriculture, soil, water, ecology, vegetation, forestry, geoscience, climate,

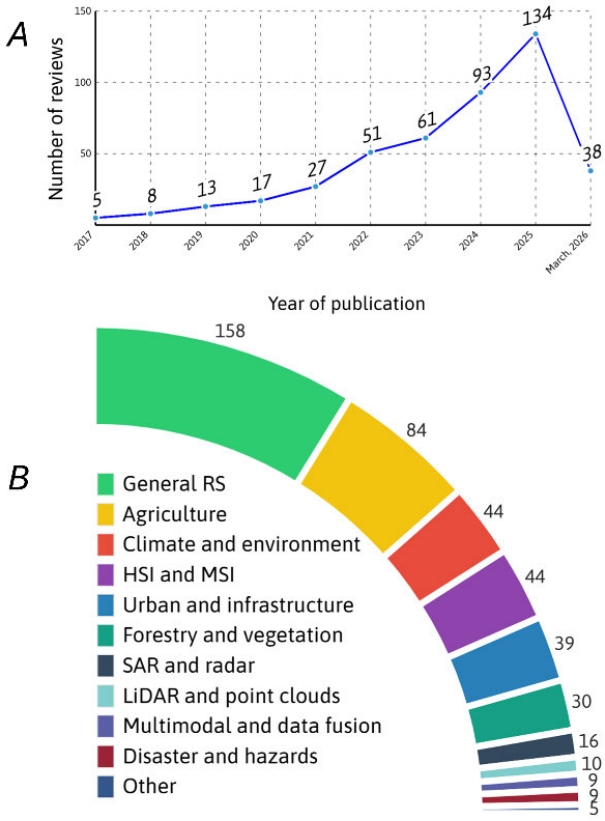

FIGURE 3. Main characteristics of the included reviews from 2016 to 2026, where *A* is the number of review articles per year and *B* is subject areas of the papers.

and urban environments. Several emerging clusters, including marine science, RS video analysis, 3D point cloud processing, and multimodal fusion, have developed into distinct research subfields.

Agriculture and precision farming. Agriculture remains the most prolific DL in RS domain. Reviews covering more than 650 publications identify CNNs as the dominant architecture for crop monitoring tasks including disease detection, yield forecasting, and field classification, while highlighting persistent challenges such as annotated dataset scarcity, cross-environment generalization, and multi-source data fusion [21]. Semantic segmentation and transformer-based field boundary delineation have emerged as a specialized subfield with growing emphasis on domain generalization and FMs for large scale mapping [22].

Crop specific reviews document DL applications for cotton [23], wheat [24], [25], potato [26], sugarcane [27], and rubber [28]. Diffusion models have been applied to agricultural data augmentation and pest detection [29]. Soil monitoring has been reviewed extensively across organic carbon estimation [30], [31], moisture retrieval [32], and nutrient mapping using Hyperspectral imaging (HSI) [33]. Reported yield prediction models under abnormal climate conditions achieve RMSE values of 0.2–0.5 t/ha and R^2 values of 0.4–0.9 across Machine Learning (ML), DL, ensemble, and XAI approaches [34]. UAV based HSI for nitrogen monitoring consistently identifies near infrared and red edge bands as most informative [35]. Federated learning has been proposed for privacy preserving collaborative model training across distributed farm networks [36].

Forestry, vegetation, and ecosystems. Forestry has become a major growth domain, driven by UAV platforms, LiDAR sensors, and multi-temporal satellite imagery. Reviews of 181 studies document a shift from structure from motion photogrammetry toward integrated frameworks combining canopy metrics, multispectral or LiDAR data, and DL models for biomass and volume estimation [37]. Additional reviews cover forest carbon stock estimation [38], mangrove species mapping [39], woody plant encroachment [40], forest structural parameter modeling [41], and eucalypt dieback detection [42], confirming DL as a viable complement to allometric methods.

Urban and built environments. Urban RS has expanded beyond land cover mapping to include building geometry, ecological quality, thermal environments, and hydrological hazards. A review of 243 studies on building extraction identifies five persistent challenges, complex boundaries, spectral variability, limited training samples, tree and shadow interference, and lightweight model requirements, and documents DL solutions including semantic segmentation and Mamba based architectures [43]. Urban flood prediction has progressed from physical hydrodynamic models to AI systems integrating RS, LiDAR, IoT, and crowdsourced data, with LSTM and Random Forest models achieving 95–98% accuracy and reducing computation time from hours to seconds [44]. Additional reviews cover urban ecological quality [45], heat island monitoring [46], nighttime light analysis [47], and blue space mapping across 1369 publications [48].

Disaster monitoring and hazards. Disaster monitoring is one of the fastest growing DL-RS clusters. Wildfire detection reviews covering 27 detection and 16 emissions studies report CNN and YOLO accuracies of 79.4–99.9% and document end-to-end uncertainty chains from detection to atmospheric impact assessment [49]. Comprehensive wildfire risk assessments identify hybrid multi-source models as most effective [50]. Landslide identification has expanded to more than 400 studies using optical, Synthetic Aperture Radar (SAR), and DEM inputs for susceptibility mapping and multi-temporal change analysis [51]. Multi hazard detection using satellite-based DL has also been reviewed with emphasis on near-real-time architectures [52].

Climate and environmental monitoring. Climate and environmental applications have broadened significantly. Air pollution monitoring reviews covering 241 studies document improvements in spatial resolution from 10 km to sub-kilometer scales and report a mean 28% reduction in major pollutants during COVID-19 lockdowns [53]. Evapotranspiration estimation reviews identify hybrid physics DL models as offering the best balance of interpretability and accuracy [54]. Sea ice concentration monitoring using SAR demonstrates that DL outperforms

conventional algorithms [55]. Plastic waste monitoring in rivers and coastal zones bridges the gap between satellite scale coverage and field level detection [56], [57].

Marine and coastal applications. Marine and coastal RS has consolidated into a distinct research cluster. Coastal aquaculture mapping reviews covering 304 studies report that DL accounts for 43.5% of approaches [58]. Additional reviews synthesize ocean color ML [59], coastal water quality monitoring [60], harmful algal bloom detection [61], and real-time marine pollution monitoring including oil spills using U-Net, Mask R-CNN, and YOLOs [62]. DL based climate ocean studies report U-Net LSTM achieving root mean squared errors of 0.48–0.63 degrees Celsius for sea surface temperature prediction [63].

Other domains. SAR specific DL has emerged as a semi-independent cluster, with reviews covering ship interpretation [64], optical SAR fusion [65], and SAR scene classification, target recognition, and change detection [66]. HSI reviews document UAV based classification from classical ML to GCN and CNN transformer hybrids [67], hyperspectral unmixing using autoencoder and variational architectures [68], and LULC classification frameworks confirming 2D plus 3D CNNs as state of the art [69], [70]. Change detection reviews identify sample scarcity, multimodal fusion, and semantic complexity as key challenges, with FMs such as SAM CD and ChangeViT demonstrating improved accuracy [71]. RS video analysis remains an emerging domain with ten open challenges including lightweight tracker design and cross domain transfer [72].

4) MAJOR KEYWORDS AND THEMATIC CLUSTERS

Keyword co-occurrence analysis revealed several dominant research themes. We first selected the 100 most frequent author keywords and constructed a symmetric co-occurrence matrix, where edges represent keyword pairs that appeared together in at least two review articles. Based on this matrix, we built an undirected weighted graph and applied the Louvain community detection algorithm to identify coherent thematic groups. The largest connected component was visualized using a stress-based layout. Node size reflects keyword frequency, and edge width represents the strength of co-occurrence. The four main communities are shown in Figure 4. The most frequent terms were DL (235 occurrences), RS (194), ML (127), AI (56), Uncrewed Aerial Vehicle (UAV) (33), CNN (38), semantic segmentation (20), computer vision (19) and HSI (39). These high-frequency terms emphasize the pivotal role of AI-driven image analysis, sensor technologies, and environmental monitoring in modern RS research. The co-occurrence network formed four major thematic clusters, each representing a distinct but interconnected research direction.

The first cluster, marked in red, is the largest and most cohesive. It is dominated by DL (235), CNN (38), object detection (27), image classification (16), SAR (15) and GAN (12). This cluster reflects the intensive use of DNNs for classification, segmentation, change detection and high-resolution image reconstruction. Topics such as super-resolution, optical RS, instance segmentation, weakly supervised learning, and data augmentation demonstrate a strong methodological focus on enhancing model accuracy and robustness. Applications include crop monitoring, forestry, ship detection and building extraction, demonstrating the broad applicability of DL techniques across environmental and geospatial domains.

The second cluster in blue centers on RS (194), satellite imagery (23), and semantic segmentation (20). It reflects the broader Earth Observation (EO) landscape, including few-shot learning, water quality, vegetation mapping, and land use/land cover analysis. The presence of self-supervised learning, multimodal approaches, and satellite data suggests an emerging shift toward FMs and hybrid architectures for large-scale EO tasks. This cluster captures the methodological diversity of EO research, from classical mapping to advanced transformer-based architectures.

The third cluster, represented by the green color, is centered around ML (127), AI (56) and UAVs (33). This cluster highlights the integration of AI with drone-based sensing, precision agriculture (26), climate change (10), environmental monitoring (4) and mineral exploration (4). The presence of NNs, image processing, and image segmentation demonstrates that ML remains fundamental to UAV-based analytics and agricultural decision support.

The fourth cluster in yellow focuses on advanced sensing technologies, led by HSI (39), computer vision (19), MSI (14), and LiDAR (11). These modalities support applications in land cover mapping, feature extraction, and environmental assessment. Keywords such as sensor fusion (3) and GNSS (3) highlight the integration of multiple data sources to enhance spatial accuracy and spectral richness. This cluster represents the technological backbone of modern RS, emphasizing high-dimensional data and multimodal fusion.

B. POPULAR DEEP LEARNING ARCHITECTURES IN REMOTE SENSING

The landscape of DL architectures applied to RS has expanded substantially over the past decade. While early applications relied predominantly on standard CNNs adapted from general-purpose computer vision, the field has evolved toward specialized and heterogeneous architectures tailored to the unique properties of RS data, such as high spatial resolution, spectral diversity, multi-source sensor configurations, and large geographic coverage.

The organization of this section follows the taxonomy of DL architectures commonly used across the included review articles. Because the reviews differ widely in scope and do not consistently report temporal or architecture-specific statistics, synthesizing year-by-year architectural trends is not feasible within a review-of-reviews framework.

The overall DL workflow for RS data, comprising six main stages from data collection to model updating, is illustrated in Figure 5.

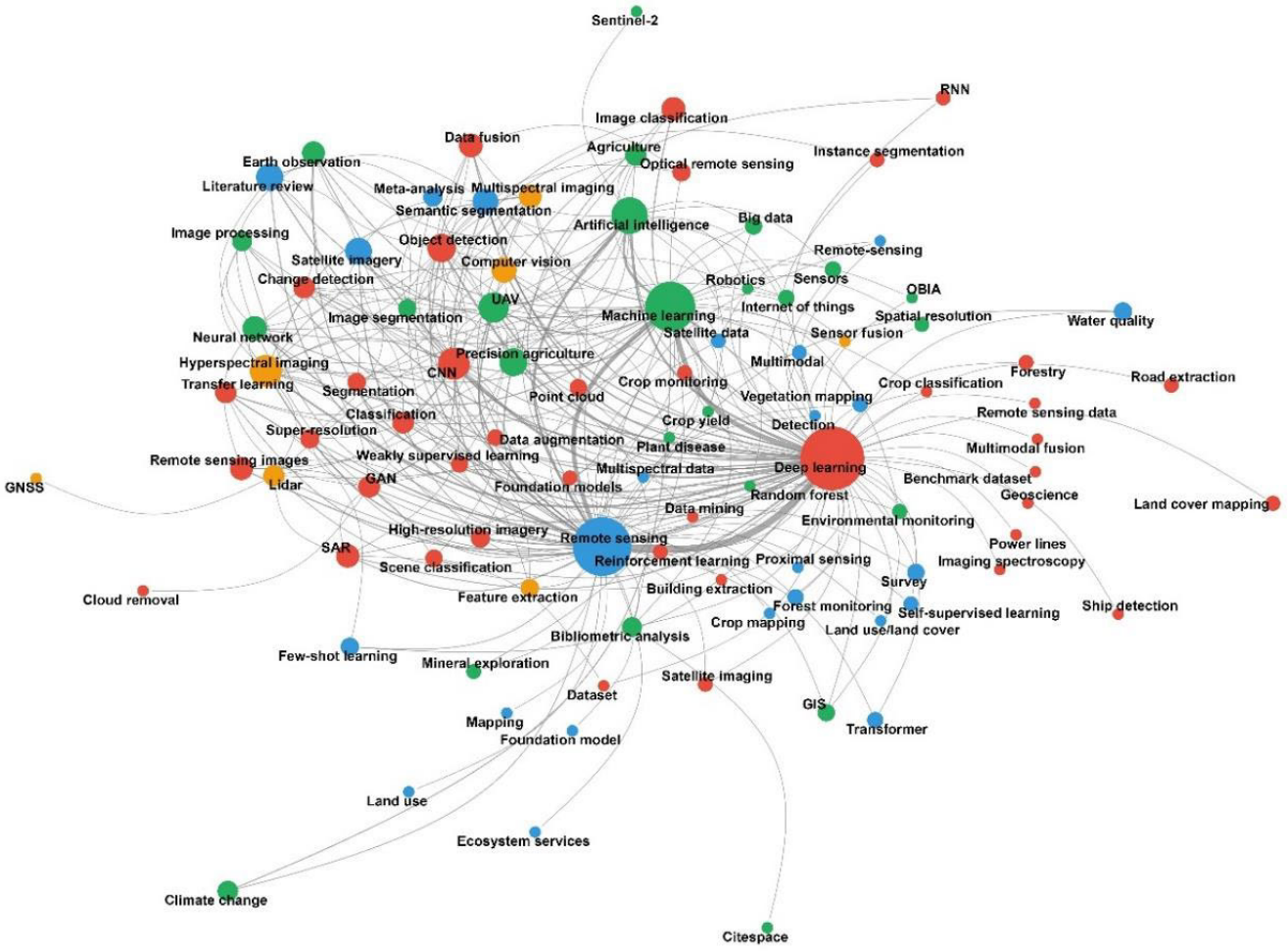

FIGURE 4. Keyword co-occurrence knowledge graph.

1) CONVOLUTIONAL NEURAL NETWORKS

CNNs remain the most widely deployed DL architecture in remote sensing (RS). Their use has evolved from direct adaptation of general-purpose backbones toward architectures tailored to the structural, spectral, and geometric properties of RS data. Four major design directions have emerged: (i) lightweight CNNs optimized for resource-constrained deployment [73], [74], [75]; (ii) domain-specific architectures targeting particular sensors or applications [76], [77], [78]; (iii) multi-scale feature extraction networks for objects spanning diverse spatial extents [79], [80], [81]; and (iv) attention-augmented CNNs that selectively weight spatial or channel-wise representations [82], [83], [84].

For HSI, three main CNN strategies are used: spectral-only models, now relatively uncommon; spatial CNNs, which dominate object classification; and combined spectral–spatial models, which achieve the highest accuracy [85].

One-, two-, and three-dimensional CNNs account for approximately 8%, 88%, and 4% of HSI studies, respectively [1]. A combined 2D+3D CNN framework reports the highest overall accuracy, average accuracy, and kappa coefficient across the Indian Pines, Pavia, and Salinas benchmarks [70].

CNNs are also widely applied to SAR data, with publications increasing by approximately 88% since 2016. CNN–GEOBIA hybrids (GeoCNN) improve object segmentation and classification by integrating spatial context with learned features [86].

Lightweight CNNs have become increasingly important for onboard and edge-computing scenarios, where power, memory, and latency constraints limit model complexity. For HSI processing on resource-constrained platforms, 1D CNNs offer a feasible alternative to 2D and 3D variants, and hardware accelerators such as FPGAs enable deployment under strict operational budgets [75]. The Phi-Sat-1 mission demonstrated the viability of lightweight CNNs for real-time cloud detection in space.

Object detection in RS has been strongly influenced by the YOLO family of architecture. Versions 1 through 26 incorporate redesigned feature pyramids, improved multi-scale representations, and class-balanced enhancement mechanisms tailored to the diverse object sizes found in RS scenes [87].

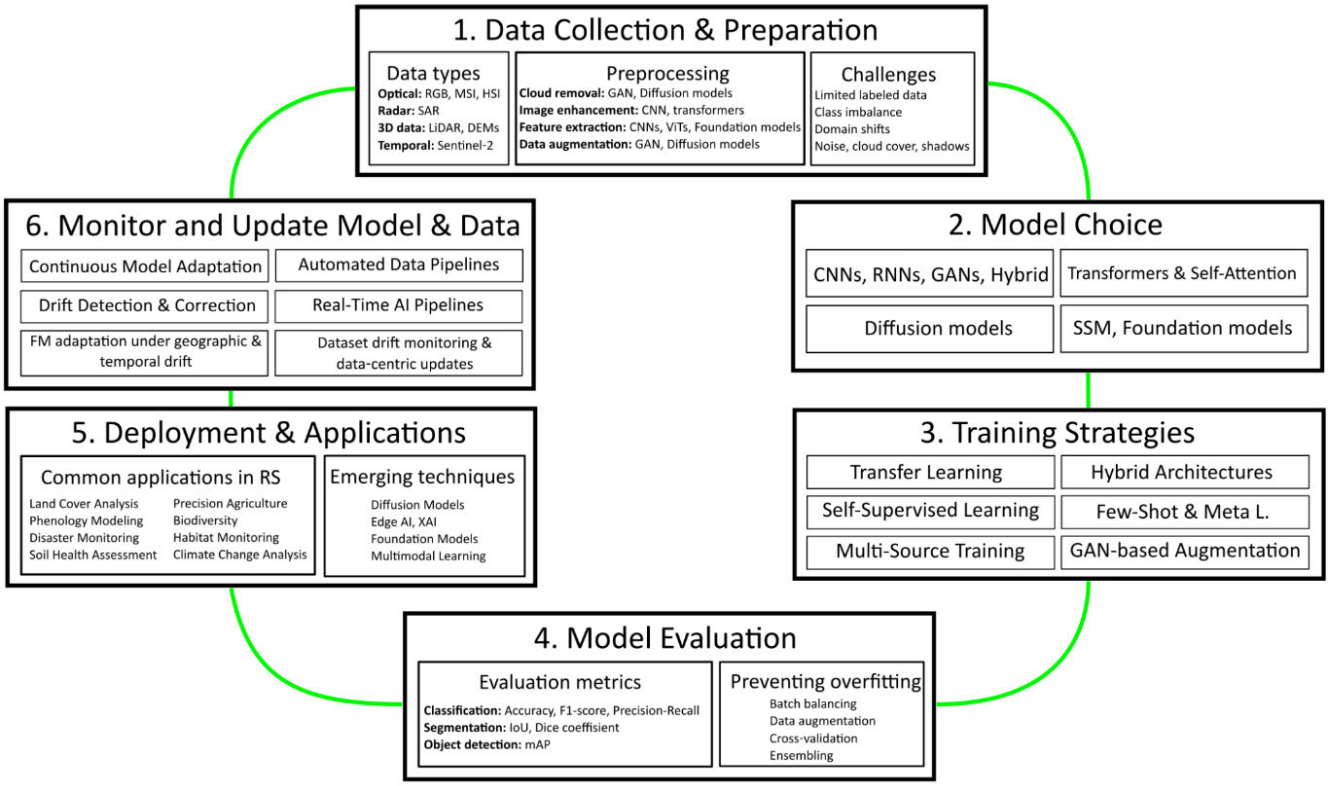

FIGURE 5. DL workflow overview for RS data.

These models achieve mean average precision gains of 5–9 percentage points over conventional architectures for small-object categories on DIOR and NWPU VHR-10 benchmarks [88].

Persistent challenges remain, including domain shift, limited annotated data for rare object categories, and the difficulty of transferring models trained on natural images to RS-specific spectral and geometric configurations [89].

2) TRANSFORMERS AND SELF-ATTENTION

Transformers have emerged as a prominent architecture for RS due to their ability to model long-range spatial dependencies through self-attention. Vision Transformers (ViTs) demonstrate competitive or superior performance to CNNs in image classification and related tasks [90], [91], while hybrid CNN–transformer models have been used to capture subtle phenological variations in vegetation [92]. Attention-enhanced CNNs such as SENet and CBAM remain widely used in SAR applications for improved noise robustness [93].

A major recent development is the rise of FMs in RS, which extend the transformer paradigm through large-scale pretraining and enable zero-shot or few-shot generalization. Existing work identifies four main families: vision FMs adapted from general visual pretraining, text-prompted models leveraging natural language supervision, visually prompted models using sparse annotations, and heterogeneous models integrating multiple data modalities [94]. RS-specific FMs such as SatMAE, GFM, and mmEarth adapt masked autoencoder strategies to multispectral and multi-temporal imagery [95]. Despite their promise, performance varies substantially across benchmarks and modalities, with accuracy dropping from 92.58% to 60.06% across two optical change-detection datasets and degrading further for heterogeneous modality pairs such as optical–SAR [94].

Transformers and FM pretraining have reshaped downstream RS tasks, including classification, detection, segmentation, change detection, and land-cover analysis [96], [97]. Large-scale pretraining particularly benefits few-shot semantic segmentation, where foundation-assisted approaches outperform meta-learning and conditioning-based methods as dataset scale increases [98].

An emerging architectural alternative is the class of State Space Models, exemplified by the Mamba architecture. CNNs are limited by local receptive fields, while ViTs incur quadratic complexity with respect to sequence length, both problematic for high-resolution RS data. Mamba provides linear-complexity global context modeling and has been applied to object detection, semantic segmentation, and change detection. Hybrid CNN–Transformer–Mamba architectures and frequency-domain variants represent the current frontier, though the causal design of the Mamba S6 block, optimized for 1D sequences, can lead to spatial information loss that adaptive scanning strategies only partially mitigate [99].

3) GENERATIVE ADVERSARIAL NETWORKS

GANs are widely used in RS for data augmentation, spectral modeling, and image-to-image translation. Architectures such as DCGAN, WGAN, CycleGAN, and conditional GANs address limited training data, particularly for SAR and HSI, by learning continuous latent representations that interpolate between sparse samples and mitigate class imbalance [93], [100]. In HSI analysis, GANs support spectral unmixing, anomaly detection, classification, and the generation of minority-class samples, with specialized variants tailored to the statistical properties of hyperspectral data [101].

Cloud removal has become one of the most prominent application domains for GAN-based models. Approaches such as SpA-GAN, RGAN, and CycleGAN generate visually realistic cloud-free imagery through adversarial learning, while SAR–optical fusion models (e.g., Simulation-Fusion GAN) leverage microwave observations to recover occluded regions [102]. However, multi-stage fusion pipelines increase computational and memory demands. Recent studies show that diffusion models, including DDPM-CR and DiffCR, achieve superior spatial and structural fidelity through iterative denoising, outperforming GANs in preserving fine-grained textures and geometric consistency.

A broader methodological shift is underway: although GANs have long dominated generative modeling in RS, diffusion models increasingly demonstrate greater robustness to noise, improved generalization across geographic domains, and more stable training dynamics [103]. Their advantages are particularly evident in cloud removal, super-resolution, and image generation, while GANs retain benefits in faster inference. Persistent challenges include limited geographic diversity in benchmarks such as SEN12MS-CR and RICE, which restricts model generalization to unseen cloud conditions and regional contexts [102].

4) FROM RNNS TO HYBRID MODELS

DL models are widely used for capturing sequential relationships in multivariate RS time series. The most common architectures include RNN, CNN, transformers, and hybrid variants. RNN were the dominant choice until 2018, after which transformers and fusion models that combine convolutional and recurrent components became more prevalent. Temporal convolutional networks and 3D CNN generally outperform 2D CNN and classical ML methods in temporal classification tasks [104].

RNN based models continue to play an important role in long term dependency modeling. Architectures such as LSTM networks [105], gated recurrent units [106], bidirectional LSTM [107], and improved bidirectional LSTM [108] address vanishing and exploding gradient problems that affect standard RNNs. Transformers, which can process long sequences exceeding thirty-time steps, frequently achieve the highest classification accuracy in RS time series.

Hybrid architectures have become a widely adopted solution for spatio-temporal modeling. Models such as U-Net combined with LSTM networks, convolutional gated recurrent units, and convolutional LSTM jointly capture spatial structure and temporal evolution and consistently outperform single model baselines in accuracy and stability [104]. These approaches have been applied across a broad range of sensors, most frequently optical systems including Landsat, MODIS, Sentinel-2, WorldView, and SPOT, followed by radar, LiDAR, and hyperspectral instruments.

Recently, graph neural networks (GNNs) have expanded the set of temporal modeling tools in RS. GNNs are particularly effective when spatial relationships cannot be adequately represented by grid-based CNNs. In UAV hyperspectral classification, GNN-based and CNN-based methods outperform RNN-based approaches. Hybrid graph convolutional network (HGCN) and convolutional network configurations also improve accuracy [67]. In surface water management, GNNs are used with LSTM networks and transformers to forecast streamflow in spatially connected river networks. In ocean monitoring, a U-Net combined with LSTM networks achieves root mean squared errors between 0.48 and 0.63 degrees Celsius for sea surface temperature prediction. Meanwhile, spatial-temporal transformer networks maintain reliable El Niño Southern Oscillation forecasting up to 18 months ahead [63].

State space models represent a new architectural direction for high-resolution RS. The Mamba architecture provides linear computational scaling while preserving global context modeling. This enables the processing of large images without the quadratic complexity of vision transformers or the limited receptive field of CNNs. Hybrid convolutional transformer Mamba configurations have become increasingly common. In UAV-based crop disease detection, for example, these models achieve F1-scores above 95% and represent the dominant architectural paradigm as of 2024. In inland water quality retrieval, hybrid networks integrating sequence modeling, attention mechanisms, and autoencoder-based feature compression constitute a recognized architectural family [109].

C. LEARNING PARADIGMS IN REMOTE SENSING

1) TRANSFER LEARNING

Transfer Learning (TL) enables the reuse of DL models that were trained on large datasets and allows downstream RS tasks to benefit from previously learned feature representations [110]. TL transfers knowledge between related tasks within the same domain and is particularly valuable in RS, where labeled data are limited [111]. In standard TL workflows, models are initialized with pretrained weights and fine-tuned through backpropagation. For small datasets, early CNN layers are kept fixed to preserve generic features, while deeper layers adapt to the target task [112].

Dataset size strongly influences TL performance in RS. Small datasets increase the risk of overfitting, whereas larger datasets support effective adaptation. TL has been applied successfully in crop classification with HSI, disease detection, yield prediction, weed identification, and the mapping of abandoned farmland [113]. However, TL performance

varies across sensing modalities. Models trained on optical data often generalize poorly to SAR imagery [93]. Optical sensors capture reflected spectral information that depends on illumination, while SAR sensors actively transmit microwave signals that record structural surface properties independent of lighting. These differences produce incompatible feature representations, since low level CNN features such as edges, textures, and gradients are modality specific. As a result, direct transfer between optical and SAR data frequently yields suboptimal results. Effective TL for heterogeneous RS requires alignment of modality specific features and reconciliation of spectral and structural differences. Even so, fine tuning only the fully connected layers of large pretrained models can still achieve high classification accuracy.

Modern TL in RS now spans a broad spectrum of strategies that extend beyond classical weight reuse. Contemporary approaches include feature level transfer, layer wise fine tuning, domain adaptive alignment, and efficient parameter adaptation of large FMs [114]. Parameter efficient methods such as low rank adaptation, adapter modules, prompt tuning, and reparameterization based techniques achieve competitive performance while reducing computational and data requirements for downstream tasks. Domain adaptation methods formalize the alignment of source and target distributions across sensors, acquisition seasons, and geographic regions [115]. Current RS systems increasingly combine SSL pretraining with parameter efficient fine tuning and domain adaptive alignment to improve generalization across diverse sensing conditions [114].

2) UNSUPERVISED AND SELF-SUPERVISED LEARNING

Unsupervised and SSL methods address the scarcity of annotated RS data by learning directly from unlabeled samples and generating supervisory signals automatically. In ML and CV, SSL approaches are commonly grouped into generative, predictive such as RotNet [116], contrastive such as SimCLR [117], MoCo [118], SwAV [119], non-contrastive such as BYOL [120], SimSiam [121], Barlow Twins [122], VICReg [123], DINO [124], and reconstructive methods such as masked autoencoders [125].

Recent RS research focuses primarily on contrastive and non-contrastive SSL for scene classification. Frameworks such as SimCLR, MoCo, and SwAV are widely used, while more recent studies increasingly adopt BYOL and DINO. SSL pretraining followed by fine tuning improves performance on RS datasets relative to random initialization and provides a strong alternative when labeled data are limited [126], [127]. Models including MoCo, BYOL, SimCLR, and Barlow Twins achieve strong results on datasets such as RESISC 45 and EuroSAT, although standard augmentations are not always suitable for RS imagery. Cropping, for example, may remove spatially distributed discriminative regions, which makes augmentation design more critical for RS tasks [126].

The RS domain introduces challenges that are not present in natural image SSL. Geographic autocorrelation causes spatially proximate patches to share similar spectral and structural properties, which can weaken contrastive objectives [22]. Seasonal variability, differences in viewing geometry, and sensor specific characteristics can introduce excessive dissimilarity between positive pairs, destabilizing training when standard augmentation pipelines are used [128]. These factors have motivated RS aware augmentation strategies and geospatially informed sampling protocols.

Masked image modeling provides an alternative SSL paradigm that is particularly effective for MSI and HSI. Local spectral structure offers a strong self-supervisory signal, and temporal or spectral masking strategies incorporated into FM architectures leverage the multi-temporal and multi-band nature of Earth observation data [95]. RS FMs represent an emerging class of SSL methods that integrate FM principles with geospatial knowledge. These models support classification, segmentation, and change detection across multimodal inputs and are expected to play an increasingly central role in RS research [126].

D. BEYOND PREDICTION: EXPLAINABILITY AND UNCERTAINTY QUANTIFICATION

Although most DL applications in RS prioritize predictive accuracy, explainability and uncertainty quantification have become increasingly important. Early work relied on feature importance of ranking and sensitivity analysis to interpret model behavior and reduce dimensionality, approaches that were particularly valuable for HSI analysis [129]. As DL architectures became more complex, XAI methods expanded from feature level interpretation toward spatial, temporal, and sample-based reasoning [130].

Uncertainty quantification is now a critical component of DL for RS, especially in high stakes applications such as disaster monitoring and environmental assessment. Bayesian DL, Monte Carlo dropout, and ensemble modeling are commonly used to estimate different forms of uncertainty, assess model confidence, and identify out of distribution samples [131]. Active learning further leverages uncertainty estimates to guide data selection and improve model performance when labeled data are limited [112].

Recent RS studies highlight the need for explainability and uncertainty analysis in operational settings. In inland water quality retrieval, the propagation of spectral measurement uncertainty through DL inference pipelines presents a fundamental challenge for the reliability of RS products [98]. Gradient based attribution, attention visualization, and concept-based explanation techniques are increasingly applied in RS, driven by scientific requirements for model interpretability and by regulatory expectations in environmental monitoring and decision support systems [54].

E. OPENLY AVAILABLE DATASETS

A wide range of openly available datasets supports DL research in RS and enables progress across tasks such as image classification, object detection, semantic

segmentation, and change detection. These datasets span multiple modalities including MSI, optical imagery, HSI, SAR, and UAV based data. They provide essential resources for benchmarking, promote reproducibility, and allow fair comparisons across algorithms and sensing conditions. To improve accessibility and support systematic dataset selection, a comprehensive and regularly updated catalog of publicly available RS datasets has been compiled. The catalog includes dataset descriptions, sensor types, spatial resolutions, and intended applications, and is available at DOI: https://doi.org/10.5281/zenodo.19470643. This centralized reference is designed to facilitate dataset discovery and to support continued advancement in geospatial DL research.

IV. OVERALL CHALLENGES

This section presents an overview of the principal challenges that the field is currently facing, organized into four categories: data-related issues, model-related issues, application and deployment issues, and AI security and ethics concerns.

A. DATA-RELATED ISSUES

Limited labeled RS data remains one of the most persistent barriers. Supervised DL models continue to outperform unsupervised and SSL approaches when sufficient annotated data are available, yet labeling requires domain expertise and substantial time. Although large unlabeled datasets are increasingly accessible, manual annotation remains a bottleneck, and SSL methods still depend on representative labeled samples for model adaptation and evaluation [132].

Dataset quality and diversity present additional constraints. Many publicly available datasets consist of preprocessed samples with limited class diversity or narrow application scope, a problem described as problematic ground truth [132]. High resolution data are often scarce, and noise, cloud cover, shadows, and other artifacts degrade sample quality. Class imbalance arises from geographic variation, annotation errors, sensor characteristics, or environmental conditions. UAV datasets are constrained by weather dependent flight schedules, while satellite imagery is affected by cloud contamination and illumination variability [104]. SAR imagery provides all weather and day night acquisition [93], but sample diversity remains limited due to sensitivity to polarization, incidence angle, and target scattering.

Benchmark fragmentation further limits cumulative progress. Inconsistent annotation protocols, heterogeneous radiometric calibration, and the absence of shared validation frameworks hinder direct comparison across studies [133]. Multispectral UAV datasets exhibit inconsistent annotation practices and strong geographic concentration in temperate agricultural regions [134]. The primary obstacle is not the absolute volume of data but the lack of coherent, comparable, and standardized datasets.

Geographic and demographic bias introduces structural fairness concerns. Most benchmark datasets originate from a small number of countries and agro climatic zones [89]. Widely used agricultural disease datasets encode region specific disease expression patterns as if universally representative [26]. This concentration produces an inverse relationship between model capability and territorial relevance: the most accurate models are often least reliable in regions with the greatest operational need for trustworthy RS intelligence [135].

B. MODEL-RELATED ISSUES

Generalization remains a central challenge for DL models in RS. Domain shift across geographic regions, sensor types, seasons, and environmental conditions leads to substantial performance variability. A model that performs well in one region may fail in another due to differences in feature distributions, annotation practices, or category definitions. Fine tuning models pre-trained on natural image datasets such as ImageNet remains difficult because RS data exhibit unique spectral, spatial, and temporal characteristics.

Integrating advanced DL architectures into RS is nontrivial. Many architectures have not been fully evaluated or optimized for RS data [93]. Transferring natural image FMs to RS is hindered by the high heterogeneity of RS modalities and spatial resolutions [94]. SAR imagery contains speckle, coherence artifacts, and geometry dependent scattering that are poorly represented in optical pretraining corpora [115]. HSI requires specialized tokenization strategies before transformer-based models can be applied effectively [95].

Scalability and computational efficiency remain limiting factors. High resolution satellite, UAV, and radar imagery require substantial computational resources, and large-scale training is often infeasible in operational settings [132]. SSL and unsupervised approaches require further evaluation across diverse RS scenarios. Cloud computing platforms used for large area land cover analysis often lack native support for advanced DL frameworks, resulting in continued reliance on simpler statistical classifiers despite the demonstrated advantages of DL in controlled experiments [136]. The increasing computational cost of training large FMs imposes practical limits on model sophistication at operational scale.

Figure 6 summarizes core technical issues of the most frequently cited data and model challenges in DL for RS.

C. APPLICATION AND DEPLOYMENT ISSUES

Translating DL models from controlled benchmarks to operational RS deployments introduces challenges that accuracy metrics alone do not capture. A tradeoff exists between detection accuracy and inference speed. Single stage object detection architectures such as YOLO and Single-Shot Detector (SSD) provide the low latency required for operational use but rely on bounding box representations that reduce spatial precision [137]. Semantic segmentation offers higher boundary accuracy but incurs computational overhead unsuitable for real-time processing on resource constrained platforms.

Environmental variability further complicates deployment. In wildfire detection, DL systems achieve inference speeds above fifteen frames per second on edge hardware, yet real

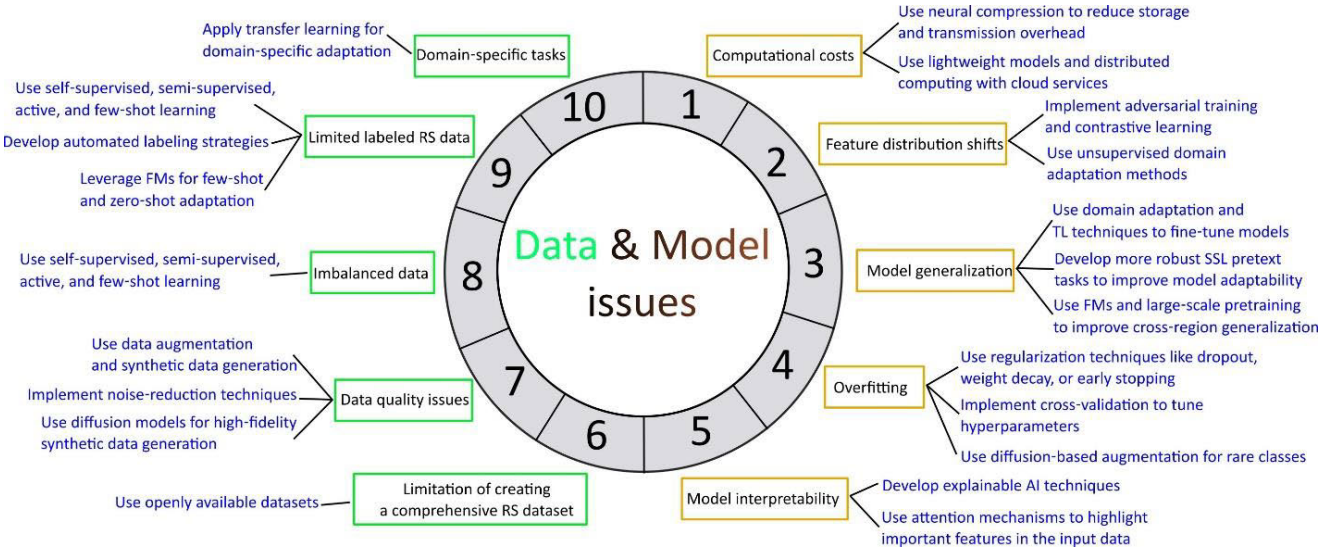

FIGURE 6. The top 10 data and DL model challenges and possible solutions (in blue).

world conditions degrade reliability in ways that curated benchmarks seldom reflect. Uncertainties in detection outputs propagate into downstream tasks such as emission quantification [49].

Edge device constraints limit UAV based RS applications. Processing HSI or high-resolution imagery on board requires hardware capable of supporting substantial computational loads, while transmitting data to ground stations introduces latency incompatible with real-time decision making [25]. Model compression, pruning, and knowledge distillation offer pathways to lightweight architectures [138], but performance penalties often disproportionately affect rare or fine-grained classes. Recent work on neural compression for geospatial analytics highlights that data transmission and storage bottlenecks are becoming a critical limiting factor for large-scale RS pipelines, motivating the development of learned lossy compression methods that enable efficient onboard processing and downlink reduction [139].

However, domain specific deployment barriers remain significant. Models trained on benchmark datasets from major crop producing regions often fail to transfer to other crops, cultivars, or agro climatic zones [134]. The absence of standardized evaluation frameworks for tasks such as crop leaf disease detection limits systematic comparison [140]. In forest biomass estimation, strong performance in uniform orchards does not generalize to mixed forests, where structural complexity reduces accuracy [37]. Cross geographic generalization gaps remain one of the most entrenched challenges, with benchmark performance concentrated in regions with established data infrastructure and degraded reliability in underrepresented environments [135].

D. AI SECURITY AND ETHICS

The deployment of DL systems in RS raises security and ethical concerns that remain insufficiently addressed. DL models trained on RS imagery are susceptible to adversarial perturbations that induce misclassification without visibly altering image quality. In contexts where satellite or UAV data support automated monitoring or disaster response, adversarial manipulation could have substantial operational consequences. Data poisoning risks are particularly relevant in open or collaborative data environments where training corpora are assembled from heterogeneous and partially unverified sources [141].

Privacy concerns are especially acute in UAV based RS. Increasing sensor resolution, extended flight endurance, and autonomous navigation create conditions for incidental or deliberate collection of personally identifiable information. Regulatory frameworks governing UAV flight altitude, proximity to populated areas, and data retention vary widely across jurisdictions, and ethical obligations for researchers conducting aerial surveys in inhabited areas are not yet standardized [142]. UAV based livestock monitoring illustrates a related concern, as persistent low altitude surveillance can induce physiological stress in monitored animals [138]. The dual use nature of RS AI, where systems developed for agricultural or environmental monitoring can be repurposed for population surveillance, requires explicit ethical scrutiny [88].

Model interpretability is central to responsible RS AI. When DL systems inform decisions related to land management, disaster response, resource allocation, or environmental compliance, the opacity of model predictions creates accountability gaps that post hoc explanation methods cannot fully resolve [67]. Calibrated uncertainty quantification is essential for high stakes applications, and establishing interpretability and uncertainty reporting as primary evaluation criteria is a necessary step toward accountable governance [98].

V. FUTURE DIRECTIONS

Future research in DL for RS is shaped by the need to address persistent challenges while leveraging emerging

methodological advances. The synthesis of recent literature indicates a field transitioning toward large scale pretraining, multimodal integration, and trustworthy AI. Six interrelated directions define the near-term research frontier.

A. FOUNDATION MODELS AND LARGE-SCALE PRETRAINING

FMs represent one of the most significant paradigm shifts in RS-DL. Unlike domain specific models trained from scratch, FMs learn broadly transferable representations from large heterogeneous corpora and can be adapted to downstream tasks with limited supervision [94]. This approach is particularly valuable in RS, where sensor diversity and limited annotated data constrain traditional supervised learning [95].

Recent RS specific FMs such as SatMAE, GFM, and mmEarth extend masked autoencoder strategies to MSI and multi-temporal satellite imagery [95]. These models demonstrate strong generalization across land cover classification, change detection, and scene understanding tasks, although a substantial performance gap remains relative to natural image FMs due to domain mismatch [94].

Efficient fine tuning and domain adaptation remain central challenges. Existing adaptation frameworks are typically designed for single source to single target transfer, whereas operational RS involves multi-source multi-target scenarios [115]. Parameter efficient fine-tuning methods including low rank adaptation, adapter modules, and prompt tuning reduce computational cost while maintaining competitive performance [114]. Future work must also address the deployment of FMs in data scarce regions such as polar environments and developing world geographies that are systematically underrepresented in global benchmarks [128].

B. SELF-SUPERVISED AND CONTRASTIVE LEARNING

SSL continues to expand as a promising direction for RS representation learning [127]. Contrastive methods such as SimCLR and MoCo have shown strong performance, but RS introduces unique challenges including geographic autocorrelation [22] and substantial variability across seasons, viewing geometries, and sensor configurations [128].

Masked image modeling is particularly effective for MSI and HSI, where local spectral structure provides a rich self-supervisory signal [128]. Cross modal SSL, which aligns representations from different sensors observing the same area, offers a pathway to improved generalization and reduced annotation requirements, especially for optical SAR pairs [25]. Designing effective cross modal objectives requires careful treatment of modality specific noise and precise geometric co registration [13].

A key priority is the standardization of SSL evaluation protocols. Consistent benchmark datasets and evaluation splits that reflect realistic deployment conditions are essential for rigorous progress [13].

C. MULTIMODAL AND CROSS-SENSOR FUSION

The integration of heterogeneous RS data sources remains a central research direction. Transformer based fusion architectures have expanded the ability to combine optical, SAR, HSI, LiDAR, and geospatial vector data within unified DL frameworks. 3D geospatial modeling increasingly integrates point clouds, RS imagery, GIS layers, and auxiliary non spatial data for urban reconstruction, terrain analysis, and infrastructure mapping [143].

Optical SAR fusion is particularly impactful, leveraging the complementarity between SAR backscatter and optical spectral information [102]. Diffusion based cloud removal methods guided by SAR auxiliary data have demonstrated superior spatial fidelity relative to GAN based approaches. Fusion of HSI and LiDAR data improves vegetation mapping by combining spectral discrimination with structural geometry. Multi-temporal fusion has proven essential for global land cover mapping, especially in ecotonal regions where single season imagery is insufficient [133].

D. TRUSTWORTHY AI: EXPLAINABILITY, UNCERTAINTY, AND ROBUSTNESS

Safety critical RS applications require DL systems that are interpretable, reliable, and robust. Wildfire monitoring, landslide assessment, infrastructure surveillance, and ecological evaluation demand calibrated uncertainty estimates and transparent model behavior [49], [51]. Bayesian DL, ensemble methods, and conformal prediction frameworks provide mechanisms for generating confidence estimates [95]. Quantifying distribution shift is equally important in domain adaptation scenarios [115].

Explainability methods including gradient based attribution, attention visualization, and concept-based explanations are increasingly applied in RS [54]. Adapting these methods to RS characteristics such as high spectral dimensionality, spatial autocorrelation, and multi-temporal inputs remains challenging. Adversarial robustness is an emerging concern in security sensitive applications including target recognition and critical infrastructure monitoring [88]. Trustworthy AI in RS therefore extends beyond technical optimization to encompass the broader social and operational context of deployment.

E. BENCHMARKS, DATASETS, AND EVALUATION PROTOCOLS

High quality standardized datasets and rigorous evaluation protocols are essential for reproducible DL in RS research, yet persistent deficiencies remain [133], [134]. Global land cover datasets suffer from inconsistent legend harmonization and the absence of shared validation frameworks. Precision agriculture datasets exhibit heterogeneous radiometric calibration, inconsistent annotation practices, and geographic concentration in temperate regions. The primary obstacle is not data volume but dataset incoherence.

Few shot segmentation research illustrates downstream consequences: inconsistent episodic evaluation designs,

limited geographic diversity, and the absence of held out test sets for genuinely novel classes constrain interpretability and generalizability [98]. Cloud native geospatial platforms have improved data accessibility [133], but few shot and zero shot evaluation protocols for underrepresented regions and rare land cover classes remain largely absent [136]. Coordinated community efforts to establish shared benchmark suites covering diverse sensors, applications, and geographies would substantially accelerate progress.

F. DOMAIN-SPECIFIC ARCHITECTURES AND EMERGING APPLICATIONS

The diversity of RS platforms and sensors motivates the development of architectures tailored to specific data types. UAV mounted HSI sensors exhibit spectral characteristics distinct from spaceborne systems, requiring models calibrated to UAV acquisition conditions [67]. Progress from classical methods to CNNs, transformers, and hybrid architectures has improved UAV HSI classification, yet real-time constraints and payload limits necessitate lightweight edge deployable models [25].

Emerging RS applications highlight the need for specialized architectures. Marine and coastal monitoring, including oil spill detection and water quality assessment, benefits from multi-sensor DL but remains constrained by data scarcity and sensor inconsistency [62]. Wildfire detection systems achieve high accuracy, yet end-to-end integration of detection, emission quantification, and uncertainty characterization remains incomplete [49]. Inland water quality retrieval using spectral, spatial, and temporal inputs demonstrates significant environmental value [109]. UAV based biomass estimation requires standardized uncertainty quantification and scaling protocols [37]. Satellite video analysis for moving object detection and dynamic event recognition remains underdeveloped due to limited benchmarks [72]. Evapotranspiration estimation and land surface flux retrieval represent a frontier where geospatial FMs, physics-based constraints, and XAI converge [54].

Data centric ML is an emerging direction with substantial potential. While FMs optimize architectures and pretraining strategies, data centric approaches focus on improving dataset quality and diversity [144]. Integrating FM techniques with data centric principles may yield computationally efficient, task independent models capable of robust generalization in data scarce RS environments.

VI. CONCLUSION

The first decade of DL in RS has reshaped Earth observation science, enabling advances in land cover mapping, object detection, multimodal fusion, time series analysis, and environmental monitoring. This review synthesizes 448 peer reviewed studies and highlights the architectural, methodological, and operational developments that have defined the evolution of DL in RS.

CNN based pipelines established the foundation for early progress, while the transition toward transformers, diffusion models, and state space architectures has expanded the capacity to model long range spatial, spectral, and temporal dependencies. Multimodal integration of optical, SAR, LiDAR, HSI, and ancillary geospatial data has become a defining characteristic of high-performance systems, supported by cloud-based platforms and UAV enabled ultra-high-resolution monitoring. SSL, few shot learning, and domain adaptation have emerged as principled responses to annotation scarcity, and FMs now represent the most promising pathway toward reducing label dependency and improving cross domain generalization.

Despite these advances, persistent challenges remain. Data scarcity continues to limit generalization across crops, ecosystems, and underrepresented geographic regions. Benchmark fragmentation and inconsistent evaluation protocols hinder cumulative scientific progress. Computational constraints restrict real-time deployment on UAVs and edge devices, and domain shift across sensors and regions remains a fundamental obstacle. Trustworthy AI principles, including calibrated uncertainty estimation, interpretability, and robustness to distribution shift, are increasingly recognized as essential for safety critical RS applications.

Looking ahead, progress will depend on the integration of FMs with multimodal fusion pipelines, parameter efficient adaptation strategies, and explainable reasoning frameworks. Data centric approaches, including rigorous dataset curation and coordinated benchmark development, must be elevated to the same priority as architectural innovation. Lightweight and efficient models will be necessary to meet the growing demand for real-time geospatial intelligence, while hybrid physical and data driven modeling offers a promising direction for improving performance in label scarce and environmentally complex domains.

Collectively, the evidence reviewed here indicates that DL in RS has reached an inflection point. The field is moving beyond demonstrating technical feasibility toward engineering the reliability, scalability, and societal accountability required for operational deployment on a global scale. This review aims to support that transition by providing a comprehensive synthesis of current capabilities, unresolved challenges, and the research directions most likely to shape the next decade of DL in RS.

ACKNOWLEDGMENT

They are very grateful to the reviewers for their valuable comments, which helped to improve the article.

CONFLICT OF INTEREST

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

REFERENCES

[1] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 24–49, Mar. 2021, doi: 10.1016/j.isprsjprs.2020.12.010.

[2] M. Weiss, F. Jacob, and G. Duveiller, "Remote sensing for agricultural applications: A meta-review," *Remote Sens. Environ.*, vol. 236, Jan. 2020, Art. no. 111402, doi: 10.1016/j.rse.2019.111402.
[3] Q. Yuan, H. Shen, T. Li, Z. Li, S. Li, Y. Jiang, H. Xu, W. Tan, Q. Yang, J. Wang, J. Gao, and L. Zhang, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sens. Environ.*, vol. 241, May 2020, Art. no. 111716, doi: 10.1016/j.rse.2020.111716.
[4] B. Huang, B. Zhao, and Y. Song, "Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery," *Remote Sens. Environ.*, vol. 214, pp. 73–86, Sep. 2018, doi: 10.1016/j.rse.2018.04.050.
[5] N. Kerle, F. Nex, M. Gerke, D. Duarte, and A. Vetrivel, "UAV-based structural damage mapping: A review," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 1, p. 14, Dec. 2019, doi: 10.3390/ijgi9010014.
[6] D. J. Søren, "70 years of machine learning in geoscience in review," *Adv. Geophys.*, vol. 61, pp. 1–55, Sep. 2020.
[7] M. A. Savelonas, C. N. Veinidis, and T. K. Bartsokas, "Computer vision and pattern recognition for the analysis of 2D/3D remote sensing data in geoscience: A survey," *Remote Sens.*, vol. 14, no. 23, p. 6017, Nov. 2022, doi: 10.3390/rs14236017.
[8] C. Bai, X. Bai, and K. Wu, "A review: Remote sensing image object detection algorithm based on deep learning," *Electronics*, vol. 12, no. 24, p. 4902, Dec. 2023, doi: 10.3390/electronics12244902.
[9] J. Li, Y. Cai, Q. Li, M. Kou, and T. Zhang, "A review of remote sensing image segmentation by deep learning methods," *Int. J. Digit. Earth*, vol. 17, no. 1, Dec. 2024, Art. no. 2328827, doi: 10.1080/17538947.2024.2328827.
[10] S.-H. Tsang. (Nov. 12, 2019). *Review: AlexNet, CaffeNet—Winner of ILSVRC 2012 (Image Classification)*. [Online]. Available: https://medium.com/coinmonks/paper-review-of-alexnet-caffenet-winner-in-ilsvrc-2012-image-classification-b93598314160
[11] G. Dudek, S. Smyl, and P. Pełka, "Recurrent neural networks for forecasting time series with multiple seasonality: A comparative study," 2022, *arXiv:2203.09170*.
[12] W. Jiang, "Object-based deep convolutional autoencoders for high-resolution remote sensing image classification," *J. Appl. Remote Sens.*, vol. 12, no. 3, Jul. 2018, Art. no. 035002, doi: 10.1117/1.jrs.12.035002.
[13] H. S. Altahainah, O. M. Alyasiri, and M. H. M. Noor, "A survey of generative adversarial networks for satellite imagery: Applications, image types, tasks, and challenges," *IEEE Access*, vol. 13, pp. 182192–182219, 2025, doi: 10.1109/ACCESS.2025.3623067.
[14] B. Lindemann, T. Müller, H. Vietz, N. Jazdi, and M. Weyrich, "A survey on long short-term memory networks for time series prediction," *Proc. CIRP*, vol. 99, pp. 650–655, Jan. 2021, doi: 10.1016/j.procir.2021.03.088.
[15] H. Arksey and L. O'Malley, "Scoping studies: Towards a methodological framework," *Int. J. Social Res. Methodology*, vol. 8, no. 1, pp. 19–32, Feb. 2005, doi: 10.1080/1364557032000119616.
[16] D. Levac, H. Colquhoun, and K. K. O'Brien, "Scoping studies: Advancing the methodology," *Implement. Sci.*, vol. 5, no. 1, p. 69, Sep. 2010, doi: 10.1186/1748-5908-5-69.
[17] M. D. J. Peters, C. M. Godfrey, H. Khalil, P. McInerney, D. Parker, and C. B. Soares, "Guidance for conducting systematic scoping reviews," *Int. J. Evidence-Based Healthcare*, vol. 13, no. 3, pp. 141–146, Sep. 2015, doi: 10.1097/xeb.0000000000000050.
[18] K. K. Westphaln, W. Regoeczi, M. Masotya, B. Vazquez-Westphaln, K. Lounsbury, L. McDavid, H. Lee, J. Johnson, and S. D. Ronis, "From arksey and O'Malley and beyond: Customizations to enhance a team-based, mixed approach to scoping review methodology," *MethodsX*, vol. 8, Jan. 2021, Art. no. 101375, doi: 10.1016/j.mex.2021.101375.
[19] M. L. Rethlefsen et al., "PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews," *Systematic Rev.*, vol. 10, no. 1, p. 39, Jan. 2021, doi: 10.1186/s13643-020-01542-z.
[20] N. R. Haddaway, M. J. Page, C. C. Pritchard, and L. A. McGuinness, "PRISMA2020 : An r package and shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis," *Campbell Systematic Rev.*, vol. 18, no. 2, p. 1230, Jun. 2022, doi: 10.1002/cl2.1230.
[21] R. Zhang, X. Wu, J. Li, P. Zhao, Q. Zhang, L. Wuri, D. Zhang, Z. Zhang, and L. Yang, "A bibliometric review of deep learning in crop monitoring: Trends, challenges, and future perspectives," *Frontiers Artif. Intell.*, vol. 8, Sep. 2025, Art. no. 1636898, doi: 10.3389/frai.2025.1636898.
[22] J. Zheng, Z. Ye, Y. Wen, J. Huang, Z. Zhang, Q. Li, Q. Hu, B. Xu, L. Zhao, and H. Fu, "A comprehensive review of agricultural parcel and boundary delineation from remote sensing images: Recent progress and future perspectives," 2025, *arXiv:2508.14558*.
[23] Z.-Y. Yang, W.-K. Xia, H.-Q. Chu, W.-H. Su, R.-F. Wang, and H. Wang, "A comprehensive review of deep learning applications in cotton industry: From field monitoring to smart processing," *Plants*, vol. 14, no. 10, p. 1481, May 2025, doi: 10.3390/plants14101481.
[24] J. G. A. Barbedo, "A review of artificial intelligence techniques for wheat crop monitoring and management," *Agronomy*, vol. 15, no. 5, p. 1157, May 2025, doi: 10.3390/agronomy15051157.
[25] F. A. Zidi, A. Ouafi, F. Bougourzi, C. Distante, and A. Taleb-Ahmed, "Advancing wheat crop analysis: A survey of deep learning approaches using hyperspectral imaging," *Comput. Electron. Agricult.*, vol. 238, Nov. 2025, Art. no. 110770, doi: 10.1016/j.compag.2025.110770.
[26] C. M. Johnson, J. S. Estrada, and F. A. Cheein, "Potato plant phenotyping and characterisation utilising machine learning techniques: A state-of-the-art review and current trends," *Comput. Electron. Agricult.*, vol. 234, Jul. 2025, Art. no. 110304, doi: 10.1016/j.compag.2025.110304.
[27] Vinayaka, P. R. C. Prasad, G. Avinash, R. A. Kumar, P. Murali, C. Palaniswami, and P. Govindaraj, "Harnessing AI and remote sensing for precision sugarcane farming: Tackling water stress, salinity, and nitrogen challenges," *Frontiers Agronomy*, vol. 7, Nov. 2025, Art. no. 1681294, doi: 10.3389/fagro.2025.1681294.
[28] Z. Yue and C. Xiao, "A meta-review of remote sensing for rubber plantations," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 141, Jul. 2025, Art. no. 104625, doi: 10.1016/j.jag.2025.104625.
[29] X. Hu, H. Chen, Q. Duan, and D. Zhang, "A comprehensive review of diffusion models in smart agriculture: Progress, applications, and challenges," 2025, *arXiv:2507.18376*.
[30] Z. Ding, K. Liu, S. Grunwald, P. Smith, P. Ciais, B. Wang, A. M. J. Wadoux, C. Ferreira, S. Karunaratne, N. Shurpali, X. Yin, D. Roberts, O. Madgett, S. A. Duncan, M. Zhou, Z. Liu, and M. T. Harrison, "Advancing soil organic carbon prediction: A comprehensive review of technologies, AI, process-based and hybrid modelling approaches," *Adv. Sci.*, vol. 12, no. 31, 2025, Art. no. e04152, doi: 10.1002/advs.202504152.
[31] A. A. J. Lima, J. C. Lopes, R. P. Lopes, T. de Figueiredo, E. Vidal-Vázquez, and Z. Hernández, "Soil organic carbon assessment using remote-sensing data and machine learning: A systematic literature review," *Remote Sens.*, vol. 17, no. 5, p. 882, Mar. 2025, doi: 10.3390/rs17050882.
[32] S. C. Haokip, Y. A. Rajwade, K. V. R. Rao, S. P. Kumar, A. B. Marak, and A. Srivastava, "Approaches for assessment of soil moisture with conventional methods, remote sensing, UAV, and machine learning methods—A review," *Water*, vol. 17, no. 16, p. 2388, Aug. 2025, doi: 10.3390/w17162388.
[33] D. M. Sobhy and A. Anandhi, "Soil nutrient monitoring technologies for sustainable agriculture: A systematic review," *Sustainability*, vol. 17, no. 18, p. 8477, Sep. 2025, doi: 10.3390/su17188477.
[34] J. W. Choi, M. S. Hidayat, S. B. Cho, W.-H. Hwang, H. Lee, B.-K. Cho, M. S. Kim, I. Baek, and G. Kim, "Recent trends in machine learning, deep learning, ensemble learning, and explainable artificial intelligence techniques for evaluating crop yields under abnormal climate conditions," *Plants*, vol. 14, no. 18, p. 2841, Sep. 2025, doi: 10.3390/plants14182841.
[35] K. Yu, A. Belwalkar, W. Wang, Y. Hu, A. Hunegnaw, A. Nurunnabi, T. Ruf, F. Li, L. Jia, L. Kooistra, Y. Miao, and F. N. Teferle, "UAV hyperspectral remote sensing for crop nitrogen monitoring: Progress, challenges, and perspectives," *Smart Agricult. Technol.*, vol. 12, Dec. 2025, Art. no. 101507, doi: 10.1016/j.atech.2025.101507.
[36] S. M. Shawon, I. A. Adan, M. Moniruzzaman, O. H. Sinan, M. J. Abed, M. Al Qwaid, M. T. Sarker, and H. B. A. Karim, "Federated learning approaches in precision agriculture: A systematic review of current trends, challenges, and opportunities," *IEEE Access*, vol. 14, pp. 30513–30538, 2026, doi: 10.1109/ACCESS.2026.3667548.
[37] D. Munteanu, S. Moldovanu, G. Murariu, and L. Dinca, "Sustainable estimation of tree biomass and volume using UAV imagery: A comprehensive review," *Sustainability*, vol. 18, no. 2, p. 1095, Jan. 2026, doi: 10.3390/su18021095.
[38] W. Xu, Y. Cheng, M. Luo, X. Mai, W. Wang, W. Zhang, and Y. Wang, "Progress and limitations in forest carbon stock estimation using remote sensing technologies: A comprehensive review," *Forests*, vol. 16, no. 3, p. 449, Mar. 2025, doi: 10.3390/f16030449.

[39] Y. Wu, C. Lu, K. Wu, W. Gao, N. Yang, and J. Lin, "Advancements and trends in mangrove species mapping based on remote sensing: A comprehensive review and knowledge visualization," *Global Ecology Conservation*, vol. 57, Jan. 2025, Art. no. e03408, doi: 10.1016/j.gecco.2025.e03408.

[40] A. Toqeer, A. Hall, A. Horta, and S. Wassens, "Remote sensing of woody plant encroachment: A global systematic review of drivers, ecological impacts, methods, and emerging innovations," *Remote Sens.*, vol. 18, no. 3, p. 390, Jan. 2026, doi: 10.3390/rs18030390.

[41] Q. Ye and Z. Zhao, "Remote sensing based modeling of forest structural parameters: Advances and challenges," *Forests*, vol. 17, no. 2, p. 209, Feb. 2026, doi: 10.3390/f17020209.

[42] W. Charerntantanakul, J. T. Burley, M. Yebra, A. B. Nicotra, S. A. Cunningham, and M. T. Brookhouse, "Remote-sensing methods for mapping eucalypt dieback: A systematic review," *Remote Sens. Appl., Soc. Environ.*, vol. 41, Jan. 2026, Art. no. 101908, doi: 10.1016/j.rsase.2026.101908.

[43] D. Yang, X. Gao, Y. Yang, K. Guo, K. Han, and L. Xu, "Advances and future prospects in building extraction from high-resolution remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 18, pp. 6994–7016, 2025, doi: 10.1109/JSTARS.2025.3538662.

[44] M. S. B. M. Anik, C. An, and S. S. Li, "Evolution from the physical process-based approaches to machine learning approaches to predicting urban floods: A literature review," *Environ. Syst. Res.*, vol. 14, no. 1, p. 15, Jul. 2025, doi: 10.1186/s40068-025-00409-3.

[45] L. B. Lima, W. J. S. F. Rocha, D. T. M. Souza, J. S. B. Lobão, M. M. M. de Santana, E. C. B. Cambui, and R. N. Vasconcelos, "Urban quality: A remote-sensing-perspective review," *Urban Sci.*, vol. 9, no. 2, p. 31, Jan. 2025, doi: 10.3390/urbansci9020031.

[46] S. Darmawan, R. Hernawati, and S. Rahmani, "Satellite-based urban heat island study: A prisma-based systematic literature review," *Geomatics Environ. Eng.*, vol. 19, no. 6, pp. 5–40, Nov. 2025, doi: 10.7494/geom.2025.19.6.5.

[47] B. Dong, R. Zhang, S. Li, Y. Ye, and C. Huang, "A meta-analysis for the nighttime light remote sensing data applied in urban research: Key topics, hotspot study areas and new trends," *Sci. Remote Sens.*, vol. 11, Jun. 2025, Art. no. 100186, doi: 10.1016/j.srs.2024.100186.

[48] Y. Xu, Y. Yang, Q. Liu, and C. Song, "Advancing urban blue space monitoring and management: A review of remote sensing applications and interdisciplinary impact assessment," *Ecol. Indicators*, vol. 182, Jan. 2026, Art. no. 114563, doi: 10.1016/j.ecolind.2025.114563.

[49] K. Das, J. Poovvancheri, S. Flesca, C. R. Calidonna, and D. Chen, "Emerging trends in wildfire detection through the lens of computer vision and wildfire emission quantification: A comprehensive survey," *IEEE Access*, vol. 14, pp. 20201–20228, 2026, doi: 10.1109/ACCESS.2026.3660843.

[50] N. Ejaz and S. Choudhury, "A comprehensive survey of the machine learning pipeline for wildfire risk prediction and assessment," *Ecol. Informat.*, vol. 90, Dec. 2025, Art. no. 103325, doi: 10.1016/j.ecoinf.2025.103325.

[51] P. Jiang, Z. Ma, and G. Mei, "Review article: Deep learning for potential landslide identification: Data, models, applications, challenges, and opportunities," *Natural Hazards Earth Syst. Sci.*, vol. 26, no. 1, pp. 487–529, Jan. 2026, doi: 10.5194/nhess-26-487-2026.

[52] M. Nurtas, S. Nurakynov, A. Altaibek, A. Mergembayeva, and M. A. Mohammed, "Satellite based deep learning approaches for detecting environmental disasters across Kazakhstan," *Discover Appl. Sci.*, vol. 8, no. 2, p. 144, Dec. 2025, doi: 10.1007/s42452-025-08133-4.

[53] F. D. Javan, F. Samadzadegan, and A. Toosi, "Air pollution observation—Bridging spaceborne to unmanned airborne remote sensing: A systematic review and meta-analysis," *Air Qual., Atmos. Health*, vol. 18, no. 8, pp. 2481–2549, Aug. 2025, doi: 10.1007/s11869-025-01771-y.

[54] T. Y. Fong, Y. F. Huang, R. J. Chin, and C. H. Koo, "Advancing evapotranspiration estimation with remote sensing and artificial intelligence—A review," *Agricult. Water Manage.*, vol. 322, Dec. 2025, Art. no. 110023, doi: 10.1016/j.agwat.2025.110023.

[55] M. M. Stofa, S. R. Abdani, A. M. Moubark, M. A. A. M. Zainuri, A. A. Ibrahim, N. A. M. Kamari, and M. A. Zulkifley, "Recent developments of artificial intelligence methods for sea ice concentration monitoring using high-resolution imaging datasets," *Ecol. Informat.*, vol. 87, Jul. 2025, Art. no. 103132, doi: 10.1016/j.ecoinf.2025.103132.

[56] A. T. Marye, C. Caramiello, D. De Nardi, D. Miglino, G. Proietti, K. C. Saddi, C. Biscarini, S. Manfreda, M. Poggi, and F. Tauro, "Remote sensing for monitoring macroplastics in rivers: A review," *WIREs Water*, vol. 12, no. 2, Mar. 2025, Art. no. e70020, doi: 10.1002/wat2.70020.

[57] S. Kako, T. Kataoka, D. Matsuoka, Y. Takahashi, M. Hidaka, S. Aliani, U. Andriolo, H. Dierssen, T. H. M. van Emmerik, G. Gonçalves, V. Martinez-Vicente, P. Mishra, J. G. Monteiro, K. Topouzelis, and A. Isobe, "Remote sensing and image analysis of macro-plastic litter: A review," *Mar. Pollut. Bull.*, vol. 222, Jan. 2026, Art. no. 118630, doi: 10.1016/j.marpolbul.2025.118630.

[58] Y. Xu, L. Lu, L. Yan, and H. K. Zhang, "A review of remote sensing in coastal aquaculture: Data, geographic hotspots, methods, and challenges," *GISci. Remote Sens.*, vol. 62, no. 1, Dec. 2025, Art. no. 2573529, doi: 10.1080/15481603.2025.2573529.

[59] Z. Zhang, P. Chen, S. Zhang, H. Huang, Y. Pan, and D. Pan, "A review of machine learning applications in ocean color remote sensing," *Remote Sens.*, vol. 17, no. 10, p. 1776, May 2025, doi: 10.3390/rs17101776.

[60] Z. Wu, J. Pang, J. Li, Y. Wang, J. Ruan, X. Zhang, L. Yang, Y. Pang, and Y. Gao, "A review of remote sensing-based water quality monitoring in turbid coastal waters," *Intell. Mar. Technol. Syst.*, vol. 3, no. 1, p. 24, Aug. 2025, doi: 10.1007/s44295-025-00075-2.

[61] Z. Wu and K. L. E. Law, "Remote sensing and deep learning for algal Blooms: A review," *IEEE Access*, vol. 13, pp. 180891–180908, 2025, doi: 10.1109/ACCESS.2025.3622243.

[62] N. Prakash and O. Zielinski, "AI-enhanced real-time monitoring of marine pollution: Part 1-A state-of-the-art and scoping review," *Frontiers Mar. Sci.*, vol. 12, Apr. 2025, Art. no. 1486615, doi: 10.3389/fmars.2025.1486615.

[63] X. Qiao, K. Zhang, and W. Huang, "Impacts of climate change on oceans and ocean-based solutions: A comprehensive review from the deep learning perspective," *Remote Sens.*, vol. 17, no. 13, p. 2306, Jul. 2025, doi: 10.3390/rs17132306.

[64] S. Qiao, Q. Zhang, and Z. Wang, "A review of deep-learning-based SAR image ship interpretation technology: The latest advances," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 18, pp. 26152–26185, 2025, doi: 10.1109/JSTARS.2025.3614504.

[65] R. Zhang, Y. Yang, Z. Li, P. Li, and H. Wang, "Optical and SAR image fusion: A review of theories, methods, and applications," *Remote Sens.*, vol. 18, no. 1, p. 73, Dec. 2025, doi: 10.3390/rs18010073.

[66] E. Peter, L.-M. Ang, K. P. Seng, and S. Srivastava, "Recent advances in deep learning for SAR images: Overview of methods, challenges, and future directions," *Sensors*, vol. 26, no. 4, p. 1143, Feb. 2026, doi: 10.3390/s26041143.

[67] Z. Zhang, L. Huang, Q. Wang, L. Jiang, Y. Qi, S. Wang, T. Shen, B.-H. Tang, and Y. Gu, "UAV hyperspectral remote sensing image classification: A systematic review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 18, pp. 3099–3124, 2025, doi: 10.1109/JSTARS.2024.3522318.

[68] J. Zou, H. Qu, and P. Zhang, "Conventional to deep learning methods for hyperspectral unmixing: A review," *Remote Sens.*, vol. 17, no. 17, p. 2968, Aug. 2025, doi: 10.3390/rs17172968.

[69] C. Lou, M. A. A. Al-Qaness, D. Al-Alimi, A. Dahou, M. A. Elaziz, L. Abualigah, and A. A. Ewees, "Land use/land cover (LULC) classification using hyperspectral images: A review," *Geo-Spatial Inf. Sci.*, vol. 28, no. 2, pp. 345–386, Mar. 2025, doi: 10.1080/10095020.2024.2332638.

[70] V. Sangeetha and L. Agilandeeswari, "Artificial intelligence enabled spectral–spatial feature extraction techniques for land use and land cover classification using hyperspectral images—An inclusive review," *Egyptian J. Remote Sens. Space Sci.*, vol. 28, no. 3, pp. 455–467, Sep. 2025, doi: 10.1016/j.ejrs.2025.06.004.

[71] C. Yu, H. Yang, L. Ma, J. Yang, Y. Jin, W. Zhang, K. Wang, and Q. Zhao, "Deep learning-based change detection in remote sensing: A comprehensive review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 18, pp. 24415–24437, 2025, doi: 10.1109/JSTARS.2025.3605170.

[72] F. Liu, J. Wang, L. Jiao, J. Zhang, H. Wang, S. Li, L. Li, P. Chen, X. Liu, W. Ma, S. Wang, S. Yang, X. Zhang, Y. Du, Q. Bao, L. Sun, and B. Hou, "Remote sensing video tracking: Current status, challenges, and future," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 18, pp. 14338–14367, 2025, doi: 10.1109/JSTARS.2025.3573572.

[73] J. Liu, J. Xiang, Y. Jin, R. Liu, J. Yan, and L. Wang, "Boost precision agriculture with unmanned aerial vehicle remote sensing and edge intelligence: A survey," *Remote Sens.*, vol. 13, no. 21, p. 4387, Oct. 2021, doi: 10.3390/rs13214387.

[74] L. P. Osco, J. M. Junior, A. P. M. Ramos, L. A. de Castro Jorge, S. N. Fatholahi, J. de Andrade Silva, E. T. Matsubara, H. Pistori, W. N. Gonçalves, and J. Li, "A review on deep learning in UAV remote sensing," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 102, Oct. 2021, Art. no. 102456, doi: 10.1016/j.jag.2021.102456.

[75] N. Ghasemi, J. A. Justo, M. Celesti, L. Despoisse, and J. Nieke, "Onboard processing of hyperspectral imagery: Deep learning advancements, methodologies, challenges, and emerging trends," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 18, pp. 4780–4790, 2025, doi: 10.1109/JSTARS.2025.3527898.

[76] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "Deep learning classifiers for hyperspectral imaging: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 158, pp. 279–317, Dec. 2019, doi: 10.1016/j.isprsjprs.2019.09.006.

[77] C. Liu, Y. Sun, Y. Xu, Z. Sun, X. Zhang, L. Lei, and G. Kuang, "A review of optical and SAR image deep feature fusion in semantic segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 12910–12930, 2024, doi: 10.1109/JSTARS.2024.3424831.

[78] N. Li, Z. Wang, and F. A. Cheikh, "Discriminating spectral–spatial feature extraction for hyperspectral image classification: A review," *Sensors*, vol. 24, no. 10, p. 2987, May 2024, doi: 10.3390/s24102987.

[79] N. Zang, Y. Cao, Y. Wang, B. Huang, L. Zhang, and P. T. Mathiopoulos, "Land-use mapping for high-spatial resolution remote sensing image via deep learning: A review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 5372–5391, 2021, doi: 10.1109/JSTARS.2021.3078631.

[80] L. Huang, B. Jiang, S. Lv, Y. Liu, and Y. Fu, "Deep-learning-based semantic segmentation of remote sensing images: A survey," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 8370–8396, 2024, doi: 10.1109/JSTARS.2023.3335891.

[81] K. R. Hasan, A. B. Tuli, M. A.-M. Khan, S.-H. Kee, M. A. Samad, and A.-A. Nahid, "Deep-learning-based semantic segmentation for remote sensing: A bibliometric literature review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 1390–1418, 2024, doi: 10.1109/JSTARS.2023.3328315.

[82] S. Ghaffarian, J. Valente, M. van der Voort, and B. Tekinerdogan, "Effect of attention mechanism in deep learning-based remote sensing image processing: A systematic literature review," *Remote Sens.*, vol. 13, no. 15, p. 2965, Jul. 2021, doi: 10.3390/rs13152965.

[83] B. Mishra, A. Dahal, N. Luintel, T. B. Shahi, S. Panthi, S. Pariyar, and B. R. Ghimire, "Methods in the spatial deep learning: Current status and future direction," *Spatial Inf. Res.*, vol. 30, no. 2, pp. 215–232, Apr. 2022, doi: 10.1007/s41324-021-00425-2.

[84] L. Yu, L. Tang, and L. Mu, "A review of detection transformer: From basic architecture to advanced developments and visual perception applications," *Sensors*, vol. 25, no. 13, p. 3952, Jun. 2025, doi: 10.3390/s25133952.

[85] S. Bera, V. K. Shrivastava, and S. Chandra Satapathy, "Advances in hyperspectral image classification based on convolutional neural networks: A review," *Comput. Model. Eng. Sci.*, vol. 133, no. 2, pp. 219–250, 2022, doi: 10.32604/cmes.2022.020601.

[86] G. Chen, Q. Weng, G. J. Hay, and Y. He, "Geographic object-based image analysis (GEOBIA): Emerging trends and future opportunities," *GIScience Remote Sens.*, vol. 55, no. 2, pp. 159–182, Mar. 2018, doi: 10.1080/15481603.2018.1426092.

[87] A. Safonova, A. Campanera-Moliné, J. P. Silva, D. F. N. Sarró, M. C. Guerra, J. Serra-Sagristà, and A. Santangeli, "Deep learning for UAV thermal bird detection: Benchmarking YOLOv8–YOLOv26," *Remote Sens. Appl., Soc. Environ.*, vol. 42, Apr. 2026, Art. no. 101981, doi: 10.1016/j.rsase.2026.101981.

[88] A. Raza and F. Hanif, "Chronological review and performance analysis of YOLO-based deep learning frameworks in complex geospatial environments," *Appl. Artif. Intell.*, vol. 40, no. 1, Feb. 2026, Art. no. 2626114, doi: 10.1080/08839514.2026.2626114.

[89] M. Nikouei, B. Baroutian, S. Nabavi, F. Taraghi, A. Aghaei, A. Sajedi, and M. E. Moghaddam, "Small object detection: A comprehensive survey on challenges, techniques and real-world applications," *Intell. Syst. Appl.*, vol. 27, Sep. 2025, Art. no. 200561, doi: 10.1016/j.iswa.2025.200561.

[90] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," 2021, *arXiv:2010.11929*.

[91] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 11966–11976, doi: 10.1109/CVPR52688.2022.01167.

[92] I. Teixeira, R. Morais, J. J. Sousa, and A. Cunha, "Deep learning models for the classification of crops in aerial imagery: A review," *Agriculture*, vol. 13, no. 5, p. 965, Apr. 2023, doi: 10.3390/agriculture13050965.

[93] J. Li, Z. Yu, L. Yu, P. Cheng, J. Chen, and C. Chi, "A comprehensive survey on SAR ATR in deep-learning era," *Remote Sens.*, vol. 15, no. 5, p. 1454, Mar. 2023, doi: 10.3390/rs15051454.

[94] C. Huo, K. Chen, S. Zhang, Z. Wang, H. Yan, J. Shen, Y. Hong, G. Qi, H. Fang, and Z. Wang, "When remote sensing meets foundation model: A survey and beyond," *Remote Sens.*, vol. 17, no. 2, p. 179, Jan. 2025, doi: 10.3390/rs17020179.

[95] Z. Yu, M. Y. I. Idris, H. Wang, P. Wang, J. Chen, and K. Wang, "From physics to foundation models: A review of AI-driven quantitative remote sensing inversion," 2025, *arXiv:2507.09081*.

[96] A. A. Aleissaee, A. Kumar, R. M. Anwer, S. Khan, H. Cholakkal, G.-S. Xia, and F. S. Khan, "Transformers in remote sensing: A survey," *Remote Sens.*, vol. 15, no. 7, p. 1860, Mar. 2023, doi: 10.3390/rs15071860.

[97] W. Li and C.-Y. Hsu, "GeoAI for large-scale image analysis and machine vision: Recent progress of artificial intelligence in geography," *ISPRS Int. J. Geo-Information*, vol. 11, no. 7, p. 385, Jul. 2022, doi: 10.3390/ijgi11070385.

[98] M. Petrov, E. Pandilova, I. Dimitrovski, D. Trajanov, V. Spasev, and I. Kitanovski, "Few-shot semantic segmentation in remote sensing: A review on definitions, methods, datasets, advances and future trends," *Remote Sens.*, vol. 18, no. 4, p. 637, Feb. 2026, doi: 10.3390/rs18040637.

[99] M. Bao, S. Lyu, Z. Xu, H. Zhou, J. Ren, S. Xiang, X. Li, and G. Cheng, "Vision mamba in remote sensing: A comprehensive survey of techniques, applications and outlook," *Remote Sens.*, vol. 18, no. 4, p. 594, Feb. 2026, doi: 10.3390/rs18040594.

[100] V. Sampath, I. Maurtua, J. J. A. Martín, and A. Gutierrez, "A survey on generative adversarial networks for imbalance problems in computer vision tasks," *J. Big Data*, vol. 8, no. 1, p. 27, Jan. 2021, doi: 10.1186/s40537-021-00414-0.

[101] P. Ranjan, A. Nandal, S. Agarwal, and R. Kumar, "A dive into generative adversarial networks in the world of hyperspectral imaging: A survey of the state of the art," *Remote Sens.*, vol. 18, no. 2, p. 196, Jan. 2026, doi: 10.3390/rs18020196.

[102] S. Edirisinghe, B. Schoen-Phelan, and S. Hensman, "Generative deep learning models for cloud removal in satellite imagery: A comparative review of GANs and diffusion methods," *ISPRS Open J. Photogramm. Remote Sens.*, vol. 19, Jan. 2026, Art. no. 100110, doi: 10.1016/j.ophoto.2025.100110.

[103] Y. Liu, J. Yue, S. Xia, P. Ghamisi, W. Xie, and L. Fang, "Diffusion models meet remote sensing: Principles, methods, and perspectives," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 10684806, doi: 10.1109/TGRS.2024.3464685.

[104] X. Cheng, Y. Sun, W. Zhang, Y. Wang, X. Cao, and Y. Wang, "Application of deep learning in multitemporal remote sensing image classification," *Remote Sens.*, vol. 15, no. 15, p. 3859, Aug. 2023, doi: 10.3390/rs15153859.

[105] X. Li, L. Peng, X. Yao, S. Cui, Y. Hu, C. You, and T. Chi, "Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation," *Environ. Pollut.*, vol. 231, pp. 997–1004, Dec. 2017, doi: 10.1016/j.envpol.2017.08.114.

[106] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014, *arXiv:1412.3555*.

[107] Z. Huang, W. Xu, and K. Yu, "Bidirectional LSTM-CRF models for sequence tagging," 2015, *arXiv:1508.01991*.

[108] B. Chen, H. Zheng, L. Wang, O. Hellwich, C. Chn, L. Yang, T. Liu, G. Luo, A. Bao, and X. Chen, "A joint learning im-BiLSTM model for incomplete time-series Sentinel-2A data imputation and crop classification," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 108, Apr. 2022, Art. no. 102762, doi: 10.1016/j.jag.2022.102762.

[109] Z. Pang, Z. Zhou, J. Fu, W. Jiang, X. Qin, and M. Sun, "Deep learning-based remote sensing retrieval of inland water quality: A review," *J. Hydrol., Regional Stud.*, vol. 61, Oct. 2025, Art. no. 102759, doi: 10.1016/j.ejrh.2025.102759.

[110] M. Pashaei, H. Kamangir, M. J. Starek, and P. Tissot, "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.*, vol. 12, no. 6, p. 959, Mar. 2020, doi: 10.3390/rs12060959.
[111] A. Safonova, G. Ghazaryan, S. Stiller, M. Main-Knorn, C. Nendel, and M. Ryo, "Ten deep learning techniques to address small data problems with remote sensing," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 125, Dec. 2023, Art. no. 103569, doi: 10.1016/j.jag.2023.103569.
[112] S. Yu and J. Ma, "Deep learning for geophysics: Current and future trends," *Rev. Geophysics*, vol. 59, no. 3, p. 2021, Sep. 2021, doi: 10.1029/2021rg000742.
[113] J. A. García-Berná, S. Ouhbi, B. Benmouna, G. García-Mateos, J. L. Fernández-Alemán, and J. M. Molina-Martínez, "Systematic mapping study on remote sensing in agriculture," *Appl. Sci.*, vol. 10, no. 10, p. 3456, May 2020, doi: 10.3390/app10103456.
[114] D. Yin, T.-F. Zhao, D.-P. Fan, S. Li, B. Du, X. Sun, and S.-M. Hu, "Remote sensing tuning: A survey," *Comput. Vis. Media*, vol. 11, no. 5, pp. 897–937, Oct. 2025, doi: 10.26599/cvm.2025.9450490.
[115] S. Lyu, Q. Zhao, Z. Zhou, M. Li, Y. Zhou, D. Yao, G. Cheng, H. Zhou, and Z. Shi, "Deep learning based domain adaptation methods in remote sensing: A comprehensive survey," 2025, *arXiv:2510.15615*.
[116] H. Tu, Y. Han, Z. Wang, A. Chen, K. Tao, S. Ye, S. Wang, Z. Wei, and J. Li, "RotNet: A rotationally invariant graph neural network for quantum mechanical calculations," *Small Methods*, vol. 8, no. 1, Jan. 2024, Art. no. 2300534, doi: 10.1002/smtd.202300534.
[117] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," 2020, *arXiv:2002.05709*.
[118] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 9726–9735, doi: 10.1109/CVPR42600.2020.00975.
[119] M. Caron, I. Misra, J. Mairal, P. Goyal, P. Bojanowski, and A. Joulin, "Unsupervised learning of visual features by contrasting cluster assignments," 2026, *arxiv:2006.09882*.
[120] J.-B. Grill, F. Strub, F. Altché, C. Tallec, P. H. Richemond, E. Buchatskaya, C. Doersch, B. Avila Pires, Z. D. Guo, M. G. Azar, B. Piot, K. Kavukcuoglu, R. Munos, and M. Valko, "Bootstrap your own latent: A new approach to self-supervised learning," 2020, *arXiv:2006.07733*.
[121] X. Chen and K. He, "Exploring simple Siamese representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 15745–15753, doi: 10.1109/CVPR46437.2021.01549.
[122] J. Zbontar, L. Jing, I. Misra, Y. LeCun, and S. Deny, "Barlow twins: Self-supervised learning via redundancy reduction," 2021, *arXiv:2103.03230*.
[123] A. Bardes, J. Ponce, and Y. LeCun, "VICReg: Variance-invariance-covariance regularization for self-supervised learning," 2021, *arXiv:2105.04906v3*.
[124] M. Caron, H. Touvron, I. Misra, H. Jegou, J. Mairal, P. Bojanowski, and A. Joulin, "Emerging properties in self-supervised vision transformers," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 9630–9640, doi: 10.1109/ICCV48922.2021.00951.
[125] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick, "Masked autoencoders are scalable vision learners," presented at the IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2022, doi: 10.1109/CVPR52688.2022.01553.
[126] P. Berg, M.-T. Pham, and N. Courty, "Self-supervised learning for scene classification in remote sensing: Current state of the art and perspectives," *Remote Sens.*, vol. 14, no. 16, p. 3995, Aug. 2022, doi: 10.3390/rs14163995.
[127] A. Safonova, S. Stiller, M. Yordanov, and M. Ryo, "Self-supervised learning outperforms supervised learning for crop classification by annotating only 5% of images," *Precis. Agricult.*, vol. 27, no. 1, p. 4, Dec. 2025, doi: 10.1007/s11119-025-10302-9.
[128] B. Gorry, J. Sandino, P. Moghadam, F. Gonzalez, and J. Roberts, "Advances in machine learning approaches for UAV-based remote sensing in data-deficient Antarctic environments," *Remote Sens.*, vol. 18, no. 3, p. 459, Feb. 2026, doi: 10.3390/rs18030459.
[129] D. Chutia, D. K. Bhattacharyya, K. K. Sarma, R. Kalita, and S. Sudhakar, "Hyperspectral remote sensing classifications: A perspective survey," *Trans. GIS*, vol. 20, no. 4, pp. 463–490, Aug. 2016, doi: 10.1111/tgis.12164.
[130] R. Roscher, B. Bohn, M. F. Duarte, and J. Garcke, "Explain it to me—Facing remote sensing challenges in the bio- and geosciences with explainable machine learning," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 3, pp. 817–824, Aug. 2020, doi: 10.5194/isprs-annals-v-3-2020-817-2020.
[131] J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng, A. Kruspe, R. Triebel, P. Jung, R. Roscher, M. Shahzad, W. Yang, R. Bamler, and X. X. Zhu, "A survey of uncertainty in deep neural networks," *Artif. Intell. Rev.*, vol. 56, no. S1, pp. 1513–1589, Oct. 2023, doi: 10.1007/s10462-023-10562-9.
[132] S. Zhao, K. Tu, S. Ye, H. Tang, Y. Hu, and C. Xie, "Land use and land cover classification meets deep learning: A review," *Sensors*, vol. 23, no. 21, p. 8966, Nov. 2023, doi: 10.3390/s23218966.
[133] Q. Xu, V. Yordanov, L. Bruzzone, and M. A. Brovelli, "High-resolution global land cover maps and their assessment strategies," *ISPRS Int. J. Geo-Information*, vol. 14, no. 6, p. 235, Jun. 2025, doi: 10.3390/ijgi14060235.
[134] A. Caroppo, G. Diraco, and A. Leone, "A systematic review of available multispectral UAV image datasets for precision agriculture applications," *Remote Sens.*, vol. 18, no. 4, p. 659, Feb. 2026, doi: 10.3390/rs18040659.
[135] J. Botero-Valencia, V. García-Pineda, A. Valencia-Arias, J. Valencia, E. Reyes-Vera, M. Mejia-Herrera, and R. Hernández-García, "Machine learning in sustainable agriculture: Systematic review and research perspectives," *Agriculture*, vol. 15, no. 4, p. 377, Feb. 2025, doi: 10.3390/agriculture15040377.
[136] B. Alshehri, Z. Zhang, and X. Liu, "A review of Google Earth engine for land use and land cover change analysis: Trends, applications, and challenges," *ISPRS Int. J. Geo-Information*, vol. 14, no. 11, p. 416, Oct. 2025, doi: 10.3390/ijgi14110416.
[137] P. Singh, B. Zhao, and Y. Shi, "Computer vision for site-specific weed management in precision agriculture: A review," *Agriculture*, vol. 15, no. 21, p. 2296, Nov. 2025, doi: 10.3390/agriculture15212296.
[138] M. Adam, J. Song, W. Yu, and Q. Li, "Deep learning approaches for automatic livestock detection in UAV imagery: State-of-the-art and future directions," *Future Internet*, vol. 17, no. 9, p. 431, Sep. 2025, doi: 10.3390/fi17090431.
[139] C. Gomes et al., "Lossy neural compression for geospatial analytics: A review," *IEEE Geosci. Remote Sens. Mag. (replaces Newsletter)*, vol. 13, no. 3, pp. 97–135, Sep. 2025, doi: 10.1109/MGRS.2025.3546527.
[140] D. Radočaj, P. Radočaj, I. Plaščak, and M. Jurišić, "Evolution of deep learning approaches in UAV-based crop leaf disease detection: A web of science review," *Appl. Sci.*, vol. 15, no. 19, p. 10778, Oct. 2025, doi: 10.3390/app151910778.
[141] T. Anukiruthika and D. S. Jayas, "AI-driven grain storage solutions: Exploring current technologies, applications, and future trends," *J. Stored Products Res.*, vol. 111, May 2025, Art. no. 102588, doi: 10.1016/j.jspr.2025.102588.
[142] P. Thapa and M. Thapa, "Drone in hyperspectral remote sensing data analytics: Techniques and applications," *Geocarto Int.*, vol. 40, no. 1, Dec. 2025, Art. no. 2572670, doi: 10.1080/10106049.2025.2572670.
[143] A. A. Mohamad, U. Ujang, S. Azri, and U. Mehmood, "Enhancing 3D geospatial modelling through multimodal data and machine learning: A systematic literature review," *Spatial Inf. Res.*, vol. 34, no. 1, p. 12, Jan. 2026, doi: 10.1007/s41324-025-00659-4.
[144] R. Roscher, M. Russwurm, C. Gevaert, M. Kampffmeyer, J. A. Dos Santos, M. Vakalopoulou, R. Hänsch, S. Hansen, K. Nogueira, J. Prexl, and D. Tuia, "Better, not just more: Data-centric machine learning for Earth observation," *IEEE Geosci. Remote Sens. Mag.*, vol. 12, no. 4, pp. 335–355, Dec. 2024, doi: 10.1109/MGRS.2024.3470986.

ANASTASIIA SAFONOVA received the Ph.D. degree in remote sensing from Siberian Federal University, Krasnoyarsk, Russia, in 2020, and the Ph.D. degree in data mining from the University of Granada, Granada, Spain, in 2021. She is currently affiliated with the Leibniz Centre for Agricultural Landscape Research (ZALF), Germany, and holds a Marie Skłodowska-Curie Postdoctoral Fellowship at the Autonomous University of Barcelona, Spain. Her research interests include remote sensing, deep learning, self-supervised learning, multimodal data fusion, and AI for environmental and agricultural applications.

RONNY HÄNSCH (Senior Member, IEEE) received the Diploma degree in computer science and the Ph.D. degree from Technische Universität Berlin, Berlin, Germany, in 2007 and 2014, respectively. He is currently a Senior Scientist with the Microwave and Radar Institute, German Aerospace Center (DLR), where he leads the Machine Learning Team in the Signal Processing Group of the SAR Technology Department. His research interests include machine learning for SAR and optical remote sensing, with emphasis on semantic segmentation, multi-modal data, and ensemble learning for large-scale Earth observation. He has held several leadership roles, including the Chair of the IEEE GRSS Image Analysis and Data Fusion (IADF) Technical Committee, the Chair of the GRSS German Chapter, the GRSS Membership Chair, and GRSS Representative within SpaceNet, and the Co-Chair of the ISPRS Working Group on Image Orientation and Sensor Fusion. He co-organizes the CVPR Workshop EarthVision and the IGARSS Tutorial on Machine Learning in Remote Sensing and was a Technical Lead of the SpaceNet 8 and 9 Challenges. He is the Editor-in-Chief of IEEE Geoscience and Remote Sensing Letters and an Associate Editor of *ISPRS Journal of Photogrammetry and Remote Sensing* and *IEEE Open Journal of the Computer Society*.

RIBANA ROSCHER received the Dipl.-Ing. and Ph.D. degrees in geodesy from the University of Bonn, Bonn, Germany, in 2008 and 2012, respectively. From 2015 to 2022, she was a Junior Professor of remote sensing with the University of Bonn. From 2022 to 2025, she was a Professor of data science for crop systems with the University of Bonn and led the Data Science for Crop Systems Group, Institute of Bio- and Geosciences (IBG)-2, Forschungszentrum Jülich, Jülich, Germany. Since 2025, she has been a Professor of machine learning in agriculture with the Institute of Geodesy and Geoinformation, University of Bonn.

GILBERTO CAMARA received the Ph.D. degree in computer science from Brazil's National Institute for Space Research (INPE), São José dos Campos, Brazil, in 1995. He has been a Researcher and a Professor of geoinformatics, spatial databases, and land use change with INPE, where he was the General Director, from 2006 to 2012. He was the Brazil Chair of the Institute for Geoinformatics, University of Münster, Münster, Germany, from 2013 to 2015. He advised 50 graduate students and authored 150 articles that have been cited more than 7500 times. He was a recipient of the Dr. Honoris Causa from the University of Münster, and the William T. Pecora Award from NASA and USGS for "leadership to the broad and open access to remote sensing data."

BEGÜM DEMIR (Senior Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electronic and telecommunication engineering from Kocaeli University, Kocaeli, Türkiye, in 2005, 2007, and 2010, respectively. She is currently a Full Professor and the Founder Head of the Remote Sensing Image Analysis (RSiM) Group, Faculty of Electrical Engineering and Computer Science, TU Berlin, and the Head of the Big Data Analytics for Earth Observation Research Group, Berlin Institute for the Foundations of Learning and Data (BIFOLD). Specifically, she performs research in the field of processing and analysis of large-scale Earth observation data acquired by airborne and satellite-borne systems. Her research interests include the intersection of machine learning, remote sensing, and signal processing. She is a fellow of the European Laboratory for Learning and Intelligent Systems (ELLIS). She is a scientific committee member of several international conferences and workshops. She was awarded the prestigious "2018 Early Career Award" by the IEEE Geoscience and Remote Sensing Society for her research contributions in machine learning for information retrieval in remote sensing. In 2018, she received a Starting Grant from the European Research Council (ERC) for her project "BigEarth: Accurate and Scalable Processing of Big Data in Earth Observation." She is a Referee for several journals, such as *Proceedings of the IEEE*, IEEE Transactions on Geoscience and Remote Sensing, IEEE Geoscience and Remote Sensing Letters, IEEE Transactions on Image Processing, *Pattern Recognition*, IEEE Transactions on Circuits and Systems for Video Technology, IEEE Journal of Selected Topics in Signal Processing, and *International Journal of Remote Sensing*, and several international conferences. She is an Associate Editor of *IEEE Geoscience and Remote Sensing Magazine*.

JOAN SERRA-SAGRISTÀ (Senior Member, IEEE) received the Ph.D. degree in computer science from the Autonomous University of Barcelona (UAB), in 1999. From 1997 to 1998, he was with the University of Bonn, Bonn, Germany, funded by DAAD. He is currently a Full Professor with the Department of Information and Communications Engineering, UAB. He has co-authored more than 100 articles. His research interest includes data compression, especially image coding for remote sensing applications.

MASAHIRO RYO received the Ph.D. degree in civil engineering from Tokyo Institute of Technology, in 2015. He is currently a Professor of environmental data science and a Leader with Brandenburg University of Technology Cottbus–Senftenberg, Germany. His research interest includes the application of artificial intelligence for agricultural and biodiversity problems.