



Advancing horizons in remote sensing: a comprehensive survey of deep learning models and applications in image classification and beyond

Sidike Paheding¹  · Ashraf Saleem² · Mohammad Faridul Haque Siddiqui³ · Nathir Rawashdeh² · Almabrok Essa⁴ · Abel A. Reyes²

Received: 19 December 2023 / Accepted: 19 June 2024 / Published online: 2 August 2024

© The Author(s) 2024

Abstract

In recent years, deep learning has significantly reshaped numerous fields and applications, fundamentally altering how we tackle a variety of challenges. Areas such as natural language processing (NLP), computer vision, healthcare, network security, wide-area surveillance, and precision agriculture have leveraged the merits of the deep learning era. Particularly, deep learning has significantly improved the analysis of remote sensing images, with a continuous increase in the number of researchers and contributions to the field. The high impact of deep learning development is complemented by rapid advancements and the availability of data from a variety of sensors, including high-resolution RGB, thermal, LiDAR, and multi-/hyperspectral cameras, as well as emerging sensing platforms such as satellites and aerial vehicles that can be captured by multi-temporal, multi-sensor, and sensing devices with a wider view. This study aims to present an extensive survey that encapsulates widely used deep learning strategies for tackling image classification challenges in remote sensing. It encompasses an exploration of remote sensing imaging platforms, sensor varieties, practical applications, and prospective developments in the field.

Keywords Deep learning · Remote sensing · Image classification

✉ Sidike Paheding
spaheding@fairfield.edu

Ashraf Saleem
ashraf@mtu.edu

Mohammad Faridul Haque Siddiqui
msiddiqui@wtamu.edu

Nathir Rawashdeh
narawash@mtu.edu

Almabrok Essa
aessa@jcu.edu

Abel A. Reyes
areyesan@mtu.edu

1 Introduction

Remote sensing (RS) images are valuable resources of data to quantify and observe intricate formations on the Earth's surface.

Remote sensing image classification (RSIC), which task is to automatically assign a semantic label for a given remote sensing image, has been a fast-growing research topic in recent years, and it has significant contributions to monitoring and understanding key environmental processes. Thanks to a large volume of remote sensing data availability, sensor development, and ever-increasing computing powers, rapid advancement in RSIC has been witnessed by its real-world applications, such as natural hazard detection [1, 2], precision agriculture [3, 4], landscape mapping [5], urban planning [6], and climate changes [7]. The enabler of this wide range of applications of RSIC is also attributed to the ability of RS images to capture multi-scale, multi-dimensional, and multi-temporal information. Hence, one of the challenging but important

¹ Department of Computer Science and Engineering, Fairfield University, 1073 N Benson Rd, Fairfield, CT 06824, USA

² Department of Applied Computing, Michigan Technological University, 1400 Townsend Dr, Houghton, MI 49931, USA

³ Department of Computer Science, West Texas AM University, 2501 4th Ave, Canyon, TX 79016, USA

⁴ Department of Mathematics, Computer Science, and Data Science, John Carroll University, 1 John Carroll Blvd, University Heights, OH 44118, USA

tasks in RSIC is to effectively extract valuable information from the various kinds of RS data to aid further image analysis and interpretation.

Traditional approaches to exploiting features from RS images heavily rely on feature extraction and/or feature selection. The former process produces new features to describe specific spatial and spectral attributes of RS data using a transformation matrix or a set of filtering processes. For instance, knowledge-based approaches depend on spectral characteristics and supervised and unsupervised methods such as canonical analysis and principal component analysis. The latter identifies a subset of feature candidates from a feature pool via selection criteria. Although feature extraction or selection provides useful information to improve the accuracy of RSIC, most of those methods are suboptimal for comprehensively representing original data for given applications [8]. Particularly, their suitability can be even degraded when comes to big data with multi-Sensors, since RS images can vary greatly in terms of time, geo-location, atmospheric conditions, and imaging platform [9, 10]. Therefore, an effective and unified approach is needed to automatically extract pertinent features from diverse RS data.

Deep learning (DL) [11], as a subset of machine learning, has demonstrated unprecedented performance in feature representation and is capable of performing end-to-end learning in various vision tasks, including image classification [12], object detection [13], semantic segmentation [14], and natural language processing [15]. Since the astonishing accuracy was produced by a deep convolutional neural network (CNN) in the large-scale visual recognition challenge [12], CNN and its variants prevailed in many fields [16], along with tremendous successes including those most important yet unsolved issues of modern science, such as AlphaFold [17] which can accurately predict 3D models of protein structures.

Over the last few years, researchers have made an effort to utilize the most advanced techniques in AI for RSIC, from traditional methods in machine learning, all the way through deep learning techniques such as the use of CNN and its variants.

This study aims to fill the gap in the existing literature in RSIC with the following key contributions:

- An extensive examination of cutting-edge deep learning models through a systematic review, covering a brief description of architectures and frameworks for RSIC. Our review includes concise descriptions of the architectural nuances and frameworks that have shown promise in this domain.
- A summary of remote sensing datasets, modalities, as well as corresponding applications. This comprehensive resource will serve as a reference for researchers and

practitioners navigating the rich landscape of remote sensing data.

- Suggestions for promising research direction and insights around RSIC. These recommendations aim to catalyze innovation and drive the field forward.

This study serves as evidence of the growing impact of deep learning within the domain of remote sensing image analysis. It sheds light on how deep learning can be a powerful tool for addressing persistent challenges in RSIC and seeks to stimulate further research in this dynamic and essential field.

1.1 Review statistics

Deep learning techniques have been actively implemented for RS-related tasks (i.e., image classification) in the last lustrum. Statistical analysis was conducted using the latest Scopus data for a literature search on the most popular keywords in RS publications: “deep learning” AND “remote sensing” AND “image classification,” ranging the year from 2017 to 2023 (mid-September). The detailed research setting can be found in Table 1. The research results are analyzed from two aspects: (1) the number of conferences and journals published, as shown in Fig. 1, and (2) the distribution of publications across subject areas, as depicted in Fig. 2.

From Fig. 1, it can be clearly seen that the total number of publications (i.e., conference papers + journal articles) is consistently increased yearly in the past 5 years and a particularly significant difference is observed comparing 2021 with 2017, demonstrating fast growth of this research. While journal articles also exhibit a yearly increase, the number of conference papers published in 2021 is reduced compared to the previous year. In addition, it is worth noting that the number of published journal articles in 2021 greatly exceeded the number of conference papers, indicating remarkable development of the aforementioned topic. In Fig. 2, the subject of “*Earth and planetary science*” turns out to be the most widely applied area of deep learning applications in RSIC, while the computer science field is ranked second with a minor margin in terms of publication count.

The remainder of this article is organized as follows. In Sect. 2, we delve into an overview of existing surveys on deep learning-based Remote Sensing Image Classification (RSIC). Section 3 provides a comprehensive summary of Deep Learning (DL) applications in remote sensing. The various deep learning models pertinent to RSIC are reviewed briefly in Sect. 4. Detailed descriptions of remote sensing datasets and their associated applications are presented in Sect. 5. Section 6 addresses current challenges and outlines potential future research directions. Finally, Sect. 7 offers concluding remarks on this survey.

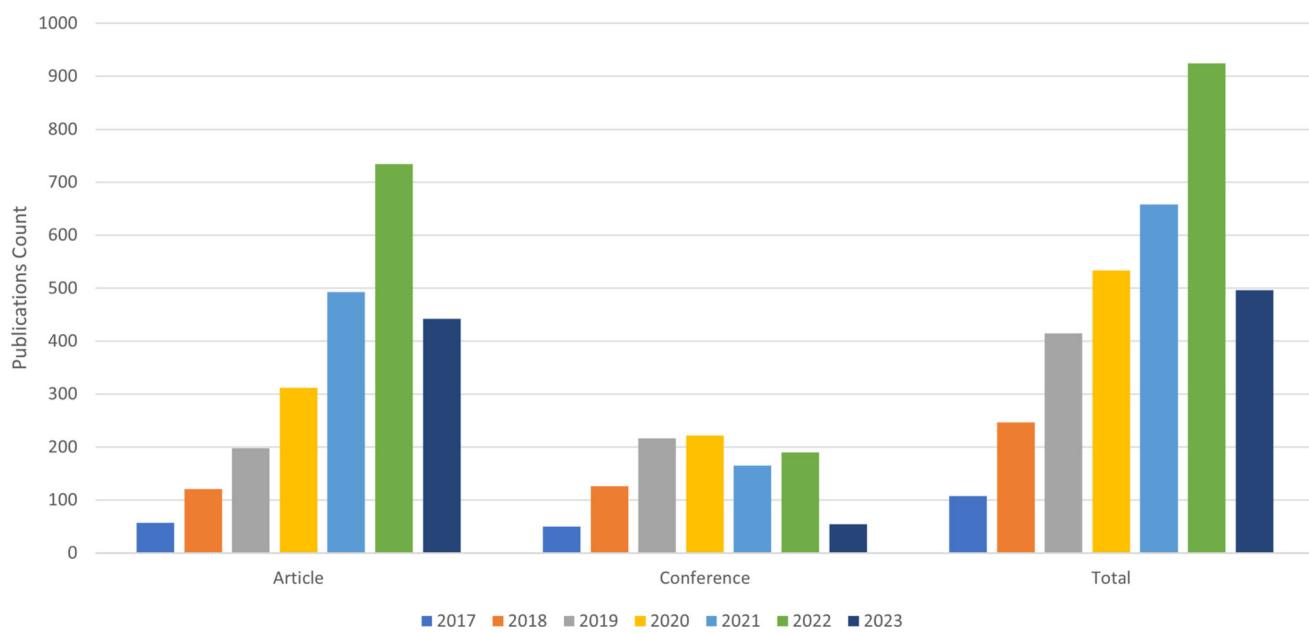


Fig. 1 Related conference papers and journal articles published per year from 2017 to 2023. Source: Scopus database accessed on September 15, 2023. The search results obtained by searching on “deep learning” AND “remote sensing” AND “image classification.”

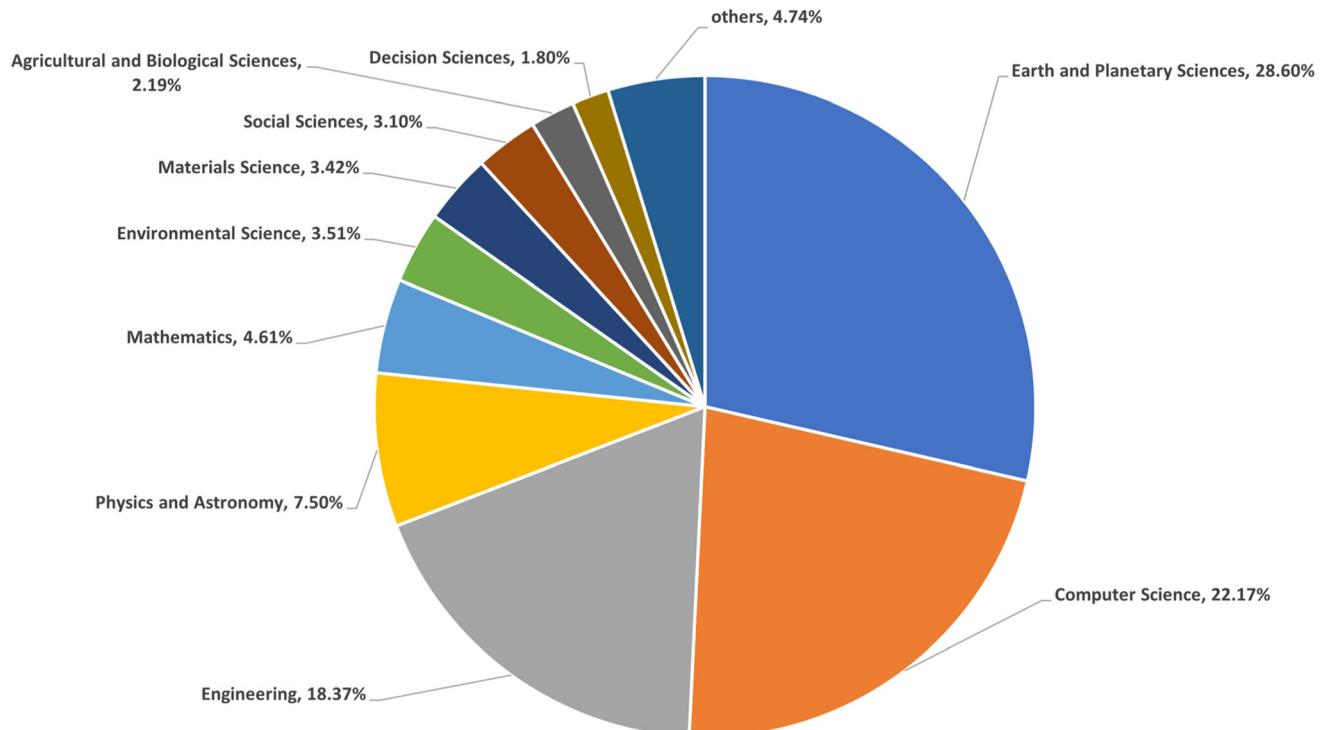


Fig. 2 The distribution of relevant papers based on the top 10 subject areas

2 Related work

In the past several years, driven by DL, a great number of RSIC methods sprung up, and consequently, many related survey or review papers have been published, which are summarized in Table 2. These surveys cover various

aspects of the field, including the methods used and the content they focus on.

In chronological order, starting in Yao et al. [18] conducted a survey focused on providing data sources for RS and current deep learning-based classification methods. Moving to Li et al. [19] conducted a comprehensive review

Table 1 Literature research settings in Scopus database

Search field	Search selections
Search Documents	“Deep Learning” AND “Remote Sensing” AND “Image Classification”
Document journals	Article, Conference paper
Keyword	Remote Sensing, Deep Learning, Image Classification
Source Type	Journal, Conference Proceeding
Year	2017–2023
Search within	Article title, Abstract, Keywords

Table 2 A summary of the number of surveys for deep learning-based remote sensing image classification (RSIC)

References	Title	Year	Methods	Content
[18]	A review on image classification of remote sensing using deep learning	2017	CNN SAE DBN	Provides data sources of remote sensing and current deep learning-based classification methods
[19]	Deep learning for remote sensing image classification: A survey	2018	CNN SAE DBN	A systematic review and comparative analysis of pixel-wise and scene-wise deep learning approaches for RS image classification
[20]	Deep learning for hyperspectral image classification: An overview	2019	DBN SAE CNNs GANs TL	Review of deep learning methods in hyperspectral image scene classification and provides guideline to improve the classification performance
[21]	Deep learning classifiers for hyperspectral imaging: A review	2019	CNNs LSTM GRU SSL AL TL	Detailed review of deep learning algorithms, frameworks, normalization methods for hyperspectral image Classification
[22]	A survey of remote sensing image classification based on CNNs	2019	CNNs	Summary of CNN-based RS scene classification methods, and provides challenges and suggestions for CNN-based RS image classification
[23]	Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities	2020	SAE CNNs GANs	Discusses the main challenges of RS image scene classification and comprehensively compares popular deep learning architectures based on CNN, GAN, and SAE.
[24]	Deep learning methods for land-cover and land-use classification in remote sensing: a review	2020	CNN TL	Comparative analysis of deep CNN models on diverse RS datasets
[25]	Deep Learning for Land-Use and Land-Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review	2020	CNNs TL	Review of deep learning in RS scene classification and provide a framework from multispectral and hyperspectral images perspective
[26]	Hyperspectral and Lidar data applied to the urban land-cover machine learning and neural-network-based classification: a review	2021	CNN CRNN	Review of hyperspectral and LiDAR data fusion approaches for urban land-cover classification
[27]	Land-use mapping for high spatial resolution remote sensing image via deep learning: A review	2021	CNN DBN SAE FCN TL	Review of deep learning-based land-use mapping methods consists of supervised learning, semi-supervised learning, or unsupervised learning, as well as pixel-based or object-based approaches
[28]	Remote Sensing Image Classification: A Comprehensive Review and Applications	2022	CNN SSL RL	Overview of trending DL models and training techniques for RSIC task. The review maps those trending with corresponding benchmark datasets and provides a sense of potential future research directions in remote sensing image analysis

CNN, Convolutional Neural Network; CRNN, Convolutional Recurrent Neural Network; RL, Reinforcement Learning; DBN, Deep Belief Network; SAE, Stacked Autoencoder; GAN, Generative Adversarial Network; TL, Transfer Learning; LSTM, Long Short-Term Memory; GRU, Gated Recurrent Units; SSL, Self-Supervised Learning; AL, Active Learning; and FCN, Fully Convolutional Network

and comparative analysis of deep learning approaches for RSIC, considering both pixel-wise and scene-wise strategies. The year 2019 witnessed the presentation of several

related surveys. Li et al. [20] delved into the realm of hyperspectral image scene classification, revising deep learning methods and offering guidance on how to enhance

classification performance. Paoletti et al. [21] offered a detailed review of deep learning algorithms, frameworks, and normalization methods tailored to hyperspectral image classification. Additionally, Song et al. [22] summarized methods based on CNNs for remote sensing scene classification and highlighted challenges and recommendations for CNN-based classification. Cheng et al. [23] discussed the challenges, methods, benchmarks, and opportunities in remote sensing image scene classification, in addition to comparing popular deep learning architectures, including CNN, GAN, and SAE. On the other hand, Alem et al. [24] conducted a comparative analysis of deep CNN models on diverse remote sensing datasets, while Vali et al. [25] reviewed deep learning in remote sensing scene classification from the perspective of multispectral and hyperspectral images. Kuras et al. [26] conducted a review of hyperspectral and LiDAR data fusion approaches for urban land-cover classification, focusing on the use of CNN and CRNN. Zang et al. [27] reviewed deep learning-based land-use mapping methods, including supervised, semi-supervised, and unsupervised learning, as well as pixel-based and object-based approaches. In [28], an overarching view of contemporary deep learning models and some hybrid methodologies for RSIC is presented.

Although the abovementioned surveys have substantially contributed to the literature by reviewing various methodologies and aspects of RSIC, there remains a compelling need for a survey that encapsulates the latest advancements and trends in this rapidly evolving field. The existing surveys, while thorough, often focus on specific subdomains or are somewhat dated given the fast pace of technological progress in deep learning applications for RSIC. This survey is necessitated by several critical factors. (1) Since the publication of the last major surveys, numerous new deep learning architectures have been developed, each accompanied by innovative applications. This necessitates an updated review that cohesively synthesizes these advancements. (2) Recent advancements in publicly available data sources, coupled with their corresponding applications in RSIC, have not been fully addressed in prior surveys. Our work seeks to fill this gap by providing a comprehensive review and categorization of these datasets. (3) Emerging Challenges and Solutions, as the application areas of RSIC expand, new challenges arise, including those related to scalability, data heterogeneity, and scarcity. Addressing these challenges requires a fresh look at the state-of-the-art, which our survey provides. In conclusion, this survey does not merely aggregate existing knowledge but critically analyzes recent innovations and trends, thereby setting a new benchmark for research in RSIC. It aims to catalyze further research and development in a field that is crucial for a wide array of applications, from environmental monitoring to disaster

response. This work is intended to serve as a cornerstone for future explorations and technological advancements in remote sensing image analysis.

3 DL in Remote sensing applications

Deep learning methods have been remarkably utilized by the research community in the recent years in RSIC due to their important role in a wide range of applications, such as agriculture, urban and forestry [29–34], environment monitoring [35–38], land mapping and management [39–43], disaster response [44–50], ecology [51, 52], mining [53, 54], oceanography [55, 56], hydrology [57, 58], archaeology [59, 60], among others. By exploring the Scopus database, it is found that agriculture and forestry are the most RS applications that researchers have used deep learning methods for data analysis. This is followed by environment monitoring, land mapping and management, and disaster response as shown in Fig. 3. Based on these initial statistics, we will focus in this survey paper on the top four RS applications on the list.

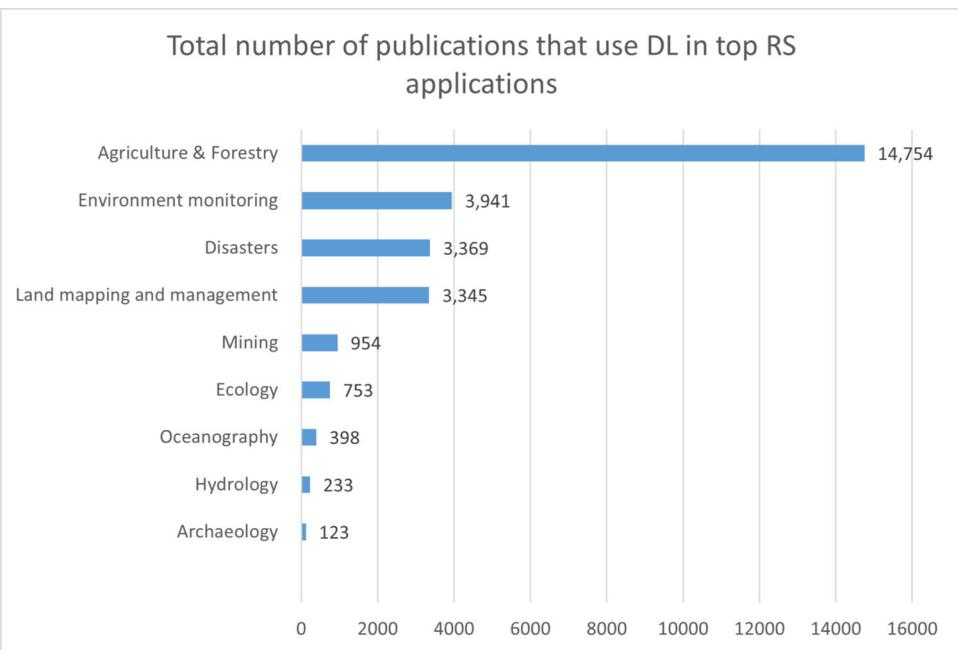
3.1 Agriculture and forestry

Countries worldwide are investing billions of dollars in precision agriculture and forestry in order to increase production efficiency while reducing environmental impact.

RS satellite and aerial images are considered very useful sources of information for many agriculture and forestry applications such as:

- *Crop monitoring:* Deep learning and RS technology are used widely by researchers and the agriculture industry to provide real-time monitoring of crop growth [61], plant morphology [62], and plantation monitoring [63]. The main advantage of using real-time intelligent RS technology for crop monitoring is providing an accurate understanding of the growth environment that leads to environment optimization and consequently improved production efficiency and quality [64]. It also helps in detecting variations in several parameters of the field such as biomass, nitrogen status, and yield estimation of the crop which determine the need for fertilizer or other actions.
- *Diseases detection:* crop health monitoring is a crucial step in avoiding economic loss and low production quality. Traditionally, disease detection and avoiding its spread in crops is done manually which takes days or months of continuous work to inspect the entire crop. Moreover, these methods lack detection accuracy and do not provide real-time monitoring, especially in large

Fig. 3 Total number of publications on using DL methods for RS data analysis in top remote sensing applications
 Source: Scopus database accessed on Sep 20, 2023. The search results were obtained by searching on “deep learning” AND “application name OR sub-application category”



crop areas [65]. RS technology with DL algorithms provided practitioners with real-time monitoring capabilities for large crops with high detection accuracy, especially when using UAVs. Further, they improved the ability to control the spread of diseases at critical times which led to reduced losses and improved product quality in precision agriculture. Recent research efforts have focused on improving existing methods in crop disease detection [66, 67], Pest identification and tracking [68, 69], and plant disease classification [67, 70, 71].

- **Weed control:** weeds detection and removal is considered one of the most important factors in improving product quality and critical to the development of precision agriculture. Accurate mapping and localization of the weeds lead to accurate pesticide spraying of the weed location without contaminating crops, humans, and water resources. Researchers have put great efforts into using RS technology and deep learning in weed detection [72–76] and weed mapping [77–80].
- **Precision irrigation:** One of the most important applications of modern precision farming where RS technology and DL algorithms play a crucial role in the efficient use of water at the right time, location, and quantity. Aerial and satellite data analysis using efficient DL algorithms helped in soil moisture estimation [81–85] and prediction [86, 87], mapping of center pivot irrigation [88–91], and estimation of soil indicators [92, 93].
- **Forest planning and management:** The modern forest management utilizes RS platforms such as UAVs,

airplanes, and satellites to provide crucial data at different spatial and spectral band resolutions. This data is mainly used in creating forest models for monitoring, conservation, and restoration [94]. In recent years, researchers focused on creating deep learning models to analyze data collected from RS platforms in land-cover and forest mapping [95–99], species classification [33, 100–102], and forest disaster management [103–105].

It is evident that RS technology and platforms along with deep learning methods have recently played an integral role in enabling precision agriculture.

3.2 Disaster response and recovery

Natural disasters such as floods, earthquakes, landslides, tsunamis, hurricanes, and wildfires have a devastating impact on the environment, cities' infrastructure, and living beings. The modern disaster management cycle consists of the following phases: (1) prevention and mitigation, (2) preparedness, (3) response, and (4) recovery [106]. RS technology and DL methods are widely applied in disaster response and recovery. While disaster response aims to immediately reduce the impact and damages caused by the disaster through damage mapping and estimation, disaster recovery is concerned with bringing life back to normal through reconstruction monitoring and wreckage clearance. Having said that, AI algorithms are still being used in disaster detection and forecasting.

Response to sudden-onset disasters requires spatial information that should be updated in real-time due to their

high dynamics. Thanks to the recent advancements in RS technologies and platforms that are capable of providing high levels of spatial and temporal resolution data. Not to mention the recent developments in DL methods that provide real-time analysis for RS data. In the last 5 years, researchers have focused on developing DL methods to analyze RS data in order to provide governments and the research community with tools that help in managing sudden-onset disasters such as floods [48, 107, 108], earthquakes [109–111], landslides [112–114], tsunamis [111, 115], hurricanes [116, 117], and wildfires [50, 118, 119].

3.3 Environment monitoring

The extraordinary level of air, land, and water pollution has led governments and researchers worldwide to take immediate action to allocate financial resources and efforts toward creating technologies that ensure ongoing and universal surveillance of the environment. Traditionally, governments use a large number of distributed fixed station that consists of advanced sensors and instruments to monitor the environment. With the advancement in the Internet of Things technology, wireless sensors network (WSN) with millions of tiny distributed sensors is widely used to monitor the environment. Recently, crowdsensing platforms, including vehicles like cars, buses, taxis, bicycles, and trains have been equipped with sensors and measurement systems that collect, process, and store data about the environment at practically zero cost. Fascista [120] stated, based on an in-depth review of the literature, that although WSNs offer an attractive solution for environmental monitoring, they suffer from several technical drawbacks including poor data quality, low communication range, reliability, and power limitation. On the other hand, crowdsensing poses some implementation challenges including incentive mechanisms, task allocation, workload balancing, data trustworthiness, and user privacy.

RS technology and platforms offer an attractive solution to these challenges by providing rich data about the environment ranging from RGB images and LiDAR to thermal and hyperspectral data. DL algorithms have also provided reliable tools for extracting information about the environment from the collected RS data. Looking at the literature, it is found that recent studies have developed DL algorithms to analyze RS data for land environment monitoring [121–124], air monitoring [35, 125–127], and marine and water monitoring [36, 128–130].

3.4 Land-cover/land-use mapping

Urban growth has historically influenced alterations in regional and global climates by impacting both

biogeochemical and biophysical processes. Therefore, remote sensing is widely used for land-cover mapping, land management, and the spatial distribution of landforms to examine earth surface processes and landscape evolution. Land-use classification using remote sensing images and DL methods [131–135] has played a crucial role in effectively identifying diverse land uses, which in turn improved urban environment monitoring, planning, and designing. RS and DL have also been used in land-cover mapping and change detection [136–140] which are employed in natural resource management, urban planning, and agricultural management.

4 DL methods for RS image classification

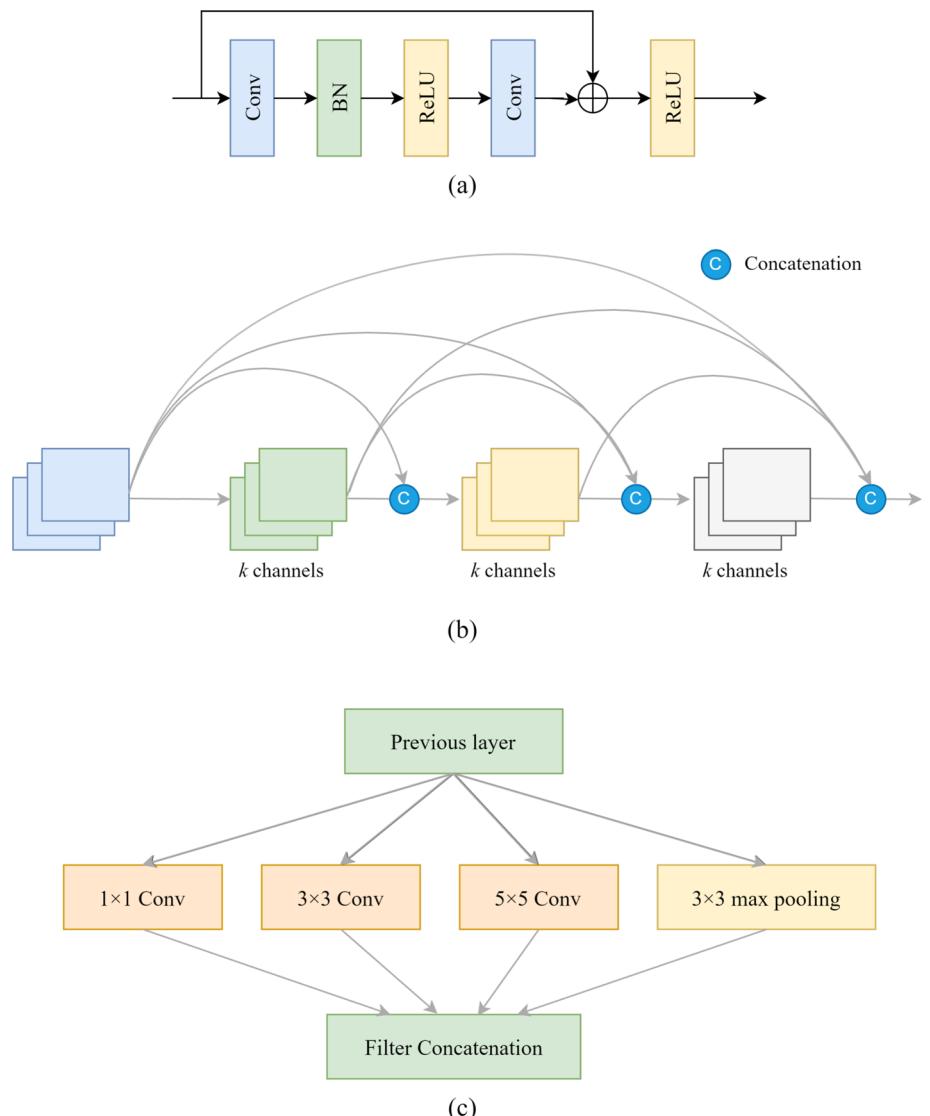
4.1 Learning approach

4.1.1 Convolutional neural networks (CNNs)

Convolutional neural network (CNN) and its variants have been widely applied to RS applications [22]. The key component in CNN is convolutional operation which involves trainable parameters and aims to extract pertinent features that are associated with specific tasks such as object recognition, segmentation, and tracking. Figures 4 and 5 depict commonly used CNN modules in RSIC, including residual connection [141], dense blocks in the Dense Convolutional Network (DenseNet) [142], inception module [143], squeeze and excitation inception module (SE-inception) [144], dilated convolution [145], and depth-wise separable convolution [146]. Key features of these learning modules are described as follows:

- **Residual connection:** The module presented by He et al. [141] uses skip or residual connections between layers to facilitate the gradient propagation and thus help to achieve deeper neural network architectures with better accuracy. It has been widely used in hyperspectral image classification [147–149]. This module is illustrated in Fig. 4a
- **Dense blocks:** Introduced by Huang et al. [142] dense blocks enhance the feature reuse by dense connectivity. In other words, the information flow between layers is given by direct connections from an original layer to all the subsequent layers. Some RS-related works have implemented this block with promising results [150, 151]. This block is illustrated in Fig. 4b.
- **Inception module:** Szegedy et al. [152] proposed this module, which allows the use of multiple filter sizes in parallel, instead of a single filter size in a series of connections. The motivation of this module is that multi-scale convolutional filters have the potential to

Fig. 4 CNN modules—I.
a Residual connection, **b** dense block in the Dense Convolutional Network (DenseNet), **c** inception module



enrich feature representation as the architectures go deeper into the number of layers. The illustration of this module is provided in Fig. 4c. RS-related works have successfully used this module [153–155].

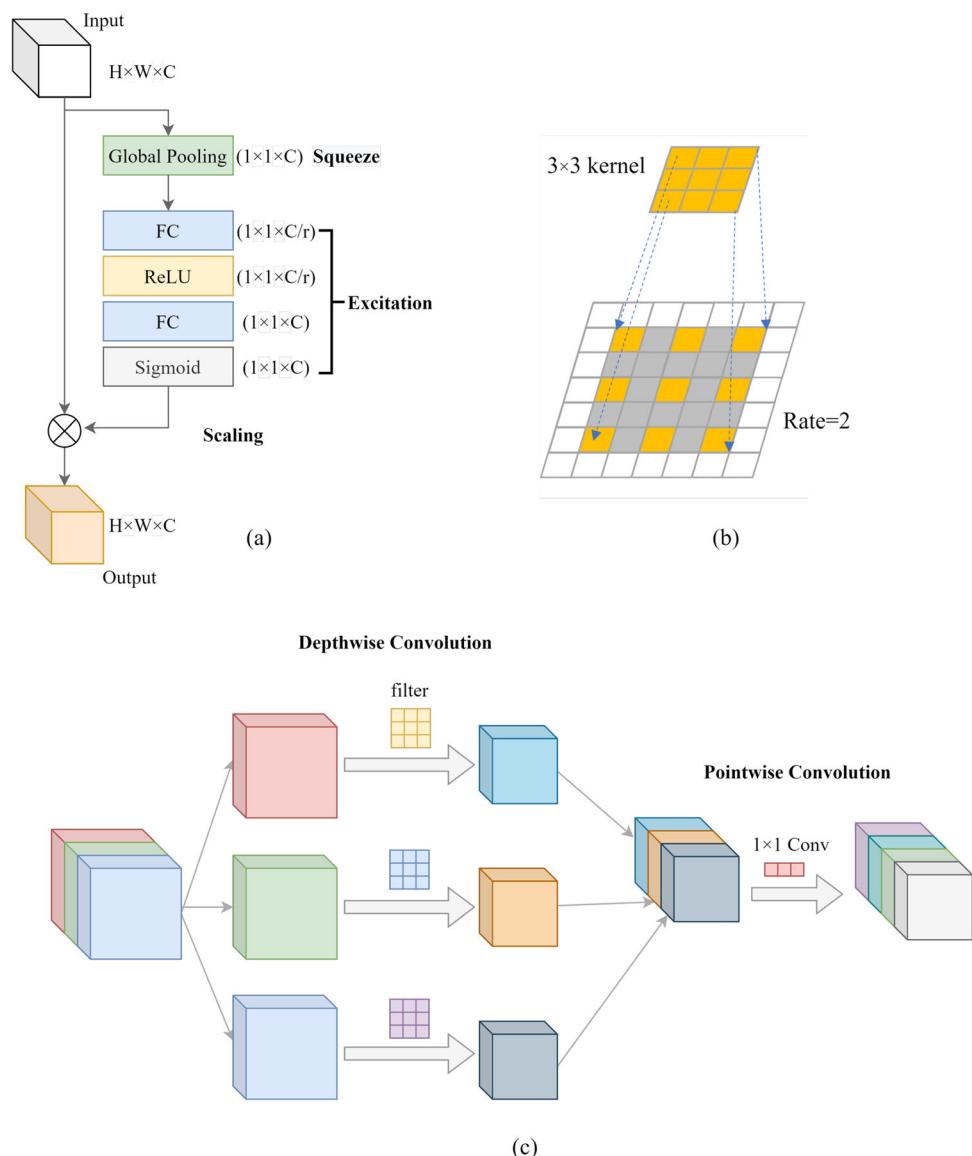
- SE-inception: The Squeeze-and-Excitation block (SE block) was introduced by Hu et al. [144] as an architectural unit that boosts the performance of a network. This architectural unit block empowers the architecture with a dynamic channel-wise feature calibration. The specifics of this architectural unit block are illustrated in Fig. 5a. The SE block has been extensively utilized for RS tasks [156–158].
- Dilated convolution: Yu et al. [145] introduced this type of convolution, which expands receptive fields by introducing gaps between the values of the filter kernel, effectively “dilating” the filter. The expansion is controlled by a dilation factor (l). Figure 5b illustrates dilated convolution with $l = 2$. This technique increases

the receptive field without increasing computation. In RS, several studies have reported the use of dilated convolutions with promising results [159–162].

- Depth-wise separable convolution: Introduced by Chollet [146], depth-wise separable convolution divides standard convolution operations into two steps. First, a depth-wise convolution is applied to each input channel independently. Second, a point-wise convolution is performed, i.e., a 1×1 convolution, mapping the outputs from the depth-wise convolution onto a new channel space. Details of this architectural module are illustrated in Fig. 5c. Experimental results show satisfactory results with the implementation of this architectural module for RS-related tasks [163–165].

Many existing works use one or more aforementioned convolutional network modules with various network connections or designs for RSIC. Zhong et al. [148] adopts

Fig. 5 CNN modules—II.
a Squeeze and excitation inception module, **b** dilated convolution with dilation factor of 2, **b** depth-wise separable convolution



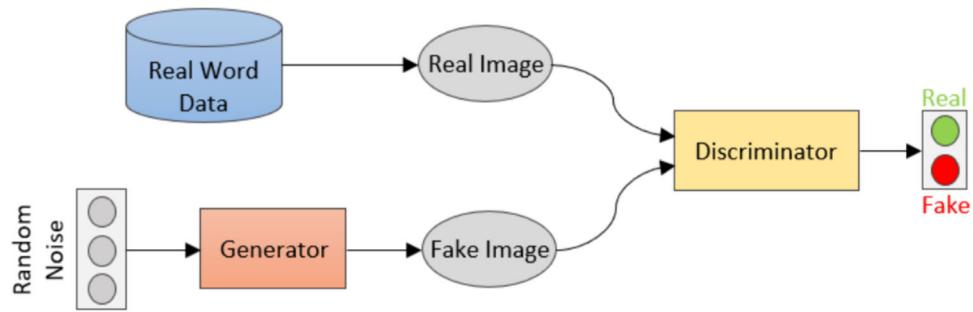
residual connections in hyperspectral data cube, aiming to extract discriminative features from both spectral signatures and spatial contexts in hyperspectral imagery, and outperformed popular classifiers such as kernel support vector machine (SVM) [166], stacked autoencoders, and 3D CNN.

4.1.2 Generative adversarial networks

A generative adversarial network (GAN) [167] has been proposed as semi-supervised and unsupervised DL models that provide a way to learn deep representations without extensively annotated training data. Generating fake data is a key component in GAN, which basically based on the two main networks that represent the GAN. A Generator (G) network tries to generate “realistic” samples and a

Discriminator (D) network distinguishes between the real and generated samples. Figure 6 shows the main concept of GAN. These semi-supervised and unsupervised DL representations have been widely applied to RS applications. Jian et al. [168] developed one class classification technique based on GAN for remote sensing image change detection aiming to train the network only with the unchanged data instead of both the changed and unchanged data. Jiang et al. [169] constructed a GAN-based edge-enhancement method for satellite imagery super-resolution reconstruction to ensure the reconstruction of sharp and clean edges with finely preserved details. Also, Ma et al. [170] introduced a GAN-based method capable of acquiring the mapping between low-resolution and high-resolution remote sensing imageries which aims to restore sharper details with fewer pseudotextures, and

Fig. 6 Conceptual of generative adversarial network



outperformed popular single-image super-resolution methods, including traditional and CNN-based techniques.

4.1.3 Autoencoders and stacked autoencoders

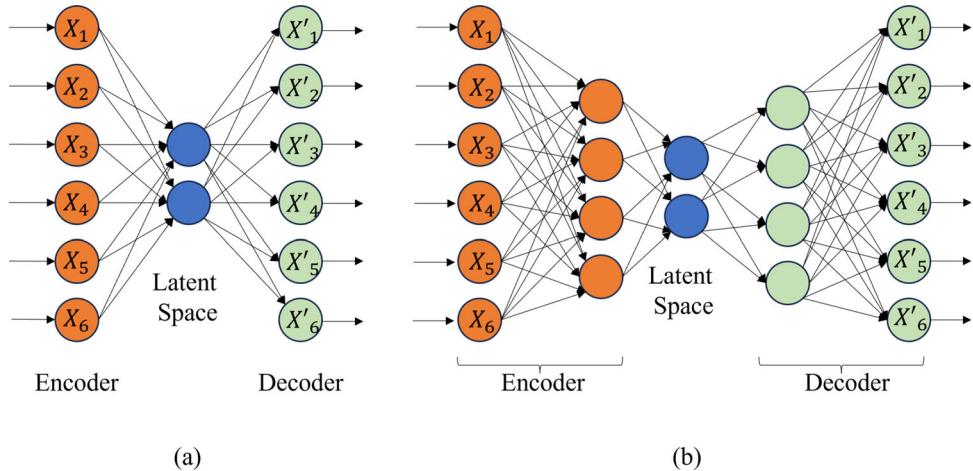
An autoencoder (AE) is a neural network that uses back-propagation to generate an output almost close to the input value in an unsupervised learning framework. As it is shown in Fig. 7a, an AE takes an input and compresses its representation into a low-dimensional latent space. This process is done by the encoder component of the AE. On the other hand, the decoder component of the AE reconstructs the input, scaling the latent space representation to the original input dimension.

Also known as deep autoencoder, stacked autoencoders (SAE) are extensions of the basic AE, consisting of several layers of encoders and decoders that are stacked on top of each other, as shown in Fig. 7b. The use of several layers, in the encoder and decoder portion of the architecture, allows the model to increasingly abstract the representation of the original input as it moves deeper into the network. This makes SAE capable of learning complex features when compared with basic AE.

AE and SAE have emerged as powerful tools for enhancing the performance of DL models for RSIC tasks. For instance, Lv et al. [171] proposed a combination of

SAE with an extreme learning machine (ELM) [172]: SAE-ELM. This ensemble-based algorithm leverages the benefit of the two key components to address challenges in RSIC, including limitation and complexity of the data. The SAE-ELM creates diverse base classifiers through feature segmentation and SAE transformations and accelerates the learning process with the use of ELM. The proposed method showed evidence of improvement in classification tasks and adaptability to different types of remote sensing images. Liang et al. [173] proposed the use of stacked denoising AE for RSIC. This model was built by stacking layers of denoising AE, using the noise input to train the algorithm in an unsupervised approach layer-wise, and turning the robust expression into characteristics by supervised learning using back-propagation. The method outperformed traditional neural networks and SVM performance. On the other hand, Zhou et al. [174] suggested a condensed and discriminative stacked AR (CDSAE) for Hyperspectral image (HSI) classification. This method consisted of two stages: The first stage is a local discriminant, and the second is an effective classifier. The CDSAE aimed to produce highly discriminative and compact feature representation from low-dimensional features. Experimental results demonstrate its effectiveness when compared to traditional methods for HSI classification. Additionally, Zhang et al. [175] introduced the use of

Fig. 7 Illustration of **a** a simple autoencoder and **b** a stacked autoencoder



recursive autoencoders (RAE) as an unsupervised method for HSI classification. This method utilizes spatial and spectral information to learn features from the interaction of the neighborhood of the targeted pixel in an HSI. This approach outperformed methods such as SVM, SVM-CK [176], and SOMP [177]. Similarly, Zhou et al. [178] presented a semi-supervised method for HSI classification with stacked autoencoders (Semis-SAE). The SAE used pre-trained hyperspectral and spatial features, followed by a fine-tuning stage prior to a classification fusion composed of the probabilities from the SAEs with a Markov random field model. The Semis-SAE outperformed state-of-the-art ML methods, such as CNN, GANs, and SVM.

4.1.4 Recurrent neural networks

Recurrent neural networks (RNNs) are a type of artificial neural network designed to recognize patterns in sequences of data that has been widely used in language modeling, text generation, and speech recognition. In RNNs, hidden layers act as the network's memory, which store information based on previous inputs, integrating not only the current input but also the knowledge accumulated from prior data. Figure 8 shows basic structures of FNN and RNN, where (x) is the input layer, (h) is the hidden layer/s, and (y) is the output layer. A , B , and C are the network parameters that are learned during the training of the model.

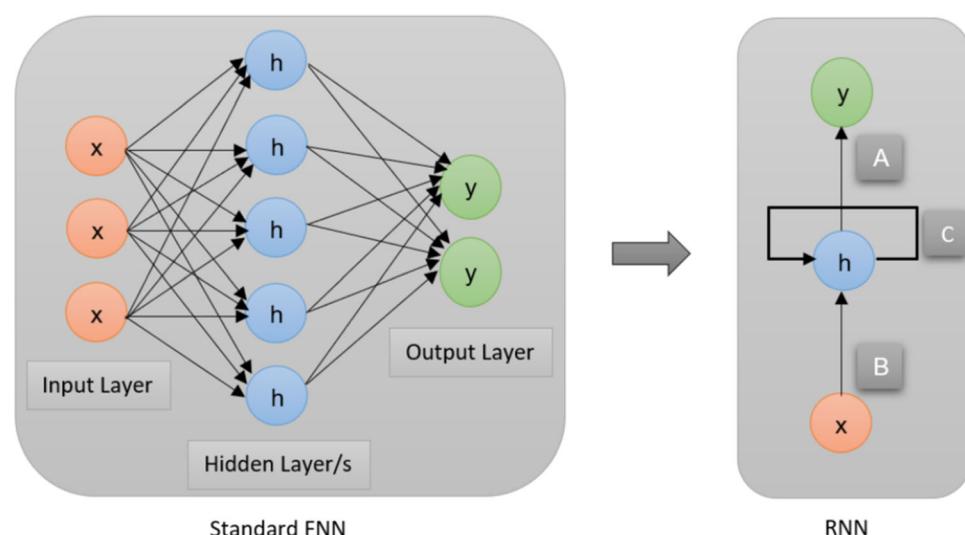
A well-known type of RNN is called long short-term memory (LSTM) [179] which was first introduced to overcome the gradient vanishing and exploding problem. Several variants of LSTM architecture have been proposed as an effective and scalable model to learn long-term dependencies [180]. LSTM has been used for the land-cover classification via multi-temporal spatial data derived

from a time series of satellite images and showed competitive results compared to the state-of-the-art classifiers with the remarkable advantage of improving the prediction quality on low-represented and/or highly mixed classes [181].

The configuration of the input and output determines the design of the RNN architecture, which can be implemented in diverse ways. Among the main such architectures: (1) One-to-one which is known as the vanilla neural network and has been used for general machine learning problems, in which a single input is used to generate a single output. (2) One-to-sequence, in which a single input is employed to produce a sequence of outputs. (3) Sequence-to-sequence, which involves taking a series of inputs and producing a corresponding sequence of outputs. (4) Sequence-to-one takes sequential data to utilize it as an input to generate a single output.

RNN and its variants have been used for RSIC. Mou et al. [182] proposed a deep RNNs architecture with a new activation function to characterize the sequential property of a hyperspectral pixel vector for the classification task. Experimental results showed promise of RNNs in capturing pertinent information for hyperspectral data analysis. RNN variant such as a Patch-based recurrent neural network (PB-RNN) system has been introduced for classifying multi-temporal remote sensing data [183]. PB-RNN is considered as a sequence-to-one architecture and used multi-temporal-spectral-spatial samples to deal with pixels contaminated by clouds/shadows present in multi-temporal data series. Recently, a bidirectional long short-term memory (Bi-LSTM)-based network with an integrated spatial-spectral attention mechanism was developed for hyperspectral image (HSI) classification, enhancing classification performance by emphasizing relevant information [184]. Experiments on three popular HSI datasets

Fig. 8 Recurrent neural network versus standard neural network



demonstrated its superiority over unidirectional RNN-based methods.

4.1.5 Vision transformer-based approach

Originally implemented to solve natural language processing (NLP) tasks, transformers [185] have crossed the threshold of a single domain with high success. Transformers-based models are getting popular in the research community for different fields, including computer vision, RS, and bio-informatics [186–188]. Transformers utilize the self-attention mechanism to handle long-range dependencies of a given sequence, providing the model with a larger “memory” in comparison with traditional recurrent neural networks. For NLP, transformers can deal with larger sequences by the use of tokens, which provide the positional information required to preserve the context of the input. In vision transformers (ViT) [189], this methodology is translated to computer vision, in which the images are divided into patches, as an analogy of tokens and sequences, and then, each patch is linearly projected along with the corresponding embedding positional information. Self-attention is known as the key component within a transformers-based framework. This component helps to capture long-range similarities between a given sequence of tokens by updating the token with aggregated global knowledge. This attention mechanism is mathematically described as follows:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (1)$$

where the vectors \mathbf{Q} , \mathbf{K} , and \mathbf{V} represent the queries, keys, and values, respectively. In this mechanism of attention, the matrices corresponding to the queries and the keys are dot multiplied, as an *attention filter* operation, and then, the output of this operation is normalized through a division operation with the $\sqrt{d_k}$ (the dimension of \mathbf{K}). The *softmax* operation provides a probability distribution for the weights that are being multiplied against the matrix corresponding to the values.

For multi-head self-attention, the aforementioned procedure is repeated in parallel h stands for heads, with different learned linear projections of \mathbf{K} , \mathbf{Q} , and \mathbf{V} ($\mathbf{W}^Q, \mathbf{W}^K, \mathbf{W}^V$). The outputs from the attention functions are concatenated and linearly projected with (\mathbf{W}^O). In summary, the multi-head self-attention mechanism can be represented as:

$$\text{Multihead}(Q, K, V) = \text{Concat}(h_1, \dots, h_i, \dots, h_h)W^O \quad (2)$$

where

$$h_i = \text{Attention}(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V) \quad (3)$$

The success of ViT-based models has increased the interest in this technology within the RS area, and several techniques have been explored in recent years for tasks involving very-high-resolution Imagery (VHR), hyperspectral imagery, and synthetic aperture radar imagery. Figure 9 illustrates the ViT architecture used with hyperspectral data.

For the remote sensing scene classification task, Deng et al. [190] presented a vision transformer-based approach in conjunction with CNN, in which two streams (one ViT and another CNN) generate concatenated features, within a joint optimized loss function framework. On the other hand, Ma et al. [191] explored the use of a transformer-based framework with a patch generation module, analyzing the effect promoted by using heterogeneous or homogeneous patches.

In the task of HSI classification, there are several efforts have been made to develop either purely transformer-based architectures or a hybrid approach that combines the merits of CNN and transformers. For instance, He et al. [192] presented HSI-BERT, a pure transformer-based architecture, with bidirectional encoders. This architecture captures the global dependencies of a target pixel, obtaining a flexible architecture that can be generalized for prediction over different regions with the pre-trained model. Another effort in pure transformers-based architecture is provided by Zhong et al. [193], proposing a spectral-spatial transformer network. Spatial attention leverages the local region feature channels with spatial kernel weights; meanwhile, spectral association leverages the integration of spatial locations for each corresponding feature map. Hybrid efforts, by combining CNN and transformers, have achieved promising outcomes for hyperspectral pixel-wise classification. For instance, the work presented by Wang et al. [194] presented a multi-scale convolutional transformer, which aims to capture spatial-spectral information effectively from a given input. Introduced by Paheding et al. [195] GAF-NAU utilizes the Gramian angular field encoder over the hyperspectral signal to produce a 2D representation for each pixel. This 2D signal is used as input in a U-Net-like framework that combines the attention mechanism with multi-level skip connections. The experimental outcomes from this proposed architecture outperform traditional approaches for pixel-wise hyperspectral classification.

Problems related to the use of SAR image interpretation have been analyzed using pure transformer-based architecture. For instance, Dong et al. [196] utilized vision transformers as a method for PolSAR (Polarimetric SAR) image classification. Each pixel is represented as a token within the architecture, and the long-range dependency is captured by the use of the self-attention mechanism. A hybrid methodology was proposed by Liu et al. [197], in

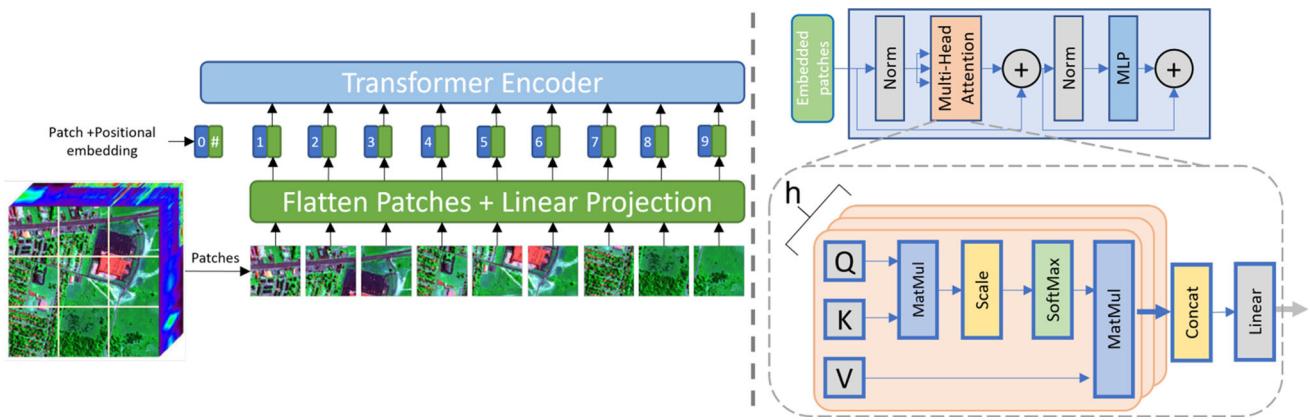


Fig. 9 Illustration of the vision transformer (ViT) [189] architecture. On the left side, the notion behind ViT is presented, including the initial embedding layer and the transformer encoder. Meanwhile, on

the right side, details about the transformer encoder with emphasis on the multi-head self-attention mechanism are provided

which the merits of CNN and transformers were combined to capture, both local and global, feature representation, for the SAR image classification task. On the other hand, the work presented by Chen et al. [198] addressed the detection task for aircraft with SAR imagery by using transformers within a geo-spatial framework composed by image decomposition, geo-spatial contextual attention in multi-scale fashion and image re-composition. Zhang et al. [199] proposed a feature relation enhancement framework, in which a fusion pyramid structure is adopted to combine feature representation at different scale levels, in addition to the use of an attention mechanism for the improvement of the position context information.

4.2 Learning type

4.2.1 Multi-task learning

The goal of the machine learning paradigm known as “Multi-task learning” (MTL) is to learn several related tasks simultaneously [200], compared to the one that learns specific tasks separately as shown in Fig. 10a. The use of MTL is to ensure that the information in one task may be used by other tasks, enhancing the generalization performance of all the involved tasks. In this context, task refers to learning an output target from a single input source [201]. Hence, MTL employs the domain knowledge in the training signals of related tasks as an inductive bias for improving the generalization [202]. This is accomplished, as shown in Fig. 10b, by learning many tasks concurrently while utilizing a common representation; what is learned for one task can aid in learning other tasks.

Multi-task learning in deep learning is often carried out with either hard or soft parameter sharing of hidden layers [203]. The method of MTL that uses hard parameter sharing is the most used one in neural networks. It is often

implemented by preserving several task-specific output layers while sharing the hidden levels across all activities. Overfitting is considerably reduced by hard parameter sharing. On the other hand, with soft parameter sharing, every task has a unique model with unique parameters. To encourage the model’s parameters to be similar, the distance between them is regularized.

The MTL architecture has been used to concurrently complete the tasks of road identification and road center-line extraction [204]. Due to its superior capacity to maintain spatial information, U-Net [205] was chosen as the MTL’s basic network. The multi-task U-Net design contains two networks, a road detection network and a center-line extraction network, which operate simultaneously during training. The hierarchical semantic features obtained from the road detection network are convoluted to create the road center-line extraction network.

4.2.2 Active learning

Active learning (AL), also referred to as query learning or optimal experimental design, is a sub-field of machine learning where the learner makes queries or selects actions that impact what data is to be added to its training set [206, 207]. It is instrumental in scenarios where the labeled data are either scarce or expensive to label the data (such as speech recognition, information extraction, and RSIC). In this model, a small training set is used to train the model initially, and then, an acquisition function decides to obtain a label for unlabeled data points from an external oracle (generally a human expert). These labeled data points are added to the training set, and the model is now trained on this updated training set. Repetitions of this process lead to an increase in the size of the training set.

Active learning for RSIC is a logical choice because it can utilize scarce labeled and abundant unlabeled data. AL

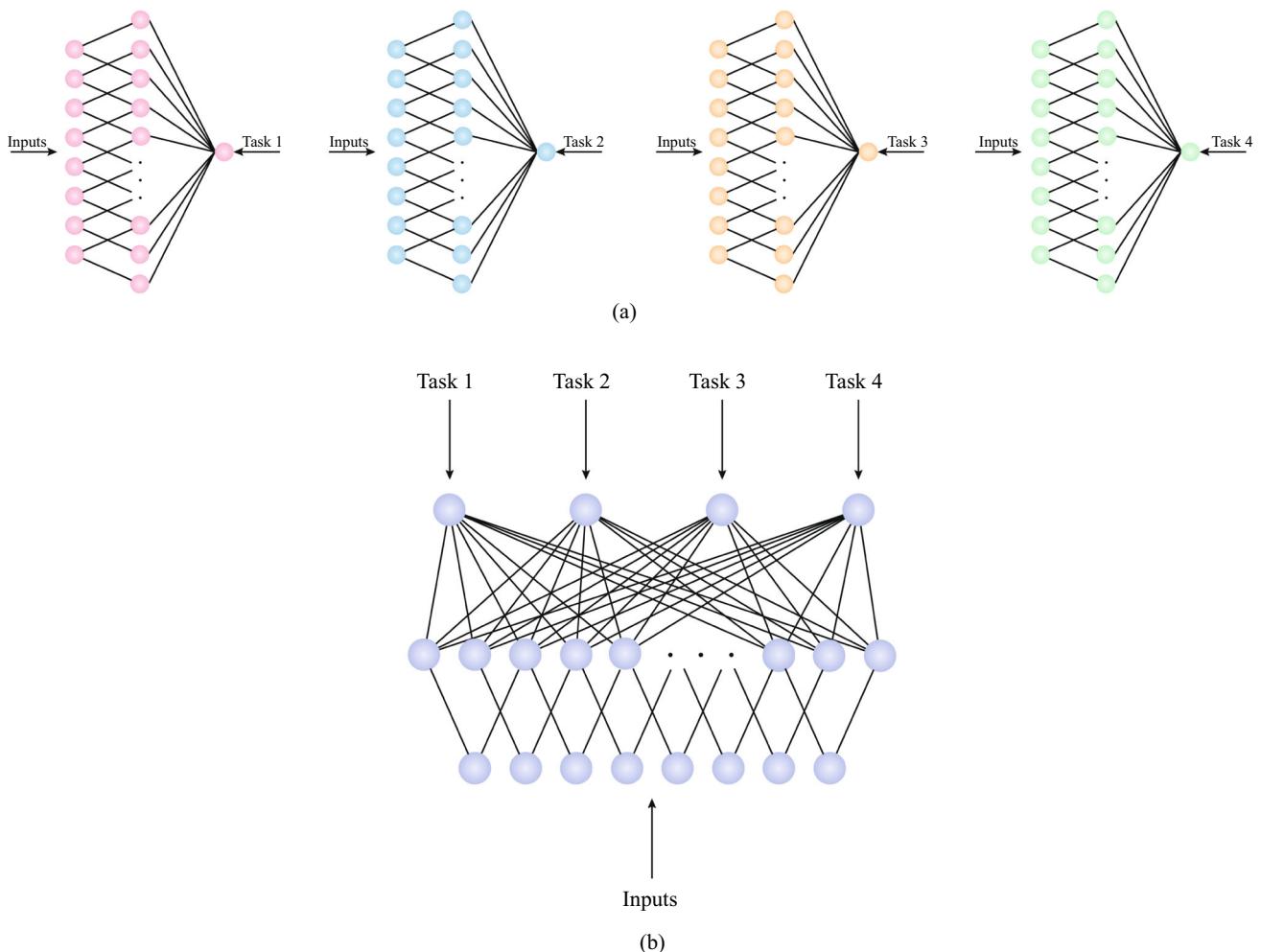


Fig. 10 **a** Traditional methods that use different models for different tasks, **b** multi-task learning via shared representation

has been used in HSI classification [208], in which DRDbSSAL (Discovering Representativeness and Discriminativeness by Semi-Supervised Active Learning) architecture was proposed to extract representative and discriminative information from unlabeled data. This architecture employs multiclass level uncertainty (MCLU), a state-of-the-art approach commonly used in RSIC to identify the most informative samples [209, 210]. Using semi-supervised active learning, it tries to identify representativeness and discriminativeness from unlabeled data using a labeling procedure. It is particularly efficient when there are only a few labeled samples and catches the overall trends of the unlabeled data while preserving the data distribution.

In a work by Haut et al. [211], AL was employed by a B-CNN (Bayesian-Convolutional Neural Network) that was based on the Bayesian Neural Networks (BNNs) [212, 213] for HSI classification. The BNN is a kind of artificial neural network (ANN) that may provide uncertainty estimates and a probabilistic interpretation of

DL models while being resistant to overfitting. They do so by inferring distributions across the models' weights, learning from small data sets, and avoiding the tendency of traditional ANNs to generate overconfident predictions in sparse data areas. Applying the same Bayesian methodology to CNNs can help them withstand overfitting on small data sets while improving their generalization capability.

4.2.3 Transfer learning

In the field of machine learning, transfer learning refers to the process where a model, initially trained for a specific task, is repurposed and utilized for a different but related task. It depicts a situation where knowledge acquired in one context is employed to improve optimization in a different context. TL is commonly employed when the new dataset, intended for training the pre-trained model, is smaller in size compared to the original dataset. TL can convey four distinct types of knowledge for target tasks: relational knowledge, feature representation, parameter

information, and instance knowledge [214]. Deep learning frequently transfers representations by reusing models built on a source model because deep learning automatically learns and keeps the feature representation on network layers and weights. Figure 11 depicts the overall TL procedure. TL typically involves the following three stages:

- Rich source domain data X_S is used to train a deep learning model $Y = f_A$ for source domain Task A until the optimal weights converge and the cost function J_A is minimized.
- The deep learning model $Y' = f_B$ for Task B is built on top of this learned model. The new model $Y' = f_B$ reuses the first n layers from the original model ($n = 3$ in Fig. 11). This ensures that f_B creates representations that adhere to the information discovered in the source domain.
- Using the sparse, labeled training data X_T , the transferred model $Y' = f_B$ is trained to minimize J_B .

The result is a deep learning model for Task B's target domain that incorporates information from the source domain.

In [215], the TL was conducted at three levels: shallow, middle, and deep. In the shallow experiment, features were extracted from the initial blocks of the base models, incorporating the small classification model. The intermediate experiment removes the block from the center of the base model. On the other hand, the deep experiment retained the original base model's blocks, excluding the last classification layers. Experimenting with two CNN models on three distinct remote sensing datasets (UC Merced, AID, and PatternNet) demonstrated that TL, especially fine-tuning, is a robust approach for classifying remote sensing images, consistently outperforming a CNN with randomly initialized weights.

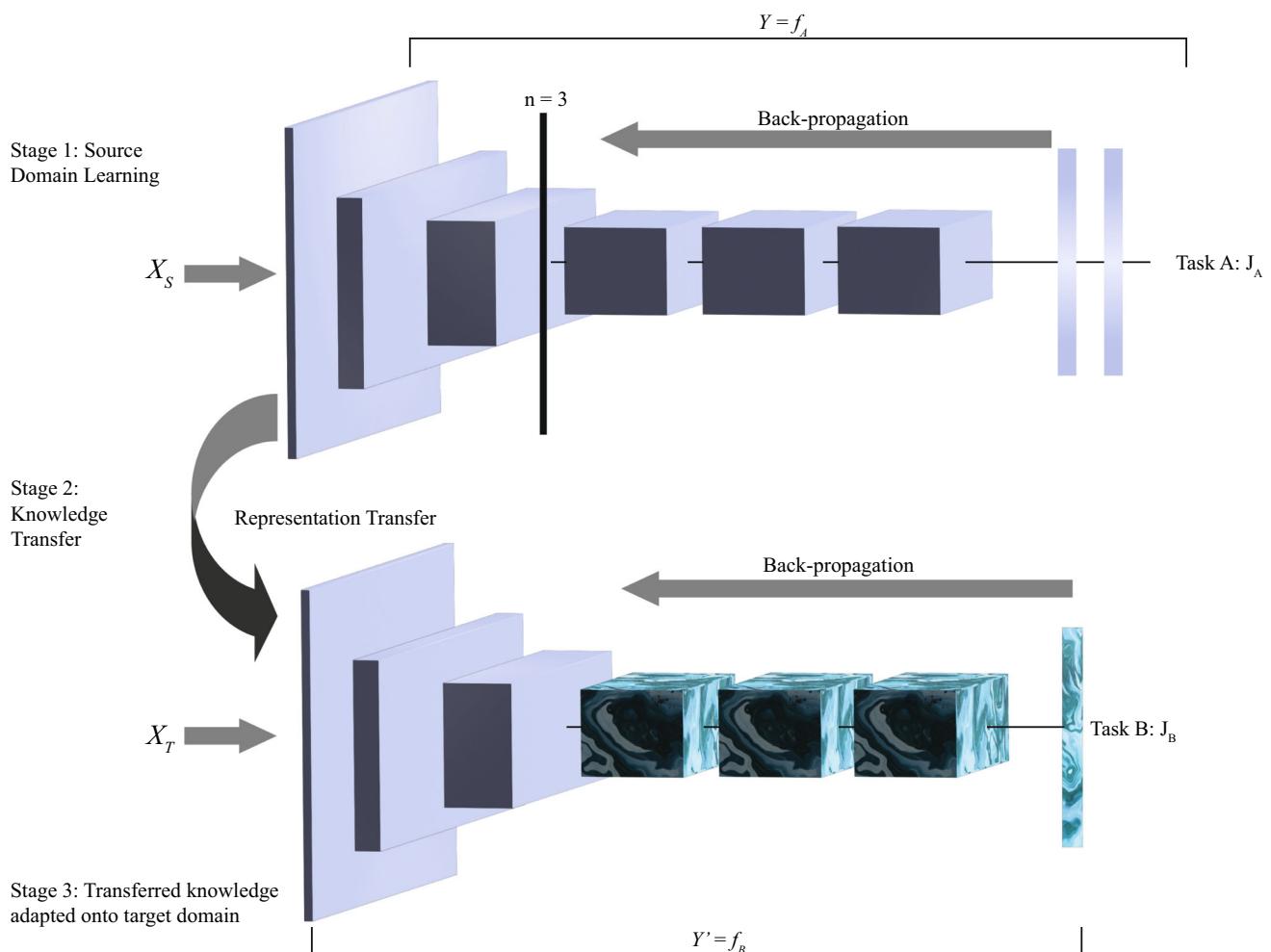


Fig. 11 A transfer learning framework

4.2.4 Ensemble learning

Ensemble learning (EL) combines outputs from multiple models to achieve superior predictive performance [216]. The four primary categories of EL techniques are as follows:

1. *Boosting*, as shown in Fig. 12a, is a technique that creates several classifiers to increase any classifier's accuracy. A classifier chooses its training set depending on how well its last classifier performed. A sample

that a prior classifier has wrongly categorized is chosen more frequently than one that has been successfully classified. As a result, boosting creates a new classifier that can successfully process the new data set.

2. *Bagging* or bootstrap aggregating, as shown in Fig. 12b, is an ensemble learning technique in machine learning designed to improve prediction accuracy by training separate models with bootstrap samples. It typically aids in reducing variance and mitigating overfitting. In classifiers that use bagging, each

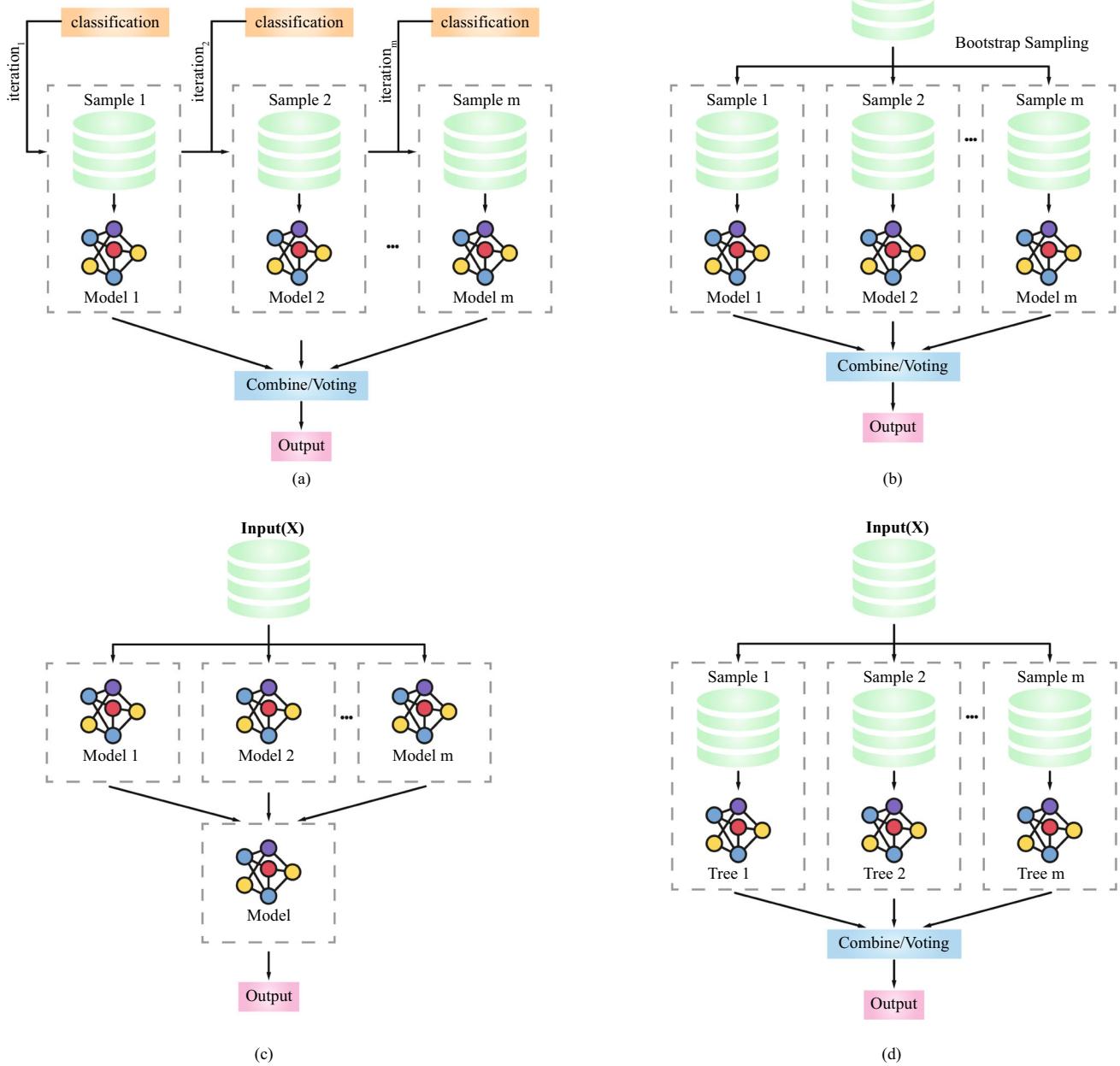


Fig. 12 Ensemble learning types: **a** boosting, **b** bagging, **c** stacking, and **d** random forest

classifier's training set is generated by drawing N instances at random with replacement from the original dataset. This process, known as bootstrap sampling, creates multiple different training sets. In this case, many of the original instances might be repeated, while others might be omitted from the training set. The learning system from a sample creates a classifier, and the final classifier is created by combining all of the classifiers created from the many trials.

3. *Stacking*, as shown in Fig. 12c, is a technique that uses a variety set of models as base learners and utilizes another model or the combiner to aggregate prediction. Here, the combiner is referred to as a meta-learner. The base models are trained first, and their predictions are aggregated as input features for the meta-learner.
4. *Random forest*, as shown in Fig. 12d, is an ensemble learning approach that includes training a large number of decision trees and combining those decision trees' predictions through voting. Instead of having just one decision tree, the random forest method uses sample data from the population to generate several decision trees. When merged, the many samples (bootstrap samples) produce many individual trees that make up the Random Forest.

The EL has been used for the semi-supervised classification of RS scenes [217]. The residual convolutional neural network (ResNet) [141] extracts initial image features. EL is used to exploit the information included in unlabeled data in order to generate discriminative picture representations. Initially, T prototype sets are generated periodically from all accessible data. Each set consists of prototype samples that serve as proxy classes for training supervised classifiers. Afterward, an ensemble feature extractor (EFE) is produced by combining T -learned classifiers. The final image representation is created by concatenating the classification scores from all T classifiers by feeding each image's preliminary ResNet feature into EFE. The experimental results on the publically available AID and Sydney datasets demonstrate that the learned features and semi-supervised technique provide improved performance.

EL has been also used to categorize multiple sensor data using a decision-level fusion technique [218]. CNN-SVM ensemble systems were used to classify Light Detection and Ranging (LiDAR), HSI, and extremely high-resolution Visible (VIS) images. A random feature selection is used to construct two CNN-SVM ensemble systems, one for LiDAR and VIS data and the other for HSI data. VIS and LiDAR data are extracted for texture and height information first. Together with hyperspectral data, these extracted features are used in a Random Features Selection technique to generate various subsets of retrieved features. All feature

spaces are provided as input layers to different deep CNN ensemble systems. Weighted majority voting (WMV) and behavior knowledge space (BKS) were applied to each CNN ensemble as the final classifier fusion approaches. The result indicated that the suggested technique produces more precision and outperforms several current methods.

4.2.5 Multi-instance learning

The conventional data description applies to single-instance learning, in which each instance of a learning object is characterized by several feature values and, perhaps, an associated output. In contrast, a bag (learning sample or object) is linked to several instances or descriptions in multiple instance learning (MIL) [219, 220]. The objective of a MIL classification is to assign unseen bags to a particular class driven by the class labels within the training data or, more precisely, to use an estimation model constructed from the labeled training bags. An instance-based algorithm's overall design is shown in Fig. 13. It is represented as a bag, $Y \in \mathbb{N}^Y$, holding n instances, $y_1, \dots, y_n \in Y$.

There are four different options to choose from:

1. K (Set of bag labels): K 's length is the number of classes.
2. Λ (Set of instance labels): Bag sub-classes or instance-level concepts might correlate to instance labels.
3. M (MIL assumption): It requires the construction of an explicit mapping between the set of instance labels and the set of bag labels. It is a function $M : \mathbb{N}^\Lambda \rightarrow \mathbb{K}$ that links the class label of instances within a bag to the class label assigned to the bag.
4. A technique for locating the instance classifier $m : \mathbb{Y} \rightarrow \Lambda$ (utilized for the classification of instances within each bag).

Since each unique instance requires a class label, single-instance learners cannot be applied directly to MIL data. (A

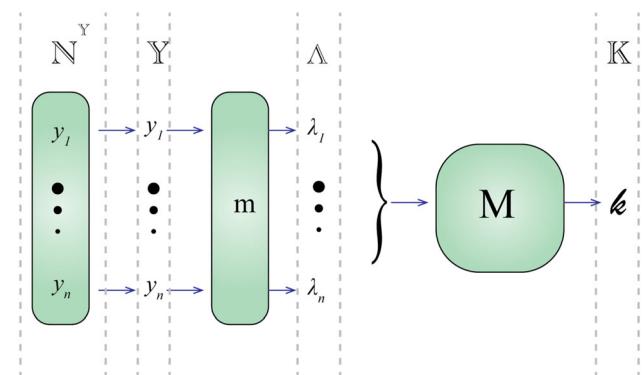


Fig. 13 MIL general architecture [219]

bag classifier is required.) A MIL hypothesis is performed over instance labels to get the bag label.

For the application of MIL to RS, the input images in an RS system are broken down into multiple sub-images, each of which is handled as a different instance of that class. The learning system will then discover which sub-image is crucial for correctly predicting the image's class. The total accuracy of the algorithm may be increased by constructing a neural network that can concentrate on the area of the picture that is more crucial to the categorization.

The MIL has been used for classifying scenes in RS [221]. Generally, one segment of a scene identifies its class, while the others are unimportant or belong to another class. The first stage of the proposed method splits the picture into five instances (the center image plus the four corners). Subsequently, a deep neural network is trained to retrieve intricate convolutional features from individual instance and ascertain the optimal weights for their fusion through weighted averaging.

Multiple instance learning has been employed as the end-to-end learning system [221]. Here, two instances were used: one to characterize the spectral information of multispectral (MS) photographs and the other to capture the spatial information of panchromatic (PAN) images. The relevant spatial information of PAN and the associated spectral information of MS are extracted using deep CNN and stacked autoencoders (SAE), respectively. The last step was joining the features from the two instances together using fully connected layers. Four aerial MS and PAN images were used in classification studies, and the results showed that the classifier offers a workable and effective solution.

4.2.6 Reinforcement learning

Reinforcement learning (RL) is a machine learning paradigm where an agent learns to make decisions by taking actions in an environment to maximize some notion of cumulative reward [222]. Unlike supervised learning, where the model learns from a dataset of input-output pairs, RL focuses on learning from the consequences of actions, guided by a reward signal. This framework makes RL particularly well suited for problems where an agent interacts with an environment, making it applicable to various remote sensing tasks.

In remote sensing, RL has been utilized for tasks such as satellite task scheduling, resource management, and dynamic path planning for unmanned aerial vehicles (UAVs). The key advantage of RL in these scenarios is its ability to handle sequential decision-making problems and adapt to changing environments.

One application of RL in remote sensing is dynamic path planning for UAVs. UAVs are increasingly used for

environmental monitoring, disaster response, and agricultural surveillance. RL algorithms, such as Q-learning or deep Q-networks (DQN) [223], can be employed to optimize the flight paths of UAVs to maximize coverage, minimize energy consumption, or avoid obstacles [224]. By learning from interactions with the environment, the UAV can adapt its path in real-time to changes in the environment, such as moving obstacles or areas of interest.

Another significant application is satellite task scheduling, where multiple satellites need to be coordinated to maximize the overall mission objectives, such as maximizing data collection or minimizing observation gaps. RL techniques can be used to optimize the scheduling of satellite observations, taking into account various constraints like limited satellite resources, orbital dynamics, and conflicting observation requests [225]. This can result in more efficient use of satellite resources and improved data acquisition strategies.

RL has also been applied to the problem of data fusion in remote sensing, where information from multiple sensors needs to be integrated to produce a comprehensive understanding of the observed environment. By treating the fusion process as a sequential decision-making problem, RL algorithms can learn optimal strategies for combining data from different sources to enhance the accuracy and reliability of the resulting information [226].

Despite its potential, the application of RL in remote sensing comes with challenges, such as the need for large amounts of data to train the models and the complexity of accurately modeling the environment. However, ongoing research is addressing these challenges [227], making RL a promising tool for advancing remote sensing technologies.

5 Sensor types and remote sensing datasets

A wide variety of datasets has been collected using an array of sensors. In this section, we will examine different types of sensors and describe different categories of remote sensing datasets.

5.1 Sensor types

In RS, sensors can be described as mechatronic instruments that comprise electrical, mechanical, and computing elements. Carried on board satellites, airborne vehicles, or installed (*in situ*) on the ground, they record electromagnetic signals as digital data to study Earth processes or atmospheric phenomena.

Satellite-mounted sensors can cover large areas of the Earth's surface, but are limited to the satellite's orbital path and are obstructed by clouds [228]. Example applications for satellite RS include monitoring forest fires [229],

drought [230], atmospheric particulate matter concentrations [231], and sea ice thickness [232].

Airplane, helicopter, and unmanned aerial vehicle (UAV) mounted sensors have the advantage of high to very-high spatial resolution, custom flight paths, and Light Detection and Ranging (LiDAR) capabilities; however, they require flight operation efforts and have relatively small area coverage. Example applications include crop monitoring and vegetation mapping (Table 3), disaster response (Table 4), and environmental monitoring (Table 5).

Ground-based remote sensing systems (GRSS) are installed on the Earth's surface, where several sensors are spatially distributed and accessed collectively. Example applications include: in situ real-time monitoring of algae blooms and water quality inland and in oceans [233]; landslide mapping and early warning [234]; distributed surface temperature; and wind speed profile measurement [235].

Sensors can be passive or active. An example of a passive sensor is a satellite-mounted infrared (IR) camera. It captures the thermal radiation emitted or reflected by objects on the Earth's surface. In this case, the heat radiation originates from the Sun and reflects into the IR camera aperture. An active sensor, such as radio detection and ranging (Radar), produces radiation energy to expose the objects it is sensing by capturing the reflected electromagnetic radiation.

Sensor resolution is an important characterization of the imaging sensor modality. Three types of resolution are meaningful in RS, namely spatial, spectral, and temporal resolution [236]. The composition of these types of resolution can affect the feasibility of RS applications as shown in Fig. 14. Spatial resolution refers to the sensor's ability to resolve small details. For example, a satellite image might have a resolution of 1 pixel per meter, whereas a UAV sensor may have twice the spatial resolution at 1 pixel per 0.5 meter, i.e., 2 pixels per meter. Spectral resolution refers to the number of discrete electromagnetic radiation bands the sensor can process, i.e., record the average power from. A high spectral resolution sensor is sensitive to narrower, and more, spectral bands. For a given spectral sensing range, a low spectral resolution sensor will have fewer, and wider, spectral bands, than a high spectral resolution sensor. For example, a color camera with red, green, and blue (RGB) channels (3 bands between 450 and 650 nm), has higher spectral resolution than a bandpass (1 band between 1150 and 1300 nm) short-wave infrared (SWIR) camera. Finally, temporal resolution in RS refers to the sensor's ability to repeat sensing the same area. For example, a UAV-mounted sensor has a much higher temporal resolution than a satellite sensor, which requires a long time to

complete the Earth orbit and return to the designated area for repeated sensing [237].

Earth's atmosphere blocks some electromagnetic wavelengths due to the presence of Ozone, water, carbon dioxide, and other particles. This protects the surface from dangerous radiation such as X-rays and high-energy ultraviolet (UV) wavelengths. RS sensors are developed to measure the radiation that is not blocked by the atmosphere, i.e., that passes through the "atmospheric window."

Ultraviolet (UV) sensors are sensitive in the range between 10 and 400 nm. RS applications that utilize UV sensors include Ozone layer detection, ocean color, and oil spill detection [238, 239].

Red-green-blue (RGB) sensors are essentially color cameras sensitive to the visible spectrum color bands 380 nm (shortest blue) to 850 nm (longest red). This is the range of wavelengths the human eye is sensitive to. The Landsat-8 satellite, for example, includes RGB sensors as follows: red (640–670 nm); green (530–590 nm); and blue (450–510 nm) [240]. Some examples of RGB use in RS include urban sprawl and drought mapping [241, 242].

Near infrared (NIR) sensors are sensitive to the electromagnetic band between 850 and 900 nm. In addition, short-wave infrared (SWIR) are sensitive between 900 and 2500 nm. These two bands measure reflected infrared radiation, as opposed to thermal radiation, which requires medium and long infrared red (MWIR, LWIR) sensors, to detect. These span the wavelengths between 3000 to 5000 nm, and 8000 to 12000 nm, respectively. RS applications utilizing infrared sensors include: NIR and SWIR in the estimation of soil carbon content [243]; vegetation canopy studies using near-infrared imaging [244]; and thermal imaging for urban climate and environmental studies [245].

The passive microwave electromagnetic range is between 1 and 200 GHz (1.5 and 300 mm). Like thermal sensors, passive microwave sensors collect radiation emitted by objects. Water and oxygen molecules in Earth's atmosphere absorb some of the shorter wavelengths. RS applications include monitoring the spatial distribution of permafrost [246] and land surface temperature [247]. Beyond the passive microwave radio are the higher-frequency detection and ranging (RADAR) waves. Synthetic aperture radar (SAR) sensors are active and send microwave pulses that reflect off of objects such as the Earth's surface back to the transmitter, usually on a satellite. An example application is the estimation of vegetation thickness for forest fire studies [248], the study of sea surface winds and waves from spaceborne SAR [249], arctic ice thickness monitoring [250].

Multispectral imagery refers to the utilization of between 3 and 10 bands in the electromagnetic spectrum.

Table 3 Public aerial datasets for precision agriculture and forestry

References	Dataset name/ website	Year	Sensors	Application	Content
[267]	Global Wheat Head Dataset 2021	2021	RGB	Crop monitoring	The dataset is composed of more than <i>6000 images</i> of 1024×1024 pixels containing 300k+ unique <i>wheat heads</i> , with the corresponding bounding boxes
[268]	RiceSeedling Datasets	2021	RGB	Crop classification	An image dataset of rice paddy for data sharing by making labeled and unlabeled data findable and accessible through domain-specific repositories
[269]	Sick Fir Tree dataset	2021	RGB	Disease detection	A dataset for the identification and categorization of individual diseased fir trees, comprising Orthomosaics, DSM, nDSM, ROI, annotations for sick firs, healthy firs, deciduous trees, and treetops. The dataset encompasses data from nine aerial missions conducted across four locations on Zao Mountain
[270]	MOPAD	2021	RGB	Plant classification	The dataset contains <i>363,877 oil palms images</i> of five categories: healthy palms, dead palms, mismanaged palms, smallish palms and yellowish palms
[254]	UAV-ARSP	2020	Multispectral	Crop monitoring	The <i>588 multispectral</i> aerial images portray an agricultural area of one square kilometer in a rural zone in Kilosa—Tanzania. Crops: <i>sesame, pigeon pea, banana, maize, mango, beans, cowpea rice and sunflower</i>
[259]	WHU-Hi	2020	Hyperspectral	Crop classification	Contains 3 datasets: WHUHi-LongKou, WHUHi-HanChuan, and WHUHi-HongHu. WHUHi-LongK: consists of aerial images of six types of crops. WHUHiHanChuan dataset consists of seven crop types. WHUHi-HongHu, consists of seventeen different cultivars of three main crop types, which are cotton, rape, and cabbage
[271]	Forest Dataset	2020	RGB	Forest Tree classification	We have seven winter mosaics capturing a blend of mountain forest and a mosaic featuring a pine tree plantation interspersed with broadleaf trees. The data has been annotated and is now accessible to the public
[272]	Northern Leaf Blight dataset	2019	RGB	Disease detection	The resultant annotated dataset comprised a total of <i>3000 resized aerial images</i> and corresponding ground truth masks containing <i>5234 lesion instances</i>
[255]	Cactus Aerial	2019	RGB	Plant classification	This dataset presents more than 16,000 examples of a columnar cacti (<i>Neobuxbaumia tetetzo</i>) for plant recognition or classification
[256]	weedNet	2018	Multispectral	Weed control	The dataset contains <i>132,243, and 90</i> annotated multispectral images of <i>crops, weeds, and crop-weed mixtures</i> . Each training image/test image consisted of near-infrared (NIR, 790 nm), Red channel (660 nm), and NDVI imagery
[256]	WeedMap	2018	Multispectral	Weed control	The datasets consist of <i>129 directories and 18,746 image files</i> . datasets from <i>sugar beet</i> fields in Eschikon, Switzerland, and Rheinbach, Germany, with a time interval of 5 months
[273]	Joint stem detection	2018	Multispectral + RGB	Weed control	The datasets contain 921 RGB + NIR and 400 RGB-only images as well as their corresponding annotations for the semantic segmentation and the stem detection task. Classes: (1) <i>soil</i> , (2) <i>sugar beet</i> , (3) <i>dicot-weed</i> , and (4) <i>grass-weed</i>
[274]	NLB Disease dataset	2018	RGB	Disease detection	The dataset contains <i>18,222 images</i> , all taken in the field, and <i>105,735 annotations</i> by one of two human experts. This is the largest publicly available collection of classified images of any single plant disease
[275]	Grass- Broadleaf	2017	RGB	Weed control	This image dataset has <i>15336 segments</i> , being 3249 of soil, 7376 of soybean, 3520 grass, and 1191 of <i>broadleaf weeds</i> . Images were segmented and the segments were annotated manually with their respective class

Table 4 List of the publicly available datasets for disaster response categorized based on the type of natural disaster

References	Dataset name/ website	Year	Sensors	Application	Content
[251]	Change detection dataset	2022	Multispectral and SAR	All/common	The dataset comprises uniform cases with both multispectral (MS) and synthetic aperture radar (SAR) images. Additionally, it includes diverse cases encompassing MS/SAR images, covering 14 instances of natural disasters, farming, and construction
[277]	FLAME Dataset	2021	RGB	Wildfire	The dataset includes 5814 fire images captured by drones during a controlled burning of piled debris in an Arizona pine forest. Both the recorded videos and images are annotated and labeled on a frame-by-frame basis
[278]	EU Flood Dataset	2019	RGB	Flood	Dataset comprises 3710 flood images that have been annotated by domain experts to assess their relevance for three tasks: identifying the flooded area, determining inundation depth, and assessing water pollution
[252]	Corsican Fire Database	2017	Multispectral	Wildfire	The database contains visible spectrum and near-infrared (NIR) images and provides a large number of images captured in this spectrum. It contains video sequences captured simultaneously in color and NIR spectrums
[279]	VisiFire	2015	RGB	Wildfire	The dataset is composed of two main parts: the first 14 videos characterized by the presence of fire and the last 17 videos that do not contain fires
[280]	wildfire smoke database	2014	RGB	Wildfire	The dataset comprises 2977 images distributed across 23 video sequences, with 16 of them featuring smoke. Smoke is discernible in various forms in 1,839 images, while the remaining images (1138) are essential for quality evaluation and encompass phenomena that might lead to false alarms

Table 5 List of the publicly available datasets for environment monitoring categorized based on the application

References	Dataset name/ website	Year	Sensors	Application	Content
[283]	FRACTURES	2021	RGB	Rock outcrops	The dataset includes filtered 500 x 500 .png tiles specifically featuring scenes with fracture data. Additionally, it provides .png binary masks for semantic segmentation and original geo-referenced shapefile annotations
[284]	IndustrialEstate	2021	RGB and Thermal	Temperature Monitoring	A collection of data from a light industrial estate in Switzerland was gathered using an eBee X fixed-wing drone equipped with a senseFly Duet T thermal mapping camera. The thermal data were processed to create a thermal 3D mesh and a temperature index map
[285]	RIVER ICE SEGMENTATION	2020	RGB	Ice segmentation	This dataset offers digital images and videos capturing surface ice conditions, collected during the winter seasons of 2016–2017 from two rivers in Alberta: the North Saskatchewan River and Peace River
[286]	LeConte Glacier Dataset	2019	RGB	Ice monitoring	The dataset contains field measurements at LeConte Glacier, Alaska, to study the impacts of subglacial discharge on plume dynamics and submarine melting. The aerial images are used to measure ice elevation and velocity at the LeConte terminus
[287]	Bottle Detection	2018	RGB	Waste plastic Detection	This dataset contains 25,407 UAV images of bottles with various kinds of backgrounds. An oriented bounding box was used to accurately and compactly annotate the bottles. The fully annotated images contain 34,791 bottles
[288]	Dune Dataset	2016	RGB	Dune monitoring	Investigating the influence of vegetation and dune size on the initial stages of dune development through the utilization of unmanned aerial vehicle (UAV) imaging

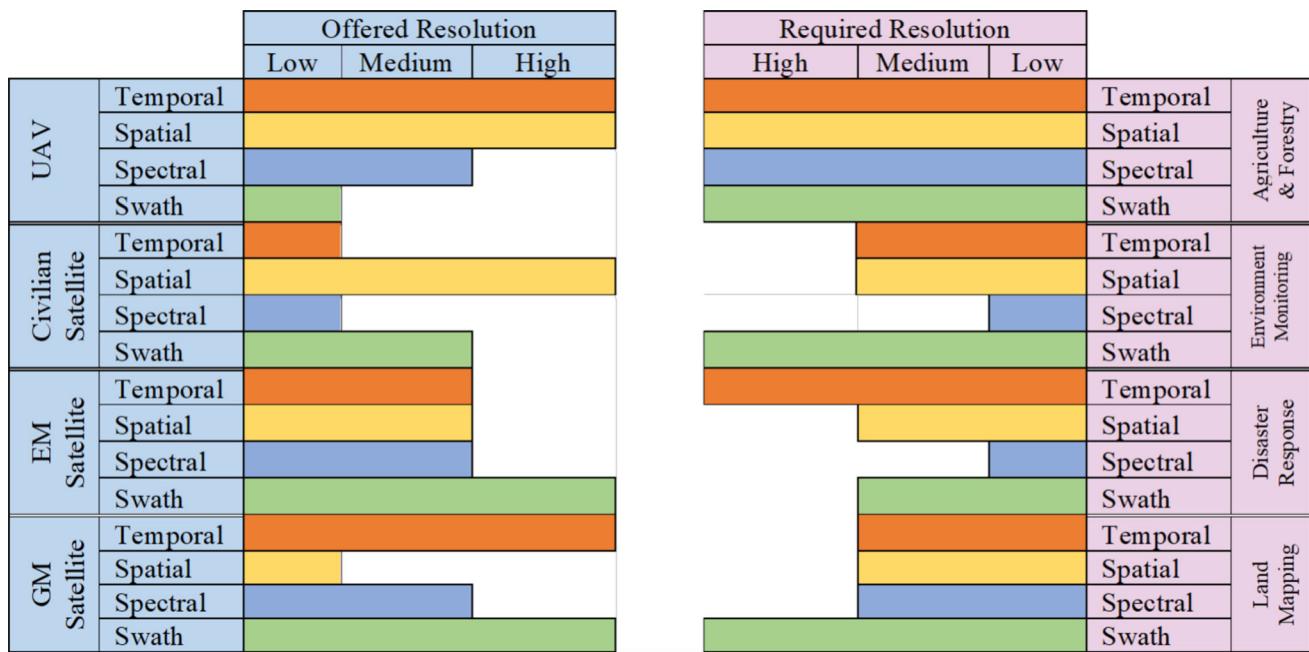


Fig. 14 Applications/RS platforms resolution matrix

RGB, for example, can be considered multispectral imaging as it collects information from three color bands. Also, the Landsat-8 satellite can measure 11 bands from indigo to Thermal IR, in roughly 40 nm steps, i.e., bands [240]. Example applications are in: disaster response with multispectral SAR [251], and multispectral RGB and NIR [252]; land mapping using multispectral LiDAR [253]; and agriculture and forestry [254–256].

Hyperspectral imaging has a much higher spectral resolution than multispectral, with narrower bands between 10 and 20 nm, as well as the measurement of from hundreds to thousands of bands [257]. The Hyperion (EO-1 satellite) [258], for example, measures 220 spectral bands between 400nm (violet) and 2500nm (SWIR). Example applications utilizing hyperspectral imaging are: land mapping [253]; agriculture and forestry [255, 259]; and reservoir water quality monitoring [260].

5.2 Aerial datasets

Unmanned aerial vehicles (UAVs) are commonly used nowadays as a remote sensing platform that holds different types of imaging devices ranging from RGB, and thermal cameras to hyperspectral and miniaturized SAR devices. Despite the fact that UAVs have limited power sources and can only cover relatively limited areas compared to their satellite counterpart, UAVs offer an attractive solution when on-demand images from low altitudes are required in time-sensitive applications. Further, with their availability, low cost, easy-to-use, and high operational capability to

capture images at high temporal and spatial resolutions, UAVs market has grown dramatically over the last decade and they are now used widely in different RS applications.

We used the data published in [261] to show the difference between using UAV and satellite platforms in terms of temporal, spatial, and spectral resolutions as well as swath. In [261], the researchers categorized the types of satellites into three categories as follows:

- *Global monitoring satellites (GM)* such as MODIS Terra work in high orbit and provide high temporal resolution and relatively high swath but offer a moderate spatial and spectral resolution.
- *Environmental monitoring satellites* such as Landsat and Sentinel-2 provide moderate temporal, spatial, and spectral resolutions and high swath.
- *Civilian satellites* such as Pleiades or Ikonos provide high spatial resolution but low temporal and spectral resolution as these satellites are at low orbit.

While different types of satellites provide different levels of resolutions, all UAV types guarantee high temporal and spatial resolution; however, they provide low swath. Nevertheless, UAVs offer an attractive solution for RS applications that require high temporal and spatial resolutions such as agriculture and disaster response. On the other hand, UAVs are not used widely in land-cover/land-use mapping due to the need for a high swath. Figure 14 illustrates the required resolution and swath in different RS applications and what resolution is offered by different RS platforms as indicated in [261].

With this big interest and growth in using UAVs as a remote sensing platform by governments and the RS research community, we present, in this section, a summary of up-to-date public UAV (aerial) datasets that were collected or synthesized over the last decade. In contrast to the very few existing review papers [262–266], we summarize the most popular and recent UAV datasets that cover the RS applications presented in section 4.2 (i.e., agriculture and forestry, environment monitoring, disaster response, land mapping). This summary of the available UAV datasets should greatly help the research community in its efforts to develop DL algorithms for aerial data analysis.

5.2.1 Datasets in agriculture and forestry

Developing reliable and robust DL methods for crop monitoring, disease detection, weed control, plant classification, and other precision agriculture and applications requires a high-quality, large-scale dataset. Practically, it is hard to build such datasets due to the cost and efforts that are needed for image acquisition, classification, and annotation. Therefore, datasets that are publicly available play an integral role in fostering remote sensing scientific progress and significantly reducing the cost and time needed for dataset preparation. In this subsection, we present a tabulated summary (Table 3) of recent publicly available datasets in the field of RS in agriculture and forestry.

The datasets are classified based on the application within precision agriculture. The table also provides the reader with a link to the dataset website as well as a brief description of the contents of the dataset. Our search was limited to aerial images that are acquired by UAVs, drones, airplanes, or any flying device. We also provide the sensor type in each dataset which is entirely dependent on the type of application [65], as indicated in Fig. 15. For example,

multispectral images are used mainly in precision irrigation and disease detection while RGB is mainly used for weed control.

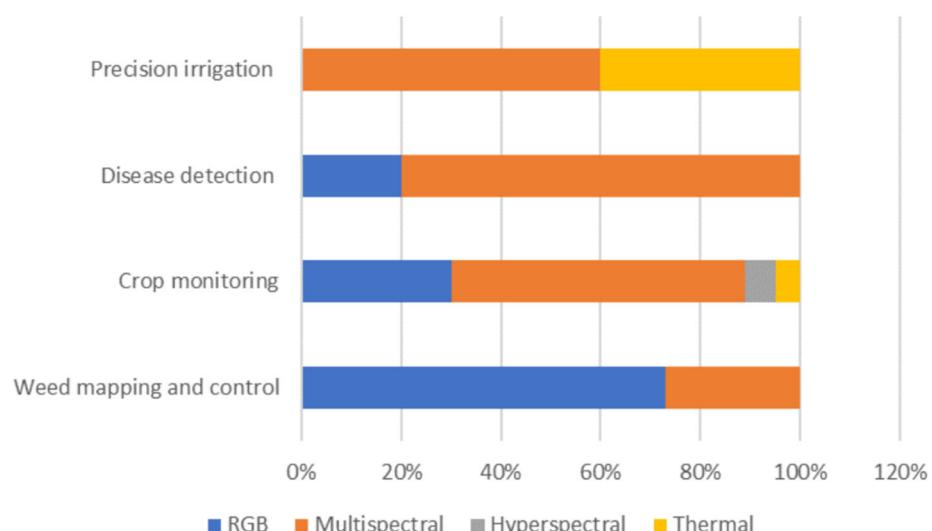
5.2.2 Aerial datasets for disaster response

Unlike agriculture and forestry, finding public datasets of aerial images for disaster response can be challenging. As shown in Fig. 16, 53% of the data sources for damage assessment as a result of natural disasters are acquired by satellites and only 21% are acquired by UAVs [276]. However, with the increased interest in using UAVs for disaster response and damage assessment over the last 5 years, it is expected that UAVs will gain more volume as a source of data than satellites due to their high temporal and spatial resolution. Therefore, we are presenting in this section the publicly available aerial image datasets which are categorized by the type of disaster and ordered by the date of the last update of the dataset as shown in Table 4. A brief description of the dataset (based on the publishing source) is also provided.

5.2.3 Aerial datasets for environment monitoring

Due to the advantages of UAVs mentioned earlier in this section, they have been increasingly used for environmental monitoring, especially in hard-to-reach areas. UAV remote sensing technology is capable of operating at different spatial resolutions while keeping a high temporal resolution. Furthermore, with the recent advancements in the miniaturized multispectral and LiDAR sensors, UAVs have become the best choice for distinguishing between natural and pollutant materials and building precise 2D/3D maps of the land surfaces [120]. In recent times, unmanned

Fig. 15 Data modality used in different precision agriculture applications



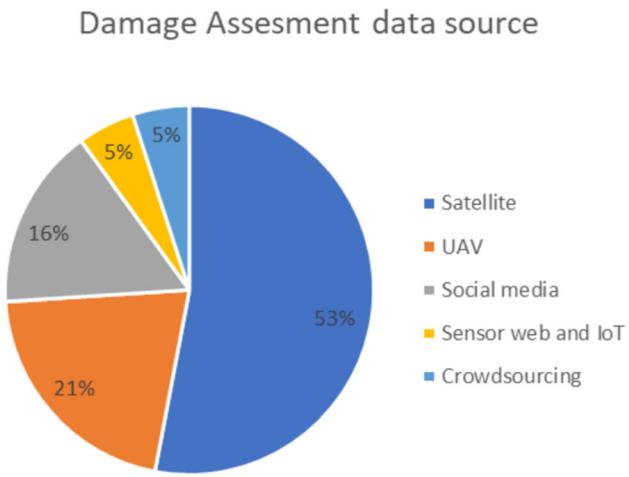


Fig. 16 Damage assessment data sources in response to different natural disaster

aerial vehicles (UAVs) have emerged as a significant transformative factor in marine monitoring. They play a crucial role in addressing biological and environmental issues, encompassing tasks such as monitoring invasive species, conducting surveys and mapping, observing marine animal activities, and monitoring marine disasters [281, 282]. Imagery data acquired by UAVs are normally analyzed using DL which requires datasets of real images collected by UAVs. In this subsection, we present the recent publicly available datasets based on our review of the literature. Table 5 provides a list of the dataset categorized based on the application as well as a description of the dataset contents.

5.2.4 Aerial datasets for land mapping

UAVs have become valuable tools for land mapping and surveying due to their ability to capture high-resolution aerial images and data efficiently. Moreover, UAVs provide rapid cover for large areas which is important for time-sensitive projects, and can access hard-to-reach or hazardous areas, making them suitable for mapping terrain, forests, cliffs, and other challenging landscapes. UAVs can also capture images from multiple angles and altitudes, enabling the creation of 3D models of the land. These models are valuable for urban planning, archaeological site preservation, and environmental assessment. Repeated UAV flights over time can be used to monitor land changes, such as urban expansion, deforestation, or erosion. For land mapping, deep learning techniques are employed in various ways to extract valuable information from aerial. Deep learning models can be used to detect changes in land cover and land use over time by comparing historical and current imagery. It can be also applied to identify and extract building footprints from high-resolution aerial

imagery and extract road networks, enabling the creation of detailed road maps. This information is essential for urban planning, disaster response, and infrastructure development. Deep learning algorithms proficiently classify land parcels or areas into different land-use types, such as residential, commercial, agricultural, or industrial, and can process LiDAR data to create accurate digital elevation models, which are essential for terrain analysis, and flood modeling. To implement deep learning in land mapping, large labeled datasets are required for model training, and specialized neural network architectures. Consequently, we present a compilation of recently accessible datasets, as per our comprehensive literature review. Table 6 illustrates a categorized list of datasets, detailing their respective applications and dataset contents.

5.3 Satellites datasets

There is a large number of active remote sensing satellites and databases, which are accessible through freely available and commercial interfacing software programs. Some of the common sources of remote sensing data are the United States Geological Survey; National Oceanic and Atmospheric Administration; National Aeronautics and Space Administration Earthdata; NASA Earth Observations; European Space Agency; Japan Aerospace Exploration Agency; Airbus Defense and Space; MAXAR Company; Planet Labs; Satellite Imaging Corporation; Apollo Mapping. We present a selection of popular and recently cited datasets of overhead imagery, mostly from remote sensing Earth satellites. The available satellite data is vast. Most datasets combine satellite imagery from multiple satellites with ground-based measurements to train artificial intelligence models to create algorithms that can be used to process new data. For example, a local 2011 study of mangrove forests in the coastal region of West Africa was used to develop a general model for mangrove detection globally [297].

We present a partial list of interesting satellite-based datasets in the four most popular applications. Agriculture and forestry datasets are included in Table 7. Such datasets commonly reply to RGB color imagery, as well as infrared to detect the extent of vegetation on the surface. Often data products for analysis via machine learning are created to include urban sprawl, human development, and water levels. The results are often global maps useful for assessment and planning.

Furthermore, disaster response satellite datasets are listed in Table 8; common sensor modalities include infrared, radar, RGB, and multispectral imagery. Applications include the study of global storms, floods, and landslide patterns, as well as volcano activity, temperature extremes, wildfires, and building damage mapping.

Table 6 List of the publicly available datasets for land mapping based on the application

References	Dataset name/ website	Year	Sensors	Application	Content
[289]	UC Merced Land Use Dataset	2010	RGB	Land-use classification	Overhead images of 21 urban structure classes. 100 images per class. 256×256 pixels with 30 cm spatial resolution.
[290]	DynamicEarthNet	2022	RGB, IR	Land-use classification	The dataset contains daily observations of 75 separate areas of interest around the globe, with 7 land-use and land-cover (LULC) semantic classes
[291]	SinkholeNet	2022	RGB	Land-use classification	The dataset comprises 467 high-resolution UAV images. The alignment of these spatially overlapped images was achieved using the Visual Structure from Motion system (VisualSfM) algorithm to produce an orthomosaic image and a Digital Elevation Model (DEM) for the study area
[292]	UAVid	2020	RGB	Land-cover classification	This dataset has 300 images and each of size e 4096 x 2160 or 3840 x 2160. In total, 30 small UAV video sequences are captured in 30 different places to bring variance to the dataset. Data collection takes place under favorable weather conditions with ample illumination
[293]	UNCD	2018	RGB	Land use	The UMCD Dataset (about 3.50 GB) is composed of two main sets (geo-referenced and not geo-reference) of challenging video sequences acquired at very low-altitude. This dataset can be used for land-use classification
[253]	IEEE GRSS DATA FUSION	2018	Fusion of multispectral LiDAR and hyperspectral data	Land management	The provided data encompasses Multispectral-LiDAR point cloud information at wavelengths 1550 nm, 1064 nm, and 532 nm; Hyperspectral data spanning a spectral range of 380–1050 nm with 48 bands at a 1-m Ground Sample Distance (GSD), along with high-resolution RGB imagery at a 5-centimeter GSD
[294]	LandCover	2017	RGB, LiDAR and ALS	Land-cover classification	The dataset comprises aerial images with ground control points and direct geo-referencing data. The data cover various areas of interest for research, such as buildings with diverse sizes and roof materials, roads, and vegetation
[295]	Pavia University Scene	2022	Hyper spectral (430–960 nm)	Land-cover classification	Contains 42,776 aerial images of 9 classes of land-cover types such as asphalt, meadow, gravel, trees, shadows, painted metal, etc.
[296]	Indian Pines	2022	Hyper spectral with 224 bands(400–2500 nm)	Vegetation classification	Contains 10,249 aerial images of 16 classes of vegetation and agricultural fields, including Alfalfa, Corn, Grass, Oats, Wheat, woods, etc.

In addition, environmental monitoring satellite datasets are listed in Table 9. Applications include monitoring wildlife habitats, ice sheet monitoring, climate predictions, atmospheric gas concentrations, freshwater reservoir assessment, ocean flux, and plastic pollution monitoring. Sensor modalities include combinations of multispectral, radar, GPS, ground data, infrared, and RGB.

Finally, satellite datasets for scene classification and object segmentation detection are listed in Table 10. Applications include semantic segmentation of overhead scene pixels into categories of land use, as well as road, car, and ship detection. Sensor modalities include RGB, infrared, radar, and multispectral imagery.

6 Discussion and future direction

6.1 Perspective from imaging and sensing

The combination of UAV and satellite imagery with deep learning algorithms in remote sensing is anticipated to persist in transforming our capacity to observe and comprehend the Earth's surface and its dynamic processes. This advancement is poised to contribute to progress in fields such as agriculture, environmental conservation, disaster management, and urban planning, among others. Drawing from our expertise and comprehensive review of this field, we present key trends and areas of development as follows:

Table 7 Selected satellite-based datasets used in forestry and agriculture

References	Dataset name/ Website	Year	Modalities	Applications	Content
[297]	Mangrove Forests Distribution Polygon, v1 (2000)	2000	RGB, NIR, SWIR	Study West African mangrove forests subject to sea level rise threat	A map of “true mangroves” based on training data from 200 km of African coastline mangrove areas. 30 m resolution.
[297]	Global Mangrove Forests Distribution, 2000	2000	RGB, NIR, SWIR	Global mangrove forest distribution modeling, land-cover change analysis, global carbon accounting	A global map of “true mangroves” based on training data from 137,760 km ² in 118 countries and territories
[298]	Deforestation, v1 (2000–2012)	2012	RGB, NIR, SWIR	Study global tree cover extent, loss, and gain	A time-series analysis of 654,178 Landsat images. 30 m resolution
[299]	MA Biodiversity, v1 (1950–2001)	2001	Moderate Resolution Imaging Spectroradiometer and Data	Global biodiversity studies	Data spanning 50 years, including disease agents, animal and fish populations, pollution, and loss of natural land cover
[300]	Gravity Recovery and Climate Experiment (GRACE), v1 (2002–2016)	2016	Gravity sensor	Global freshwater availability trends	A global gridded data set at a spatial resolution of 0.5° that represents trends in between 2022 and 2016
[301]	Croplands, v1 (2000)	2000	Moderate Resolution Imaging Spectroradiometer	Human-environment interaction: agriculture	Global maps for agricultural land distribution in 2000
[301]	Pastures, v1 (2000)	2000	Moderate Resolution Imaging Spectroradiometer	Human-environment interaction: pastureland	Global maps for animal grazing in 2000
[302]	HANPP Collection	2004	Radiometer	Human appropriation of net primary productivity (NPP) studies	Global map of the annual amount of Earth’s terrestrial net primary production humans require for food, fiber (including fabrication) and fuel, as well as the amount of NPP required
[303]	Food Insecurity Hotspots Data Set, v1 (2009–2019)	2019	Data product	Food security studies	Maps (250 m resolution) for intensity and frequency of food insecurity (2009–2019) and effects

- Enhancing the spatial and temporal resolution of satellite and UAV imagery is imperative. This enhancement will facilitate more frequent and detailed monitoring of landscapes, ecosystems, and urban areas.
- Multi-sensor integration is a critical requirement within the field of remote sensing data analysis and remote sensing, where data from a variety of sensors such as optical, thermal, LiDAR, and hyperspectral sensors must be effectively combined. This integration offers the potential for a more comprehensive understanding of the environment. However, it also presents a new challenge in the implementation of novel deep learning architectures, as they need to be capable of handling the fusion of multi-modal data for enhanced analysis. Standardize data and calibration procedures across different sensors and the development of DL architectures designed for multi-sensor integration with the

- collaboration of the sensor manufacturers, could help to mitigate the aforementioned challenges
- In dynamic scenarios applications, such as disaster response and precision agriculture, real-time and on-device processing of UAV imagery is an immediate need. However, this also requires efficient algorithms and hardware capable of handling large volumes of data with low latency. The development of lightweight DL models optimized for on-device processing could be needed for this scenario, in which edge computing and distributed processing systems are used to reduce latency problems.
- There will be an increasing trend toward tailoring deep learning models with satellite/UAV platforms to specific applications, whether it is precision agriculture, forestry management, environment monitoring, or disaster management. This requires a deep understanding

Table 8 Selected satellite-based datasets used in disaster response

References	Dataset name/ Website	Year	Modalities	Applications	Content
[304]	Geocoded Disasters (GDIS)	2021	Geocoded Data	Coded data for global disaster analysis: floods; storms; landslides; droughts; volcanoes; extreme temperatures	39,953 locations geocoded for 9,924 disasters that occurred worldwide in the years 1960–2018
[305]	Hephaestus Dataset	2022	Interferometric SAR (InSAR)	Volcano state classification, semantic segmentation, ground deformation, the atmospheric signal in InSAR imagery	Contains 19,919 manually annotated Sentinel-1 satellite SAR interferograms of 44 different volcanoes globally
[306]	WildFires	2019	RGB, NIR, MIR, and Thermal IR, and data	Predict the occurrence of wildfires using machine learning and data mining	804 rows vectors of data (386 “fire,” 418 “no fire”) collected from 2 million hectares in Canada in the period 2013–2014
[307]	VIIRS NightFire (VNF)	2013	NIR, SWIR	Detect fires and combustion sources on Earth’s surface at night	Infrared data recorded at night. Bands M10, M11, M12, and M13 bands are used to detect combustion sources
[308]	Global Fire Emissions Database (GFED4)	2013	Radiometer, RGB, IR, Rainfall/Fire Maps	Global area burning assessment	Data product of global green area loss between 1995 and 2013
[309]	xBD	2019	multi-band RGB	Assessment of building damage after wildfires, earthquakes, tsunamis, wind damage, etc.	Labeled pre and post-damage top view imagery of buildings with bounding boxes and labels of damage caused. Images of 700,000 buildings from 15 countries covering 5,000 square kilometers
[310]	Harmonized Landsat Sentinel-2 (HLS)	2018	Multispectral	Flood, mudslide, and urban destruction monitoring after weather-related disasters	A Virtual Constellation of surface reflectance data acquired by the Operational Land Imager and Multi-Spectral Instrument on Landsat-8 and Sentinel-2. 2–3 day and 30 m resolutions
[311]	NASA Disasters Program	2021	True Color RGB, SAR	Maps and data to monitor hurricane and flooding development and aftermath	Near real-time Geographic Information Systems (GIS) format producer from multiple satellites. 3–4 m resolution

of the domain and the ability to adapt to changing requirements.

- As the use of UAV/satellite and deep learning in remote sensing grows, there will be a greater focus on ethical and regulatory issues, including privacy concerns, data security, and compliance with local and international regulations. Methods such as strong data encryption and access control measures to protect sensitive data could be implemented to mitigate privacy concerns. Additionally, educate stakeholders about the importance of data privacy and security.

6.2 Perspective from learning algorithms

Supervised classification algorithms need labeled data to classify RS images correctly. However, collecting labeled samples for learning is time-consuming and costly. Accurate classification maps need sufficient and high-quality training data. One of the major limitations of these

supervised approaches is the scarcity of a dataset superior in quantity and quality for training the classifier. In remote sensing image classification, data labeling can be tricky and time-consuming. Thus, semi-supervised learning (SSL) approaches have been used to train the classifier using labeled and unlabeled data to enrich the input to the supervised learning algorithm and increase classification accuracy. This is especially beneficial in remote sensing, where gathering and classifying vast training data may be time-consuming and costly. However, SSL approach poses some challenges. It requires carefully selecting the labeled and unlabeled data and an appropriate semi-supervised learning algorithm. Different SSL algorithms may perform variably depending on the dataset characteristics and classification tasks, where there is a need for algorithms that can adapt to the specific properties of remote sensing data, such as high dimensionality and spectral variability. To address this issue, some advanced data selection methods could be an option, such as developing

Table 9 Selected satellite-based datasets used in environmental monitoring

References	Dataset name/ Website	Year	Modalities	Applications	Content
[310]	Harmonized Landsat Sentinel-2 (HLS)	2018	Multispectral	Monitoring of phenomena such as natural disasters, cyclical crop yields, and wildlife habitats	Surface reflectance data acquired by the OLI and Multi-Spectral Instrument on Landsat-8 and Sentinel-2
[312]	MERRA-2	2017	Hyperspectral, microwave, GPS	climate prediction, aerosols, and ice sheet monitoring	Timestamped ozone, wind, rain, temperature, and other data between 1980 and 2012
[313]	MODIS MYD11C3	2018	Data transformations	Earth climate monitoring, land-climate interactions	Temporal and spatial surface air temperature records (2003–2016)
[314]	CT2019B	2022	Satellite and ground data	Study concentrations of atmospheric carbon dioxide	Global carbon dioxide surface flux, with a focus on North America (2000–2018)
[315]	GRBD	2020	Multi-satellite imagery and altimetry	Global water reservoir studies and bathymetry	30 m resolution data for 347 global reservoirs, representing half of the global water reservoir capacity
[316]	CMORPH	2023	Passive MW, IR	Global precipitation estimation	Precipitation rates from 1998 to the present in mm per hour with 8 km by 8 km spatial, and 30 to 60 minute temporal resolution.
[317]	HydroSat	2022	Satellite altimetry, imagery, and gravimetry	Global surface water storage studies	Hydrological time series (2016 to 2018), based on: water level; river width estimation; water storage anomalies; river discharge
[318]	J-OFURO3	2019	Data product	Global ocean flux studies	Ocean surface heat, momentum, freshwater fluxes, and related parameters globally between 1988 and 2013
[319]	Global Surface Water	2021	Multispectral	Surface water levels studies.	Pixel level classified Landsat images (3 million): water; land; non-valid. 30-m resolution (1984–2021)
[320]	Hyperspectral Plastics Dataset	2021	Hyperspectral	Ocean plastic pollution detection	Images from 30 test sites. 13 pixel-level classes over water and land, including plastics. Plastics sub-classes: tires; plastic; greenhouse; waste site

reinforcement learning-based strategies for dynamically selecting the most informative labeled and unlabeled samples during training. In addition, future research can focus on enhancing domain adaptation techniques [330] within SSL frameworks to handle variations in data distributions across different regions and sensor types. We also suggest exploring other types of machine learning techniques like curriculum learning [331], where the model is trained on easier tasks or samples first, gradually increasing the complexity to improve learning efficiency and performance.

In addition, multi-task learning (MTL) can be helpful in RSIC when multiple related land-cover classification tasks need to be performed. It allows the model to learn shared features that may be useful for multiple tasks. One of the main advantages of MTL is that it allows a model to learn shared features that may be useful for multiple tasks, which can help improve the model's performance on each task. For example, an MTL model could be used to classify

vegetation and water bodies in a satellite image, and the model could learn useful features for both tasks. Thus, MTL is a helpful tool for RSIC, mainly when multiple related tasks need to be performed, and can help improve accuracy and efficiency. In terms of future trends in MTL development, we envision that research will continue developing robust techniques for learning the degree of relatedness between tasks in a context-aware manner, potentially using attention mechanisms or graph-based representations. Besides, as loss function is an important task of MTL, innovative approaches to enhance the functionality of task-specific layers with adaptive loss weighting strategies would be beneficial. This ensures the importance of each task's loss is dynamically adjusted based on task performance and difficulty.

Furthermore, active learning (AL) can be particularly useful when labeled data are scarce and it is not feasible to label the entire dataset. One of the main advantages of active learning is that it can help improve a model's

Table 10 Selected satellite-based datasets used in scene classification, object segmentation, and detection

References	Dataset name/ Website	Year	Modalities	Applications	Content
[321]	SEN12MS	2019	Dual Polarity SAR, multispectral images.	Image classification, semantic segmentation, and data fusion. Land- cover maps of all inhabited continents overall meteorological seasons	180,748 image triplets holding dual-polarity SAR (Sentinel-1), multispectral images (Sentinel-2), and land-cover maps (MODIS)
[322]	Massachusetts Roads Dataset	2013	RGB	Road segmentation from overhead imagery	1171 images 1500 × 1500 pixels of a 2600 square kilometer area with urban, suburban, and rural areas with 1 pixel per square meter
[323]	SpaceNet 6	2020	Quad- polarity SAR, RGB	Semantic segmentation of city buildings in all types of weather conditions	48,000 unique building rooftops in the city of Rotterdam, with height information labels.
[324, 325]	RSD46-WHU	2017	RGB	Overhead scene classification: airplane, building, bridge, oil tank, playground, etc.	117,000 images of 46 classes. Between 500 and 3000 images per class
[326]	Cars Overhead with Context (COWC)	2016	RGB, Grayscale	Deep learning-based car detection in overhead images	Overhead images from 6 cities: Toronto; Selwyn; Potsdam; Vaihingen; and Columbus. Also, 32,716 annotated car objects with 58247 unique negative examples and 15 cm per pixel resolution
[327]	SAR-Ship- Dataset	2019	SAR	Object (ship) detection in images with complex backgrounds. Scene classification	43,819 SAR images of ships, from the Gaofen-3 and Sentinel-1 satellites
[328]	HRSC2016	2017	RGB	Object (ship) recognition in high- resolution color satellite images	1061 sea and sea-land images with 3786 annotated ship samples
[329]	NEPUOWOD V1.0	2021	RGB	Oil Well detection	1192 oil wells in 432 images from Daqing City, China

performance by focusing on labeling the most informative samples. This can be particularly useful in RS image classification, where labeling the entire dataset may not be practical due to the size and complexity of the images. The future trends in active learning are likely to focus on integration with other learning methods such as meta-learning algorithms to allow the active learning system to adapt its querying behavior dynamically.

Besides, in RS image classification, transfer learning (TL) can be beneficial when there is a lack of labeled data for the specific land-cover classification task, but there is a related task for which labeled data are available. This is because the model can leverage the knowledge it has learned from the source task to better classify the data for the target task. We anticipate the development of TL will be emphasized from the following aspects: (1) Unsupervised domain adaptation, which allows models to adapt to new domains without requiring labeled data in the target domain. (2) Multi-view multi-source transfer learning. With ever-increasing remote sensing data availability, sensor and methodology advancement, ongoing research in TL is most likely to leverage multi-source remote sensing data from multiple views or perspectives to better

generalize to a target domain and thus improve performance. (3) Parameter-efficiency models or lightweight models. This makes it easier to fine-tune and deploy in resource-constrained environments. (4) Privacy-preserving and efficient transfer learning, such as developing transfer learning models in federated learning settings that are robust to adversarial attacks and compatible with varying computational resources.

Finally, ensemble learning (EL) can be used to improve the accuracy and robustness of a model's classification tasks by using the strengths of multiple models. For example, if different models are trained on different subsets of the data or with different algorithms, the EL model can use the complementary strengths of each model to achieve better results. Another advantage of EL is that it can help reduce the risk of overfitting, as it can use the predictions of multiple models rather than just one. Ongoing work is suggested in the area of adaptive ensemble methods that can dynamically adjust model aggregation based on data characteristics and tasks. Another interesting area of research is designing ensembles that work efficiently in resource-constrained environments, such as edge computing and mobile devices.

6.3 Perspective from foundation models

The rapid advancement of large language models (LLMs) has significantly impacted various natural language tasks. This progression heralds an era where visual foundation models (VFM) will become integral to numerous RS applications. We anticipate the use of VFMs in tasks like remote sensing image classification, segmentation, and captioning. Furthermore, there is potential for the development of domain-specific VFMs tailored to particular RS applications. These applications could include precision agriculture, disaster monitoring and response, and climate change analysis, among others. This trend reflects a growing synergy between advanced machine learning techniques and practical, real-world applications in remote sensing and environmental monitoring. However, the development of VFM for RS applications requires overcoming challenges related to model complexity, scalability, and interoperability. To this end, future development in this area will focus on scalable VFM methods capable of handling large volumes of RS data. Additionally, we envision that domain-specific or data-specific VFMs will become more prevalent for managing various types of optical and radar RS data. For instance, models like SpectralGPT [332] can provide deeper insights into advancing spectral RS big data applications.

6.4 Perspective from the ethical aspect of remote sensing with deep learning systems

Remote sensing technologies, when combined with deep learning systems, offer remarkable capabilities for monitoring and understanding the Earth's surface. However, the integration of these technologies raises ethical concerns that must be carefully addressed to ensure their responsible and ethical use.

One of the primary ethical considerations in remote sensing with deep learning systems is privacy. These systems can capture highly detailed information about individuals and their activities, raising concerns about unauthorized surveillance and data misuse. To address these concerns, privacy-preserving techniques should be employed [333, 334]. In addition, given the sensitivity of remote sensing data, ensuring its security is paramount. This involves implementing robust data encryption techniques to enable the efficient encryption of remote sensing data [335, 336].

Another critical issue is the potential for bias and unfairness in deep learning models used for remote sensing. If these models are trained on biased data, they may produce unfair outcomes. To mitigate these risks, it is

essential to use diverse and representative datasets and implement bias detection and mitigation techniques in the model development process [337].

Compliance with regulations and standards, as well as establishing accountability for decisions made using remote sensing data, is crucial. This includes establishing clear policies and procedures for data collection, processing, and use, providing transparency about data sources and analysis algorithms, and implementing mechanisms for accountability and oversight [338, 339].

Addressing these ethical concerns is essential for the responsible and sustainable deployment of remote sensing technologies with deep learning systems. It enables us to harness the potential of these technologies for positive impact while mitigating potential harm to individuals, society, and the environment.

7 Conclusions

This review offers a thorough examination of diverse deep learning models and frameworks, along with their architectural designs, tailored for remote sensing image classification tasks. It also delves into various sensor modalities and imaging platforms pertinent to RS image analysis. Additionally, the review encompasses both existing and potential applications in this domain. Further, it analyzes current trends in the application of deep learning within the remote sensing field and proposes predictions for ongoing and future developments. Our aim is to inspire increased interest and engagement within the vision and remote sensing community, encouraging the use of various deep learning models not only for RS image classification challenges but also for broader applications in the field.

Data availability Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

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 paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright

holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Yi Y, Zhang W (2020) A new deep-learning-based approach for earthquake-triggered landslide detection from single-temporal radiopie satellite imagery. *IEEE J Sel Top Appl Earth Observ Remote Sens* 13:6166–6176
- Toan NT, Cong PT, Hung NQV, Jo J (2019) A deep learning approach for early wildfire detection from hyperspectral satellite images. In: 2019 7th international conference on robot intelligence technology and applications (RiTA), IEEE, pp 38–45
- Maimaitijiang M, Sagan V, Sidike P, Hartling S, Esposito F, Fritschi FB (2020) Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens Environ* 237:111599
- Babaeian E, Paheding S, Siddique N, Devabhaktuni VK, Tuller M (2021) Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning. *Remote Sens Environ* 260:112434
- Sidike P, Sagan V, Maimaitijiang M, Maimaitiyiming M, Shakkor N, Burken J, Mockler T, Fritschi FB (2019) DPEN: deep progressively expanded network for mapping heterogeneous agricultural landscape using worldview-3 satellite imagery. *Remote Sens Environ* 221:756–772
- Srivastava S, Vargas-Munoz JE, Tuia D (2019) Understanding urban landuse from the above and ground perspectives: a deep learning, multimodal solution. *Remote Sens Environ* 228:129–143
- Feddema JJ, Oleson KW, Bonan GB, Mearns LO, Buja LE, Meehl GA, Washington WM (2005) The importance of land-cover change in simulating future climates. *Science* 310(5754):1674–1678
- Jia X, Kuo B-C, Crawford MM (2013) Feature mining for hyperspectral image classification. *Proc IEEE* 101(3):676–697
- Zhang L, Zhang L, Du B (2016) Deep learning for remote sensing data: a technical tutorial on the state of the art. *IEEE Geosci Remote Sens Mag* 4(2):22–40
- Zhu XX, Tuia D, Mou L, Xia G-S, Zhang L, Xu F, Fraundorfer F (2017) Deep learning in remote sensing: a comprehensive review and list of resources. *IEEE Geosci Remote Sens Mag* 5(4):8–36
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
- Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Adv Neural Inform Process Syst* 25
- Zhao Z-Q, Zheng P, Xu S-T, Wu X (2019) Object detection with deep learning: a review. *IEEE Trans Neural Netw Learn Syst* 30(11):3212–3232
- Garcia-Garcia A, Orts-Escalano S, Oprea S, Villena-Martinez V, Garcia-Rodriguez J (2017) A review on deep learning techniques applied to semantic segmentation. [arXiv:1704.06857](https://arxiv.org/abs/1704.06857)
- Young T, Hazarika D, Poria S, Cambria E (2018) Recent trends in deep learning based natural language processing. *IEEE Comput Intell Mag* 13(3):55–75
- Alom MZ, Taha TM, Yakopcic C, Westberg S, Sidike P, Nasrin MS, Hasan M, Van Essen BC, Awwal AA, Asari VK (2019) A state-of-the-art survey on deep learning theory and architectures. *Electronics* 8(3):292
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A et al (2021) Highly accurate protein structure prediction with alphafold. *Nature* 596(7873):583–589
- Yao C, Luo X, Zhao Y, Zeng W, Chen X (2017) A review on image classification of remote sensing using deep learning. In: 2017 3rd IEEE international conference on computer and communications (ICCC), <https://doi.org/10.1109/CompComm.2017.8322878>, pp 1947–1955
- Li Y, Zhang H, Xue X, Jiang Y, Shen Q (2018) Deep learning for remote sensing image classification: a survey. *Wiley Interdiscip Rev Data Min Knowl Discov* 8(6):1264
- Li S, Song W, Fang L, Chen Y, Ghamisi P, Benediktsson JA (2019) Deep learning for hyperspectral image classification: an overview. *IEEE Trans Geosci Remote Sens* 57(9):6690–6709
- Paoletti M, Haut J, Plaza J, Plaza A (2019) Deep learning classifiers for hyperspectral imaging: a review. *ISPRS J Photogramm Remote Sens* 158:279–317
- Song J, Gao S, Zhu Y, Ma C (2019) A survey of remote sensing image classification based on CNNs. *Big Earth Data* 3(3):232–254
- Cheng G, Xie X, Han J, Guo L, Xia G-S (2020) Remote sensing image scene classification meets deep learning: challenges, methods, benchmarks, and opportunities. *IEEE J Sel Top Appl Earth Observ Remote Sens* 13:3735–3756
- Alem A, Kumar S (2020) Deep learning methods for land cover and land use classification in remote sensing: a review. In: 2020 8th international conference on reliability, infocom technologies and optimization (Trends and Future Directions)(ICRITO), IEEE, pp 903–908
- Vali A, Comai S, Matteucci M (2020) Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: a review. *Remote Sens* 12(15):2495
- Kuras A, Brell M, Rizzi J, Burud I (2021) Hyperspectral and lidar data applied to the urban land cover machine learning and neural-network-based classification: A review. *Remote Sens* 13(17):3393
- Zang N, Cao Y, Wang Y, Huang B, Zhang L, Mathiopoulos PT (2021) Land-use mapping for high-spatial resolution remote sensing image via deep learning: a review. *IEEE J Sel Top Appl Earth Observ Remote Sens* 14:5372–5391
- Mehmood M, Shahzad A, Zafar B, Shabbir A, Ali N (2022) Remote sensing image classification: a comprehensive review and applications. *Math Probl Eng* 2022:1–24
- Alkhelaiwi M, Boulila W, Ahmad J, Koubaa A, Driss M (2021) An efficient approach based on privacy-preserving deep learning for satellite image classification. *Remote Sens* 13(11):2221. <https://doi.org/10.3390/rs1311221>
- Loddo A, Loddo M, Ruberto CD (2021) A novel deep learning based approach for seed image classification and retrieval. *Comput Electron Agric* 187:106269
- Raei E, Asanjan AA, Nikoo MR, Sadegh M, Pourshahabi S, Adamowski JF (2022) A deep learning image segmentation model for agricultural irrigation system classification. *Comput Electron Agric* 198:106977. <https://doi.org/10.1016/j.compag.2022.106977>
- Liu J, Wang X, Wang T (2019) Classification of tree species and stock volume estimation in ground forest images using deep learning. *Comput Electron Agric* 166:105012. <https://doi.org/10.1016/j.compag.2019.105012>
- Ferreira MP, Almeida DRA, Almeida Papa D, Minervino JBS, Veras HFP, Formighieri A, Santos CAN, Ferreira MAD, Figueiredo EO, Ferreira EJL (2020) Individual tree detection and species classification of Amazonian palms using UAV images and deep learning. *For Ecol Manage* 475:118397. <https://doi.org/10.1016/j.foreco.2020.118397>

34. Safari K, Prasad S, Labate D (2021) A multiscale deep learning approach for high-resolution hyperspectral image classification. *IEEE Geosci Remote Sens Lett* 18(1):167–171. <https://doi.org/10.1109/lgrs.2020.2966987>

35. Sorek-Hamer M, Pohle MV, Sahasrabhojanee A, Asanjan AA, Deardorff E, Suel E, Lingenfelter V, Das K, Oza NC, Ezzati M, Brauer M (2022) A deep learning approach for meter-scale air quality estimation in urban environments using very high-spatial-resolution satellite imagery. *Atmosphere* 13(5):696. <https://doi.org/10.3390/atmos13050696>

36. He Y, Lu Z, Wang W, Zhang D, Zhang Y, Qin B, Shi K, Yang X (2022) Water clarity mapping of global lakes using a novel hybrid deep-learning-based recurrent model with landsat OLI images. *Water Res* 215:118241. <https://doi.org/10.1016/j.watres.2022.118241>

37. Huang C, Wu Z, Wen J, Xu Y, Jiang Q, Wang Y (2022) Abnormal event detection using deep contrastive learning for intelligent video surveillance system. *IEEE Trans Industr Inf* 18(8):5171–5179. <https://doi.org/10.1109/tti.2021.3122801>

38. Wyatt M, Radford B, Callow N, Bennamoun M, Hickey S (2022) Using ensemble methods to improve the robustness of deep learning for image classification in marine environments. *Methods Ecol Evol* 13(6):1317–1328. <https://doi.org/10.1111/2041-210x.13841>

39. Huang Z, Dumitru CO, Pan Z, Lei B, Dateu M (2021) Classification of large-scale high-resolution SAR images with deep transfer learning. *IEEE Geosci Remote Sens Lett* 18(1):107–111. <https://doi.org/10.1109/lgrs.2020.2965558>

40. Liu S, Shi Q, Zhang L (2021) Few-shot hyperspectral image classification with unknown classes using multitask deep learning. *IEEE Trans Geosci Remote Sens* 59(6):5085–5102. <https://doi.org/10.1109/tgrs.2020.3018879>

41. Ienco D, Interdonato R, Gaetano R, Minh DHT (2019) Combining sentinel-1 and sentinel-2 satellite image time series for land cover mapping via a multi-source deep learning architecture. *ISPRS J Photogramm Remote Sens* 158:11–22. <https://doi.org/10.1016/j.isprsjprs.2019.09.016>

42. Fetai B, Grigillo D, Liseic A (2022) Revising cadastral data on land boundaries using deep learning in image-based mapping. *ISPRS Int J Geo Inf* 11(5):298. <https://doi.org/10.3390/ijgi11050298>

43. Bhosle K, Musande V (2019) Evaluation of deep learning CNN model for land use land cover classification and crop identification using hyperspectral remote sensing images. *J Indian Soc Remote Sens* 47(11):1949–1958. <https://doi.org/10.1007/s12524-019-01041-2>

44. Lu H, Ma L, Fu X, Liu C, Wang Z, Tang M, Li N (2020) Landslides information extraction using object-oriented image analysis paradigm based on deep learning and transfer learning. *Remote Sens* 12(5):752. <https://doi.org/10.3390/rs12050752>

45. Liang X (2018) Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization. *Comput Aided Civil Infrastruct Eng* 34(5):415–430. <https://doi.org/10.1111/mice.12425>

46. Vetrivel A, Gerke M, Kerle N, Nex F, Vosselman G (2018) Disaster damage detection through synergistic use of deep learning and 3D point cloud features derived from very high resolution oblique aerial images, and multiple-kernel-learning. *ISPRS J Photogramm Remote Sens* 140:45–59. <https://doi.org/10.1016/j.isprsjprs.2017.03.001>

47. Mishra BK, Thakker D, Mazumdar S, Neagu D, Gheorghe M, Simpson S (2020) A novel application of deep learning with image cropping: a smart city use case for flood monitoring. *J Reliab Intell Environ* 6(1):51–61. <https://doi.org/10.1007/s40860-020-00099-x>

48. Kyrkou C, Theocharides T, (2019) Deep-learning-based aerial image classification for emergency response applications using unmanned aerial vehicles. In: *IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW)*, IEEE. <https://doi.org/10.1109/cvprw.2019.00077>

49. Govil K, Welch ML, Ball JT, Pennypacker CR (2020) Preliminary results from a wildfire detection system using deep learning on remote camera images. *Remote Sens* 12(1):166. <https://doi.org/10.3390/rs12010166>

50. Pinto MM, Libonati R, Trigo RM, Trigo IF, DaCamara CC (2020) A deep learning approach for mapping and dating burned areas using temporal sequences of satellite images. *ISPRS J Photogramm Remote Sens* 160:260–274. <https://doi.org/10.1016/j.isprsjprs.2019.12.014>

51. Anand A, Pandey MK, Srivastava PK, Gupta A, Khan ML (2021) Integrating multi-sensors data for species distribution mapping using deep learning and envelope models. *Remote Sens* 13(16):3284. <https://doi.org/10.3390/rs13163284>

52. Alshahrani HM, Al-Wesabi FN, Duhayyim MA, Nemri N, Kadry S, Alqaralleh BAY (2021) An automated deep learning based satellite imagery analysis for ecology management. *Eco Inform* 66:101452. <https://doi.org/10.1016/j.ecoinf.2021.101452>

53. Yadav TK, Chidburee P, Mahavik N (2021) Land cover classification based on UAV photogrammetry and deep learning for supporting mine reclamation: a case study of Mae Moh mine in Lampang Province, Thailand. *Appl Environ Res*. <https://doi.org/10.35762/aer.2021.43.4.4><https://doi.org/10.35762/aer.2021.43.4.4>

54. Balaniuk R, Isupova O, Reece S (2020) Mining and tailings dam detection in satellite imagery using deep learning. *Sensors* 20(23):6936

55. Li H, Ke C-Q, Zhu Q, Li M, Shen X (2022) A deep learning approach to retrieve cold-season snow depth over arctic sea ice from AMSR2 measurements. *Remote Sens Environ* 269:112840. <https://doi.org/10.1016/j.rse.2021.112840>

56. Li X, Liu B, Zheng G, Ren Y, Zhang S, Liu Y, Gao L, Liu Y, Zhang B, Wang F (2020) Deep-learning-based information mining from ocean remote-sensing imagery. *Natl Sci Rev* 7(10):1584–1605. <https://doi.org/10.1093/nsr/nwaa047>

57. Uz M, Atman KG, Akyilmaz O, Shum CK, Keleş M, Ay T, Tandoğdu B, Zhang Y, Mercan H (2022) Bridging the gap between GRACE and GRACE-FO missions with deep learning aided water storage simulations. *Sci Total Environ* 830:154701. <https://doi.org/10.1016/j.scitotenv.2022.154701>

58. Fang K, Shen C, Kifer D, Yang X (2017) Prolongation of SMAP to spatiotemporally seamless coverage of continental U.S. using a deep learning neural network. *Geophys Res Lett*. <https://doi.org/10.1002/2017gl075619>

59. Altawee M, Khelifi A, Li Z, Squitieri A, Basmaji T, Ghazal M (2022) Automated archaeological feature detection using deep learning on optical UAV imagery: preliminary results. *Remote Sens* 14(3):553. <https://doi.org/10.3390/rs14030553>

60. Soroush M, Mehrtash A, Khazraee E, Ur JA (2020) Deep learning in archaeological remote sensing: automated qanat detection in the Kurdistan region of Iraq. *Remote Sens* 12(3):500. <https://doi.org/10.3390/rs12030500>

61. Pérez-Zavala R, Torres-Torriti M, Cheein FA, Troni G (2018) A pattern recognition strategy for visual grape bunch detection in vineyards. *Comput Electron Agric* 151:136–149. <https://doi.org/10.1016/j.compag.2018.05.019>

62. Rico-Fernández MP, Ríos-Cabrera R, Castelán M, Guerrero-Reyes H-I, Juarez-Maldonado A (2019) A contextualized approach for segmentation of foliage in different crop species. *Comput Electron Agric* 156:378–386. <https://doi.org/10.1016/j.compag.2018.11.033>

63. Fahmi F, Trianda D, Andayani U, Siregar B (2018) Image processing analysis of geospatial UAV orthophotos for palm oil plantation monitoring. *J Phys Conf Ser* 978:012064. <https://doi.org/10.1088/1742-6596/978/1/012064>

64. Tian H, Wang T, Liu Y, Qiao X, Li Y (2020) Computer vision technology in agricultural automation—a review. *Inform Process Agric* 7(1):1–19. <https://doi.org/10.1016/j.inpa.2019.09.006>

65. Tsouros DC, Bibi S, Sarigiannidis PG (2019) A review on UAV-based applications for precision agriculture. *Information* 10(11):349. <https://doi.org/10.3390/info10110349>

66. Loey M, ElSawy A, Afify M (2020) Deep learning in plant diseases detection for agricultural crops. *Int J Serv Sci Manag Eng Technol* 11(2):41–58. <https://doi.org/10.4018/ijssmet.2020040103>

67. Nandhini M, Kala KU, Thangadarshini M, Verma SM (2022) Deep learning model of sequential image classifier for crop disease detection in plantain tree cultivation. *Comput Electron Agric* 197:106915. <https://doi.org/10.1016/j.compag.2022.106915>

68. Chen C-J, Huang Y-Y, Li Y-S, Chen Y-C, Chang C-Y, Huang Y-M (2021) Identification of fruit tree pests with deep learning on embedded drone to achieve accurate pesticide spraying. *IEEE Access* 9:21986–21997. <https://doi.org/10.1109/access.2021.3056082>

69. Cheng X, Zhang Y, Chen Y, Wu Y, Yue Y (2017) Pest identification via deep residual learning in complex background. *Comput Electron Agric* 141:351–356. <https://doi.org/10.1016/j.compag.2017.08.005>

70. Too EC, Yujian L, Njuki S, Yingchun L (2019) A comparative study of fine-tuning deep learning models for plant disease identification. *Comput Electron Agric* 161:272–279. <https://doi.org/10.1016/j.compag.2018.03.032>

71. Brahim M, Boukhalfa K, Moussaoui A (2017) Deep learning for tomato diseases: classification and symptoms visualization. *Appl Artif Intell* 31(4):299–315. <https://doi.org/10.1080/08839514.2017.1315516>

72. Sivakumar ANV, Li J, Scott S, Psota E, Jhala AJ, Luck JD, Shi Y (2020) Comparison of object detection and patch-based classification deep learning models on mid- to late-season weed detection in UAV imagery. *Remote Sens* 12(13):2136. <https://doi.org/10.3390/rs12132136>

73. Bah MD, Hafiane A, Canals R (2020) Crownet: deep network for crop row detection in UAV images. *IEEE Access* 8:5189–5200. <https://doi.org/10.1109/access.2019.2960873>

74. Hasan ASMM, Sohel F, Diepeveen D, Laga H, Jones MGK (2021) A survey of deep learning techniques for weed detection from images. *Comput Electron Agric* 184:106067. <https://doi.org/10.1016/j.compag.2021.106067>

75. Reedha R, Dericquebourg E, Canals R, Hafiane A (2022) Transformer neural network for weed and crop classification of high resolution UAV images. *Remote Sens* 14(3):592. <https://doi.org/10.3390/rs14030592>

76. Bah M, Hafiane A, Canals R (2018) Deep learning with unsupervised data labeling for weed detection in line crops in UAV images. *Remote Sens* 10(11):1690. <https://doi.org/10.3390/rs10111690>

77. Hamylton SM, Morris RH, Carvalho RC, Roder N, Barlow P, Mills K, Wang L (2020) Evaluating techniques for mapping island vegetation from unmanned aerial vehicle (UAV) images: pixel classification, visual interpretation and machine learning approaches. *Int J Appl Earth Obs Geoinf* 89:102085. <https://doi.org/10.1016/j.jag.2020.102085>

78. Pearse GD, Tan AYS, Watt MS, Franz MO, Dash JP (2020) Detecting and mapping tree seedlings in UAV imagery using convolutional neural networks and field-verified data. *ISPRS J Photogramm Remote Sens* 168:156–169. <https://doi.org/10.1016/j.isprsjprs.2020.08.005>

79. Huang H, Lan Y, Yang A, Zhang Y, Wen S, Deng J (2020) Deep learning versus object-based image analysis (OBIA) in weed mapping of UAV imagery. *Int J Remote Sens* 41(9):3446–3479. <https://doi.org/10.1080/01431161.2019.1706112>

80. Sidike P, Sagan V, Maimaitijiang M, Maimaitiyiming M, Shakoor N, Burken J, Mockler T, Fritschi FB (2019) dopen: deep progressively expanded network for mapping heterogeneous agricultural landscape using worldview-3 satellite imagery. *Remote Sens Environ* 221:756–772. <https://doi.org/10.1016/j.rse.2018.11.031>

81. Lee C, Sohn E, Park JD, Jang J-D (2018) Estimation of soil moisture using deep learning based on satellite data: a case study of South Korea. *GIScience Remote Sens* 56(1):43–67. <https://doi.org/10.1080/15481603.2018.1489943>

82. Fang K, Pan M, Shen C (2019) The value of SMAP for long-term soil moisture estimation with the help of deep learning. *IEEE Trans Geosci Remote Sens* 57(4):2221–2233. <https://doi.org/10.1109/tgrs.2018.2872131>

83. Fang K, Shen C (2020) Near-real-time forecast of satellite-based soil moisture using long short-term memory with an adaptive data integration kernel. *J Hydrometeorol* 21(3):399–413. <https://doi.org/10.1175/jhm-d-19-0169.1>

84. Xu M, Yao N, Yang H, Xu J, Hu A, Goncalves LGG, Liu G (2022) Downscaling SMAP soil moisture using a wide & deep learning method over the continental united states. *J Hydrol* 609:127784. <https://doi.org/10.1016/j.jhydrol.2022.127784>

85. Yinglan A, Wang G, Hu P, Lai X, Xue B, Fang Q (2022) Root-zone soil moisture estimation based on remote sensing data and deep learning. *Environm Res* 212:113278. <https://doi.org/10.1016/j.envres.2022.113278>

86. Ahmed AAM, Deo RC, Raj N, Ghahramani A, Feng Q, Yin Z, Yang L (2021) Deep learning forecasts of soil moisture: convolutional neural network and gated recurrent unit models coupled with satellite-derived MODIS, observations and synoptic-scale climate index data. *Remote Sens* 13(4):554. <https://doi.org/10.3390/rs13040554>

87. Fang K, Kifer D, Lawson K, Shen C (2020) Evaluating the potential and challenges of an uncertainty quantification method for long short-term memory models for soil moisture predictions. *Water Resour Res*. <https://doi.org/10.1029/2020wr028095>

88. Colligan T, Ketchum D, Brinkerhoff D, Maneta M (2022) A deep learning approach to mapping irrigation using landsat: irrmapper u-net. *IEEE Trans Geosci Remote Sens* 60:1–11. <https://doi.org/10.1109/tgrs.2022.3175635>

89. Tang J, Arvor D, Corpetti T, Tang P (2021) Mapping center pivot irrigation systems in the southern amazon from sentinel-2 images. *Water* 13(3):298. <https://doi.org/10.3390/w13030298>

90. Saraiwa M, Protas Salgado M, Souza C (2020) Automatic mapping of center pivot irrigation systems from satellite images using deep learning. *Remote Sens* 12(3):558. <https://doi.org/10.3390/rs12030558>

91. Albuquerque AO, Carvalho Júnior OA, Carvalho OLF, Bem PP, Ferreira PHG, Moura R, Silva CR, Gomes RAT, Guimarães RF (2020) Deep semantic segmentation of center pivot irrigation systems from remotely sensed data. *Remote Sens* 12(13):2159. <https://doi.org/10.3390/rs12132159>

92. Diaz-Gonzalez FA, Vuelvas J, Correa CA, Vallejo VE, Patino D (2022) Machine learning and remote sensing techniques applied to estimate soil indicators—review. *Ecol Ind* 135:108517. <https://doi.org/10.1016/j.ecolind.2021.108517>

93. Wang N, Peng J, Xue J, Zhang X, Huang J, Biswas A, He Y, Shi Z (2022) A framework for determining the total salt content of soil profiles using time-series sentinel-2 images and a random

forest-temporal convolution network. *Geoderma* 409:115656. <https://doi.org/10.1016/j.geoderma.2021.115656>

94. Dainelli R, Toscano P, Gennaro SFD, Matese A (2021) Recent advances in unmanned aerial vehicle forest remote sensing-a systematic review. part I: a general framework. *Forests* 12(3):327. <https://doi.org/10.3390/f12030327>

95. Kussul N, Lavreniuk M, Skakun S, Shelestov A (2017) Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geosci Remote Sens Lett* 14(5):778–782. <https://doi.org/10.1109/lgrs.2017.2681128>

96. Wagner FH, Dalagnol R, Casapia XT, Streher AS, Phillips OL, Gloor E, Aragão LEOC (2020) Regional mapping and spatial distribution analysis of canopy palms in an amazon forest using deep learning and VHR images. *Remote Sens* 12(14):2225. <https://doi.org/10.3390/rs12142225>

97. Lee S-H, Han K-J, Lee K, Lee K-J, Oh K-Y, Lee M-J (2020) Classification of landscape affected by deforestation using high-resolution remote sensing data and deep-learning techniques. *Remote Sens* 12(20):3372. <https://doi.org/10.3390/rs12203372>

98. Lummitz S, Devisscher T, Mayaud JR, Radic V, Coops NC, Griess VC (2021) Mapping trees along urban street networks with deep learning and street-level imagery. *ISPRS J Photogramm Remote Sens* 175:144–157. <https://doi.org/10.1016/j.isprsjprs.2021.01.016>

99. Onishi M, Ise T (2021) Explainable identification and mapping of trees using UAV RGB image and deep learning. *Sci Rep*. <https://doi.org/10.1038/s41598-020-79653-9>

100. Onishi M, Watanabe S, Nakashima T, Ise T (2022) Practicality and robustness of tree species identification using UAV RGB image and deep learning in temperate forest in japan. *Remote Sens* 14(7):1710. <https://doi.org/10.3390/rs14071710>

101. Marin W, Mondragon IF, Colorado JD (2022) Aerial identification of Amazonian palms in high-density forest using deep learning. *Forests* 13(5):655. <https://doi.org/10.3390/f13050655>

102. Miyoshi GT, Santos Arruda M, Osco LP, Junior JM, Gonçalves DN, Imai NN, Tommaselli AMG, Honkavaara E, Gonçalves WN (2020) A novel deep learning method to identify single tree species in UAV-based hyperspectral images. *Remote Sens* 12(8):1294. <https://doi.org/10.3390/rs12081294>

103. Priya RS, Vani K (2019) Deep learning based forest fire classification and detection in satellite images. In: (2019) 11th international conference on advanced computing (ICoAC). IEEE. <https://doi.org/10.1109/icoac48765.2019.246817>

104. Jiao Z, Zhang Y, Xin J, Mu L, Yi Y, Liu H, Liu D (2019) A deep learning based forest fire detection approach using UAV and yolov3. In: 1st international conference on industrial artificial intelligence (IAI). IEEE. <https://doi.org/10.1109/iciai.2019.8850815>

105. Safonova A, Tabik S, Alcaraz-Segura D, Rubtsov A, Maglinets Y, Herrera F (2019) Detection of fir trees (*abies sibirica*) damaged by the bark beetle in unmanned aerial vehicle images with deep learning. *Remote Sensing* 11(6):643. <https://doi.org/10.3390/rs11060643>

106. Iqbal U, Perez P, Barthelemy J (2021) A process-driven and need-oriented framework for review of technological contributions to disaster management. *Heliyon* 7(11):08405. <https://doi.org/10.1016/j.heliyon.2021.e08405>

107. Jiang X, Liang S, He X, Ziegler AD, Lin P, Pan M, Wang D, Zou J, Hao D, Mao G, Zeng Y, Yin J, Feng L, Miao C, Wood EF, Zeng Z (2021) Rapid and large-scale mapping of flood inundation via integrating spaceborne synthetic aperture radar imagery with unsupervised deep learning. *ISPRS J Photogramm Remote Sens* 178:36–50. <https://doi.org/10.1016/j.isprsjprs.2021.05.019>

108. Munawar HS, Ullah F, Qayyum S, Heravi A (2021) Application of deep learning on UAV-based aerial images for flood detection. *Smart Cities* 4(3):1220–1243. <https://doi.org/10.3390/smartcities4030065>

109. Xiong P, Long C, Zhou H, Zhang X, Shen X (2022) GNSS tec-based earthquake ionospheric perturbation detection using a novel deep learning framework. *IEEE J Sel Top Appl Earth Observ Remote Sensng* 15:4248–4263. <https://doi.org/10.1109/jstars.2022.3175961>

110. Zhao X, Wang C, Zhang H, Tang Y, Zhang B, Li L (2021) Inversion of seismic source parameters from satellite insar data based on deep learning. *Tectonophysics* 821:229140. <https://doi.org/10.1016/j.tecto.2021.229140>

111. Lin JT, Melgar D, Thomas AM, Searcy J (2021) Early warning for great earthquakes from characterization of crustal deformation patterns with deep learning. *J Geophys Res Solid*. <https://doi.org/10.1029/2021jb022703>

112. Ghorbanzadeh O, Blaschke T, Gholamnia K, Meena S, Tiede D, Aryal J (2019) Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sens* 11(2):196. <https://doi.org/10.3390/rs11020196>

113. Ghorbanzadeh O, Shahabi H, Crivellari A, Homayouni S, Blaschke T, Ghamisi P (2022) Landslide detection using deep learning and object-based image analysis. *Landslides* 19(4):929–939. <https://doi.org/10.1007/s10346-021-01843-x>

114. Lattari F, Rucci A, Matteucci M (2022) A deep learning approach for change points detection in insar time series. *IEEE Trans Geosci Remote Sens* 60:1–16. <https://doi.org/10.1109/tgrs.2022.3155969>

115. Sublime J, Kalinicheva E (2019) Automatic post-disaster damage mapping using deep-learning techniques for change detection: case study of the tohoku tsunami. *Remote Sens* 11(9):1123. <https://doi.org/10.3390/rs11091123>

116. Devaraj J, Ganesan S, Elavarasan R, Subramaniam U (2021) A novel deep learning based model for tropical intensity estimation and post-disaster management of hurricanes. *Appl Sci* 11(9):4129. <https://doi.org/10.3390/app11094129>

117. Zhao L, Chen Y, Sheng VS (2019) A real-time typhoon eye detection method based on deep learning for meteorological information forensics. *J Real-Time Image Proc* 17(1):95–102. <https://doi.org/10.1007/s11554-019-00899-2>

118. Viseras A, Meissner M, Marchal J (2021) Wildfire front monitoring with multiple UAVs using deep q-learning. *IEEE Access*. <https://doi.org/10.1109/access.2021.3055651>

119. McCarthy NF, Tohidi A, Aziz Y, Dennie M, Valero MM, Hu N (2021) A deep learning approach to downscale geostationary satellite imagery for decision support in high impact wildfires. *Forests* 12(3):294. <https://doi.org/10.3390/f12030294>

120. Fascista A (2022) Toward integrated large-scale environmental monitoring using wsn/uav/crowdsensing: A review of applications, signal processing, and future perspectives. *Sensors* 22(5):1824. <https://doi.org/10.3390/s22051824>

121. Huang X, Han X, Ma S, Lin T, Gong J (2019) Monitoring ecosystem service change in the city of Shenzhen by the use of high-resolution remotely sensed imagery and deep learning. *Land Degrad Dev* 30(12):1490–1501. <https://doi.org/10.1002/lrd.3337>

122. Ba R, Chen C, Yuan J, Song W, Lo S (2019) Smokenet: satellite smoke scene detection using convolutional neural network with spatial and channel-wise attention. *Remote Sens* 11(14):1702. <https://doi.org/10.3390/rs11141702>

123. Wang S, Huo Y, Mu X, Jiang P, Xun S, He B, Wu W, Liu L, Wang Y (2022) A high-performance convolutional neural network for ground-level ozone estimation in eastern China. *Remote Sens* 14(7):1640. <https://doi.org/10.3390/rs14071640>

124. Camalan S, Cui K, Pauca VP, Alqahtani S, Silman M, Chan R, Plemmons RJ, Dethier EN, Fernandez LE, Lutz DA (2022)

Change detection of Amazonian alluvial gold mining using deep learning and sentinel-2 imagery. *Remote Sens* 14(7):1746. <https://doi.org/10.3390/rs14071746>

125. Li T, Shen H, Yuan Q, Zhang X, Zhang L (2017) Estimating ground-level PM2.5 by fusing satellite and station observations: a geo-intelligent deep learning approach. *Geophys Res Lett* 44(23):11985–11993. <https://doi.org/10.1002/2017gl075710>

126. Cui Q, Zhang F, Fu S, Wei X, Ma Y, Wu K (2022) High spatiotemporal resolution PM2.5 concentration estimation with machine learning algorithm: a case study for wildfire in California. *Remote Sens* 14(7):1635. <https://doi.org/10.3390/rs14071635>

127. Gupta H, Verma OP (2021) Monitoring and surveillance of urban road traffic using low altitude drone images: a deep learning approach. *Multimed Tools Appl* 81(14):19683–19703. <https://doi.org/10.1007/s11042-021-11146-x>

128. Lin Z, Ji K, Leng X, Kuang G (2019) Squeeze and excitation rank faster R-CNN for ship detection in SAR images. *IEEE Geosci Remote Sens Lett* 16(5):751–755. <https://doi.org/10.1109/lgrs.2018.2882551>

129. Li J, Ma R, Cao Z, Xue K, Xiong J, Hu M, Feng X (2022) Satellite detection of surface water extent: a review of methodology. *Water* 14(7):1148. <https://doi.org/10.3390/w14071148>

130. Huang X, Zhang B, Perrie W, Lu Y, Wang C (2022) A novel deep learning method for marine oil spill detection from satellite synthetic aperture radar imagery. *Mar Pollut Bull* 179:113666. <https://doi.org/10.1016/j.marpolbul.2022.113666>

131. Wang J, Bretz M, Dewan MAA, Delavar MA (2022) Machine learning in modelling land-use and land cover-change (LULCC): current status, challenges and prospects. *Sci Total Environ* 822:153559. <https://doi.org/10.1016/j.scitotenv.2022.153559>

132. Kavhu B, Mashimbye ZE, Luvuno L (2022) Characterising social-ecological drivers of landuse/cover change in a complex transboundary basin using singular or ensemble machine learning. *Remote Sens Appl Soc Environ* 27:100773. <https://doi.org/10.1016/j.rsase.2022.100773>

133. Chandler BMP, Lovell H, Boston CM, Lukas S, Barr ID, Benediktsson Benn DI, Clark CD, Darvill CM, Evans DJA, Ewertowski MW, Loibl D, Margold M, Otto J-C, Roberts DH, Stokes CR, Storrar RD, Stroeven AP (2018) Glacial geomorphological mapping: a review of approaches and frameworks for best practice. *Earth Sci Rev* 185:806–846. <https://doi.org/10.1016/j.earscirev.2018.07.015>

134. Yuan F, Sawaya KE, Loeffelholz BC, Bauer ME (2005) Land cover classification and change analysis of the twin cities (minnesota) metropolitan area by multitemporal landsat remote sensing. *Remote Sens Environ* 98(2–3):317–328. <https://doi.org/10.1016/j.rse.2005.08.006>

135. Huang B, Zhao B, Song Y (2018) Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sens Environ* 214:73–86. <https://doi.org/10.1016/j.rse.2018.04.050>

136. Abdi AM (2019) Land cover and land use classification performance of machine learning algorithms in a boreal landscape using sentinel-2 data. *GIScience Remote Sens* 57(1):1–20. <https://doi.org/10.1080/15481603.2019.1650447>

137. Tong X-Y, Xia G-S, Lu Q, Shen H, Li S, You S, Zhang L (2020) Land-cover classification with high-resolution remote sensing images using transferable deep models. *Remote Sens Environ* 237:111322. <https://doi.org/10.1016/j.rse.2019.111322>

138. Mahdianpari M, Salehi B, Rezaee M, Mohammadimanesh F, Zhang Y (2018) Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. *Remote Sens* 10(7):1119. <https://doi.org/10.3390/rs10071119>

139. Zhu Q, Guo X, Deng W, Shi S, Guan Q, Zhong Y, Zhang L, Li D (2022) Land-use/land-cover change detection based on a siamese global learning framework for high spatial resolution remote sensing imagery. *ISPRS J Photogramm Remote Sens* 184:63–78. <https://doi.org/10.1016/j.isprsjprs.2021.12.005>

140. Li Y, Zhou Y, Zhang Y, Zhong L, Wang J, Chen J (2022) DKDFN: domain knowledge-guided deep collaborative fusion network for multimodal unitemporal remote sensing land cover classification. *ISPRS J Photogramm Remote Sens* 186:170–189. <https://doi.org/10.1016/j.isprsjprs.2022.02.013>

141. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778

142. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708

143. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1–9

144. Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7132–7141

145. Yu F, Koltun V (2015) Multi-scale context aggregation by dilated convolutions. [arXiv:1511.07122](https://arxiv.org/abs/1511.07122)

146. Chollet F (2017) Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1251–1258

147. Zhong Z, Li J, Ma L, Jiang H, Zhao H (2017) Deep residual networks for hyperspectral image classification. In: 2017 IEEE international geoscience and remote sensing symposium (IGARSS), IEEE, pp 1824–1827

148. Zhong Z, Li J, Luo Z, Chapman M (2017) Spectral-spatial residual network for hyperspectral image classification: a 3-d deep learning framework. *IEEE Trans Geosci Remote Sens* 56(2):847–858

149. Paoletti ME, Haut JM, Fernandez-Beltran R, Plaza J, Plaza AJ, Pla F (2018) Deep pyramidal residual networks for spectral-spatial hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 57(2):740–754

150. Yang H, Wu P, Yao X, Wu Y, Wang B, Xu Y (2018) Building extraction in very high resolution imagery by dense-attention networks. *Remote Sens* 10(11):1768

151. Zhang C, Li G, Du S (2019) Multi-scale dense networks for hyperspectral remote sensing image classification. *IEEE Trans Geosci Remote Sens* 57(11):9201–9222

152. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: 2015 IEEE conference on computer vision and pattern recognition (CVPR), pp 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>

153. Lei P, Liu C (2020) Inception residual attention network for remote sensing image super-resolution. *Int J Remote Sens* 41(24):9565–9587

154. Firat H, Asker ME, Bayindir Mİ, Hanbay D (2022) Hybrid 3d/2d complete inception module and convolutional neural network for hyperspectral remote sensing image classification. *Neural Process Lett* 55(2):1087–1130

155. Kumthekar A, Reddy GR (2021) An integrated deep learning framework of u-net and inception module for cloud detection of remote sensing images. *Arab J Geosci* 14(18):1–13

156. Han Y, Wei C, Zhou R, Hong Z, Zhang Y, Yang S (2020) Combining 3d-CNN and squeeze-and-excitation networks for

remote sensing sea ice image classification. *Math Probl Eng* 1:8065396

157. Hu Q, Zhen L, Mao Y, Zhou X, Zhou G (2021) Automated building extraction using satellite remote sensing imagery. *Autom Constr* 123:103509

158. Li G, Zhang C, Lei R, Zhang X, Ye Z, Li X (2020) Hyperspectral remote sensing image classification using three-dimensional-squeeze-and-excitation-densenet (3d-se-densenet). *Remote Sens Lett* 11(2):195–203

159. Hamaguchi R, Fujita A, Nemoto K, Imaizumi T, Hikosaka S (2018) Effective use of dilated convolutions for segmenting small object instances in remote sensing imagery. In: 2018 IEEE winter conference on applications of computer vision (WACV), IEEE, pp 1442–1450

160. Liu Q, Kampffmeyer M, Janssen R, Salberg A-B (2020) Dense dilated convolutions' merging network for land cover classification. *IEEE Trans Geosci Remote Sens* 58(9):6309–6320

161. Liu R, Cai W, Li G, Ning X, Jiang Y (2021) Hybrid dilated convolution guided feature filtering and enhancement strategy for hyperspectral image classification. *IEEE Geosci Remote Sens Lett* 19:1–5

162. Qu J, Su C, Zhang Z, Razi A (2020) Dilated convolution and feature fusion SSD network for small object detection in remote sensing images. *IEEE Access* 8:82832–82843

163. Li W, Chen H, Liu Q, Liu H, Wang Y, Gui G (2022) Attention mechanism and depthwise separable convolution aided 3dcnn for hyperspectral remote sensing image classification. *Remote Sens* 14(9):2215

164. Dang L, Pang P, Lee J (2020) Depth-wise separable convolution neural network with residual connection for hyperspectral image classification. *Remote Sens* 12(20):3408

165. Zhang T, Zhang X, Shi J, Wei S (2019) Depthwise separable convolution neural network for high-speed SAR ship detection. *Remote Sens* 11(21):2483

166. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20:273–297

167. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. *Adv Neural Inform Process Syst* 27

168. Jian P, Chen K, Cheng W (2021) Gan-based one-class classification for remote-sensing image change detection. *IEEE Geosci Remote Sens Lett* 19:1–5

169. Jiang K, Wang Z, Yi P, Wang G, Lu T, Jiang J (2019) Edge-enhanced GAN for remote sensing image superresolution. *IEEE Trans Geosci Remote Sens* 57(8):5799–5812

170. Ma J, Zhang L, Zhang J (2019) Sd-GAN: saliency-discriminated GAN for remote sensing image superresolution. *IEEE Geosci Remote Sens Lett* 17(11):1973–1977

171. Lv F, Han M, Qiu T (2017) Remote sensing image classification based on ensemble extreme learning machine with stacked autoencoder. *IEEE Access* 5:9021–9031

172. Huang G-B, Zhu Q-Y, Siew C-K (2006) Extreme learning machine: theory and applications. *Neurocomputing* 70(1–3):489–501

173. Liang P, Shi W, Zhang X (2017) Remote sensing image classification based on stacked denoising autoencoder. *Remote Sens* 10(1):16

174. Zhou P, Han J, Cheng G, Zhang B (2019) Learning compact and discriminative stacked autoencoder for hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 57(7):4823–4833

175. Zhang X, Liang Y, Li C, Huyan N, Jiao L, Zhou H (2017) Recursive autoencoders-based unsupervised feature learning for hyperspectral image classification. *IEEE Geosci Remote Sens Lett* 14(11):1928–1932

176. Li J, Marpu PR, Plaza A, Bioucas-Dias JM, Benediktsson JA (2013) Generalized composite kernel framework for hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 51(9):4816–4829

177. Chen Y, Nasrabadi NM, Tran TD (2011) Hyperspectral image classification using dictionary-based sparse representation. *IEEE Trans Geosci Remote Sens* 49(10):3973–3985

178. Zhou S, Xue Z, Du P (2019) Semisupervised stacked autoencoder with cotraining for hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 57(6):3813–3826

179. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780

180. Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J (2016) Lstm: a search space odyssey. *IEEE Trans Neural Netw Learn Syst* 28(10):2222–2232

181. Ienco D, Gaetano R, Dupacquier C, Maurel P (2017) Land cover classification via multitemporal spatial data by deep recurrent neural networks. *IEEE Geosci Remote Sens Lett* 14(10):1685–1689

182. Mou L, Ghamisi P, Zhu XX (2017) Deep recurrent neural networks for hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 55(7):3639–3655. <https://doi.org/10.1109/TGRS.2016.2636241>

183. Sharma A, Liu X, Yang X (2018) Land cover classification from multi-temporal, multi-spectral remotely sensed imagery using patch-based recurrent neural networks. *Neural Netw* 105:346–355

184. Mei S, Li X, Liu X, Cai H, Du Q (2022) Hyperspectral image classification using attention-based bidirectional long short-term memory network. *IEEE Trans Geosci Remote Sens* 60:1–12. <https://doi.org/10.1109/TGRS.2021.3102034>

185. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. *Adv Neural Inform Process Syst* 30

186. Rives A, Meier J, Sercu T, Goyal S, Lin Z, Liu J, Guo D, Ott M, Zitnick CL, Ma J et al (2021) Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proc Natl Acad Sci* 118(15):2016239118

187. Nambiar A, Heflin M, Liu S, Maslov S, Hopkins M, Ritz A (2020) Transforming the language of life: Transformer neural networks for protein prediction tasks. In: Proceedings of the 11th ACM international conference on bioinformatics, computational biology and health informatics. BCB '20. Association for Computing Machinery, New York, NY. <https://doi.org/10.1145/3388440.3412467>

188. Rao R, Bhattacharya N, Thomas N, Duan Y, Chen P, Canny J, Abbeel P, Song Y (2019) Evaluating protein transfer learning with tape. *Advances Neural Inform Process Syst* 32

189. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S et al (2020) An image is worth 16x16 words: transformers for image recognition at scale. [arXiv:2010.11929](https://arxiv.org/abs/2010.11929)

190. Deng P, Xu K, Huang H (2021) When CNNs meet vision transformer: a joint framework for remote sensing scene classification. *IEEE Geosci Remote Sens Lett* 19:1–5

191. Ma J, Li M, Tang X, Zhang X, Liu F, Jiao L (2022) Homogeneous transformer learning framework for RS scene classification. *IEEE J Sel Top Appl Earth Observ Remote Sens* 15:2223–2239

192. He J, Zhao L, Yang H, Zhang M, Li W (2019) Hsi-bert: hyperspectral image classification using the bidirectional encoder representation from transformers. *IEEE Trans Geosci Remote Sens* 58(1):165–178

193. Zhong Z, Li Y, Ma L, Li J, Zheng W-S (2021) Spectral-spatial transformer network for hyperspectral image classification: a

factorized architecture search framework. *IEEE Trans Geosci Remote Sens* 60:1–15

194. Wang Y, Jia S, Zhang Z (2022) Multiscale convolutional transformer with center mask pretraining for hyperspectral image classification. [arXiv:2203.04771](https://arxiv.org/abs/2203.04771)

195. Paheding S, Reyes AA, Kasaragod A, Oomen T (2022) Gaf-nau: Gramian angular field encoded neighborhood attention u-net for pixel-wise hyperspectral image classification. In: 2022 IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW), pp 408–416

196. Dong H, Zhang L, Zou B (2021) Exploring vision transformers for polarimetric SAR image classification. *IEEE Trans Geosci Remote Sens* 60:1–15

197. Liu X, Wu Y, Liang W, Cao Y, Li M (2022) High resolution SAR image classification using global-local network structure based on vision transformer and CNN. *IEEE Geosci Remote Sens Lett* 19:1–5

198. Chen L, Luo R, Xing J, Li Z, Yuan Z, Cai X (2022) Geospatial transformer is what you need for aircraft detection in SAR imagery. *IEEE Trans Geosci Remote Sens* 60:1–15

199. Zhang P, Xu H, Tian T, Gao P, Tian J (2022) Sfre-net: scattering feature relation enhancement network for aircraft detection in SAR images. *Remote Sens* 14(9):2076

200. Zhang Y, Yang Q (2021) A survey on multi-task learning. *IEEE Trans Knowled Data Eng* 34(12):5586–5609

201. Thung K-H, Wee C-Y (2018) A brief review on multi-task learning. *Multimed Tools Appl* 77(22):29705–29725

202. Caruana R (1997) Multitask learning. *Mach Learn* 28(1):41–75

203. Ruder S (2017) An overview of multi-task learning in deep neural networks. [arXiv:1706.05098](https://arxiv.org/abs/1706.05098)

204. Lu X, Zhong Y, Zheng Z, Liu Y, Zhao J, Ma A, Yang J (2019) Multi-scale and multi-task deep learning framework for automatic road extraction. *IEEE Trans Geosci Remote Sens* 57(11):9362–9377

205. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18, Springer, pp 234–241

206. Cohn DA, Ghahramani Z, Jordan MI (1996) Active learning with statistical models. *J Art Intell Res* 4:129–145

207. Settles B (2009) Active learning literature survey

208. Wang Z, Du B, Zhang L, Zhang L, Jia X (2017) A novel semisupervised active-learning algorithm for hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 55(6):3071–3083

209. Shi Q, Du B, Zhang L (2015) Spatial coherence-based batch-mode active learning for remote sensing image classification. *IEEE Trans Image Process* 24(7):2037–2050

210. Demir B, Persello C, Bruzzone L (2010) Batch-mode active-learning methods for the interactive classification of remote sensing images. *IEEE Trans Geosci Remote Sens* 49(3):1014–1031

211. Haut JM, Paoletti ME, Plaza J, Li J, Plaza A (2018) Active learning with convolutional neural networks for hyperspectral image classification using a new Bayesian approach. *IEEE Trans Geosci Remote Sens* 56(11):6440–6461

212. MacKay DJ (1992) A practical bayesian framework for back-propagation networks. *Neural Comput* 4(3):448–472

213. Neal RM (2012) Bayesian learning for neural networks, vol 118. Springer, Berlin

214. Zhu H, Samtani S, Chen H, Nunamaker JF Jr (2020) Human identification for activities of daily living: a deep transfer learning approach. *J Manag Inf Syst* 37(2):457–483

215. Lima R, Marfurt K (2019) Convolutional neural network for remote-sensing scene classification: transfer learning analysis. *Remote Sens* 12(1):86

216. Pal M (2007) Ensemble learning with decision tree for remote sensing classification. *World Acad Sci Eng Technol* 36:258–260

217. Dai X, Wu X, Wang B, Zhang L (2019) Semisupervised scene classification for remote sensing images: a method based on convolutional neural networks and ensemble learning. *IEEE Geosci Remote Sens Lett* 16(6):869–873

218. Bigdeli B, Pahlavani P, Amirkolaee HA (2021) An ensemble deep learning method as data fusion system for remote sensing multisensor classification. *Appl Soft Comput* 110:107563

219. Herrera F, Ventura S, Cornelis C, Zafra A, Sánchez-Tarragó D, Vluymans S, Herrera F, Ventura S, Bello R, Bello R et al (2016) Multiple instance learning. Springer, Berlin

220. Dietterich TG, Lathrop RH, Lozano-Pérez T (1997) Solving the multiple instance problem with axis-parallel rectangles. *Artif Intell* 89(1–2):31–71

221. Ajjaji DA, Alsaeed MA, Alswayed AS, Alhichri HS (2019) Multi-instance neural network architecture for scene classification in remote sensing. In: 2019 international conference on computer and information sciences (ICCIS), IEEE, pp 1–5

222. Sutton RS, Barto AG (2018) Reinforcement learning: an introduction. MIT press, Cambridge

223. Jang B, Kim M, Harerimana G, Kim JW (2019) Q-learning algorithms: a comprehensive classification and applications. *IEEE Access* 7:133653–133667

224. Pham HX, La HM, Feil-Seifer D, Nguyen LV (2018) Autonomous UAV navigation using reinforcement learning. [arXiv: 1801.05086](https://arxiv.org/abs/1801.05086)

225. Lin Z, Ni Z, Kuang L, Jiang C, Huang Z (2024) Satellite-terrestrial coordinated multi-satellite beam hopping scheduling based on multi-agent deep reinforcement learning. *IEEE Trans Wirel Commun*. <https://doi.org/10.1109/TWC.2024.3368689>

226. Zhou T, Chen M, Zou J (2020) Reinforcement learning based data fusion method for multi-sensors. *IEEE/CAA J Autom Sin* 7(6):1489–1497

227. Nguyen TT, Nguyen ND, Nahavandi S (2020) Deep reinforcement learning for multiagent systems: a review of challenges, solutions, and applications. *IEEE Trans Cybern* 50(9):3826–3839

228. Dubovik O, Schuster GL, Xu F, Hu Y, Bösch H, Landgraf J, Li Z (2021) Grand challenges in satellite remote sensing. *Front Remote Sens* 2:619818

229. Wooster MJ, Roberts GJ, Giglio L, Roy DP, Freeborn PH, Boschetti L, Justice C, Ichoku C, Schroeder W, Davies D et al (2021) Satellite remote sensing of active fires: history and current status, applications and future requirements. *Remote Sens Environ* 267:112694

230. Shahzaman M, Zhu W, Ullah I, Mustafa F, Bilal M, Ishfaq S, Nisar S, Arshad M, Iqbal R, Aslam RW (2021) Comparison of multi-year reanalysis, models, and satellite remote sensing products for agricultural drought monitoring over south Asian countries. *Remote Sens* 13(16):3294

231. Zhang Y, Li Z, Bai K, Wei Y, Xie Y, Zhang Y, Ou Y, Cohen J, Zhang Y, Peng Z et al (2021) Satellite remote sensing of atmospheric particulate matter mass concentration: advances, challenges, and perspectives. *Fundam Res* 1(3):240–258

232. Mu L, Losch M, Yang Q, Ricker R, Losa SN, Nerger L (2018) Arctic-wide sea ice thickness estimates from combining satellite remote sensing data and a dynamic ice-ocean model with data assimilation during the cryosat-2 period. *J Geophys Res Oceans* 123(11):7763–7780

233. Wang W, Shi K, Zhang Y, Li N, Sun X, Zhang D, Zhang Y, Qin B, Zhu G (2022) A ground-based remote sensing system for

high-frequency and real-time monitoring of phytoplankton blooms. *J Hazard Mater* 439:129623

234. Casagli N, Frodella W, Morelli S, Tofani V, Ciampalini A, Intrieri E, Raspini F, Rossi G, Tanteri L, Lu P (2017) Spaceborne, UAV and ground-based remote sensing techniques for landslide mapping, monitoring and early warning. *Geoenvironmental Disasters* 4(1):1–23

235. Guaraglia DO, Pousa JL (2014) 8 Ground-based remote sensing systems. De Gruyter Open Poland, Warsaw, pp 273–324

236. Shao Z, Wu W, Li D (2021) Spatio-temporal-spectral observation model for urban remote sensing. *Geo-Spatial Inform Sci* 24(3):372–386

237. Martin F-M, Müllerová J, Borgniet L, Dommange F, Breton V, Etvette A (2018) Using single-and multi-date UAV and satellite imagery to accurately monitor invasive knotweed species. *Remote Sens* 10(10):1662

238. Liu H, He X, Li Q, Kratzer S, Wang J, Shi T, Hu Z, Yang C, Hu S, Zhou Q et al (2021) Estimating ultraviolet reflectance from visible bands in ocean colour remote sensing. *Remote Sens Environ* 258:112404

239. Suo Z, Lu Y, Liu J, Ding J, Yin D, Xu F, Jiao J (2021) Ultraviolet remote sensing of marine oil spills: a new approach of Haiyang-1c satellite. *Opt Express* 29(9):13486–13495

240. Ihlen V, Zanter K (2019) Landsat 8 (l8) data users handbook. US Geological Survey, pp 54–55

241. Audebert N, Le Saux B, Lefèvre S (2018) Beyond RGB: very high resolution urban remote sensing with multimodal deep networks. *ISPRS J Photogramm Remote Sens* 140:20–32

242. Su J, Coombes M, Liu C, Zhu Y, Song X, Fang S, Guo L, Chen W-H (2020) Machine learning-based crop drought mapping system by UAV remote sensing RGB imagery. *Unmanned Syst* 8(01):71–83

243. Angelopoulou T, Tziolas N, Balafoutis A, Zalidis G, Bochtis D (2019) Remote sensing techniques for soil organic carbon estimation: a review. *Remote Sens* 11(6):676

244. Tol C, Vilfan N, Dauwe D, Cendrero-Mateo MP, Yang P (2019) The scattering and re-absorption of red and near-infrared chlorophyll fluorescence in the models fluspect and scope. *Remote Sens Environ* 232:111292

245. Weng Q (2009) Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *ISPRS J Photogramm Remote Sens* 64(4):335–344

246. Gao H, Nie N, Zhang W, Chen H (2020) Monitoring the spatial distribution and changes in permafrost with passive microwave remote sensing. *ISPRS J Photogramm Remote Sens* 170:142–155

247. Duan S-B, Han X-J, Huang C, Li Z-L, Wu H, Qian Y, Gao M, Leng P (2020) Land surface temperature retrieval from passive microwave satellite observations: state-of-the-art and future directions. *Remote Sens* 12(16):2573

248. Fernández-Guisuraga JM, Suárez-Seoane S, Calvo L (2022) Radar and multispectral remote sensing data accurately estimate vegetation vertical structure diversity as a fire resilience indicator. *Remote Sens Ecol Conserv* 9(1):117–132

249. Xiaofeng L, Biao Z, Xiaofeng Y (2020) Remote sensing of sea surface wind and wave from spaceborne synthetic aperture radar. *J Radars* 9(3):425–443

250. Landy JC, Dawson GJ, Tsamados M, Bushuk M, Stroeve JC, Howell SE, Krumpen T, Babb DG, Komarov AS, Heorton HD et al (2022) A year-round satellite sea-ice thickness record from cryosat-2. *Nature* 609(7927):517–522

251. Jimenez-Sierra DA, Quintero-Olaya DA, Alvear-Munoz JC, Benitez-Restrepo HD, Florez-Ospina JF, Chanussot J (2022) Graph learning based on signal smoothness representation for homogeneous and heterogeneous change detection. *IEEE Trans Geosci Remote Sens* 60:1–16. <https://doi.org/10.1109/tgrs.2022.3168126>

252. Toulouse T, Rossi L, Campana A, Celik T, Akhloouf MA (2017) Computer vision for wildfire research: an evolving image dataset for processing and analysis. *Fire Saf J* 92:188–194. <https://doi.org/10.1016/j.firesaf.2017.06.012>

253. Prasad S, Saux BL, Yokoya N, Hansch R (2020) 2018 IEEE GRSS data fusion challenge—fusion of multispectral LiDAR and hyperspectral data. <https://doi.org/10.21227/jnh9-nz89>

254. Cucho-Padin G, Loayza H, Palacios S, Balcazar M, Carbajal M, Quiroz R (2019) Development of low-cost remote sensing tools and methods for supporting smallholder agriculture. *Appl Geomat* 12(3):247–263. <https://doi.org/10.1007/s12518-019-00292-5>

255. López-Jiménez E, Vasquez-Gomez JI, Sanchez-Acevedo MA, Herrera-Lozada JC, Uriarte-Arcia AV (2019) Columnar cactus recognition in aerial images using a deep learning approach. *Eco Inform* 52:131–138. <https://doi.org/10.1016/j.ecoinf.2019.05.005>

256. Sa I, Chen Z, Popovic M, Khanna R, Liebisch F, Nieto J, Siegwart R (2018) weednet: dense semantic weed classification using multispectral images and MAV for smart farming. *IEEE Rob Autom Lett* 3(1):588–595. <https://doi.org/10.1109/lra.2017.2774979>

257. Chang C-I (2007) Hyperspectral data exploitation: theory and applications. Wiley, London

258. Middleton EM, Ungar SG, Mandl DJ, Ong L, Frye SW, Campbell PE, Landis DR, Young JP, Pollack NH (2013) The earth observing one (eo-1) satellite mission: over a decade in space. *IEEE J Sel Top Appl Earth Observ Remote Sens* 6(2):243–256

259. Zhong Y, Hu X, Luo C, Wang X, Zhao J, Zhang L (2020) Whuhi: Uav-borne hyperspectral with high spatial resolution (H2) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF. *Remote Sens Environ* 250:112012. <https://doi.org/10.1016/j.rse.2020.112012>

260. Abdelal Q, Assaf MN, Al-Rawabdeh A, Arabasi S, Rawashdeh NA (2022) Assessment of sentinel-2 and landsat-8 oli for small-scale inland water quality modeling and monitoring based on handheld hyperspectral ground truthing. *J Sens* 2022:4643924

261. Alvarez-Vanhard E, Corpetti T, Houet T (2021) UAV & satellite synergies for optical remote sensing applications: a literature review. *Sci Remote Sens* 3:100019. <https://doi.org/10.1016/j.srs.2021.100019>

262. Mittal P, Singh R, Sharma A (2020) Deep learning-based object detection in low-altitude UAV datasets: a survey. *Image Vis Comput* 104:104046. <https://doi.org/10.1016/j.imavis.2020.104046>

263. Lu Y, Young S (2020) A survey of public datasets for computer vision tasks in precision agriculture. *Comput Electron Agric* 178:105760. <https://doi.org/10.1016/j.compag.2020.105760>

264. Abdollahi A, Pradhan B, Shukla N, Chakraborty S, Alamri A (2020) Deep learning approaches applied to remote sensing datasets for road extraction: a state-of-the-art review. *Remote Sens* 12(9):1444. <https://doi.org/10.3390/rs12091444>

265. Long Y, Xia G-S, Li S, Yang W, Yang MY, Zhu XX, Zhang L, Li D (2021) On creating benchmark dataset for aerial image interpretation: reviews, guidances, and million-aid. *IEEE J Sel Top Appl Earth Observ Remote Sens* 14:4205–4230. <https://doi.org/10.1109/jstars.2021.3070368>

266. Kang J, Tariq S, Oh H, Woo SS (2022) A survey of deep learning-based object detection methods and datasets for overhead imagery. *IEEE Access* 10:20118–20134. <https://doi.org/10.1109/access.2022.3149052>

267. David E, Madec S, Sadeghi-Tehran P, Aasen H, Zheng B, Liu S, Kirchgessner N, Ishikawa G, Nagasawa K, Badhon MA, Pozniak C, Solan B, Hund A, Chapman SC, Baret F, Stavness I, Guo, W (2020) Global wheat head detection (GWHD) dataset: a large and diverse dataset of high resolution RGB labelled images to develop and benchmark wheat head detection methods. [arXiv:2005.02162v2](https://arxiv.org/abs/2005.02162v2)

268. Yang M-D, Tseng H-H, Hsu Y-C, Yang C-Y, Lai M-H, Wu D-H (2021) A UAV open dataset of rice paddies for deep learning practice. *Remote Sens* 13(7):1358. <https://doi.org/10.3390/rs13071358>

269. Nguyen HT, Caceres MLL, Moritake K, Kentsch S, Shu H, Diez Y (2021) Individual sick fir tree (*Abies mariesii*) identification in insect infested forests by means of UAV images and deep learning. *Remote Sens* 13(2):260. <https://doi.org/10.3390/rs13020260>

270. Zheng J, Fu H, Li W, Wu W, Yu L, Yuan S, Tao WYW, Pang TK, Kanniah KD (2021) Growing status observation for oil palm trees using unmanned aerial vehicle (UAV) images. *ISPRS J Photogramm Remote Sens* 173:95–121. <https://doi.org/10.1016/j.isprsjprs.2021.01.008>

271. Kentsch S, Caceres MLL, Serrano D, Roure F, Diez Y (2020) Computer vision and deep learning techniques for the analysis of drone-acquired forest images, a transfer learning study. *Remote Sens* 12(8):1287. <https://doi.org/10.3390/rs12081287>

272. Stewart EL, Wiesner-Hanks T, Kaczmar N, DeChant C, Wu H, Lipson H, Nelson RJ, Gore MA (2019) Quantitative phenotyping of northern leaf blight in UAV images using deep learning. *Remote Sens* 11(19):2209. <https://doi.org/10.3390/rs11192209>

273. Lottes P, Behley J, Chebrolu N, Milioto A, Stachniss C, (2018) Joint stem detection and crop-weed classification for plant-specific treatment in precision farming. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), IEEE. <https://doi.org/10.1109/iros.2018.8593678>

274. Wiesner-Hanks T, Stewart EL, Kaczmar N, DeChant C, Wu H, Nelson RJ, Lipson H, Gore MA (2018) Image set for deep learning: field images of maize annotated with disease symptoms. *BMC Res Notes*. <https://doi.org/10.1186/s13104-018-3548-6>

275. Santos Ferreira A, Freitas DM, Silva GG, Pistori H, Folhes MT (2017) Weed detection in soybean crops using convnets. *Comput Electron Agric* 143:314–324. <https://doi.org/10.1016/j.compag.2017.10.027>

276. Yu M, Yang C, Li Y (2018) Big data in natural disaster management: a review. *Geosciences* 8(5):165. <https://doi.org/10.3390/geosciences8050165>

277. Shamsoshoara A, Afghah F, Razi A, Zheng L, Fulé PZ, Blasch E (2021) Aerial imagery pile burn detection using deep learning: the FLAME dataset. *Comput Netw* 193:108001. <https://doi.org/10.1016/j.comnet.2021.108001>

278. Barz B, Schröter K, Münch M, Yang B, Unger A, Dransch D, Denzler J (2019) Enhancing flood impact analysis using interactive retrieval of social media images. *Archives of data science, series A*, 5, 1, 2018 [https://doi.org/10.5445/KSP/1000087327/06arXiv:1908.03361v1 \[cs.IR\]](https://doi.org/10.5445/KSP/1000087327/06arXiv:1908.03361v1)

279. Foglia P, Saggese A, Vento M (2015) Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion. *IEEE Trans Circuits Syst Video Technol* 25(9):1545–1556. <https://doi.org/10.1109/tcsvt.2015.2392531>

280. Bugarić M, Jakovčević T, Stipanićev D (2014) Adaptive estimation of visual smoke detection parameters based on spatial data and fire risk index. *Comput Vis Image Underst* 118:184–196. <https://doi.org/10.1016/j.cviu.2013.10.003>

281. Mahrad BE, Newton A, Icely J, Kacimi I, Abalansa S, Snoussi M (2020) Contribution of remote sensing technologies to a holistic coastal and marine environmental management framework: a review. *Remote Sens* 12(14):2313. <https://doi.org/10.3390/rs12142313>

282. Yang Z, Yu X, Dedman S, Rosso M, Zhu J, Yang J, Xia Y, Tian Y, Zhang G, Wang J (2022) UAV remote sensing applications in marine monitoring: knowledge visualization and review. *Sci Total Environ* 838:155939. <https://doi.org/10.1016/j.scitotenv.2022.155939>

283. Savastano VLM, Batista DB (2021) Fractures in UAV imagery for segmentation. <https://doi.org/10.21227/5939-y446>

284. SenseFly: Industrial Estate Dataset. <https://www.sensefly.com/>

285. Singh A, Kalke H, Loewen M, Ray N (2020) River ice segmentation with deep learning. *IEEE Trans Geosci Remote Sens* 58(11):7570–7579. <https://doi.org/10.1109/tgrs.2020.2981082>

286. Amundson J (2019) LeConte glacier unmanned aerial vehicle (UAV) imagery, LeConte Glacier, Alaska, 2018. <https://doi.org/10.18739/A2445HC19>

287. Wang J, Guo W, Pan T, Yu H, Duan L, Yang W (2018) Bottle detection in the wild using low-altitude unmanned aerial vehicles. In: 2018 21st international conference on information fusion (FUSION). IEEE. <https://doi.org/10.23919/icif.2018.8455565>

288. Puijenbroek MEB, Nolet C, Groot AV, Suomalainen JM, Riksen MJPM, Berendse F, Limpens J (2017) Exploring the contributions of vegetation and dune size to early dune development using unmanned aerial vehicle (UAV) imaging. *Biogeosciences* 14(23):5533–5549. <https://doi.org/10.5194/bg-14-5533-2017>

289. Yang Y, Newsam S (2010) Bag-of-visual-words and spatial extensions for land-use classification. In: Proceedings of the 18th SIGSPATIAL International conference on advances in geographic information systems. GIS '10, pp. 270–279. Association for Computing Machinery, New York, NY. <https://doi.org/10.1145/1869790.1869829>

290. Toker A, Kondmann L, Weber M, Eisenberger M, Camero A, Hu J, Hoderlein AP, Şenaras c, Davis T, Cremers D, Marchisio G, Zhu XX, Leal-Taixé L (2022) Dynamicearthnet: daily multispectral satellite dataset for semantic change segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 21158–21167

291. Yavariabdi A, Kusetogullari H, Orhan O, Uray E, Demir V, Celik T SinkholeNet: a novel deep learning framework for sinkhole classification and localization in high-resolution aerial images. <https://github.com/sinkholenet/sinkholenet/>

292. Lyu Y, Vosselman G, Xia G-S, Yilmaz A, Yang MY (2020) Uavid: a semantic segmentation dataset for UAV imagery. *ISPRS J Photogramm Remote Sens* 165:108–119. <https://doi.org/10.1016/j.isprsjprs.2020.05.009>

293. Avola D, Cinque L, Foresti GL, Martinel N, Pannone D, Piciarelli C (2020) A UAV video dataset for mosaicking and change detection from low-altitude flights. *IEEE Trans Syst Man Cybern Syst* 50(6):2139–2149. <https://doi.org/10.1109/tsmc.2018.2804766>

294. Tommaselli AMG, Galo M, Reis TT, Silva Ruy R, Moraes MVA, Matricardi WV (2018) Development and assessment of a data set containing frame images and dense airborne laser scanning point clouds. *IEEE Geosci Remote Sens Lett* 15(2):192–196. <https://doi.org/10.1109/lgrs.2017.2779559>

295. Pavia University Scene. https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Pavia_Centre_and_University

296. Indian Pines. https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Pavia_Centre_and_University

297. Giri C, Ochieng E, Tieszen LL, Zhu Z, Singh A, Loveland T, Masek J, Duke N (2011) Status and distribution of mangrove

forests of the world using earth observation satellite data. *Glob Ecol Biogeogr* 20(1):154–159

298. Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR et al (2013) High-resolution global maps of 21st-century forest cover change. *science* 342(6160):850–853

299. Millennium Ecosystem Assessment (2005) Millennium ecosystem assessment: MA biodiversity. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, New York. <https://doi.org/10.7927/H4V9860B>

300. Rodell M, Famiglietti JS, Wiese DN, Reager J, Beaudoing HK, Landauer FW, Lo M-H (2018) Emerging trends in global freshwater availability. *Nature* 557(7707):651–659

301. Ramankutty N, Evan AT, Monfreda C, Foley JA (2008) Farming the planet: 1. geographic distribution of global agricultural lands in the year 2000. *Global Biogeochem Cycles*. <https://doi.org/10.1029/2007GB002952>

302. Imhoff ML, Bounoua L (2006) Exploring global patterns of net primary production carbon supply and demand using satellite observations and statistical data. *J Geophys Res Atmos*. <https://doi.org/10.1029/2006JD007377>

303. Center for International Earth Science Information Network—CIESIN—Columbia University (2020) Food Insecurity Hotspots Data Set. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, New York. <https://doi.org/10.7927/cx02-2587>

304. Rosvold EL, Buhaug H (2021) Gdis, a global dataset of geocoded disaster locations. *Scientific Data* 8(1):61

305. Bountos NI, Papoutsis I, Michail D, Karavias A, Elias P, Parcharidis I (2022) Hephaestus: A large scale multitask dataset towards insar understanding. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR) workshops, pp 1453–1462

306. Sayad YO, Mousannif H, Al Moatassime H (2019) Predictive modeling of wildfires: a new dataset and machine learning approach. *Fire Saf J* 104:130–146

307. Elvidge CD, Zhizhin M, Hsu F-C, Baugh KE (2013) Viirs nightfire: satellite pyrometry at night. *Remote Sens* 5(9):4423–4449

308. Giglio L, Randerson JT, Van Der Werf GR (2013) Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (gfed4). *J Geophys Res Biogeosci* 118(1):317–328

309. Gupta R, Goodman B, Patel N, Hosfelt R, Sajeev S, Heim E, Doshi J, Lucas K, Choset H, Gaston M (2019) Creating xbd: a dataset for assessing building damage from satellite imagery. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR) workshops

310. Claverie M, Ju J, Masek JG, Dungan JL, Vermote EF, Roger J-C, Skakun SV, Justice C (2018) The harmonized landsat and sentinel-2 surface reflectance data set. *Remote Sens Environ* 219:145–161

311. Green D (2021) Nasa disasters program: Earth observation for actionable knowledge. In: 2021 IEEE international geoscience and remote sensing symposium IGARSS, IEEE, pp 1709–1710

312. Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, Randles CA, Darmenov A, Bosilovich MG, Reichle R, Wargan K, Coy L, Cullather R, Draper C, Akella S, Buchard V, Conaty A, Silva AM, Gu W, Kim G-K, Koster R, Lucchesi R, Merkova D, Nielsen JE, Partyka G, Pawson S, Putman W, Rienecker M, Schubert SD, Sienkiewicz M, Zhao B (2017) The modern-era retrospective analysis for research and applications, version 2 (merra-2). *J Clim* 30(14):5419–5454. <https://doi.org/10.1175/JCLI-D-16-0758.1>

313. Hooker J, Duveiller G, Cescatti A (2018) A global dataset of air temperature derived from satellite remote sensing and weather stations. *Scientific Data* 5(1):1–11

314. Jacobson AR, Schuldt KN, Miller JB, Oda T, Tans P, Andrews A, Mund J, Ott L, Collatz GJ, Aalto T et al (2020) Carbon-tracker documentation ct2019 release. Global Monitoring Laboratory-Carbon Cycle Greenhouse Gases

315. Li Y, Gao H, Zhao G, Tseng K-H (2020) A high-resolution bathymetry dataset for global reservoirs using multi-source satellite imagery and altimetry. *Remote Sens Environ* 244:111831

316. Joyce RJ, Janowiak JE, Arkin PA, Xie P (2004) Cmorph: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J Hydrometeorol* 5(3):487–503

317. Tourian MJ, Elmí O, Shafaghi Y, Behnia S, Saemian P, Schlesinger R, Sneeuw N (2022) Hydrosat: geometric quantities of the global water cycle from geodetic satellites. *Earth Syst Sci Data* 14(5):2463–2486

318. Tomita H, Hihara T, Kako S, Kubota M, Kutsuwada K (2019) An introduction to j-ofuro3, a third-generation Japanese ocean flux data set using remote-sensing observations. *J Oceanogr* 75(2):171–194

319. Pekel J-F, Cottam A, Gorelick N, Belward AS (2016) High-resolution mapping of global surface water and its long-term changes. *Nature* 540(7633):418–422

320. Tasseron P, Van Emmerik T, Peller J, Schreyers L, Biermann L (2021) Advancing floating macroplastic detection from space using experimental hyperspectral imagery. *Remote Sens* 13(12):2335

321. Schmitt M, Hughes LH, Qiu C, Zhu XX (2019) Sen12ms—a curated dataset of georeferenced multi-spectral sentinel-1/2 imagery for deep learning and data fusion. [arXiv:1906.07789](https://arxiv.org/abs/1906.07789)

322. Mnih V (2013) Machine learning for aerial image labeling. PhD thesis, University of Toronto

323. Shermeyer J, Hogan D, Brown J, Van Etten A, Weir N, Pacifici F, Hansch R, Bastidas A, Soenen S, Bacastow T et al (2020) Spacenet 6: Multi-sensor all weather mapping dataset. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp 196–197

324. Long Y, Gong Y, Xiao Z, Liu Q (2017) Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Trans Geosci Remote Sens* 55(5):2486–2498

325. Xiao Z, Long Y, Li D, Wei C, Tang G, Liu J (2017) High-resolution remote sensing image retrieval based on CNNs from a dimensional perspective. *Remote Sens* 9(7):725

326. Mundhenk TN, Konjevod G, Sakla WA, Boakye K (2016) A large contextual dataset for classification, detection and counting of cars with deep learning. In: Computer vision–ECCV 2016: 14th European conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part III 14, Springer, pp 785–800

327. Wang Y, Wang C, Zhang H, Dong Y, Wei S (2019) A sar dataset of ship detection for deep learning under complex backgrounds. *Remote Sens*. <https://doi.org/10.3390/rs11070765>

328. Liu Z, Yuan L, Weng L, Yang Y (2017) A high resolution optical satellite image dataset for ship recognition and some new baselines. In: International conference on pattern recognition applications and methods, vol 2. SciTePress, pp 324–331

329. Wang Z, Bai L, Song G, Zhang J, Tao J, Mulvenna MD, Bond RR, Chen L (2021) An oil well dataset derived from satellite-based remote sensing. *Remote Sens* 13(6):1132

330. Wang M, Deng W (2018) Deep visual domain adaptation: a survey. *Neurocomputing* 312:135–153

331. Bengio Y, Louradour J, Collobert R, Weston J (2009) Curriculum learning. In: Proceedings of the 26th annual international conference on machine learning, pp 41–48

332. Hong D, Zhang B, Li X, Li Y, Li C, Yao J, Yokoya N, Li H, Ghamisi P, Jia X et al (2024) Spectralgpt: spectral remote sensing foundation model. *IEEE Trans Pattern Anal Mach Intell* 46:5227–5244

333. Michler JD, Josephson A, Kilic T, Murray S (2022) Privacy protection, measurement error, and the integration of remote sensing and socioeconomic survey data. *J Dev Econ* 158:102927

334. Alkhelaiwi M, Boulila W, Ahmad J, Koubaa A, Driss M (2021) An efficient approach based on privacy-preserving deep learning for satellite image classification. *Remote Sens* 13(11):2221

335. Gao P, Zhang H, Yu J, Lin J, Wang X, Yang M, Kong F (2020) Secure cloud-aided object recognition on hyperspectral remote sensing images. *IEEE Internet Things J* 8(5):3287–3299

336. Zhang D, Ren L, Shafiq M, Gu Z (2022) A lightweight privacy-preserving system for the security of remote sensing images on iot. *Remote Sens* 14(24):6371

337. McGovern A, Bostrom A, McGraw M, Chase RJ, Gagne DJ, Ebert-Uphoff I, Musgrave KD, Schumacher A (2024) Identifying and categorizing bias in ai/ml for earth sciences. *Bull Am Meteorol Soc* 105(3):E567–E583

338. Garcia-del-Real J, Alcaráz M (2024) Unlocking the future of space resource management through satellite remote sensing and ai integration. *Resour Policy* 91:104947

339. Zhang B, Wu Y, Zhao B, Chanussot J, Hong D, Yao J, Gao L (2022) Progress and challenges in intelligent remote sensing satellite systems. *IEEE J Sel Top Appl Earth Observ Remote Sens* 15:1814–1822. <https://doi.org/10.1109/JSTARS.2022.3148139>

Publisher's Note Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.