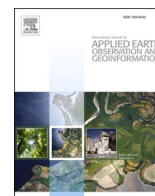


Title	Status, advancements and prospects of deep learning methods applied in forest studies
Authors	Yun, Ting;Li, Jian;Ma, Lingfei;Zhou, Ji;Wang, Ruisheng;Eichhorn, Markus P.;Zhang, Huaqing
Publication date	2024-06-04
Original Citation	Yun, T., Li, J., Ma, L., Zhou, J., Wang, R., Eichhorn, M. P. and Zhang, H. (2024) 'Status, advancements and prospects of deep learning methods applied in forest studies', International Journal of Applied Earth Observation and Geoinformation, 131, 103938 (20pp). https://doi.org/10.1016/j.jag.2024.103938
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1016/j.jag.2024.103938
Rights	© 2024, the Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-10-29 00:55:12
Item downloaded from	https://hdl.handle.net/10468/16905



UCC

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh



Status, advancements and prospects of deep learning methods applied in forest studies

Ting Yun^{a,b,1}, Jian Li^{a,1}, Lingfei Ma^c, Ji Zhou^d, Ruisheng Wang^e, Markus P. Eichhorn^f,
Huaiqing Zhang^{g,*}

^a College of Information Science and Technology & Artificial Intelligence, Nanjing Forestry University, Nanjing 210037, China

^b Co-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing 210037, China

^c School of Statistics and Mathematics, Central University of Finance and Economics, Beijing 102206, China

^d Cambridge Crop Research, National Institute of Agricultural Botany (NIAB), Cambridge CB3 0LE, United Kingdom

^e Department of Geomatics Engineering, University of Calgary, Alberta T2N 1N4, Canada

^f School of Biological, Earth and Environmental Sciences, University College Cork, Distillery Fields, North Mall, Cork T23 N73K, Ireland

^g Research Institute of Forest Resources Information Techniques, Chinese Academy of Forestry, Beijing 100091, China

ARTICLE INFO

Keywords:

Deep learning network
Forest application
Remote sensing
Point cloud
Satellite imagery
Aerial photography

ABSTRACT

Deep learning, which has exhibited considerable potential and effectiveness in forest resource assessment, is vital for comprehending and managing forest resources and ecosystems. However, extensive assessment of forest resources is highly challenging due to the complex and varied nature of forest types sourced from diverse remote sensing platforms, which include images, point clouds, and fusion data. To facilitate further study, we systematically review the current status, applications and prospects of deep learning technologies for different types of forest remote sensing data. After considering more than two hundred forest-related papers published over the past decade, we introduce sensors and devices for forest data acquisition, classify deep learning methods based on their data processing methods and operational principles, and categorize diverse instances of these methods with various forest applications. Moreover, we summarize available datasets related primarily to forest data and examine the global geographic distribution of the relevant literature. Comprehensive insights into the advantages and limitations of each method are described, offering a forward-looking perspective on the trend of applying deep learning technology to forest research. In this paper, we aim to provide an overview of the current trends and challenges of deep learning techniques applied to forest research, creating a comprehensive picture for use as a reference by both academia and industry professionals.

1. Introduction

Forest resource assessment provides essential quantitative information for understanding the extent of forest resources, perceptions of forest growing conditions, awareness of forest management efficacy, and forest degradation or deforestation. Automatically examining the trends in many forest-related variables and territories over time through intelligent, traceable and automatic computational processes is essential for precision forest management, automated forest monitoring and integrated forest retrieval. The motivations for forest resource assessment and management are multifaceted, reflecting the multiple functions of forest protection and exploitation as well as societal appreciation for the

quality and stability of forest ecosystems. These motivations are beneficial for farmer poverty alleviation (van Noordwijk, 2019), climate change mitigation (McEwan et al., 2020) and global biodiversity conservation (Lelli et al., 2019). The current challenges in the assessment of forest resources are predominantly related to the collection of appropriate remote sensing data and the methods used for analysing these data. Selecting optimal data sources and resolutions for specific forest objectives is not trivial and depends on multiple factors, including forest structure, stem density, measurement accessibility, atmospheric conditions, and sensor characteristics (Fassnacht et al., 2023). Furthermore, advanced techniques and algorithms are needed to synergistically interpret forest remote sensing data to extract valuable indications and

* Corresponding author.

E-mail address: zhang@caf.ac.cn (H. Zhang).

¹ These authors are the co-first authors who contributed equally to this work.

reduce measurement errors (Hamedianfar et al., 2022).

Modern remote sensing technologies can obtain high-precision temporal and spatial forest data with appearance and morphogenesis information. Many remote sensing sensors, such as RGB digital cameras, multispectral or hyperspectral cameras, chlorophyll fluorescence sensors and light detection and ranging (LiDAR) sensors, have been applied in practice for task-requiring data acquisition of forest resources (Zhang et al., 2021a). These modern and advanced remote sensing devices enrich the presentation of forest-related data supported by research and multiscale forest landscape monitoring. However, forest remote sensing data, often represented as images or point clouds, frequently lack comprehensive and interpretable semantic information. For instance, images may exhibit limited structural or geometrical representation of tree crowns, while point clouds may exhibit reduced spatial permutation relationships; thus, processing these data effectively is challenging. Therefore, an efficient and operable computational strategy focused on various remote sensing data is indispensable for detecting and monitoring the physiological or biogeochemical characteristics of forest areas.

The effective processing of forest remote sensing data is crucial for modern forest management and cultivation practices. Machine learning is a branch of artificial intelligence and computer science that uses algorithms and models to learn from data, similar to human learning, with the goal of teaching a machine to accurately perform a task by identifying patterns through mathematical optimization, computational statistics, and data mining. Various machine learning algorithms (Yang et al., 2024), e.g., multilinear regression methods, support vector machines, random forests, and decision trees (Gao et al., 2023), have been employed in forest biomass estimation (Yang et al., 2023), mixed tree species classification (Wagner et al., 2019), and wildfire susceptibility mapping (Gholamnia et al., 2020), respectively. Certain drawbacks still exist when using machine learning algorithms to analyse complicated forest morphological traits and crown appearance variances under different phenological periods; e.g., the results heavily depend on manual algorithm parameter tuning, and these results exhibit algorithm performance degeneracy when facing nonuniform and irregular forest remote sensing data, weak reproducibility with overfitting or underfitting phenomena, and difficult bias detection.

Deep learning is a subset of machine learning that can automatically

learn pivotal features from raw data without requiring manual feature engineering. Compared with machine learning, deep learning models are better suited for handling high-dimensional, nonlinear and complex data in forest applications (Lai et al., 2022), with the advantages of automatic feature extraction and the ability to model complex mathematical relationships between cascaded networks and build robust and adaptive frameworks for diverse dataset analysis (Schmidhuber, 2015). These advantages can help overcome some of the limitations and challenges of machine learning methods in forest applications. Advancements in remote sensing technologies have greatly improved the assessment of forest resources. However, some challenges still exist in selecting the best data sources and developing reliable algorithms for forest data processing and analysis. Therefore, it is important to investigate the specific advances of deep learning effectively and responsibly in relation to forest problems in recent years and to understand the trade-offs, knowledge gaps and restrictions involved in related works. To address this issue, we sort and classify the existing works and construct a systematic diagram describing the current status, practical obstacles and future prospects of deep learning techniques in forest applications. The predominant contributions and framework (Fig. 1) of this study are as follows:

1) The current status of deep learning technology in forest resource research is reviewed, and by analysing relevant publications, various forest application scenarios are synthesized and organized into four categories. In light of this, the historical and current trends of forest application scenarios are investigated and chronologically illustrated (Section 2).

2) According to the data type and algorithm principle, deep learning methods are divided into forest images (Section 3) and forest point clouds (Section 4) for quantitative performance comparisons, and the applicability and restrictions in the realm of the respective forest fields are evaluated. Moreover, works based on the fusion of forest images and point clouds are similarly surveyed and described (Section 5).

3) The sources of the primary public datasets containing forest data are summarized in Section 6, and the global distribution of forest-related literature over the last five years is examined in Section 7, which includes schematic representations of regional academic variances and corresponding thematic trends.

4) Adopting current studies as the context, existing obstacles and

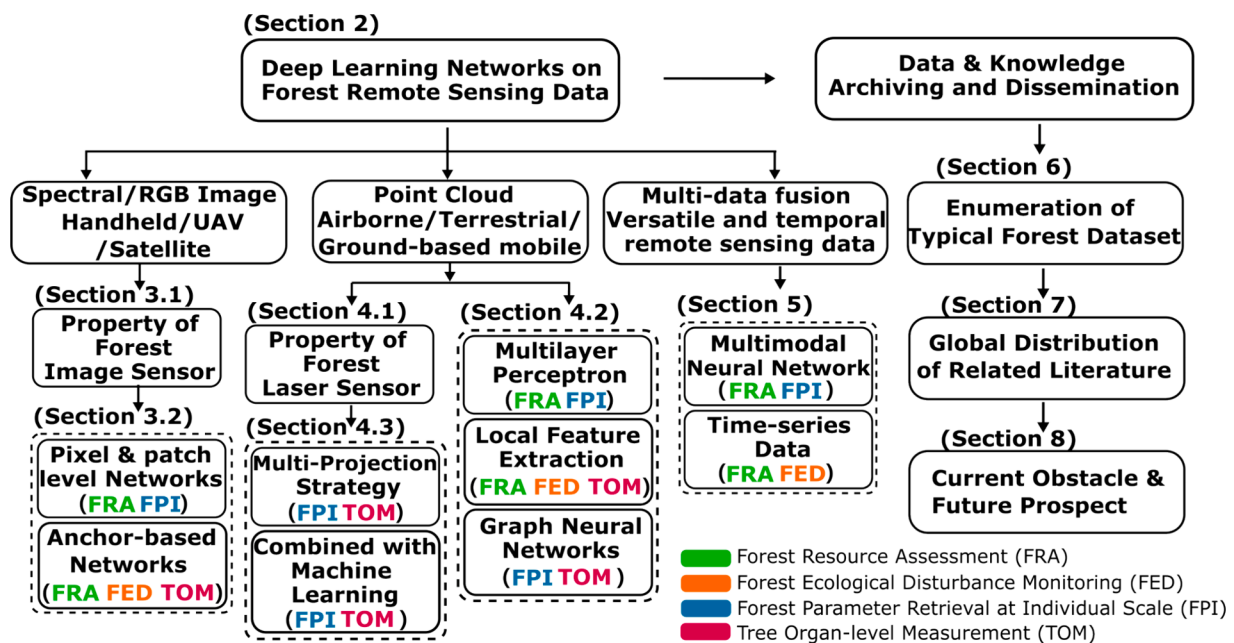


Fig. 1. Document map of this paper, comprising a taxonomy of deep learning methods for forest remote sensing data and the shared data and knowledge for forest research with a discussion on further prospects.

future prospects for advancing deep learning techniques in forest applications are established and explained (Section 8).

2. Chronological sequence of deep learning technology in forest applications

We divided deep learning methods into four categories based on the specific application of deep learning techniques in forests, i.e., forest resource assessment (FRA), forest ecological disturbance monitoring (FED), forest parameter retrieval at the individual scale (FPI) and tree organ-level measurement (TOM), which are shown in detail in Fig. 2. FRA reflects the present conditions and merging patterns of current forests and involves detecting tree species distributions, mapping forest coverage and estimating aboveground biomass. Generally, the main objective of FRA is to accommodate holistic forest stewardship through the use of large-scale forest remote sensing images and advanced algorithms for macroscopic analyses. FED concentrates exclusively on tracking biologic or abiotic stresses that affect forest wellness, metabolism, profitability, and diversity; the majority of those stresses, such as pest/disease damage, natural disasters and habitat transitions, are not anthropogenic. FED could contribute to a better quantification of forest degradation and disturbance, providing significant insights into the forest evolutionary process under interactions with environmental variations. In addition to the large-scale forest community observations above, some studies at the individual tree scale have likewise been split into two categories. One category concerns forest parameter retrieval of individual trees from multiple remote sensing data, termed “FPI”, which refers to the extraction and characterization of the intrinsic structure and growth attributes of single stands, such as tree species recognition, canopy and subcanopy tree classification, LAI estimates and crown feature evaluation. The other category is termed “TOM” and focuses on recognizing the constituents of trees, including photosynthetic and nonphotosynthetic component separation, fruit and yield appraisal, and skeleton and tree modelling. TOM is generally employed to support a preprocessing or auxiliary step for further investigation at finer scales. Generally, FPI and TOM assist in analysing structural and functional connectivity for trees, scrutinizing exterior appearance, tracking genetic performance and investigating detailed situations during tree

development under various silvicultural interventions. Through the aforementioned categorization, we can define the implementation of deep learning technology in forest application scenarios, which benefits the taxonomic alignment of related literature and bolsters the consistent and coherent presentation of the contents of this paper.

Since their development, deep learning techniques were initially deployed for industrial or traffic applications, e.g., robotics, remote measurement and automatic pilot technology, and substantial progress has been achieved in related fields, which has spurred the potential of deep learning technology to be extended to forest domains and revolutionized traditional forest survey methods using novel concepts. We present a chronological overview (Fig. 3) of the primary deep learning models applied in forest studies over the last five years; these models encompass deep learning frameworks applied at various scales of remote sensing data, from the comprehensive forest regional scale to meticulous details at the individual tree scale.

Forest resource assessment (FRA), as demonstrated in Fig. 3 and denoted in green, is a fundamental application of particular investigations into forests. FRA is usually distributed before 2022 and is modest in quantity; this type of assessment occurs predominantly in forest biomass estimation (Ghosh and Behera, 2021) and species distribution monitoring (Huang et al., 2023) from satellite and aerial images (dark blue arrow) employing machine learning tactics and black-box-like deep learning frameworks. With the increased risk of natural accidents, studies on forest ecological disturbance monitoring (FED) using deep learning techniques, marked in orange, have emerged and increased gradually since 2021. FED is employed to mitigate and lessen the impacts of natural factors that jeopardize the health and stability of forests; this approach relies primarily on large-scale satellite images (dark blue arrow) similar to FRA, and its applications include tracking soil erosion from frequent and unstable natural disasters (Senanayake et al., 2022), preventing forest mortality from pest/disease outbreaks (Yu et al., 2021), alerting for destructive forest wildfires (Yaloveha et al., 2019), and identifying hurricane onslaught on the forest in coastal areas using innovative deep learning networks. The other two research categories, forest parameter retrieval at the individual scale (FPI), labelled in blue, and tree organ-level measurement (TOM), labelled in purple, across divergent remote sensing data (purple arrow), were used to

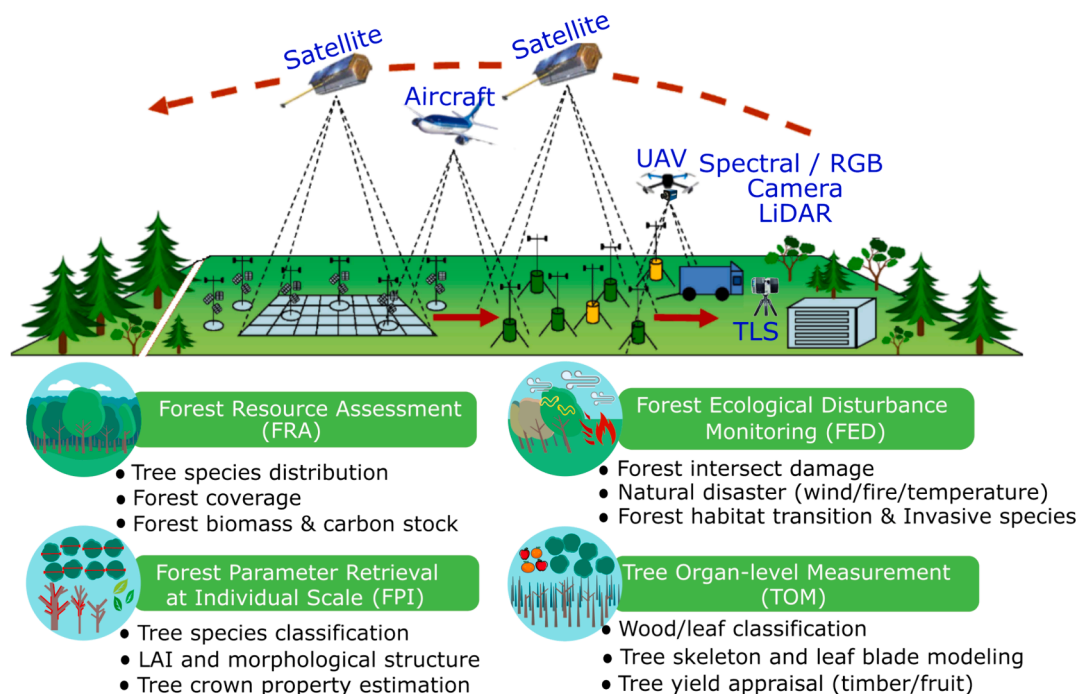


Fig. 2. Four categories of forest applications based on multiple remote sensing data and deep learning techniques.

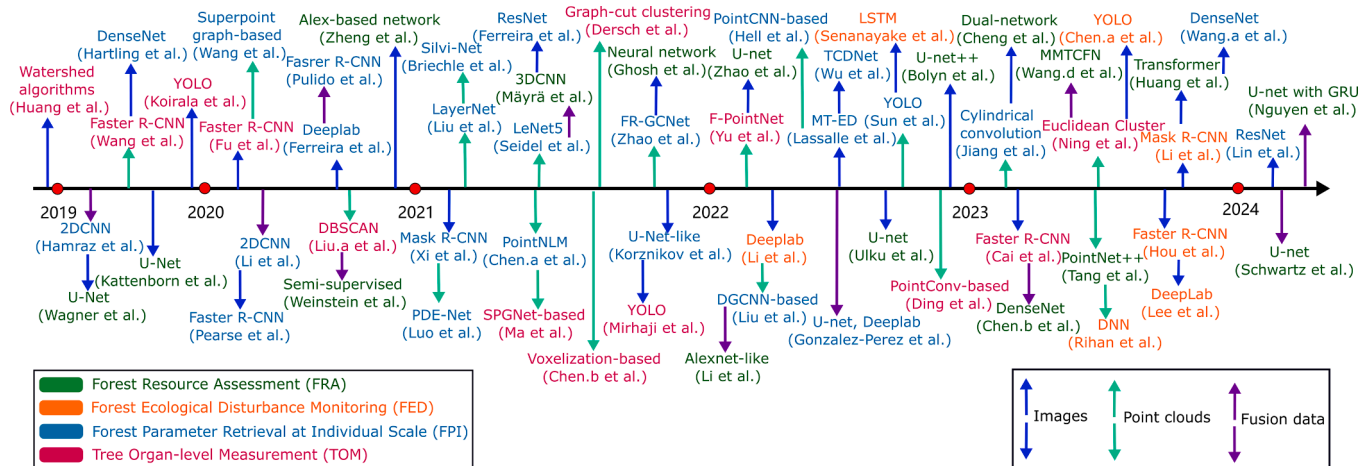


Fig. 3. Diagram of the chronological applications of deep learning algorithms to forest research over the past five years.

conduct in-depth forest analysis from 2018 to 2023. Herein, parameter retrieval is employed to accurately identify and recognize the traits of individual trees or small-scale woodland scenes by harnessing various deep learning models, such as DenseNet (Hartling et al., 2019) and DGCNN (Liu et al., 2022), to conduct fine-scale tasks, e.g., individual tree species recognition (Ferreira et al., 2020) and canopy property estimation (Sun et al., 2022) from point clouds (light blue arrow) and imagery (dark blue arrow) captured in enclosed areas, providing insights into their morphological skeleton structures, seasonal growth status, and exterior and interior characteristics. Moreover, tree organ-level measurements involve exploiting target localization networks, e.g., You Only Look Once (YOLO) (Mirhaji et al., 2021) and Faster-RCNN (Cai et al., 2023), to construct and segment respective tree organs, e.g., through fruit location positioning (Yu et al., 2022), bole and skeleton modelling (Ding et al., 2022), and orchard fruit load estimation (Koirala et al., 2019) from handheld camera images and point clouds, thereby transforming how tree structures are scrutinized and offering researchers novel perceptions of the anatomy and physiology of trees.

3. Deep learning methods for forest spectral and RGB images

3.1. Sensors for forest image data collection

Forest image data can be divided into spectral and RGB images depending on the sensor used; remote sensing equipment primarily employs hyperspectral/multispectral sensors loaded on satellites and unmanned aerial vehicles (UAVs), which serve as primary piggyback platforms. In forest research applications, common satellites include Sentinel and Landsat, as well as the Moderate Resolution Imaging Spectroradiometer (MODIS), Hyperion, and Synthetic Aperture Radar (SAR). The mesoscale satellites used include the Worldview and QuickBird satellites, whose relevant parameters are listed in Table 1. China possesses a variety of earth observation satellites, such as the Gaofen (GF) (Yao et al., 2021) series. These satellites provide high-precision spectral images with sharp spatial and temporal resolutions at the submeter and centimetre levels for monitoring forest resource management. Additionally, UAVs are widely utilized due to their rapid manoeuvrability, flexibility, and economic consumption; moreover, UAVs are loaded with hyperspectral/multispectral cameras, providing real-time and precise spectral data for forest research. The specific

Table 1

Parameters of the hyperspectral/multispectral sensors carried aboard satellites or aircraft utilized for forest research.

Satellite	Sensors	Obtained bands	Resolution	Revisit cycle
Sentinel-2	Multispectral Instrument	C/A; B; G; R; RE; NIR; WVP; SWIR (3)	C/A, WVP, SWIR: 60 m; B, G, R, NIR: 10 m; RE, SWIR (2) : 20 m	5 days
WorldView-2	WorldView-110 camera	C/A; B; G; Y; R; PAN; RE; NIR (2)	Panchromatic: 0.46 m; Multispectral: 1.84 m	1.1 days
WorldView-3	Multispectral Sensor	C/A; B; G; Y; R; PAN; RE; NIR (2); SWIR (8)	Panchromatic: 0.31 m; Multispectral: 1.24 m; SWIR: 3.7 m	1.0 days
Landsat 8	Operational Land Imager, Thermal Infrared Sensor	PAN; C/A; B; G; R; NIR; SWIR (3); TIR (2)	Panchromatic: 15 m; Multispectral: 30 m; TIR: 100 m	16 days
QuickBird	Ball Aerospace's Global Imaging System	PAN; B; G; R; NIR	Panchromatic: 0.61–0.72 m; Multispectral: 2.44–2.88 m	1–6 days
GF-II	Panchromatic, Multispectral Sensor	PAN; B; G; R; NIR	Panchromatic: 0.8 m; Multispectral: 3.2 m	5 days
GEO-eye	Optical Camera	G; R; B; NIR	Panchromatic: 0.41 m; Multispectral: 1.65 m	3 days
Portable hyperspectral sensors	Company and location	Bands range and quantity		Spectral resolution
FigSpec FS-2X	Zhejiang CHNSpec Technology Co., Ltd.	900–1700 nm (254)		8 nm
GaiaField-Pro	Sichuan Dualix Spectral Imaging Technology Co., Ltd.	900–1700 nm (300)		4 nm
SPECIM IQ	Finland Specim, Spectral Imaging Ltd.	400–1000 nm (204)		7 nm
xiSpec	Germany XIMEA SOFTARD Technologies	470–900 nm (150)		3 nm
Headwall-Nano	USA Headwall Photonics	400–1000 nm (340)		1.9 nm

Note: C/A (Coastal/Aerosol), B (Blue), G (Green), Y (Yellow), R (Red), and PAN (Panchromatic); RE (Red edge), NIR (Near-Infrared), WVA (Water vapour), SWIR (Short-Wave Infrared), and TIR (Thermal Infrared). SWIR (3) indicates that the SWIR band is divided into 3 different SWIR bands.

instrument parameters of several well-known brands of hyperspectral/multispectral cameras are listed in Table 1.

3.2. Deep learning techniques for forest image processing

The acquired images of trees/forests provide complex and diverse sets of spectral and texture features defined by divergent crown reflectance, various branch structures, and complex leaf appearance. These complexities inhibit the extraction of valuable features from forest images through traditional image processing methods (image segmentation, enhancement, object recognition and feature extraction). To address these challenges, researchers have increasingly employed deep learning techniques, which have proven highly effective in identifying and analysing fine-grained details in forest images with assistance in excavating a bounty of details stemming from forest pictures. These methods are broadly categorized into two categories, i.e., pixel- and patch-level and anchor-based methods. Fig. 4 outlines several prominent deep learning frameworks founded on the principles of pixel- and patch-level and anchor-based networks, demonstrating their applicability across various forest-related applications. Table 2 shows a list of deep learning methodologies based on forest image research conducted in recent years.

(1) Pixel- and patch-level networks are applied to forest images mostly collected from satellite observations and UAV collections. The original intention of “pixel- and patch-level networks” was to process each pixel or small patch in an image separately by multilayer convolutions, allowing these networks to capture slight variations in each pixel or patch and perform user-desired operations suitable for meticulous classification of target trees within feed-in images. The frequently used deep neural networks include DenseNet, U-Net and DeepLab for pixelwise classification tasks and ResNet for obtaining patch-level classification labels (Chen et al., 2017). All of these networks execute feature extraction based on convolution and pooling operations with distinct operations for tree crown identification and species classification at the individual tree scale towards creating a whole-tree picture. Select typical neural networks are shown in Fig. 4, and their forest implementation cases are addressed below.

The U-Net framework comprises two parts: a contracting path for spatial information reduction and feature information enhancement and an expansive pathway combining feature and spatial information through a sequence of upconvolution and concatenation operations with high-resolution features from the contracting path. Accordingly, U-net and its extended network architectures have enabled various successful forest applications, including mapping of forest types and disturbances in the Atlantic rainforest from WorldView-3 satellite images, yielding a classification accuracy over 95 % (Wagner et al., 2019), classification of *Pinus radiata* from UAV images of central Chile forests, providing an accuracy exceeding 84 % (Kattenborn et al., 2019), classification of Mongolian poplar and certain coniferous trees on the basis of GeoEye-1 satellite RGB images of the Primorsky Region of the Russian Far East, with an accuracy score of 96 % (Korznikov et al., 2021), determination of tree cover areas from satellite and aerial images of Hamlin Beach State Park, New York, with pixel input features containing RGB and vegetation indices, yielding an identification rate of approximately 80 % (Ulku et al., 2022), and mapping of the proportion and composition of ten tree species from Sentinel-2 satellite images of southern Belgium, with an average classification precision of 73 % (Bolyn et al., 2022).

A semantic segmentation model of DeepLab provided atrous convolution embedded in a network cascade and involved the use of a spatial pyramid pooling module for better encoding of the global context and improving the performance. Many networks similar in structure to DeepLab have been employed for multiple forest tasks, e.g., individual tree crown detection and species classification of three Amazonian palm species from UAV images with the assistance of morphological image processing (Ferreira et al., 2020); pine wilt disease identification based on multispectral images acquired by a UAV in the southern Korean

Peninsula, with an accuracy of 74.2 % (Lee et al., 2023); and Brazil mangrove crown delineation from WorldView satellite images, with an accuracy higher than 90 % (Lassalle et al., 2022).

ResNet adopts image patches as input and introduces a skip connection pattern into the network architecture that allows deep layers to directly learn features derived from shallow layers, yielding the predicted category label of the input image patch. Some examples include processing UAV hyperspectral images of Taiwanese mountainous forests for different types of forest classifications (Lin et al., 2024), mapping the crowns of *Bertholletia excelsa* trees in Brazil from WorldView-3 satellite images, with an accuracy of approximately 93 % (Ferreira et al., 2021), and realizing woody vegetation segmentation from aerial images of the Iberian Peninsula in Spain by adopting the Resnet50 architecture as the foundational framework, with an accuracy of approximately 80 %.

The DenseNet network contains a dense block in which a dense connectivity pattern is provided that links each layer to each subsequent layer, which exhibits the merits of feature propagation enhancement and efficient feature reuse. DenseNet generates a categorical label as output, indicating the class of the input image, e.g., mountainous and urban tree species classification based on UAV high-resolution photos of ecological forests (Wang et al., 2023a) and multispectral pictures of public parks in USA (Hartling et al., 2019), achieving precision values of 90.1 % and 82.6 %, respectively.

Other network architectures, e.g., the AlexNet network combined with a multilevel attention mechanism for identifying image patches containing individual Malaysia oil palm trees from UAV images, achieved an accuracy higher than 80 % (Zheng et al., 2020); GoogLeNet (Tao et al., 2020) was employed for determining the spatial distribution of dead pine trees in mountainous forests in southeastern China, achieving favourable results in practical applications. All these endeavours reflect the notable evolution of deep learning networks in forest image processing, as this domain has been increasingly investigated.

(2) Anchor-based networks, in contrast to pixel-level networks, employ predefined reference boxes (anchors) of different sizes and aspect ratios to generate candidate boxes and aim to predict whether the candidate boxes contain objects and to determine their categories and locations. Anchor-based networks are always combined with manually annotated datasets for middle-scale tree crown localization and distribution analysis. These networks can be classified as one- or two-stage networks.

One-stage networks, e.g., YOLO (Koirala et al., 2019), which detects and classifies objects within a single neural network, skip the intermediate step of region proposals and class probability calculations for object localization and recognition. In recent years, YOLO has become a very popular tool for incorporating the latest concepts in the forest analysis field. For instance, the application of a lightweight model based on YOLO for detecting forest fire smoke in UAV images (Chen et al., 2023a) and TCDNet, which is based on an attention-enhancement path aggregation module with YOLOX and an occlusion-aware loss concentrated on tree crown detection, achieved an accuracy of 87.9 % and limited the influence of crown overlap from a UAV perspective (Wu et al., 2022).

Two-stage networks, e.g., Faster R-CNN (Hou et al., 2023) and Mask R-CNN (He et al., 2017), include a first stage of generating region proposals that contain objects and a second stage of extracting features and categorizing targets. Due to their superior generalization performance and detection accuracy, these networks are particularly well suited for middle-scale tree crown detection, parameter measurement, and fruit and pathological area localization in compact scenes. The applications of mainstream networks such as Faster R-CNN include the following: locating and mapping seedlings of Monterey pine in New Zealand based on UAV orthomosaic imagery, with an accuracy up to 96.5 % (Pearse et al., 2020); embedding an inception squeezed-excitation ResNet (ISResNet) network for apple leaf disease discrimination in Chinese

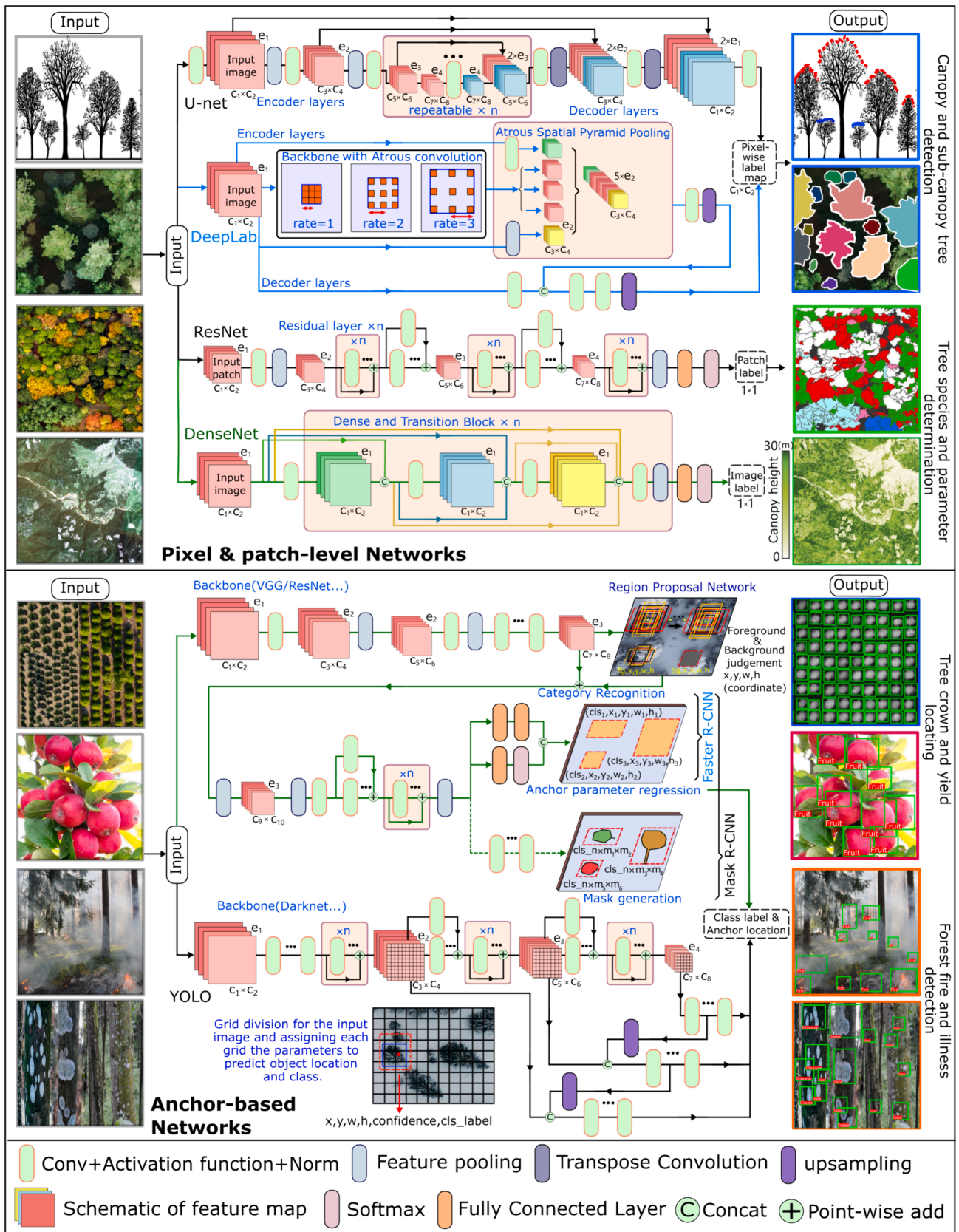


Fig. 4. Visualization of the prevailing deep learning frameworks based on the concepts of pixel-level and anchor-based networks for various forest applications.

Table 2

Applications of deep learning technology in forest image research.

Processing Strategy	Network Backbone	Remote sensing Source	Forest Type	Eva OA/P	Objective	Reference	Research location
Pixel & patch –level	U-Net-Based	WorldView3 (M)	MF	>95.0	FRA	(Wagner et al., 2019)	Atlantic Forest biome, Brazil
		UAV (R)	CF	>84.0	FPI	(Kattenborn et al., 2019)	Mount Cook Park, New Zealand
		GeoEye-1 (R)	MF	96.0	FPI	(Korzniakov et al., 2021)	Forest, Primorsky, Russian
		WorldView3 (M)	MF	79.7	FPI	(Ulku et al., 2022)	Beach park, New York, USA
		Sentinel-2 (M)	MF	73.0	FPI	(Boly et al., 2022)	Public forest, Wallonia, Belgium
	DeepLab-based	UAV (R)	DF	>85.0	FPI	(Ferreira et al., 2020)	Amazon, Rio Branco, Brazil
		UAV (M)	CF	99.7	FED	(Lee et al., 2023)	South of the Korean Peninsula
	ResNet	WorldView3 (M)	MF	>93.0	FPI	(Lassalle et al., 2022)	Brazil, USA, Australia, Gabon
		UAV (H)	MF	>93.0	FPI	(Lin et al., 2024)	Mountain, Taiwan, China
		WorldView3 (M)	MF	93.9	FPI	(Ferreira et al., 2021)	Forest, Amazon, Brazil
	DenseNet	UAV (R)	MF	81.5	FRA	(Cheng et al., 2023),	Highland, Iberian, Spain
		UAV (R)	MF	90.1	FPI	(Wang et al., 2023a)	Mountain, Chuzhou, China
	AlexNet-based	WorldView3 (M)	MF	82.6	FPI	(Hartling et al., 2019)	Forest park, St. Louis, USA
		QuickBird (M)Google Earth (R)	DF	84.8	FRA	(Zheng et al., 2020)	Plantation, Peninsular Malaysia
	GoogLeNet	UAV(R)	MF	97.3	FED	(Tao et al., 2020)	Jinjiang, Fujian, China
Anchor-based	YOLO-based	Depth Camera	MF	98.3	TOM	(Koirala et al., 2019)	Orchards, Queensland, Australia
		UAV (R)	MF	80.2	FED	(Chen et al., 2023a)	Online forest fire images
		UAV (R)	MF	87.9	TOM	(Wu et al., 2022)	Forest park, Yongjia, China
	Faster R-CNN	UAV (R)	CF	>96.5	FRA	(Pearse et al., 2020)	Rangip, Kaingaroa, New Zealand
		Camera (R)	DF	98.8	FED	(Hou et al., 2023)	Apple orchard, Enhe, China
		Depth Camera	DF	89.3	TOM	(Fu et al., 2020)	Orchard, Washington, USA
		UAV (R)	CF	95.3	TOM	(Cai et al., 2023)	Loblolly pine, Rusk, USA
	Mask R-CNN	UAV (M)	MF	95.3	TOM	(Xi et al., 2021)	Campus, Hangzhou, China
		UAV (H)	CF	83.5	FED	(Li et al., 2023)	Pine plantation, Chaohu, China

Note: H (hyperspectral), M (multispectral), R (RGB), CNN (convolutional neural network), CF (coniferous forest), DF (deciduous forest), MF (mixed forest), Eva (evaluation), OA (overall accuracy), P (precision). OA/P is directly derived from the corresponding literature, but in actual experiments, the magnitude may vary due to several factors including the captured image resolution, training dataset, number of categories, forest type and occlusion effect. FRA (forest resource assessment), FED (forest ecological disturbance monitoring), FPI (forest parameter retrieval at the individual scale), TOM (tree organ-level measurement).

orchards, with an accuracy of 98 % (Hou et al., 2023); analysing RGB images aligned with depth images for background filtering and apple detection in foreground-RGB images (Fu et al., 2020); and merging ResNet and the feature pyramid network (FPN) for loblolly pine canopy detection in the USA, with an accuracy of 95.3 % (Cai et al., 2023). Moreover, Mask R-CNN is an extension of Faster R-CNN that introduces a mask prediction branch to realize pixel-level segmentation for each instance. Therefore, Mask R-CNN, with higher-resolution UAV RGB and spectral images, can be utilized for ginkgo canopy edge delineation (Xi et al., 2021) and identification of pine trees infected by wilt disease (Li et al., 2023).

Considering the aforementioned works, deep learning methodologies have demonstrated significant capabilities in analysing RGB and spectral forest images of large- and medium-scale woodland scenes. Canopy and dominant trees can be detected more easily due to their elevated and apparent positions, which can be readily observed from an aerial view, but subcanopy or suppressed trees are always impeded by the upper tree canopy, providing limited visible information. Moreover, the captured RGB images of tree crowns exhibit inconsistent differences in appearance due to the varying capture positions between the UAV-loaded cameras and the target trees. In hyperspectral images, different tree species or vegetation types may exhibit similar reflection properties, while the same tree species may produce varying spectral signatures due to variations in solar illumination conditions, atmospheric turbidity, overlapping crowns and varying phenological periods of the target tree crowns. All these variable factors likely increase the complexity of deep learning networks for accurately identifying specific forests because of their complex surrounding environments.

4. Deep learning methods for forest point clouds

4.1. Point cloud data and acquisition equipment

Laser scanning, also known as LiDAR, which encompasses terrestrial,

mobile terrestrial and airborne laser scanning configurations, can capture spatial attributes and complex forest structures (Zhang et al., 2021b). Terrestrial laser scanning provides high angular and scanning resolutions for each scan (Zhu et al., 2021), albeit with a limited coverage range, thus requiring additional scanning positions; moreover, challenges associated with scanning hard-to-reach areas are encountered, and labour-intensive efforts are needed to overcome such issues. In contrast, mobile terrestrial LiDAR offers flexibility and a relatively high resolution, particularly for capturing features of the lower-middle parts of forests. However, their applications are restricted to drivable areas, may produce motion artefacts, and require data compensation for the uppermost section of the target. Airborne laser scanning offers expansive coverage, high operational efficiency, and the ability to scan high-altitude and difficult-to-access regions. Nevertheless, they have lower resolutions, are susceptible to weather conditions, and may miss fine details. To address the unique challenges posed by the irregular and permutation-invariant structure of tree point cloud data, deep learning networks are crucial for effectively extracting distinctive features for forest analysis. Point cloud networks have emerged as a new way to process and utilize LiDAR data for forest research, providing a solid foundation for the development and study of forests with deep learning techniques (Guo et al., 2020). Table 3 lists the specifications of several common laser scanners.

4.2. Specific applications of point cloud-related deep learning methods in forest research

A comprehensive overview of recent advancements in deep learning technology for extracting forest stand parameters from point clouds is provided in Table 4.

In research involving point cloud data, many innovative networks have emerged and been implemented in practical applications. To process forest point cloud data via deep learning, researchers require the selection of suitable networks and methods. The different categories of

Table 3

Characteristics of the laser scanners commonly utilized in forest remote sensing.

Type	Equipment	Wavelength/ Frequency	FOV Angle	Spot Size
Ground-based	Leica HDS6100	785 nm/79 MHz	H: 360°; V: 310°	3 mm + 0.22 mard
	Optech Ilris HD	1535 nm/10 kHz	H: 40°; V: 40°	3 mm + 0.22 mard
	Velodyne HDL-32E	905 nm/10 kHz	H: 360°; V: 40°	3 mm + 0.22 mard
	UTM-30LX 2D	905 nm/40 kHz	H: 270°; V: 4°	3 mm + 0.22 mard
	Faro Focus 3D 120	1500 nm/976 kHz	H: 360°; V: 300°	3 mm + 0.22 mard
Mobile terrestrial	Leica Scanstation C10	532 nm/50 kHz	H: 360°; V: 270°	4.5 mm + 0.22 mard
	Optech Lynx	1064 nm/1200 kHz	H: 360°	3 mm + 0.22 mard
	Trimble Mx8	1550 nm/400 kHz	H: 360°; V: 58°	3 mm + 0.22 mard
	VLP-16 LiDAR	903 nm/10 kHz	H: 360°; V: 30°	3 mm + 0.22 mard
	Leica ALS60	1064 nm/150 kHz	H: 75°	3 mm + 0.22 mard
Airborne	RIEGL VQ 1560i	532 nm/2 MHz	H: 360°; V: 58°	3 mm + 0.22 mard
	RIEGL LMS-Q680i	1064 nm/2 MHz	H: 60°	3 mm + 0.22 mard
	RIEGL mini VUX-1	1064 nm/400 kHz	H: 360°	3 mm + 0.22 mard
		905 nm/100 kHz	H: 360°	3 mm + 0.22 mard
				0.22 mard

Note: H (horizontal angle), V (vertical angle).

deep learning for forest applications are described below, and

Table 4

Different deep learning models used in forest applications from point cloud data.

Processing strategy		Device	Forest Type	Eva		Organ	Objective	Reference	Locations
				OA	F				
Features extraction using MLPs	LayerNet	U	MF	92.5	—	W	FPI	(Liu et al., 2021)	Forest park, Chengde, China
	PDE-Net	M	MF	—	95.0	W	TOM	(Luo et al., 2021)	Columbus, USA. Beijing, China
	F-PointNet	M	DF	—	—	B	TOM	(Yu et al., 2022)	Orchard, Nanjing, China
Local feature extraction	PointNLM	M	DF	96.3	92.7	W	FPI	(Chen et al., 2021a)	Urban area, France, Germany, Europe
	PointCNN	U	MF	87.0	94.0	W	FPI	(Hell et al., 2022)	Forest Park, Bavarian, Germany
	PointConv-based	M	DF	93.2	—	B, L	FPI	(Ding et al., 2022)	Plantation, Danzhou, China
Graph-based method	Cylindrical	M	DF	84.0	—	W	FPI	(Jiang et al., 2023)	Urban area, Wuhan, Shanghai, China
	Voxelized-based	U	MF	—	86.0	C	TOM	(Chen et al., 2021b)	Forest park, Chizhou, China
	Superpoint	S	MF	88.8	—	B, L	FPI	(Wang, 2020)	Coniferous forest, Evo, Finland
	FR-GCNet	U	MF	93.4	78.6	W	FPI	(Zhao et al., 2021)	Town, Whitchurch, Canada
	DGCNN	M	MF	—	96.5	W	FPI	(Liu et al., 2022)	Forest, Greater Khingan, Huailai,Forest farm, Gaofeng, China
	SPGNet	S	DF	90.0	—	B	TOM	(Ma et al., 2021a)	Xingkai Farm, Yiwu, China
	Graph-cut	U	MF	—	73.0	T	TOM	(Dersch et al., 2021)	Mixed forest, Styria, Austria
Projection onto the 2D plane	Silvi-Net	U	MF	96.1	—	W	FPI	(Briechle et al., 2021)	Forest Park, Bavarian, Germany
	Deep CNN	U	MF	82.0	—	W	FPI	(Hamraz et al., 2019)	Robinson Forest, Kentucky, USA
	LeNet5	S	MF	86.0	—	W	FPI	(Seidel et al., 2021)	Park, Germany. Oregon, USA
	YOLO	U	MF	81.4	—	C	FPI	(Sun et al., 2022)	Campus, Nanjing, China
	Faster R-CNN	M	DF	—	98.0	T	TOM	(Wang et al., 2019)	Plantation, Danzhou, China
Fusion with machine learning	PointNet	M	MF	90.0	92.08	T	TOM	(Ning et al., 2023)	Urban area, Paris, Lille, France
	Mask R-CNN	S	DF	89.7	82.5	L	TOM	(Liu et al., 2020a)	Laboratory, Nanjing, China
	PointNet++	M	MD	80.7	—	W	FRA	(Tang et al., 2023)	Urban area, Fuzhou, China
	DNN	S	MF	97.7	—	W	FED	(Rihan et al., 2023)	Hilly region, Western Himalayas

Note: U (UAV-loaded), M (mobile terrestrial), S (static), CF (coniferous forest), DF (deciduous forest), MF (mixed forest), OA (overall accuracy), F (F score), B (branch), T (trunk), L (leaf), C (canopy), and W (whole tree). OA/F is directly derived from the corresponding literature, but in actual experiments, the magnitude may vary due to several factors including the collected point density, training dataset, number of categories, forest type and occlusion effect.

representative examples are shown in Fig. 5.

(1) Feature extraction using a multilayer perceptron (MLP)

Forest point cloud data have unique characteristics (e.g., unordered permutation) that make processing such data using traditional image processing or geometric methods challenging because the data are distributed in space. However, MLP neural networks directly manage the raw data, and clouds learn complex nonlinear relationships between the point attributes, enabling scalable and adaptive features to be extracted. The MLP neural network is mathematically expressed as follows.

$$X' = f(WX + B) \quad (1)$$

where $W \in R^{n \times n}$, $X \in R^{n \times m}$, n denotes the number of input points and m denotes the corresponding features of the scanned points, i.e., colour information, point position, normal vector, and sparse tensor (Wang et al., 2023b). Moreover, $B \in R^{n \times m}$ is the bias matrix, and f is the activation function, which encompasses multiple options, e.g., sigmoid, tanh, ReLU, maxout and ELU.

Typical deep learning networks using MLPs aim to address LiDAR data point-by-point. For instance, PointNet (Chen et al., 2021b) and PointNet++ (Qi et al., 2017) perform better when operating on tree parameter retrieval at the individual tree scale than at the organ level because of the crisscrossing wood-leaf structure and the limitations in localizing intraorgan features, as indicated by their mathematical expression. Derivative networks, such as F-PointNet (Yu et al., 2022), fuse features extracted from point clouds and images with MLPs to implement pomegranate segmentation, achieving a precision of 82.6 %; grouped tree point clouds are hierarchically posted on LayerNet (Liu et al., 2021), which employs fully connected networks and an MLP

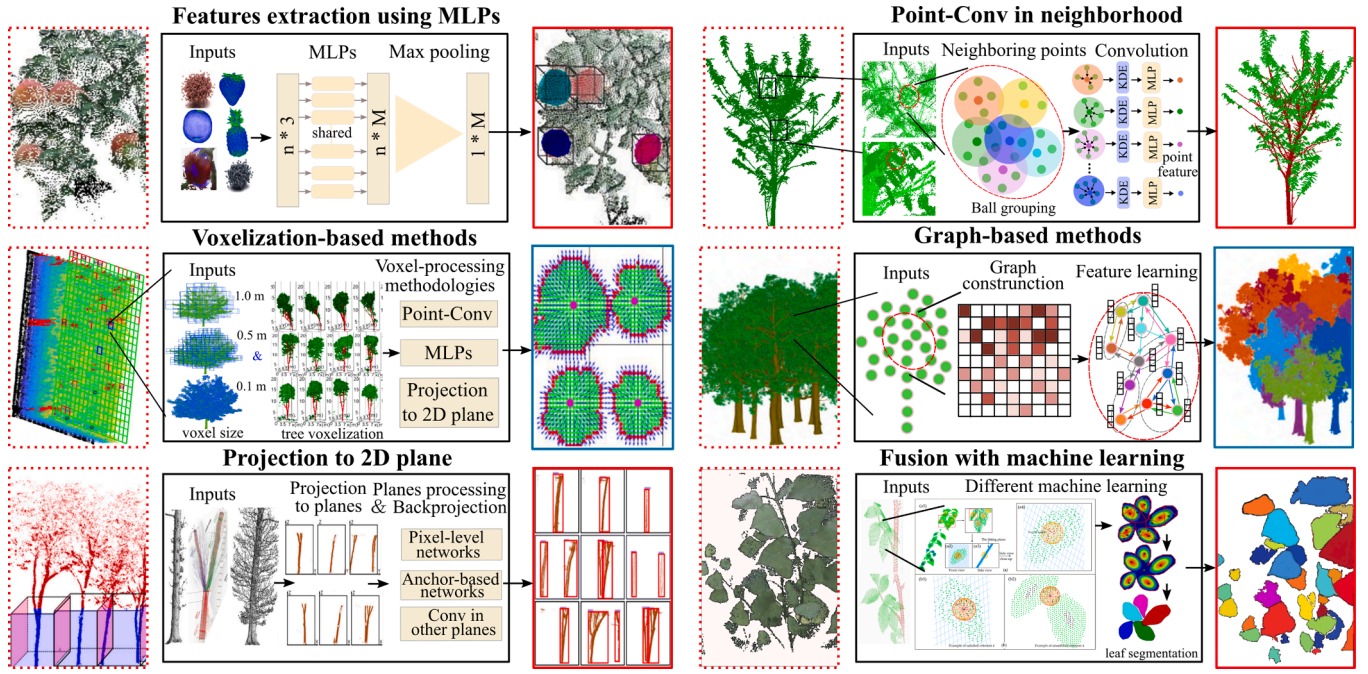


Fig. 5. Illustration of the typical deep learning methods for point cloud processing in forest applications, with the primary concepts, inputs and outputs of each method shown.

decoding architecture to obtain tree structural features for birch and larch classification, achieving a precision of 92.5 %. Additionally, the pointwise direction embedding deep network (PDE-Net) (Luo et al., 2021) predicts the direction vectors between each point and the selected tree centres for multitree cluster separation, revealing divergent cogitation for individual street tree extraction with an exceptional precision of 96.0 %.

The MLP serves as the fundamental structure in a deep learning network. Although it provides a nonlinear modelling capability and functions as a universal approximator, there are certain aspects that require additional improvement and reinforcement. The first aspect is the susceptibility to local minima, which embodies situations in which the network settles for suboptimal solutions regarding classification or recognition results in the training process; the second aspect is the deficiency of feature scaling calculations, which require proper standardization or subdivision preprocessing of the input data to effectively mine local semantic features of the target trees.

(2) Concepts of local feature extraction

Among local feature extraction models for forest point clouds, various approaches that utilize distinct geometrical features of branches, fruits and leaves, known as “point neighbourhood-based convolutional neural networks”, have shown promising results. These methods conjointly leverage the inherent positional relationships between the target and neighbouring points to extract spatial association features and semantic features of each point and capture local geometrical information, such as multivariate distances, spatial malleability, and principle curvatures and directions for detecting different forest component organizations and stand organ characteristics. The shapes of these local neighbourhoods typically differ and include a variety of modalities, such as spherical neighbourhoods and voxel block neighbourhoods.

Local feature extraction-based neural networks are mathematically formulated as follows:

$$(f * g)(p_i) = \frac{1}{hN_{p_i}} \sum_{p_{ij} \in N_{p_i}} f(p_{ij}) \cdot \varphi\left(\frac{\nu(p_i - p_{ij})}{h}\right) \quad (2)$$

where $\nu(p_i - p_{ij})$ denotes the relationship between the centre point p_i and its neighbouring points p_{ij} , i.e., ν comprises be the L1 and L2 norms, φ is the kernel function, e.g., parabolic (Epanechnikov), tricubic or Gaussian kernel functions (Altman, 1992), and h denotes the kernel radius or bandwidth. Moreover, $f(p_{ij})$ is the input feature of the neighbourhood region surrounding p_i . Generally, Eq. (2) serves as a local feature extractor that fully explores local geometric or positional information via successive convolution operations or self-attention mechanisms (Ding et al., 2022).

To efficiently perform neighbourhood searches, various data structures, such as 3D grids, octrees, Kd-trees, and binary space partitioning (BSP) trees, are often employed to organize point cloud data. These structures effectively partition and index the point cloud, enabling faster retrieval of neighbouring points. Hence, various shape neighbourhoods, i.e., spherical, cylindrical and cuboid neighbourhoods based on different distance metrics, can be employed. For example, a deep learning network containing the optimized PointConv module was combined with kernel density estimation of local spherical point features of rubber trees for wood-leaf separation (Ding et al., 2022). Wu et al. (2020) extracted geometrical traits in spherical neighbourhoods and obtained laser return intensities for leaf-wood separation, with average accuracy levels of 95.6 % for broadleaf forests and 94.2 % for coniferous forests. Another application entailed the employment of a convolution network based on cylindrical neighbourhoods to segment individual trees from urban ground-mobile point clouds, with an accuracy exceeding 84 % (Jiang et al., 2023). Chen et al., (2021b) employed the voxelization strategy and PointNet to dissect forest point clouds and analyse the entirety of forest morphological characteristics, which enabled thorough voxel-scale tree recognition and segmentation of individual canopies; Chen et al., (2021a) proposed the PointNLM network coupled with a supervoxelization subdivision strategy for urban vegetation segmentation, with a segmentation precision of 96.3 %.

Local feature extraction models can extract the geometric relationships and semantic information associated with edges, corners, and planar or column regions within tree crowns. However, the network performance always decreases due to the unique morphologies of interlaced branches and interlocking leaf clusters, and determining

appropriate values for many crucial parameters, e.g., the designated neighbourhood range and kernel radius, can be challenging due to the varying leaf area density and scanning patterns. Hence, scholars expect deep learning techniques to continue to evolve to enhance the discriminative power when applied to the study of trees with complex geometric details.

(3) Graph neural network for forest point cloud processing

Organs of forest or individual stands present topological connectivity and growth affiliation relationships, which are mapped by the points and their neighbourhoods with biological organizations, e.g., first-order branches emanating from the trunk, secondary branches extending from the primary ones, and leaves growing from the branches. Pioneer studies (Dersch et al., 2021) adopted graph-related algorithms from computer vision theory for trunk positioning, crown edge delineation and wood-leaf component segmentation. More recently, graph-based networks have been introduced for processing forest point clouds (Wang et al., 2023c).

The core of the graph-based network can be mathematically represented as $G = (V, E, W)$, where V is a finite set containing k neighbouring points and $p_{ij} \in N_{p_i}$ in the local spherical domain centred on point p_i . Moreover, E denotes the set of edges between p_i and p_{ij} , and the symmetric matrix $W \in \mathbb{N}^{k \times k}$ denotes the relationship weight corresponding to each edge between centre points p_i and p_{ij} . Element $w_{ij} \in W$ of matrix W can be expressed as:

$$w_{ij} = \begin{cases} \exp(-a_{ij}/(\sigma^2 + \varepsilon)), & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases} \quad (3)$$

where σ is the adaptive normalization parameter, ε is a small positive number, and $a_{ij} = \|p_i - p_{ij}\|^2$. Moreover, element d_{ij} of the corresponding degree matrix D for each centre point p_i can be calculated as

$$d_{ij} = \begin{cases} \sum_{l=1}^k w_{il}, & \text{if } i = l \\ 0, & \text{if } i \neq l \end{cases}$$

Then, a symmetrically normalized Laplacian matrix $L = I - D^{-1/2} W D^{-1/2}$ can be obtained, where I is the identity matrix. Consequently, $L \in \mathbb{N}^{k \times k}$ incorporates the neighbouring point features $F \in \mathbb{N}^{k \times m}$ as an important component for feature propagation.

Two types of graph convolution operations exist. (1) The node features and edges can be directly convolved by weighting and summing the features of neighbouring points based on their attributes and positions (Zhang et al., 2020). (2) Spectral domain-based convolution employs the graph Laplacian matrix and Fourier transform to convert features into the frequency domain with learnable filters and transforms them back to the spatial domain for further feature propagation (Shuman et al., 2013).

In contrast to local feature extraction, graph neural networks present data as a graph structure with complex relationships and interdependencies between objects; these networks mainly focus on growth parameter retrieval and organ property measurements at the individual tree level. Several recent networks, e.g., FR-GCNet, which was proposed by Zhao et al. (2021) and incorporates an edge attention and convolution module for each node to aggregate features, implemented the classification task in large-scale scenes containing trees with a classification accuracy of up to 93.55 %. SPGNet combines superpoint graph embedding with geometrically homogeneous partitioning for contextual segmentation and was successfully applied to segment different-order branches of jujube trees, achieving a segmentation rate greater than 85.5 % (Ma et al., 2021a).

Graph-based convolution methods are promising for processing point clouds but are limited by their high computational costs and low adaptability to data of various qualities. These methods are sensitive to the chosen graph construction and graph pooling strategies, making it challenging for the graph-based convolution method to achieve consistent and accurate results across different datasets. Hence, developing a practical and theoretically justified graph-based framework for

balancing intrinsic and extrinsic considerations requires insights from graph theory and specialized experience in tree geometry processing. Nevertheless, despite the aforementioned limitations, these methods hold immense promise and broaden the perspectives for analysing the relevance of forest point clouds, which is worthy of further investigation.

4.3. Other processing strategies for forest point clouds

(1) Projection onto the 2D plane

Multiview deep neural networks (DNNs) (Su et al., 2015) first project point clouds of 3D objects onto multiple view planes, extract and merge visual features through CNNs, and classify diverse stereoscopic objects. Accordingly, numerous strategies have been used to investigate forest point clouds, focusing on forest species and status detection at the individual scale. For example, Silvi-Net (Briechele et al., 2021) rendered multiple side views grounded on a single tree point cloud and combined those views with a dual-ResNet18 for tree categorization and dead-tree discrimination. Ground-based forest point clouds were employed to construct projection images from ten different viewing angles, which were then converted to many images for tree species recognition via LeNet5 (Seidel et al., 2021). Moreover, Wang et al. (2019) introduced a voxelization strategy for rubber tree point clouds to generate projected depth images from front and side perspectives in tandem with Faster R-CNN for main trunk detection and individual tree crown segmentation. Sun et al. (2022) converted airborne LiDAR data from forest canopies into height maps. Generative adversarial networks (GANs) were subsequently integrated with YOLO for individual tree crown segmentation and parameter retrieval. Additionally, Xu constructed a deep learning network to learn the optimal projection angle for terrestrial laser points of trees with the aim of conducting wood and leaf point segmentation (Xu et al., 2023).

The projection of forest point clouds onto multiple view planes leverages advanced image processing techniques to analyse 3D point clouds, but this strategy results in intrinsic spatial information loss and geometrical morphology distortion during the projection process. Moreover, 2D projection images are deficient in expressing depth information and sensitivity to the projection angles and orientations of multiple views. To overcome the restrictive portrayal of the phenotypic traits of tree crowns, auxiliary approaches are preferred as conjoint methods for capturing complex tree structures and environment-driven tree body deformation.

(2) Fusion of deep learning and machine learning concepts

Deep learning methods often involve machine learning concepts to reprocess forest point clouds for further mining of other advantages. Certain typical works, such as semantic segmentation, used PointNet for road tree point extraction, and a distance-based cluster algorithm was utilized to distinguish the intermediate and boundary points of tree crowns for individual tree segmentation (Ning et al., 2023). Liu et al. (2020a) combined a Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method based on the manifold distance with Mask R-CNN to separate overlapping poplar seedling leaves, obtaining an accuracy of 89.7 %. Similarly, a DBSCAN method, along with PointNet++, was employed to process ground mobile laser point clouds for extracting street trees and green belts, and a virtual camera was used to estimate the green view index in urban scenarios (Tang et al., 2023). Rihan et al. (2023) conducted fire susceptibility mapping of Himalayan forests by fusing logistic regression and random forest models with DNNs to assess parameter sensitivity and uncertainty, achieving an accuracy of 97.7 %. By combining deep learning with traditional machine learning methods, scholars can create hybrid models that leverage the best of both methodologies with greater flexibility and extensibility.

5. Multimodal deep learning neural network for forest studies

Multimodal deep learning networks aim to train deep learning

models to process and find relationships between different types of data (modalities), e.g., images, point clouds, time span data, and text. Multimodal networks exhibit immense potential for completing complex tasks by effectively fusing information acquired from diverse sources (Shen et al., 2020), optimizing comprehensive feature presentation and compensating for the constraints of information derived from a single data source.

For instance, tree height and crown area information was extracted from airborne LiDAR data, which served as guiding region division information for preprocessing spectral (Mäyrä et al., 2021) and RGB images (Weinstein et al., 2020) to generate training samples for tree species classification. In addition, multimodal cascaded fusion networks, with the input of stacked joint features containing a height map stemming from scanned points and spectral or RGB information from captured images, are gradually emerging for the purpose of achieving certain tasks, e.g., vegetation species classification (Gonzalez-Perez et al., 2022), ground target classification (Wang et al., 2023d) and forest canopy fraction estimation (Li et al., 2022). The results show that the combination of LiDAR-based digital surface models (DSMs) and aerial images provides highly accurate morphological and elevation features associated with tree spectral textures, effectively improving the accuracy by approximately 8 % (Chen et al., 2023b). Moreover, deep learning networks can be used for superresolution tasks, i.e., learning mapping relationships from low-resolution images and correlated internal relationships to produce high-quality interpolated results. Hence, scholars have employed Sentinel multiband images as auxiliary information to compensate for the discrete height information of Global Ecosystem Dynamics Investigation (GEDI) data and have adopted deep learning models to derive high-resolution canopy height maps of Europe (Schwartz et al., 2024).

In contrast, the joint input of multitemporal data into deep learning networks for future prediction and impending phenomena analysis is gradually emerging. These methodologies allow models to manage sequential data with temporal dependencies more effectively, thereby providing a more comprehensive understanding of the correlations and mutual influences in chronological sequences to guide decision making and result refinement. Successful examples can be summarized as follows: the convolutional gated recurrent unit (ConvGRU) module was utilized to resolve irregular time series data of aerial images for monitoring the changes in Swiss Alps forests over the past 70 years in response to climate and land use changes (Nguyen et al., 2024). A modified transformer network with time-series Sentinel-2 multispectral imagery from 2018 to 2020 was designed to explore the key spectral-temporal features and phenological stages of tree species mapping in two representative plantation forests (Huang et al., 2023). Temporally continuous aerial images and Sentinel satellite data were fused into U-net networks to map the monthly changes in forests under the influence of timber harvesting from 2017 to 2019 (Zhao et al., 2022).

Forest ecosystems are inherently heterogeneous, with varying conditions and species compositions across different regions and phenological spans. Multimodal deep learning networks could address this heterogeneity by learning advanced semantic information from a diverse set of spatial-temporal input data but still face several drawbacks at present. First, simultaneously merging multimodal information with distinct properties and structures must consider data registration for inconsistent character fusion and data equilibration to reach a weight extraction balance between each dataset, but this issue remains underexplored. Second, multimodal deep learning networks may suffer from data bias issues. Networks generally direct disproportionate attention towards certain types of data, leading to reduced performance in resolving diverse impact modalities. Hence, certain strategies, e.g., regularization (Zhu et al., 2022) or data augmentation, to mitigate unfavourable impacts are important.

6. Available datasets for forest research

The available dataset benchmarks for training deep learning models are essential in forest research for obtaining optimal coefficient adjustment and regression results. Such datasets are manually collected by investigators or created using existing remote sensing data, helping data-demanding methods reach their full potential and learn rich forest information. Hence, many altruistic academics and institutions are willing to disseminate valuable forest datasets, thus substantially propelling the domain forwards. Table 5 summarizes some published forest datasets and publicly available datasets from many research institutes, with the intention of broadening the impacts of these datasets and encouraging more researchers to utilize and conduct data sharing to facilitate the dissemination of knowledge and collaboration.

The forest datasets included growth equations for various tree species and forest data types. For instance, the PANGAEA database contains many in situ destructive sampling measurements of forest biomass encompassing different tree organs and coarse woody debris in Eurasia (Schepaschenko et al., 2017) and UAV orthoimages of trees and shrubs in Siberia, with individually labelled information such as height, species, and crown diameter (van Geffen et al., 2021). Many satellite datasets for estimating global land cover and land use variance have been established; these datasets include information on the extent of forest spread and seasonal disturbance monitoring at the regional scale. For example, the DeepGlobe dataset (Demir et al., 2018), a CVPR image analysis challenge based on the DigitalGlobe satellite, contains many forest coverage images; additionally, EuroSAT (Helber et al., 2019), which stems from Sentinel-2 satellite images, was proposed by the Technische Universität Kaiserslautern. Global Land Analysis & Discovery (GLAD) (Potapov et al., 2022), which was derived from the Landsat satellite, was released by Maryland University; WH-MAVS (Yuan et al., 2022), a multitemporal and georeferenced scene classification dataset derived from GF-2, was dispensed by the Wuhan Sigma Laboratory; TreeSatAI (Ahlsvede et al., 2023), a benchmark forest archive modelled on Lower Saxony Forest, was developed by the Technische Universität Berlin; and TiSeLaC, which offers multispectral satellite data from nine classes, including urban areas, grasslands, crops and forests, etc., for categorization tasks, was organized by Ienco et al (2019). The NeonTreeCrowns dataset (Weinstein et al., 2020) consists of 100 million tree images acquired by an airborne observation platform from 37 geographic sites across the United States, with many tree crown properties provided. Regarding laser points of forests, the University of Yaoundé I supplied the aboveground biomass of the woody components of tropical trees accompanied by terrestrial point clouds with wood-leaf tags for semi-deciduous forests in eastern Cameroon (Momo Takoudjou et al., 2018); similarly, Wan et al. (2021) manually classified woody components of white birch, Dahurian larch, and Chinese scholar trees obtained from terrestrial laser scanning (TLS) point clouds, which could also serve as training information for wood-leaf classification tasks. Several researchers (Graves and Marconi, 2020) of the IDTreeS group provided a fused dataset of images and LiDAR data for professional competitions. The competition aims to use remote sensing data to quantify the locations, sizes, and species identities of millions of trees and to test how well existing deep learning methods generalize to diverse forests. Despite the existence of relevant forest datasets containing large areas of forests and individual trees, forest researchers generally encounter barriers and restrictions in data collection due to data ownership, legal restrictions, privacy issues and proprietary sources, making many forest datasets unavailable for public research purposes.

In contrast, computer science and remote sensing researchers are more willing to share their datasets, some of which contain individual tree data and forest scenes. For example, the ImageNet dataset (Deng et al., 2009) contains numerous annotated images with location and category tags. This dataset contains many partial trees and crop images from various viewing angles, and they can be used as augmented training samples for tree detection and classification. Similarly, the

Table 5

Summary of shared and publicly available datasets containing forest data.

Agency	Type	Website	Attribution	Source
CVPR2018	Satellite Image	deepglobe.org/	Forest cover	(Demir et al., 2018)
TU Kaiserslautern		zenodo.org/records/7711810#.ZAm3k-zMKEA	Forest cover	(Helber et al., 2019)
Maryland University		glad.umd.edu/dataset/GLCLUC2020	Forest extent	(Potapov et al., 2022)
Wuhan University		sigma.whu.edu.cn/resource.php	Forest cover	(Yuan et al., 2022)
TU Berlin		zenodo.org/records/6780578	Tree species classification	(Ahlsweide et al., 2023)
INRAE	Camera Image	sites.google.com/site/dinoienco/tiselac-time-series-land-cover-classification-challenge	Forest mapping	(Ienco et al., 2019)
AWI		doi.pangaea.de/10.1594/PANGAEA.932821	Tree top view	(van Geffen et al., 2021)
NeonTreeCrowns		https://zenodo.org/records/3765872	Tree crown	(Weinstein et al., 2020)
Princeton University		www.image-net.org/download	Tree	(Deng et al., 2009)
Berkeley University		eeecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/	Hand-labelled Tree	(Arbeláez et al., 2011)
Microsoft		cocodataset.org/	Tree	(Lin et al., 2014)
Google		opensource.google/projects/open-images-dataset	Tree	(Kuznetsova et al., 2020)
KIT and TTIC		www.cvlibs.net/datasets/kitti/raw_data.php	Street tree	(Geiger et al., 2013)
FGVC		https://www.kaggle.com/c/plant-pathology-2021-fgvc8	Foliar disease	(Belongie et al., 2021)
Kaggle Community		www.kaggle.com/datasets/mohnishsaiprasad/forest-fire-images/data	Forest fire	(Xian and Nugroho, 2022)
Yaoundé I University	Point Cloud	doi.org/10.5061/dryad.10hq7	Leaf-wood	(Momo Takoudjou et al., 2018)
CRSRI		datadryad.org/stash/dataset/https://doi.org/10.5061/dryad.rfj6q5799	Leaf-wood	(Wan et al., 2021)
Osnabrück University		kos.informatik.uni-osnabrueck.de/3Dscans/	City scene	(Nüchter et al., 2010)
ETH Zürich		www.semantic3d.net/	Urban tree	(Hackel et al., 2017)
CMU		www.cs.cmu.edu/~vmr/datasets/oakland_3d/cvpr09/doc/	Campus tree	(Munoz et al., 2009)
Sydney University	Fusion Data	https://www.acfr.usyd.edu.au/papers/SydneyUrbanObjectsDataset.shtml	Urban tree	(De Deuge et al., 2013)
Mines ParisTech		npm3d.fr/paris-lille-3d	City scene	(Roynard et al., 2018)
Wuhan University		3s.whu.edu.cn/ybs/en/benchmark.htm	Forest & park	(Dong et al., 2020)
IDTReeS		https://zenodo.org/records/3934932	Tree	(Graves and Marconi, 2020)

Note: TU Kaiserslautern (Technische Universität Kaiserslautern), TU Berlin (Technische Universität Berlin), KIT (Karlsruhe Institute of Technology), NEON (National Ecological Observation Network), CMU (Carnegie Mellon University), and ETH Zürich (Eidgenössische Technische Hochschule Zürich).

BSBD (Arbeláez et al., 2011), Microsoft COCO (Lin et al., 2014) and Google Open image (Kuznetsova et al., 2020) datasets contain many tree images for object detection and segmentation tasks. The KITTI vision benchmark suite (Geiger et al., 2013) was designed for advanced autonomous driving tasks, including many urban tree images. Many data scientists and machine learning engineers have published datasets on the online platform Kaggle, which includes forest datasets, e.g., foliage disease (Belongie et al., 2021) and forest fire (Xian and Nugroho, 2022) images.

Many publicly available forest point cloud datasets have also been released. For the TLS data, Osnabrück University (Nüchter et al., 2010) shared a repository of TLS data with ground truth object category labels, and the dataset contains images of many city scenes with trees. ETH Zürich (Hackel et al., 2017) published benchmarks for a large labelled TLS dataset, which covers a wide range of diverse natural scenes and urban forest scenarios, such as street trees, villages and castle parks, with high-density representation and fine detail mapping. Additionally, in terms of mobile terrestrial LiDAR (MT-LiDAR) data, Carnegie Mellon University (CMU) shared an MT-LiDAR dataset around the CMU campus in Oakland, including some city tree data (Munoz et al., 2009). Sydney University (De Deuge et al., 2013) shared an MT-LiDAR dataset that contained a variety of common urban road objects in the CBD of Sydney. Similarly, the Paris-Lille-3D dataset (Roynard et al., 2018) was produced by a mobile laser scanning system in two different cities in France and contains more than 50 different class labels, including urban trees, pedestrians, and traffic signs. Additionally, Wuhan University (Dong et al., 2020) presented a point cloud dataset acquired by multiple scanning patterns consisting of more than 1740 million 3D points collected from 11 different environments, e.g., mountains, forests, parks, campuses and riverbanks. All these data can be utilized for benchmark generation through hand-annotation processing.

Sharing these datasets not only provides researchers with valuable data resources but also facilitates a widening sphere of influence for forest research, promoting an environment conducive to sustainable innovation. Overall, by aggregating available forest datasets and expanding their reach, we can promote sustainable forest research and

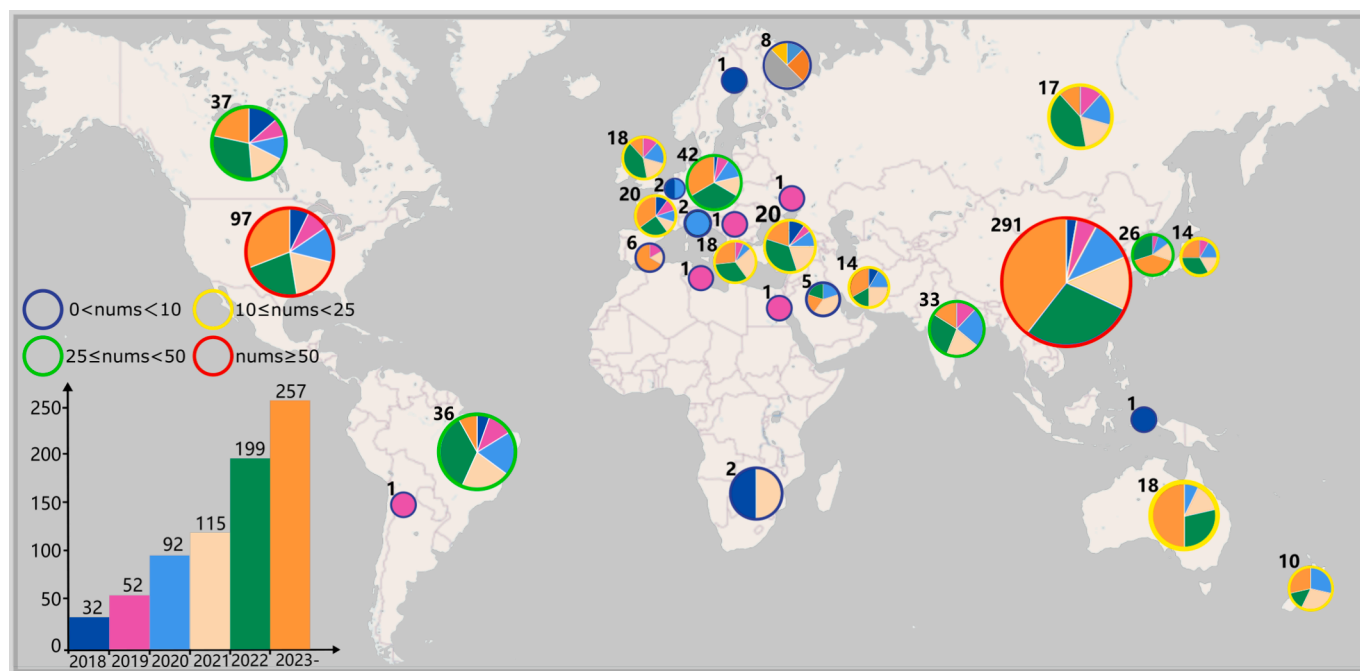
foster global partnerships and innovation. Such efforts will provide forest fields with more opportunities to address key issues and achieve sustainable forest management.

7. Global perspective of the literature distribution of forest research using deep learning techniques

In the past five years, research institutes and academics worldwide have been actively investigating further deep learning forest applications. Relevant research inputs and outcomes may vary between countries and regions due to differences in economic sector strengths, geomorphic and meteorological features, forest species heterogeneity and cultivation strategies. To investigate and derive insights into the research activities related to the applications of deep learning techniques in forests in diverse regions, we collated published studies by using related keywords, such as ‘deep learning, forest and remote sensing’, to search the ‘Web of Science’ platform. The resulting statistical graphs, which depict the global distribution of related topics across five vegetated continents (excluding Antarctica) and encompass temperate, subtropical, and tropical regions with evident geographical differences, are shown in Fig. 6.

The research shown in Fig. 6 encompasses comprehensive applications involving forests or individual trees, with the study locations primarily concentrated in the Northern Hemisphere and middle latitudes. Overall, the number of articles related to deep learning in the field of forests from 2018 to 2023 shows an upward trend. The number of articles gradually increased from 32 in 2018 to 115 in 2021, indicating steady growth followed by a significant increase to 199 in 2022. This trend will continue into 2023, with a total of 257 articles published by the end of 2023. The cumulative sum of articles across the whole period was 747. The lower left rectangular graph visually illustrates these trends.

Among the specific countries, China encompasses 291 articles, accounting for almost 43.4 % of the total studies; the United States exhibits 97 articles, accounting for 14.5 % of the total studies; and the other countries have fewer than 50 articles each. In terms of the geographical



work, herein, a general summation of the gaps in existing deep learning techniques is provided:

(1) Challenging autogeneration of training samples and collection of new forest attributes by novel sensors. To improve the performance of deep learning models for forest applications, they must be trained on large amounts of labelled data, which can be expensive and time consuming for various forest types. The GAN algorithm (Shu et al., 2019) is promising because of its innovative adversarial process and ability to autonomously generate vast volumes of high-quality training samples, including forest images and point clouds. Concurrently, the

application of large language models (LLMs), along with diffusion models (Rombach et al., 2022) and variational autoencoders (VAEs) (Liu et al., 2020b), is exhibiting constant development and renovation. Text-to-image diffusion models, including text, image, and additional data types, facilitate multimodal data integration to realize the transition from text to image or point clouds. During the training phase, VAEs and diffusion models could be exposed to a large amount of forest image data with noise, and they attempt to understand the effects of noise on the appearance and structure of forest images in pursuit of grasping the association between noise and images. During the generation phase, the

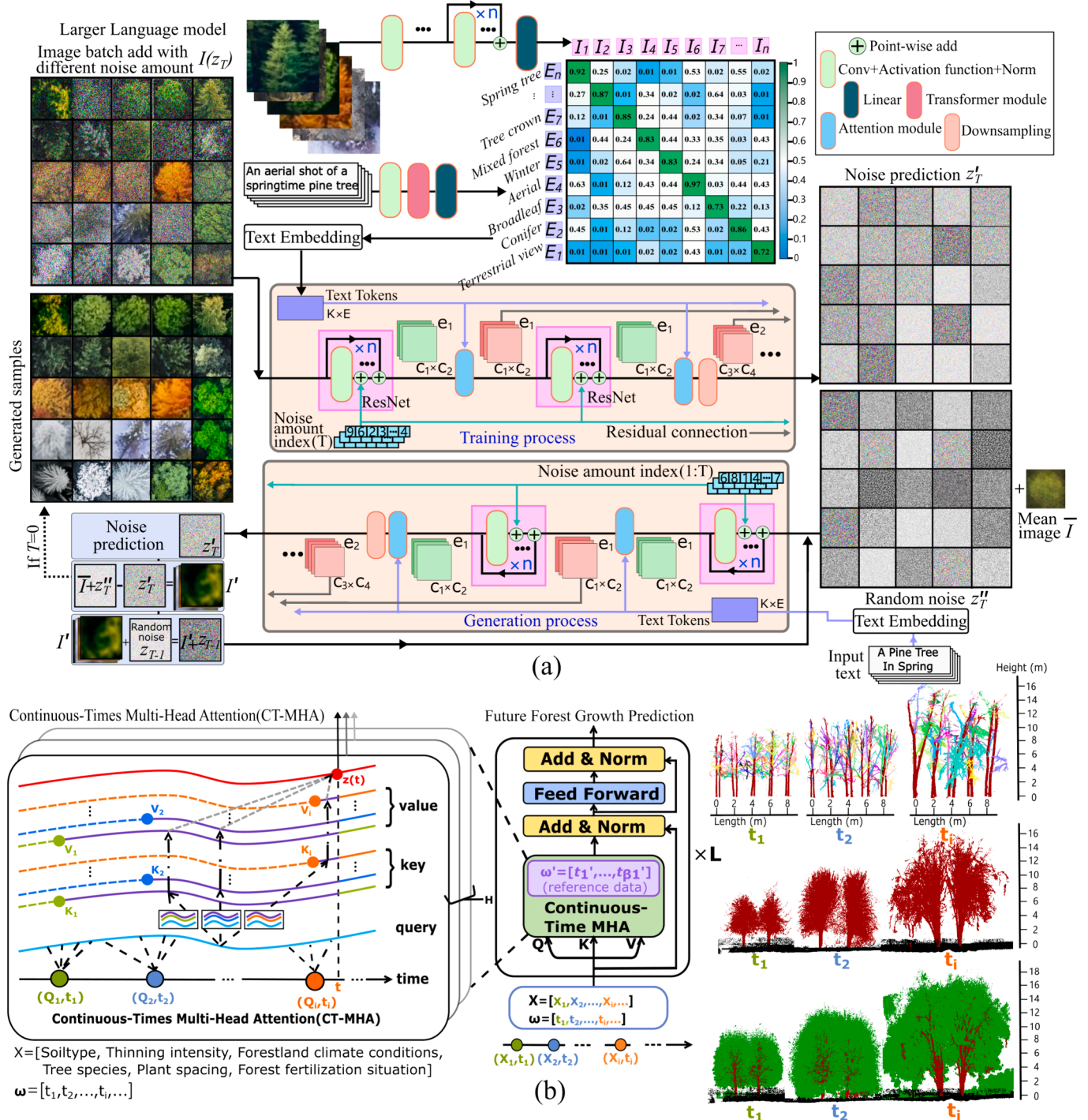


Fig. 7. Two anticipated instances of deep learning application in forest studies. (a) Diffusion models, in tandem with a large language model, can be used to forecast the association between noise and forest images to generate innovative training samples. (b) A cascaded recurrent neural network integrating past temporal forest measurement data can be employed for forecasting future tree growth parameters.

adopted models initialize an average mean forest image with an additive noise tensor and then gradually remove noise, allowing diverse forest training samples to be generated. A fundamental network case is shown in Fig. 7a. All of the above latest models could possibly overcome the barriers of training data accessibility and diversity, allowing forest researchers to overcome the traditional bottlenecks associated with training data scarcity and laborious manual labelling efforts. In contrast, the advent of advanced remote sensing sensors has increased the ability to acquire novel features in the forest domain, delivering numerous complex, comprehensive datasets that reflect the holistic and multifaceted characteristics of target trees, enabling the identification of additional biochemical and physiological processes in plants and thereby revealing a new potential for data-driven decision-making across a range of deep learning networks. For example, the chlorophyll fluorescence sensor (Damm et al., 2022) can rapidly and accurately measure plant chlorophyll levels, photosynthetic physiology, and pigment concentrations; ground-penetrating radar (Tong et al., 2020) emits electromagnetic waves that can penetrate soil and encounter roots undergoing reflection and refraction, which can capture and process the reflected signals to generate root system images in pursuit of gaining insight into the distribution, size, and depth of roots. Although the stability, data acquisition accuracy, and noise resilience of these new sensors must still be enhanced, these next-generation sensors nonetheless provide innovative perspectives and characteristics that are beneficial for deep learning to realize knowledge discovery and plant analysis.

(2) Requirements of enhanced semantic feature extraction modules and forest growth prediction networks. Forests are considered the most complex nonrigid living organisms in the world. Improving semantic feature extraction requires the development of models that better understand and interpret the context of forest morphogenesis. Encouragingly, many scholars have adopted the newest strategies in related research. For instance, the Transformer algorithm employs attention mechanisms that focus on crucial features for analysing forest ecosystems (Fayad et al., 2024), while EfficientNet automatically adjusts the network architecture with automatic scalability across different computational budgets for lightweight model design applicable to handheld devices (Atila et al., 2021). Dynamic MLPs incorporate contextual and environmental cues to capture subtleties and nuances for improving the discriminative power of image features and accomplishing fine-grained forest image classification tasks (Yang et al., 2022). We believe the consistent promotion of artificial intelligence by computer science will lead to emerging technologies playing a pivotal role in driving innovation and development in forest feature extraction. On the other hand, in contrast to crops, the slow pace of tree growth poses significant challenges in collecting reliable and comprehensive data over extended periods. Considering many factors, including the soil type, physical intervention, climate change, growth rate, seeding situation and interspecific and intraspecific competition in the chorological sequence, existing deep learning frameworks can hardly extrapolate short-term crop data to generate accurate long-term tree growth outcomes. Efficiently utilizing the latest techniques, e.g., temporal convolutional networks (Mohan et al., 2023) and recurrent neural networks (Qiao et al., 2021) shown in Fig. 7 b, for constructing architectures that learn and identify growth patterns and underlying trends, could enable temporal data dependencies to be captured and key information to be extracted for accurately predicting future growth trends of trees.

8.2. Prospects for future forest research incorporating deep learning techniques

The incorporation of innovative deep learning technologies is breaking new ground in the quickly changing field of forest research. This represents a paradigm shift towards more precise, intelligent and all-encompassing forest management strategies and signals a notable point in our comprehension and oversight of forest ecosystems. In the following section, three intriguing avenues for future forest studies are

examined.

(1) Deep learning-based forest phenotyping.

Forest phenotyping refers to the analysis and measurement of the physiological and biogeochemical characteristics of trees and forests for a better understanding of tree growth status, physiological and health conditions, and future forest production. This process includes measuring morphological traits, such as forest organism size and chronological growth development, as well as determining the relationships between genotypes and envirotyping (Wang et al., 2023e). The inherent feature extraction ability of deep learning networks facilitates the development of high-throughput automatic and semiautomatic phenotyping platforms and promotes the comprehensive analysis of multi-source perceptual data from forests, such as high-spectral images, electron microscope images and light and ranging (LiDAR) points. This approach helps characterize the structural and functional traits of plants from the cellular level to the entire plant or canopy level under controlled conditions. Deep learning has emerged as an auxiliary tool for computer graphics and computer vision algorithms. Direction-aware tree skeletons have been intelligently developed (Ding et al., 2022), and individual leaf segmentation (Zhu et al., 2023) with sophisticated appearance rendering has been achieved, benefiting automatic botanical tree modelling and forest scenario portrayal. Deep learning techniques have been employed for envirotyping, and the results have been combined with genotyping to predict genome-environment integration, which can reveal environmental factors that can alter forest growth and development through human intervention. The potential applications of these developments in the forest field are characterized in Fig. 8 below. By incorporating large amounts of new sensor data and improving semantic feature extraction capabilities, as mentioned in subsection 8.1, for analysing complex forest ecological phenomena and natural successional processes, deep learning could reinforce the performance of next-generation high-throughput phenotyping platforms with respect to culling elite tree varieties, tracing spatiotemporal phenotypic trait variations, and detecting invasive tree pests and phenotypic plasticity under environmental stress.

(2) Deep learning-based forest carbon stock estimation. Accurately estimating carbon stocks for various forest tree types is an essential step in understanding carbon neutrality and the terrestrial carbon cycle. Traditional methods often use multiregression models (McRoberts et al., 2013) or allometric growth equations (Karna et al., 2015) based on the input variables, such as the spectral band-related vegetation indices acquired by satellite sensors, LiDAR-derived forest structural parameters (basal area, LAI or tree heights, etc.), and local meteorological and geographical factors (Jayathunga et al., 2018). In addition, the space-terrene joint observation system for assessing carbon sequestration in forests, which includes gross primary production (GPP) and net primary production (NPP) data derived from MODIS aboard the Terra satellite, is often utilized, and the results are calibrated by magnitudes obtained from built eddy covariance towers (Griebel et al., 2020). Compared with existing regression-based methods, deep learning techniques applied to forest remote sensing data have better adaptive performance and more trainable parameters; as a result, these techniques can help rapidly and accurately appraise carbon storage within forest ecosystems (Sharma et al., 2022). Deep learning has robust universality and wide expandability for simulating carbon sink processes, performing biometric studies and generalizing complex ecological relationships (Huy et al., 2022). This approach could be implemented in the future to resolve heteroscedasticity for many different types of complex tree variables and samples, does not require the derivation of appropriate equation forms as traditional regression methods do, and can epitomize implicit equations for complex biometric regulations, vegetation productivity patterns and terrestrial carbon cycle systems (Ma et al., 2021b). As mentioned in subsection 8.1, new sensors, such as ground-penetrating radar, facilitate underground forest biomass and carbon storage assessments, while recurrent neural networks can contribute to forecasting future forest carbon sequestration growth. These advantages help

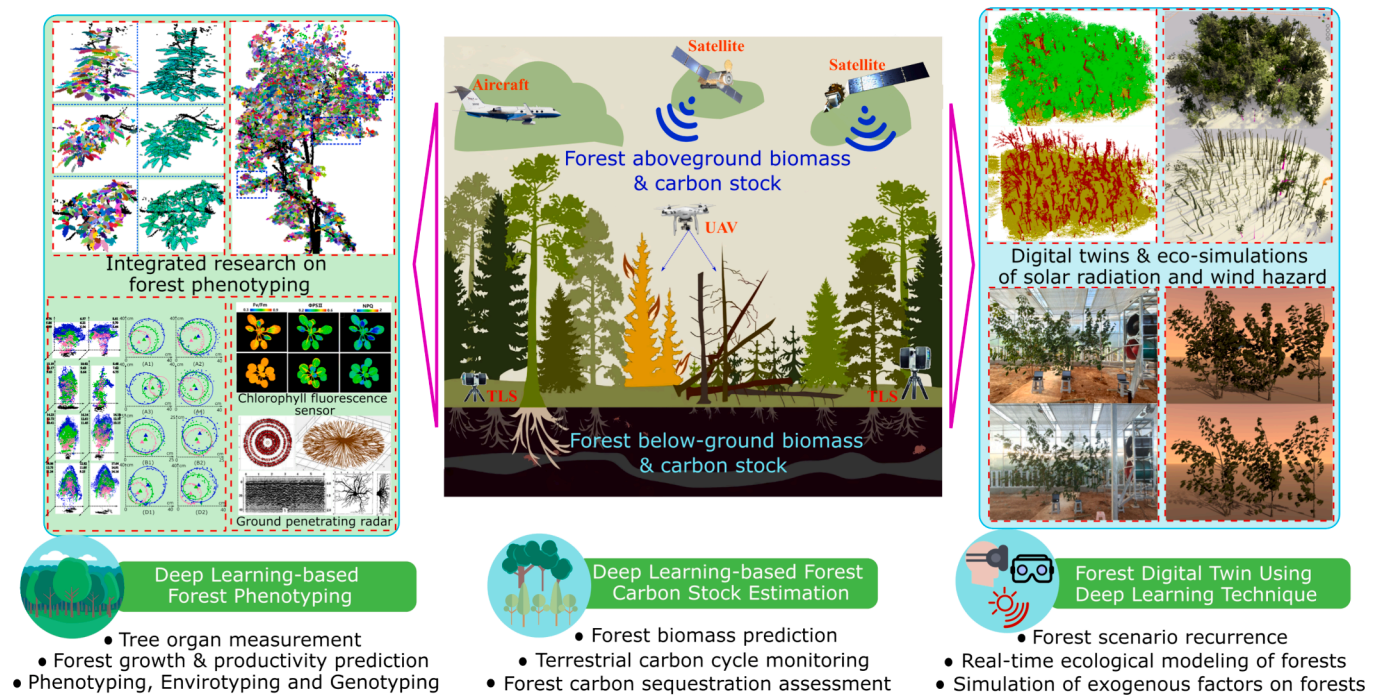


Fig. 8. A glimpse into the future: the advancement of deep learning in forests.

humans process large and complex forest carbon-related datasets for dynamic monitoring and *trans*-scale estimation of forest carbon sinks and enable policymakers to make informed decisions that reinforce the utilization of forested lands and maintain ecological balance (Reiersen et al., 2022).

(3) Deep learning-based forest digital twin creation. The forest digital twin develops a digital replica based on captured properties and phenotypic traits to contextualize the study forests in their living environments and represent a real forest or individual trees at a specific moment or over a time span (Hejtmánek et al., 2022). The digital twin employs digital embodiment to report the growth status of the forest during its life cycle, simulate the actual forest dynamics in response to natural environmental changes and conduct fidelity modalities of forest cultivation practices in virtual space, serving to ameliorate sustainable forest stewardship (Niță, 2021) and sharpen forest risk management (Buonocore et al., 2022). The interactions between life forms and their living conditions represented by digital twins comprise complex and dynamic processes involving the transmission and response of multiple signals and information through life entities, which necessitates greater interdisciplinary learning abilities and reasoning capabilities with the aid of deep learning networks. For example, by leveraging the shape, structure, texture and other morphologies from simple geometric primitives to complex scenes, deep learning networks can learn to recognize and reconstruct complex 3D structures from 2D images or point clouds, serving scene renderings and animation production with an enormous potential in applications, e.g., virtual reality, augmented reality and game development. Instances of this field include the neural radiance field (NeRF) (Hu et al., 2024), a novel deep learning-based technique for 3D scene representation and photorealistic view synthesis that combines complex forest scenes by optimizing an underlying continuous volumetric scene function with a sparse set of input views; the continuously evolving Sora model (Liu et al., 2024), with its advanced semantic analysis performance and powerful image generation capabilities, possibly accomplishing video production of different visual styles that match the needs and preferences of forest users. Moreover, deep learning models can learn to approximate the complex interactions between light and object surfaces to accelerate ray tracing calculations, rendering them suitable for real-time portrayal and

computational burden alleviation (Zuo et al., 2022). Deep learning models can also analyse vast amounts of historical data, e.g., geo-location, time of day, season, meteorological conditions and tree growth parameters, and learn to identify patterns and relationships for determining solar radiation levels influencing forest growth and productivity in different regions. In addition, deep learning models can assist in understanding and analysing the underlying mechanisms of intricate systems, which can help users study the unpredictable and chaotic motion of fluids and predict the lift, drag and other aerodynamic forces exerted on forests (Lye et al., 2020) to assess the optimal pruning shapes or planting configurations for various tree species that minimize hurricane damage or ensure safe growth in the future. Overall, the forest digital twin technique can be enhanced by deep learning techniques as a powerful tool for creating and mirroring a realistic representation of the forest ecosystem and provides the possibility to capture a virtual forest by adjusting its growth environmental factors, which offers effective tools and approaches for improving forest management and silvicultural decision-making.

9. Conclusion

In this paper, the current and future prospects of deep learning techniques in many forest applications are systematically reviewed, and the principles and frameworks of diverse deep learning methods using a wide range of forest remote sensing data types (i.e., imagery, point clouds, and fusion data), as well as their strengths and restrictions for various tissues, are detailed. Moreover, the patterns and trends of deep learning applications in forest research from 2018 to 2023 are described, focusing on the geographical distributions of the number of publications. The sharing of primarily forest-related datasets by many institutions is also discussed, with an outlook on the promising prospects of forest fields being provided. We likewise anticipate that future global forest research will expand from the interdisciplinary perspectives of computer science, remote sensing and forestry, and we hope this systematic review provides a comprehensive overview of the applications of deep learning techniques to forest research, offering valuable insights and roadmaps for academics and industry professionals in the related field, as well as for those embarking on a research journey in the forest

domain.

CRediT authorship contribution statement

Ting Yun: Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Jian Li:** Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Lingfei Ma:** Writing – review & editing. **Ji Zhou:** Supervision. **Ruisheng Wang:** Formal analysis, Data curation. **Markus P. Eichhorn:** Validation, Supervision. **Huaiqing Zhang:** Validation, Supervision.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

This study was financially supported by the National Natural Science Foundation of China (grant numbers 32371876, 32271877, and 42101451), the Natural Science Foundation of Jiangsu Province (BK20221337), China, the Jiangsu Provincial Agricultural Science and Technology Independent Innovation Fund Project (CX(22)3048) and the Key Laboratory of Land Satellite Remote Sensing Application, Ministry of Natural Resources of the People's Republic of China (KLSMNR-G202208). Moreover, we are indebted to Professor Cao Lin of Nanjing Forestry University for his guidance and support during the writing of this review.

References

- Ahlswede, S., Schulz, C., Gava, C., Helber, P., Bischke, B., Förster, M., Arias, F., Hees, J., Demir, B., Kleinschmit, B., 2023. TreeSataI Benchmark Archive: a multi-sensor, multi-label dataset for tree species classification in remote sensing. *Earth Syst. Sci. Data* 15, 681–695. <https://doi.org/10.5194/essd-15-681-2023>.
- Altman, N.S., 1992. An introduction to kernel and nearest-neighbor nonparametric regression. *Am. Stat.* 46, 175–185.
- Arbeláez, P., Maire, M., Fowlkes, C., Malik, J., 2011. Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 898–916. <https://doi.org/10.1109/TPAMI.2010.161>.
- Atila, Ü., Uçar, M., Akyol, K., Uçar, E., 2021. Plant leaf disease classification using EfficientNet deep learning model. *Ecol. Inform.* 61, 101182 <https://doi.org/10.1016/j.ecoinf.2020.101182>.
- Belongie, S., Wang, Q., Snively, N., Khan, A., 2021. The Plant Pathology 2021 Challenge dataset to classify foliar disease of apples. *Plant Pathol.* 2021 Chall. - FGVC8.
- Bolyn, C., Lejeune, P., Michez, A., Latte, N., 2022. Mapping tree species proportions from satellite imagery using spectral – spatial deep learning. *Remote Sens. Environ.* 280, 113205 <https://doi.org/10.1016/j.rse.2022.113205>.
- Briechele, S., Krzystek, P., Vosselman, G., 2021. Silvi-Net-A dual-CNN approach for combined classification of tree species and standing dead trees from remote sensing data. *Int. J. Appl. Earth Obs. Geoinf.* 98, 102292 <https://doi.org/10.1016/j.jag.2020.102292>.
- Buonocore, L., Yates, J., Valentini, R., 2022. A Proposal for a Forest Digital Twin Framework and Its Perspectives. *Forests*. <https://doi.org/10.3390/f13040498>.
- Cai, C., Xu, H., Chen, S., Yang, L., Weng, Y., Huang, S., Dong, C., Lou, X., 2023. Tree recognition and crown width extraction based on novel faster-RCNN in a dense loblolly pine environment. *Forests*. <https://doi.org/10.3390/f14050863>.
- Chen, G., Cheng, R., Lin, X., Jiao, W., Bai, D., Lin, H., 2023a. LMDFS: a lightweight model for detecting forest fire smoke in UAV images based on YOLOv7. *Remote Sens.* <https://doi.org/10.3390/rs15153790>.
- Chen, X., Jiang, K., Zhu, Y., Wang, X., Yun, T., 2021a. Individual tree crown segmentation directly from UAV-borne LiDAR data using the PointNet of deep learning. *Forests* 12, 131.
- Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* 40, 834–848.
- Chen, X.G., Shen, X., Cao, L., 2023b. Tree species classification in subtropical natural forests using high-resolution UAV RGB and superview-1 multispectral imageries based on deep learning network approaches: a case study within the Baima snow mountain national nature reserve, China. *Remote Sens.* 15 <https://doi.org/10.3390/rs15102697>.
- Chen, Y., Wu, R., Yang, C., Lin, Y., 2021b. Urban vegetation segmentation using terrestrial LiDAR point clouds based on point non-local means network. *Appl. Earth Obs. Geoinf.* 105, 102580 <https://doi.org/10.1016/j.jag.2021.102580>.
- Cheng, Y., Lan, S., Fan, X., Tjahjadi, T., Jin, S., Cao, L., 2023. A dual-branch weakly supervised learning based network for accurate mapping of woody vegetation from remote sensing images. *Int. J. Appl. Earth Obs. Geoinf.* 124, 103499 <https://doi.org/10.1016/j.jag.2023.103499>.
- Damm, A., Cogliati, S., Colombo, R., Fritsche, L., Genangeli, A., Genesio, L., Hanus, J., Peressotti, A., Rademski, P., Rascher, U., Schuettemeyer, D., Siegmann, B., Sturm, J., Miglietta, F., 2022. Response times of remote sensing measured sun-induced chlorophyll fluorescence, surface temperature and vegetation indices to evolving soil water limitation in a crop canopy. *Remote Sens. Environ.* 273, 112957 <https://doi.org/10.1016/j.rse.2022.112957>.
- De Deuge, M., Quadros, A., Hung, C., Douillard, B., 2013. Unsupervised feature learning for classification of outdoor 3D scans. *Australasian Conference on Robotics and Automation*. University of New South Wales Kensington, Australia.
- Demir, I., Koperski, K., Lindenbaum, D., Pang, G., Huang, J., Basu, S., Hughes, F., Tuia, D., Raskar, R., 2018. Deepglobe 2018: A challenge to parse the earth through satellite images. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. pp. 172–181.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. ImageNet: A large-scale hierarchical image database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 248–255. [10.1109/CVPR.2009.5206848](https://doi.org/10.1109/CVPR.2009.5206848).
- Dersch, S., Heurich, M., Krueger, N., Krzystek, P., 2021. Combining graph-cut clustering with object-based stem detection for tree segmentation in highly dense airborne lidar point clouds. *ISPRS J. Photogramm. Remote Sens.* 172, 207–222. <https://doi.org/10.1016/j.isprsjprs.2020.11.016>.
- Ding, Z., Zhou, L., Fan, J., An, F., Chen, B., Wang, M., Jiang, L., Xue, L., Yun, T., 2022. Rubber tree branch modeling and property retrieval based on laser scanning data and deep learning technique. *Trans. Chinese Soc. Agric. Eng.* 38, 187–199. <https://doi.org/10.11975/j.issn.1002-6819.2022.08.022>.
- Dong, Z., Liang, F., Yang, B., Xu, Y., Zang, Y., Li, J., Wang, Y., Dai, W., Fan, H., Hyyppä, J., Stilla, U., 2020. Registration of large-scale terrestrial laser scanner point clouds: A review and benchmark. *ISPRS J. Photogramm. Remote Sens.* 163, 327–342. <https://doi.org/10.1016/j.isprsjprs.2020.03.013>.
- Fassnacht, F.E., White, J.C., Wulder, M.A., Næsset, E., 2023. Remote sensing in forestry: current challenges, considerations and directions. *For. an Int. J. for. Res.* cpad024 <https://doi.org/10.1093/forestry/cpad024>.
- Fayad, I., Clais, P., Schwartz, M., Wigneron, J.-P., Baghdadi, N., de Truchis, A., D'Aspremont, A., Frappart, F., Saatchi, S., Sean, E., Pellissier-Tanon, A., Bazzi, H., 2024. Hy-TeC: a hybrid vision transformer model for high-resolution and large-scale mapping of canopy height. *Remote Sens. Environ.* 302, 113945 <https://doi.org/10.1016/j.rse.2023.113945>.
- Ferreira, M.P., Almeida, D.R.A. de, Papa, D. de A., Minervino, J.B.S., Veras, H.F.P., Formighieri, A., Santos, C.A.N., Ferreira, M.A.D., Figueiredo, E.O., Ferreira, E.J.L., 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>.
- Ferreira, M.P., Lotte, R.G., D'Elia, F.V., Stamatopoulos, C., Kim, D.-H., Benjamin, A.R., 2021. Accurate mapping of Brazil nut trees (*Bertholletia excelsa*) in Amazonian forests using WorldView-3 satellite images and convolutional neural networks. *Ecol. Inform.* 63, 101302 <https://doi.org/10.1016/j.ecoinf.2021.101302>.
- Fu, L., Majeed, Y., Zhang, X., Karkee, M., Zhang, Q., 2020. Faster R-CNN-based apple detection in dense-foliage fruiting-wall trees using RGB and depth features for robotic harvesting. *Biosyst. Eng.* 197, 245–256. <https://doi.org/10.1016/j.biosystemseng.2020.07.007>.
- Gao, D., Wang, H., Guo, X., Wang, L., Gui, G., Wang, W., Yin, Z., Wang, S., Liu, Y., He, T., 2023. Federated Learning Based on CTC for Heterogeneous Internet of Things. *IEEE Internet Things J.* 10, 22673–22685. <https://doi.org/10.1109/IJOT.2023.3305189>.
- Geiger, A., Lenz, P., Stiller, C., Urtasun, R., 2013. Vision meets robotics: The kitti dataset. *Int. J. Rob. Res.* 32, 1231–1237.
- Gholamnia, K., Gudiyangada Nachappa, T., Ghorbanzadeh, O., Blaschke, T., 2020. Comparisons of Diverse Machine Learning Approaches for Wildfire Susceptibility Mapping. *Symmetry (basel)*. <https://doi.org/10.3390/sym12040604>.
- Ghosh, S.M., Behera, M.D., 2021. Aboveground biomass estimates of tropical mangrove forest using Sentinel-1 SAR coherence data-The superiority of deep learning over a semi-empirical model. *Comput. Geosci.* 150, 104737.
- Gonzalez-Perez, A., Abd-Elrahman, A., Wilkinson, B., Johnson, D.J., Carthy, R.R., 2022. Deep and machine learning image classification of coastal wetlands using unpiloted aircraft system multispectral images and lidar datasets. *Remote Sens.* <https://doi.org/10.3390/rs14163937>.
- Graves, S., Marconi, S., 2020. IDTreeS 2020 Competition Data. [10.5281/zenodo.3934932](https://zenodo.org/record/3934932).
- Griebel, A., Metzner, D., Boer, M.M., Barton, C.V.M., Renchon, A.A., Andrews, H.M., Pendall, E., 2020. Using a paired tower approach and remote sensing to assess carbon sequestration and energy distribution in a heterogeneous sclerophyll forest. *Sci. Total Environ.* 699, 133918 <https://doi.org/10.1016/j.scitotenv.2019.133918>.
- Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., Bennamoun, M., 2020. Deep learning for 3d point clouds: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 43, 4338–4364.
- Hackel, T., Savinov, N., Ladicky, L., Wegner, J.D., Schindler, K., Pollefeys, M., 2017. Semantic3d. net: A new large-scale point cloud classification benchmark. *arXiv Prepr. arXiv1704.03847*.
- Hamedianfar, A., Mohamedou, C., Kangas, A., Vauhkonen, J., 2022. Deep learning for forest inventory and planning: a critical review on the remote sensing approaches so

- far and prospects for further applications. *For. An Int. J. for. Res.* 95, 451–465. <https://doi.org/10.1093/forestry/cpac002>.
- Hamraz, H., Jacobs, N.B., Contreras, M.A., Clark, C.H., 2019. Deep learning for conifer/deciduous classification of airborne LiDAR 3D point clouds representing individual trees. *ISPRS J. Photogramm. Remote Sens.* 158, 219–230. <https://doi.org/10.1016/j.isprsjprs.2019.10.011>.
- Hartling, S., Sagan, V., Sidike, P., Maimaitijiang, M., Carron, J., 2019. Urban tree species classification using a worldview-2/3 and LiDAR data fusion approach and deep learning. *Sensors (Switzerland)* 19, 1–23. <https://doi.org/10.3390/s19061284>.
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask r-cnn, in: *Proceedings of the IEEE International Conference on Computer Vision*. pp. 2961–2969.
- Hejtmánek, L., Hůla, M., Herrová, A., Surový, P., 2022. Forest digital twin as a relaxation environment: A pilot study. *Front. Virtual Real.*
- Helber, P., Bischke, B., Dengel, A., Borth, D., 2019. EuroSAT: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 12, 2217–2226.
- Hell, M., Brandmeier, M., Briechele, S., Krzystek, P., 2022. Classification of tree species and standing dead trees with lidar point clouds using two deep neural networks: PointCNN and 3DmFV - Net. *Photogramm. Remote Sens. Geoinf. Sci.* 90, 103–121. <https://doi.org/10.1007/s41064-022-00200-4>.
- Hou, J., Yang, C., He, Y., Hou, B., 2023. Detecting diseases in apple tree leaves using FPN-ISResNet-Faster RCNN. *Eur. J. Remote Sens.* 56, 2186955.
- Hu, K., Ying, W., Pan, Y., Kang, H., Chen, C., 2024. High-fidelity 3D reconstruction of plants using Neural Radiance Fields. *Comput. Electron. Agric.* 220, 108848 <https://doi.org/10.1016/j.compag.2024.108848>.
- Huang, Z., Zhong, L., Zhao, F., Wu, J., Tang, H., Lv, Z., Xu, B., Zhou, L., Sun, R., Meng, R., 2023. A spectral-temporal constrained deep learning method for tree species mapping of plantation forests using time series Sentinel-2 imagery. *ISPRS J. Photogramm. Remote Sens.* 204, 397–420. <https://doi.org/10.1016/j.isprsjprs.2023.09.009>.
- Huy, B., Truong, N.Q., Khiem, N.Q., Poudel, K.P., Temesgen, H., 2022. Deep learning models for improved reliability of tree aboveground biomass prediction in the tropical evergreen broadleaf forests. *For. Ecol. Manage.* 508, 120031 <https://doi.org/10.1016/j.foreco.2022.120031>.
- Ienco, D., Interdonato, R., Gaetano, R., Minh, D.H., 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>.
- Jayathunga, S., Owari, T., Tsuyuki, S., 2018. The use of fixed-wing UAV photogrammetry with LiDAR DTM to estimate merchantable volume and carbon stock in living biomass over a mixed conifer–broadleaf forest. *Int. J. Appl. Earth Obs. Geoinf.* 73, 767–777. <https://doi.org/10.1016/j.jag.2018.08.017>.
- Jiang, T., Liu, S., Zhang, Q., Xu, X., Sun, J., Wang, Y., 2023. Segmentation of individual trees in urban MLS point clouds using a deep learning framework based on cylindrical convolution network. *Int. J. Appl. Earth Obs. Geoinf.* 123, 103473 <https://doi.org/10.1016/j.jag.2023.103473>.
- Karna, Y.K., Hussin, Y.A., Gilani, H., Bronsveld, M.C., Murthy, M.S.R., Qamer, F.M., Karky, B.S., Bhattarai, T., Aigong, X., Baniya, C.B., 2015. Integration of WorldView-2 and airborne LiDAR data for tree species level carbon stock mapping in Kayar Khola watershed. *Nepal. Int. J. Appl. Earth Obs. Geoinf.* 38, 280–291. <https://doi.org/10.1016/j.jag.2015.01.011>.
- Kattenborn, T., Eichel, J., Fassnacht, F.E., 2019. Convolutional Neural Networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution UAV imagery. *Sci. Rep.* 9, 1–9. <https://doi.org/10.1038/s41598-019-53797-9>.
- Koira, A., Walsh, K.B., Wang, Z., McCarthy, C., 2019. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO'. *Precis. Agric.* 20, 1107–1135. <https://doi.org/10.1007/s11119-019-09642-0>.
- Korznikov, K.A., Kislov, D.E., Altman, J., Doležal, J., Vozmishcheva, A.S., Krestov, P. V., 2021. Using U-Net-Like Deep Convolutional Neural Networks for Precise Tree Recognition in Very High Resolution RGB (Red, Green, Blue) Satellite Images. *Forests*. 10.3390/f12010066.
- Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., Pont-Tuset, J., Kamali, S., Popov, S., Mallocci, M., Kolesnikov, A., 2020. The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale. *Int. J. Comput. Vis.* 128, 1956–1981.
- Lai, J., Kang, X., Lu, X., Li, S., 2022. A review of land observation satellite remote sensing application technology with new generation artificial intelligence. *Natl. Remote Sens. Bull.* 26, 1530–1546. <https://doi.org/10.11834/jrs.20221555>.
- Lassalle, G., Ferreira, M.P., La Rosa, L.E.C., de Souza Filho, C.R., 2022. Deep learning-based individual tree crown delineation in mangrove forests using very-high-resolution satellite imagery. *ISPRS J. Photogramm. Remote Sens.* 189, 220–235. <https://doi.org/10.1016/j.isprsjprs.2022.05.002>.
- Lee, M.-G., Cho, H.-B., Youm, S.-K., Kim, S.-W., 2023. Detection of Pine Wilt Disease Using Time Series UAV Imagery and Deep Learning Semantic Segmentation. *Forests*. <https://doi.org/10.3390/f14081576>.
- Lelli, C., Bruun, H.H., Chiarucci, A., Donati, D., Frascaroli, F., Fritz, Ö., Goldberg, I., Nascimbene, J., Tottup, A.P., Rahbek, C., Heilmann-Clausen, J., 2019. Biodiversity response to forest structure and management: Comparing species richness, conservation relevant species and functional diversity as metrics in forest conservation. *For. Ecol. Manage.* 432, 707–717. <https://doi.org/10.1016/j.foreco.2018.09.057>.
- Li, H.C., Chen, L., Yao, Z.Q., Li, N.W., Long, L., Zhang, X.L., 2023. Intelligent Identification of Pine Wilt Disease Infected Individual Trees Using UAV-Based Hyperspectral Imagery. *Remote Sens.* 15 <https://doi.org/10.3390/rs15133295>.
- Li, L., Mu, X., Chianucci, F., Qi, J., Jiang, J., Zhou, J., Chen, L., Huang, H., Yan, G., Liu, S., 2022. Ultrahigh-resolution boreal forest canopy mapping: Combining UAV imagery and photogrammetric point clouds in a deep-learning-based approach. *Int. J. Appl. Earth Obs. Geoinf.* 107, 102686 <https://doi.org/10.1016/j.jag.2022.102686>.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L., 2014. Microsoft coco: Common objects in context, in: *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13*. Springer, pp. 740–755.
- Lin, F.-C., Shiu, Y.-S., Wang, P.-J., Wang, U.-H., Lai, J.-S., Chuang, Y.-C., 2024. A model for forest type identification and forest regeneration monitoring based on deep learning and hyperspectral imagery. *Ecol. Inform.* 80, 102507 <https://doi.org/10.1016/j.ecoinf.2024.102507>.
- Liu, W., Li, R., Zheng, M., Karanam, S., Wu, Z., Bhanu, B., Radke, R.J., Camps, O., 2020b. Towards visually explaining variational autoencoders, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 8642–8651. <https://arxiv.org/abs/1911.07389>.
- Liu, Y., Zhang, K., Li, Y., Yan, Z., Gao, C., Chen, R., Yuan, Z., Huang, Y., Sun, H., Gao, J., 2024. Sora: A Review on Background, Technology, Limitations, and Opportunities of Large Vision Models. *arXiv Prepr. arXiv2402.17177*.
- Liu, M., Han, Z., Chen, Y., Liu, Z., Han, Y., 2021. Tree species classification of LiDAR data based on 3D deep learning. *Measurement* 177, 109301. <https://doi.org/10.1016/j.measurement.2021.109301>.
- Liu, X., Hu, C., Li, P., 2020. Automatic segmentation of overlapped poplar seedling leaves combining Mask R-CNN and DBSCAN. *Comput. Electron. Agric.* 178, 105753 <https://doi.org/10.1016/j.compag.2020.105753>.
- Liu, B., Huang, H., Su, Y., Chen, S., Li, Z., Chen, E., Tian, X., 2022. Tree Species Classification Using Ground-Based LiDAR Data by Various Point Cloud Deep Learning Methods. *Remote Sens.* <https://doi.org/10.3390/rs14225733>.
- Luo, H., Khoshelham, K., Chen, C., He, H., 2021. Individual tree extraction from urban mobile laser scanning point clouds using deep pointwise direction embedding ☆. *ISPRS J. Photogramm. Remote Sens.* 175, 326–339. <https://doi.org/10.1016/j.isprsjprs.2021.03.002>.
- Lye, K.O., Mishra, S., Ray, D., 2020. Deep learning observables in computational fluid dynamics. *J. Comput. Phys.* 410, 109339 <https://doi.org/10.1016/j.jcp.2020.109339>.
- Ma, B., Du, J., Wang, L., Jiang, H., Zhou, M., 2021a. Automatic branch detection of jujube trees based on 3D reconstruction for dormant pruning using the deep learning-based method. *Comput. Electron. Agric.* 190, 106484.
- Ma, H., Mo, L., Crowther, T.W., Maynard, D.S., van den Hoogen, J., Stocker, B.D., Terrer, C., Zohner, C.M., 2021b. The global distribution and environmental drivers of aboveground versus belowground plant biomass. *Nat. Ecol. Evol.* 5, 1110–1122. <https://doi.org/10.1038/s41559-021-01485-1>.
- Mäyrä, J., Keski-Saari, S., Kivinen, S., Tanhuanpää, T., Hurskainen, P., Kullberg, P., Poikolainen, L., Viinikka, A., Tuominen, S., Kumpula, T., Vihervaara, P., 2021. Tree species classification from airborne hyperspectral and LiDAR data using 3D convolutional neural networks. *Remote Sens. Environ.* 256 <https://doi.org/10.1016/j.rse.2021.112322>.
- McEwan, A., Marchi, E., Spinelli, R., Brink, M., 2020. Past, present and future of industrial plantation forestry and implication on future timber harvesting technology. *J. For. Res.* 31, 339–351. <https://doi.org/10.1007/s11676-019-01019-3>.
- McRoberts, R.E., Næsset, E., Gobakken, T., 2013. Inference for lidar-assisted estimation of forest growing stock volume. *Remote Sens. Environ.* 128, 268–275. <https://doi.org/10.1016/j.rse.2012.10.007>.
- Mirhaji, H., Soleymani, M., Asakereh, A., Abdanan Mehdiadeh, S., 2021. Fruit detection and load estimation of an orange orchard using the YOLO models through simple approaches in different imaging and illumination conditions. *Comput. Electron. Agric.* 191, 106533 <https://doi.org/10.1016/j.compag.2021.106533>.
- Mohan, A., Venkatesan, M., Prabhavathy, P., Jayakrishnan, A., 2023. Temporal convolutional network based rice crop yield prediction using multispectral satellite data. *Infrared Phys. Technol.* 135, 104960 <https://doi.org/10.1016/j.infrared.2023.104960>.
- Momo Takoudjou, S., Ploton, P., Sonké, B., Hackenberg, J., Griffon, S., De Coligny, F., Kamdem, N.G., Libalah, M., Mofack, G.L.L., Le Moguécé, G., 2018. Using terrestrial laser scanning data to estimate large tropical trees biomass and calibrate allometric models: A comparison with traditional destructive approach. *Methods Ecol. Evol.* 9, 905–916.
- Munoz, D., Bagnell, J.A., Vandapel, N., Hebert, M., 2009. Contextual classification with functional max-margin markov networks. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 975–982.
- Nguyen, T.-A., Rußwurm, M., Lenczner, G., Tuia, D., 2024. Multi-temporal forest monitoring in the Swiss Alps with knowledge-guided deep learning. *Remote Sens. Environ.* 305, 114109 <https://doi.org/10.1016/j.rse.2024.114109>.
- Ning, X., Ma, Y., Hou, Y., Lv, Z., Jin, H., Wang, Z., Wang, Y., 2023. Trunk-constrained and tree structure analysis method for individual tree extraction from scanned outdoor scenes. *Remote Sens.* <https://doi.org/10.3390/rs15061567>.
- Niță, M.D., 2021. Testing forestry digital twinning workflow based on mobile LiDAR scanner and AI platform. *Forests*. <https://doi.org/10.3390/f12111576>.
- Nüchter, A., Elseberg, J., Schneider, P., Paulus, D., 2010. Study of parameterizations for the rigid body transformations of the scan registration problem. *Comput. Vis. Image Underst.* 114, 963–980.
- Pearse, G.D., Tan, A.Y.S., Watt, M.S., Franz, M.O., Dash, J.P., 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>.

- Potapov, P., Hansen, M.C., Pickens, A., Hernandez-Serna, A., Tyukavina, A., Turubanova, S., Zalles, V., Li, X., Khan, A., Stolle, F., Harris, N., Song, X.-P., Baggett, A., Kommareddy, I., Kommareddy, A., 2022. The global 2000–2020 land cover and land use change dataset derived from the landsat archive: first results. *Front. Remote Sens.*
- Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Adv. Neural Inf. Process. Syst.* 30.
- Qiao, M., He, X., Cheng, X., Li, P., Luo, H., Zhang, L., Tian, Z., 2021. Crop yield prediction from multi-spectral, multi-temporal remotely sensed imagery using recurrent 3D convolutional neural networks. *Int. J. Appl. Earth Obs. Geoinf.* 102, 102436 <https://doi.org/10.1016/j.jag.2021.102436>.
- Reisersen, G., Dao, D., Lütjens, B., Klemmer, K., Amara, K., Steinegger, A., Zhang, C., Zhu, X., 2022. ReforesTree: A Dataset for Estimating Tropical Forest Carbon Stock with Deep Learning and Aerial Imagery. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 12119–12125.
- Rihan, M., Bindajam, A.A., Talukdar, S., Naikoo, M.W., Mallick, J., Rahman, A., 2023. Forest fire susceptibility mapping with sensitivity and uncertainty analysis using machine learning and deep learning algorithms. *Adv. Sp. Res.* 72, 426–443. <https://doi.org/10.1016/j.asr.2023.03.026>.
- Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B., 2022. High-resolution image synthesis with latent diffusion models, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 10684–10695. <https://arxiv.org/abs/2112.10752>.
- Roynard, X., Deschard, J.-E., Goulette, F., 2018. Paris-Lille-3D: A large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification. *Int. J. Rob. Res.* 37, 545–557. <https://doi.org/10.1177/0278364918767506>.
- Schepaschenko, D., Shvidenko, A., Usoltsev, V.A., Lakyda, P., Luo, Y., Vasylyshyn, R., Lakyda, I., Myklush, Y., See, L., McCallum, I., Fritz, S., Kraxner, F., Obersteiner, M., 2017. Biomass tree data base. Suppl. to Schepaschenko, D et al. A Dataset for. *Biomass Struct. Eurasia. Sci. Data* 4, 170070. <https://doi.org/10.1038/sdata.2017.70>. DOI: 10.1594/PANGAEA.871491.
- Schmidhuber, J., 2015. Deep learning in neural networks: An overview. *Neural Networks* 61, 85–117.
- Schwartz, M., Clais, P., Ottlé, C., De Truchis, A., Vega, C., Fayad, I., Brandt, M., Fensholt, R., Baghdadi, N., Morneau, F., Morin, D., Guyon, D., Dayau, S., Wigneron, J.-P., 2024. High-resolution canopy height map in the Landes forest (France) based on GEDI, Sentinel-1, and Sentinel-2 data with a deep learning approach. *Int. J. Appl. Earth Obs. Geoinf.* 128, 103711 <https://doi.org/10.1016/j.jag.2024.103711>.
- Seidel, D., Annighöfer, P., Thielman, A., Seifert, Q.E., Thauer, J.-H., Glatthorn, J., Ehbrecht, M., Kneib, T., Ammer, C., 2021. Predicting tree species from 3D laser scanning point clouds using deep learning. *Front. Plant Sci.* 12, 635440.
- Senanayake, S., Pradhan, B., Alamri, A., Park, H.-J., 2022. A new application of deep neural network (LSTM) and RUSLE models in soil erosion prediction. *Sci. Total Environ.* 845, 157220 <https://doi.org/10.1016/j.scitotenv.2022.157220>.
- Sharma, S., Dhal, S., Rout, T., Acharya, B.S., 2022. Drones and machine learning for estimating forest carbon storage. *Carbon Res.* 1, 21. <https://doi.org/10.1007/s44246-022-00021-5>.
- Shen, X., Cao, L., Coops, N.C., Fan, H., Wu, X., Liu, H., Wang, G., Cao, F., 2020. Quantifying vertical profiles of biochemical traits for forest plantation species using advanced remote sensing approaches. *Remote Sens. Environ.* 250, 112041 <https://doi.org/10.1016/j.rse.2020.112041>.
- Shu, D., Park, S.W., Kwon, J., 2019. 3D Point Cloud Generative Adversarial Network Based on Tree Structured Graph Convolutions. In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 3858–3867. <https://doi.org/10.1109/ICCV.2019.00396>.
- Shuman, D.I., Narang, S.K., Frossard, P., Ortega, A., Vandergheynst, P., 2013. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag.* 30, 83–98. <https://doi.org/10.1109/MSP.2012.2235192>.
- Su, H., Maji, S., Kalogerakis, E., Learned-Miller, E., 2015. Multi-view convolutional neural networks for 3d shape recognition. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 945–953.
- Sun, C., Huang, C., Zhang, H., Chen, B., An, F., 2022. Individual tree crown segmentation and crown width extraction from a Heightmap derived from aerial laser scanning data using a deep learning framework. *Front. Plant Sci.* 13, 1–23. <https://doi.org/10.3389/fpls.2022.914974>.
- Tang, L., He, J., Peng, W., Huang, H., Chen, C., Yu, C., 2023. Assessing the visibility of urban greenery using MLS LiDAR data. *Landsc. Urban Plan.* 232, 104662 <https://doi.org/10.1016/j.landurbplan.2022.104662>.
- Tao, H., Li, C., Zhao, D., Deng, S., Hu, H., Xu, X., Jing, W., 2020. Deep learning-based dead pine tree detection from unmanned aerial vehicle images. *Int. J. Remote Sens.* 41, 8238–8255. <https://doi.org/10.1080/01431161.2020.1766145>.
- Tong, Z., Gao, J., Yuan, D., 2020. Advances of deep learning applications in ground-penetrating radar: A survey. *Constr. Build. Mater.* 258, 120371 <https://doi.org/10.1016/j.conbuildmat.2020.120371>.
- Ulku, I., Akagündüz, E., Ghamisi, P., 2022. Deep semantic segmentation of trees using multispectral images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 15, 7589–7604.
- van Geffen, F., Schulte, L., Geng, R., Heim, B., Pestryakova, L.A., Herzschiuh, U., Kruse, S., 2021. SiDroForest: Individual-labelled trees acquired during the fieldwork expeditions that took place in 2018 in Central Yakutia and Chukotka. *Siberia. https://doi.org/10.1594/PANGAEA.932821*.
- van Noordwijk, M., 2019. Integrated natural resource management as pathway to poverty reduction: Innovating practices, institutions and policies. *Agric. Syst.* 172, 60–71.
- Wagner, F.H., Sanchez, A., Tarabalka, Y., Lotte, R.G., Ferreira, M.P., Aidar, M.P.M., Gloor, E., Phillips, O.L., Aragão, L.E.O.C., 2019. Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. *Remote Sens. Ecol. Conserv.* 5, 360–375. <https://doi.org/10.1002/rse2.111>.
- Wan, P., Shao, J., Jin, S., Wang, T., Yang, S., Yan, G., Zhang, W., 2021. A novel and efficient method for wood-leaf separation from terrestrial laser scanning point clouds at the forest plot level. *Methods Ecol. Evol.* 12, 2473–2486.
- Wang, D., 2020. Unsupervised semantic and instance segmentation of forest point clouds. *ISPRS J. Photogramm. Remote Sens.* 165, 86–97. <https://doi.org/10.1016/j.isprsjprs.2020.04.020>.
- Wang, J., Chen, X., Cao, L., An, F., Chen, B., Xue, L., Yun, T., 2019. Individual rubber tree segmentation based on ground-based LiDAR data and faster R-CNN of deep learning. *Forests* 10, 793.
- Wang, S., Hou, C., Chen, Y., Liu, Z., Zhang, Z., Zhang, G., 2023c. Classification of Hyperspectral and LiDAR Data Using Multi-Modal Transformer Cascaded Fusion Net. *Remote Sens.* <https://doi.org/10.3390/rs15174142>.
- Wang, N., Pu, T., Zhang, Y., Liu, Y., Zhang, Z., 2023b. More appropriate DenseNetBL classifier for small sample tree species classification using UAV-based RGB imagery. *Heliyon* 9, e20467.
- Wang, X.L., Wang, J., Lian, Z.Z., Yang, N., 2023d. Semi-Supervised Tree Species Classification for Multi-Source Remote Sensing Images Based on a Graph Convolutional Neural Network. *FORESTS* 14. <https://doi.org/10.3390/f14061211>.
- Wang, F., Yang, Y., Wu, Z., Zhou, J., Zhang, W., 2023a. Real-Time Semantic Segmentation of Point Clouds Based on an Attention Mechanism and a Sparse Tensor. *Appl. Sci.* 13, 3256.
- Wang, Y., Zhang, J., Diao, S., Xue, L., 2023e. Nonlinear model prediction of needle chlorophyll content of *Picea koraiensis* Nakai at different needle ages based on hyperspectral features. *Front. For. Glob. Chang.*
- Weinstein, B.G., Marconi, S., Bohlman, S.A., Zare, A., White, E.P., 2020. Cross-site learning in deep learning RGB tree crown detection. *Ecol. Inform.* 56, 101061 <https://doi.org/10.1016/j.ecoinf.2020.101061>.
- Wu, W., Fan, X., Qu, H., Yang, X., Tjahjadi, T., 2022. TCDNet: Tree Crown Detection From UAV Optical Images Using Uncertainty-Aware One-Stage Network. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5. <https://doi.org/10.1109/LGRS.2022.3214281>.
- Wu, B.X., Zheng, G., Chen, Y., 2020. An improved convolution neural network-based model for classifying foliage and woody components from terrestrial laser scanning data. *Remote Sens.* 12 <https://doi.org/10.3390/rs12061010>.
- Xi, X., Xia, K., Yang, Y., Du, X., Feng, H., 2021. Evaluation of dimensionality reduction methods for individual tree crown delineation using instance segmentation network and UAV multispectral imagery in urban forest. *Comput. Electron. Agric.* 191, 106506 <https://doi.org/10.1016/j.compag.2021.106506>.
- Xian, T.K., Nugroho, H., 2022. Forest Fire Detection for Edge Devices. In: *2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, pp. 1–4. <https://doi.org/10.1109/IICAIET55139.2022.9936786>.
- Xu, S., Li, X., Yang, H., Xu, S., 2023. R-ProjNet: an optimal rotated-projection neural network for wood segmentation from point clouds. *Remote Sens. Lett.* 14, 60–69. <https://doi.org/10.1080/2150704X.2022.2163203>.
- Yaloveha, V., Hlavcheva, D., Podorozhniak, A., Kuchuk, H., 2019. Fire hazard research of forest areas based on the use of convolutional and capsule neural networks. In: *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*. IEEE, pp. 828–832.
- Yang, L., Li, X., Song, R., Zhao, B., Tao, J., Zhou, S., Liang, J., Yang, J., 2022. Dynamic mlp for fine-grained image classification by leveraging geographical and temporal information. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 10945–10954. <https://arxiv.org/abs/2203.03253>.
- Yang, Q.L., Niu, C.Y., Liu, X.Q., Feng, Y.H., Ma, Q., Wang, X.J., Tang, H., Guo, Q.H., 2023. Mapping high-resolution forest aboveground biomass of China using multisource remote sensing data. *GIOSCIENCE Remote Sens.* 60 <https://doi.org/10.1080/15481603.2023.2203303>.
- Yang, X., Zhang, Z., Zhang, L., Fan, X., Ye, Q., Fu, L., 2024. Global superpixel-merging via set maximum coverage. *Eng. Appl. Artif. Intell.* 127, 107212 <https://doi.org/10.1016/j.engappai.2023.107212>.
- Yao, L., Liu, T., Qin, J., Lu, N., Zhou, C., 2021. Tree counting with high spatial-resolution satellite imagery based on deep neural networks. *Ecol. Indic.* 125, 107591 <https://doi.org/10.1016/j.ecolind.2021.107591>.
- Yu, T., Hu, C., Xie, Y., Liu, J., Li, P., 2022. Mature pomegranate fruit detection and location combining improved F-PointNet with 3D point cloud clustering in orchard. *Comput. Electron. Agric.* 200, 107233 <https://doi.org/10.1016/j.compag.2022.107233>.
- Yu, R., Luo, Y., Zhou, Q., Zhang, X., Wu, D., Ren, L., 2021. Early detection of pine wilt disease using deep learning algorithms and UAV-based multispectral imagery. *For. Ecol. Manage.* 497, 119493 <https://doi.org/10.1016/j.foreco.2021.119493>.
- Yuan, J., Ru, L., Wang, S., Wu, C., 2022. WH-MAVS: a novel dataset and deep learning benchmark for multiple land use and land cover applications. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 15, 1575–1590. <https://doi.org/10.1109/JSTARS.2022.3142898>.
- Zhang, Z., Cao, L., Liu, H., Fu, X., Shen, X., 2021b. Assessing the 3-D structure of bamboo forests using an advanced pseudo-vertical waveform approach based on airborne full-waveform LiDAR data. *IEEE Trans. Geosci. Remote Sens.* 59, 10647–10670. <https://doi.org/10.1109/TGRS.2020.3042790>.

- Zhang, J., Gu, H., Yang, Y., Zhang, H., Li, H., 2021a. Research progress and trend of high-resolution remote sensing imagery intelligent interpretation. *Natl. Remote Sens. Bull.* 25, 2198–2210.
- Zhang, J.M., Hu, X.Y., Dai, H.M., Qu, S.R., 2020. DEM Extraction from ALS Point Clouds in Forest Areas via Graph Convolution Network. *Remote Sens.* 12 <https://doi.org/10.3390/rs12010178>.
- Zhao, P., Guan, H., Li, D., Yu, Y., Wang, H., Gao, K., Junior, J.M., Li, J., 2021. Airborne multispectral LiDAR point cloud classification with a feature Reasoning-based graph convolution network. *Int. J. Appl. Earth Obs. Geoinf.* 105, 102634.
- Zhao, F., Sun, R., Zhong, L., Meng, R., Huang, C., Zeng, X., Wang, M., Li, Y., Wang, Z., 2022. Monthly mapping of forest harvesting using dense time series Sentinel-1 SAR imagery and deep learning. *Remote Sens. Environ.* 269, 112822 <https://doi.org/10.1016/j.rse.2021.112822>.
- Zheng, J., Fu, H., Li, W., Wu, W., Zhao, Y., Dong, R., Yu, L., 2020. Cross-regional oil palm tree counting and detection via a multi-level attention domain adaptation network. *ISPRS J. Photogramm. Remote Sens.* 167, 154–177. <https://doi.org/10.1016/j.isprsjprs.2020.07.002>.
- Zhu, F., Gao, J., Yang, J., Ye, N., 2022. Neighborhood linear discriminant analysis. *Pattern Recognit.* 123, 108422 <https://doi.org/10.1016/j.patcog.2021.108422>.
- Zhu, Y., Sun, G., Ding, G., Zhou, J., Wen, M., Jin, S., Zhao, Q., Colmer, J., Ding, Y., Ober, E.S., Zhou, J.i., 2021. Large-scale field phenotyping using backpack LiDAR and CropQuant-3D to measure structural variation in wheat. *Plant Physiol.* 187, 716–738. <https://doi.org/10.1093/plphys/kiab324>.
- Zhu, Y., Li, D., Fan, J., Zhang, H., Eichhorn, M.P., Wang, X., Yun, T., 2023. A reinterpretation of the gap fraction of tree crowns from the perspectives of computer graphics and porous media theory. *Front. Plant Sci.* 14, 1109443. <https://doi.org/10.3389/fpls.2023.1109443>.
- Zuo, C., Qian, J., Feng, S., Yin, W., Li, Y., Fan, P., Han, J., Qian, K., Chen, Q., 2022. Deep learning in optical metrology: a review. *Light Sci. Appl.* 11, 39. <https://doi.org/10.1038/s41377-022-00714-x>.