*Review*

# Sensing and Artificial Perception for Robots in Precision Forestry: A Survey

**João Filipe Ferreira** [1,2,*], **David Portugal** [2,3], **Maria Eduarda Andrada** [2], **Pedro Machado** [1], **Rui P. Rocha** [2,3] **and Paulo Peixoto** [2,3]

1    Computational Intelligence and Applications Research Group, Department of Computer Science, School of
     Science and Technology, Nottingham Trent University, Nottingham NG11 8NS, UK;
     pedro.machado@ntu.ac.uk
2    Institute of Systems and Robotics, University of Coimbra, 3030-290 Coimbra, Portugal;
     davidbsp@isr.uc.pt (D.P.); duda.andrada@isr.uc.pt (M.E.A.); rprocha@deec.uc.pt (R.P.R.);
     peixoto@isr.uc.pt (P.P.)
3    Department of Electrical and Computer Engineering, University of Coimbra, 3030-290 Coimbra, Portugal
*    Correspondence: joao.ferreira@ntu.ac.uk

**Abstract:** Artificial perception for robots operating in outdoor natural environments, including forest scenarios, has been the object of a substantial amount of research for decades. Regardless, this has proven to be one of the most difficult research areas in robotics and has yet to be robustly solved. This happens namely due to difficulties in dealing with environmental conditions (trees and relief, weather conditions, dust, smoke, etc.), the visual homogeneity of natural landscapes as opposed to the diversity of natural obstacles to be avoided, and the effect of vibrations or external forces such as wind, among other technical challenges. Consequently, we propose a new survey, describing the current state of the art in artificial perception and sensing for robots in precision forestry. Our goal is to provide a detailed literature review of the past few decades of active research in this field. With this review, we attempted to provide valuable insights into the current scientific outlook and identify necessary advancements in the area. We have found that the introduction of robotics in precision forestry imposes very significant scientific and technological problems in artificial sensing and perception, making this a particularly challenging field with an impact on economics, society, technology, and standards. Based on this analysis, we put forward a roadmap to address the outstanding challenges in its respective scientific and technological landscape, namely the lack of training data for perception models, open software frameworks, robust solutions for multi-robot teams, end-user involvement, use case scenarios, computational resource planning, management solutions to satisfy real-time operation constraints, and systematic field testing. We argue that following this roadmap will allow for robotics in precision forestry to fulfil its considerable potential.

**Keywords:** sensors; artificial perception; forestry robotics; precision forestry; agricultural robotics

## 1. Introduction

Forestry—the practice of creating, managing, using, conserving, and repairing forests, woodlands, and associated resources—has substantial importance in the economy of many industrial countries [1], and is often overlooked. It provides direct economic gains, but also societal and environmental benefits. Its core industry, i.e., silviculture, which involves the growing and cultivation of trees in order to provide timber and fuel wood as primary products as well as many secondary commodities (e.g., wildlife habitats, natural water quality management, recreation and tourism, landscape and community protection), is an important source of wealth and well-being. Moreover, forests also host many other subsidiaries, high-added-value industries, such as apiculture and forest farming (i.e., the cultivation of non-timber forest products, or NTFPs, such as speciality mushrooms, ginseng, decorative ferns, pine straw or strawberry trees; see [2,3]).

However, for all its potential, forestry is a risky investment with very slow returns. Wood, as a natural resource, renews itself over a long period of time, generally several dozens or even hundreds of years depending on the species, which is too long in terms of investment cycles in the modern economy [4]. Consequently, most companies and private owners convert forest lands into grazing lands or industrial plantations involving single tree species, such as oil palms, acacia mangium, or eucalyptus, which yield a higher rate of return on invested capital, rather than managing secondary forests (or second-growth forest, as opposed to an old-growth, primary or primaeval forest: a woodland area that has regrown after a timber harvest and has undergone a sufficient period of reforestation, such that the effects of the previous harvest are no longer evident) with felling cycles spanning from 20 to 40 years [4]. In fact, this is the only realistic option for private owners in countries with a lower Gross Domestic Product (GDP) per capita. Therefore, achieving widespread long-term conservation and management of forestry resources has been challenging. In many of these countries, forestry is seen more as a strategic commitment for the future rather than an investment in the present.

*1.1. Motivations*

The ongoing decline in the available workforce, primarily caused by low wages and the harshness of forestry operations, along with the gradual abandonment of rural areas and traditional practices like pastoralism, has led to the growing mechanization of the forestry sector, following a trend observed in other industries, including agriculture [5,6]. Modern forestry, therefore, relies on technologically advanced machinery to increase productivity, although its high cost remains a significant barrier for many small private forest landowners. However, as technological advancements continue and equipment prices decrease, the use of these machines is expected to become more widespread in the future [5].

The introduction of (semi-)autonomous vehicles and robots in forestry could potentially solve this problem and contribute to the achievement of several of the United Nations' sustainable development goals (SDGs) [7,8]. For example, it could help to promote sustainable economic growth by increasing productivity and reducing running costs, thereby making forestry a more viable industry for small private forest landowners (SDG 8: decent work and economic growth). It could also contribute to the goal of reducing inequalities by providing skilled job opportunities for people in rural areas (SDG 10: reduced inequalities), potentially reducing the harsh working conditions and health hazards associated with forestry work (SDG 3: good health and well-being). Additionally, the use of advanced machinery and robotics in forestry could help to promote responsible consumption and production by reducing waste, optimising the use of natural resources and protecting the environment (SDG 12: responsible consumption and production; SDG 15: life on land).

In fact, as mentioned above, obtaining direct increases in productivity and safety is not the only reason for the introduction of precision forestry (the use of advanced technologies to improve forest-management results [9]), and robots in particular, to this industry. According to data provided by the European Commission, Europe experiences approximately 65,000 fires annually [10]. More than 85% of the total burnt area is concentrated in Mediterranean countries. Portugal leads these unfortunate statistics, averaging 18,000 fires per year over the past quarter-century. Following Portugal, are Spain, Italy, and Greece, with yearly averages of 15,000, 9000, and 1500 fires per year, respectively. In 2017, Portugal was one of the most severely affected countries worldwide, with 500,000 hectares (almost 1.5 million acres) of burnt areas and more than 100 fatalities [11,12].

Unsurprisingly, wildfires exert a substantial economic impact that extends far beyond the mere depletion of wood as a primary resource. They disrupt the forest's ability to regenerate, thereby causing immeasurable harm to the environment. This sets off a detrimental vicious cycle, wherein rural abandonment hinders effective wildfire monitoring and prevention efforts. Consequently, more people migrate away from rural regions, leading to significant declines in the tourism sector and heightened unemployment rates. This in

turn perpetuates a gradual disinterest in forest management [13]. As one would expect, wildfires also profoundly impact all subsidiary industries.

While significant progress has been made in key areas, the development of fully autonomous robotic solutions for precision forestry is still in a very early stage. This arises from the substantial challenges imposed by the traversability of rough terrain [14], for example, due to steep slopes, autonomous outdoor navigation and locomotion systems [15], limited perception capabilities [16], the real-time processing of sensory data [17], and reasoning and planning under a high level of uncertainty [18]. Artificial perception for robots operating in outdoor natural environments has been the subject of study for several decades. In the case of robots operating within forest environments, research dating back to the late 1980s and early 1990s can be found; see, for instance, Gougeon et al. [19]. Nevertheless, despite many years of research, as documented in various research works conducted during the past two decades [20–24], substantial problems have yet to be robustly solved. Given the significant challenges listed above, we believe it is highly relevant and timely to produce a new survey on sensing and artificial perception in forestry robotics, as it would provide valuable insights into the current state of research, update the current scientific landscape and help identify necessary advancements in this field.

### 1.2. Contributions, Methodology, and Document Structure

In this paper, we extensively describe the current state of the art in artificial perception and sensing for robots in precision forestry. It, therefore, differentiates from research works, such as research by [20,21,24], by being more specific in its scope, and includes and represents an update on the topics reviewed in [22,23]. As explained in the previous section, two important facts have contributed strongly to the fragmentation of research in this field throughout the years: the relatively slow uptake of this kind of technology in precision forestry due to the demanding social and economic structuring of this industry and the challenges of artificial perception for forestry robotics; therefore, we needed to extend our literature review to the past two decades of active research in this field. To overcome the problem posed by such a considerable time frame, we used a wide range of methodologies to survey the scientific and technological landscape under scrutiny. These included information gathered from two workshops that we organised on the subject area, and insights obtained through condensing the collection of individual literature reviews, which were written for our own research papers in the field, following a snowball sampling strategy; we conducted systematic keyword searches corresponding to the topics of each subsequent section and respective subsection in major robotics and computer vision conference repositories and scholarly search engines, such as Google Scholar, Web of Science, Science Direct, IEEE Xplore, Wiley, SpringerLink, and Scopus, with a focus on a narrower and more recent time span of the last 5 years. This methodology enabled the survey of a total of 154 publications on original research concerning approaches, algorithms, and systems for sensing and artificial perception; a tally of this body of surveyed work is presented in Figure 1.
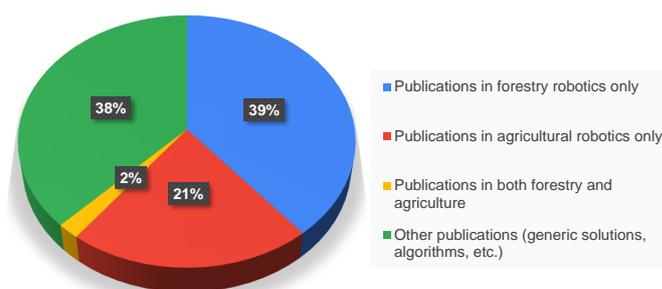


**Figure 1.** Distribution of surveyed works from 2018–2023 according to application area.

Our research work is, therefore, structured as follows:

- We start by enumerating the research groups involved in introducing robots and other autonomous systems in precision forestry and related fields (Section 2).
- Next, we present a review of the sensing technologies used in these applications and discuss their relevance (Section 3).
- We follow this by providing a survey of the algorithms and solutions underpinning artificial perception systems in this context (Section 4).
- We finish by discussing the current scientific and technological landscape, including an analysis of open research questions, and finish by drawing our final conclusions and proposing a tentative roadmap for the future (Section 5).

## 2. Research Groups Involved in the Research and Development of Robots in Precision Forestry and Agriculture

Agricultural robotics in general, and forestry robotics in particular, have attracted the interest of many research groups in all regions of the world. Several high-profile research projects have been awarded to these groups, boosting their scientific productivity and visibility. As a result, we are now able to pinpoint the key scientific players in the field.

### 2.1. Research and Development in Portugal and Spain

The Forestry Robotics group at the University of Coimbra (FRUC—https://www.youtube.com/@forestryroboticsuc, accessed on 1 July 2023), Portugal, in conjunction with the Robotics and Artificial Intelligence Initiative for a Sustainable Environment (RAISE) at Nottingham Trent University (NTU), UK, under the coordination of Ingeniarius, Ltd. (Alfena, Portugal—http://www.ingeniarius.pt, accessed on 1 July 2023), Portugal, have collaborated to produce significant research on forestry robotics in the Safety, Exploration, and Maintenance of Forests with Ecological Robotics (SEMFIRE) project [25], supported by the sister project Centre of Operations for Rethinking Engineering (CORE) [26], which proposes to combine a wide range of technologies into a multi-robot system to reduce combustible material for fire prevention, thus assisting in landscaping maintenance tasks. The key elements of the SEMFIRE solution involve two different types of mobile robotic platforms: the *Ranger* and the *Scouts*. The *Ranger* (Figure 2) is a heavy-duty, multi-purpose tracked robotic mulcher, based on the Bobcat platform. It is equipped with a forestry mulcher attachment to cut down thin trees and shred ground vegetation to grind them into mulch. It can operate in fully autonomous and semi-autonomous modes (with human control). Scouts are small assistive flying robots with swarm self-organising capabilities; they are used to explore and supervise wide forestry areas. Figure 3 illustrates the deployment of the SEMFIRE solution in the field, which was explained with greater detail in [27]. The research efforts from this project are being followed up under the currently running Semi-Autonomous Robotic System for Forest Cleaning and Fire Prevention (SAFEFOREST) project [28], incorporating the Field Robotics Center of the Robotics Institute at Carnegie Mellon University (https://www.ri.cmu.edu/, accessed on 1 July 2023) [29–31], other R&D companies, and private institutions to the consortium, with the aim of developing advanced monitoring and robotic systems to semi-automatically manage forest fuels in wildland and wildland–urban interface (WWUI) areas with complex terrains, in terms of slope and roughness. The proposed project aims to enhance and refine the technology and platforms devised in the SEMFIRE project for the execution of intricate landscaping missions, focusing on the removal of redundant vegetation and clearing fuel breaks and WWUI areas, based on preliminary mapping of the operational theatre carried out with the support of advanced drone terrain and vegetation monitoring techniques.

**Figure 2.** The *Ranger* landscape maintenance robot developed in the SEMFIRE project. For more details, please refer to [32].



**Figure 3.** SEMFIRE *Scout* UAV platform on the left. Illustrative deployment of the SEMFIRE solution on the right: (1) the heavy-duty, multi-purpose *Ranger* can autonomously mulch down the thickest brushes as well as cutting down small trees to reduce the risk of wildfires; (2) the area is explored (finding new regions of interest for landscaping) and patrolled (checking the state of these regions of interest) by *Scouts*, with the additional task of estimating the pose of each other and the *Ranger*, and supervising the area for external elements (e.g., living beings).

The Centre for Robotics in Industry and Intelligent Systems (CRIIS) at the Engineering Institute of Systems, Computers, Technology and Science (INESC-TEC) in Portugal has been working on the development of robotics solutions for field robotics. The team focuses on the localization and mapping in GNSS-denied areas, path planning (e.g., for steep slope vineyards), visual perception, manipulation, and safety research, especially for agriculture and forestry contexts. Relevant projects include BIOTECFOR [33], an Iberian partnership to increase efficiency levels in the use of forest resources. Through the use of intelligent robotic systems for the collection and processing of biomass, it promotes the bio-economy and circular economy in the northern region of Portugal and Galicia. Moreover, the team has been coordinating the SCORPION H2020 EU initiative [34], which focuses on developing a safe and autonomous precision spraying tool for agriculture, integrated into a modular unmanned ground vehicle (UGV), to increase spraying efficiency in permanent crops while

reducing human and animal exposure to pesticides, water usage, and labour costs. Steep slope vineyards have been mainly chosen for demonstration of the project (see [35]).

Spanish institutions have also been very involved in the latest developments in the field, participating in many high-profile research projects. The State Agency High Council for Scientific Research (CSIC), through its Centre for Automation and Robotics and Institutes of Agricultural Sciences and Sustainable Agriculture, has contributed to projects such as RHEA [36] (see also Figure 4), and contributed to the sub-fields of robotic fleets and swarms for agricultural applications [37,38]. The Televitis at Universidad de La Rioja and the Agricultural Robotics Laboratory (ARL) at the Universidad Politecnica de Valencia, which is involved in projects such as VINEROBOT [39], has contributed to a large-scale study on soil erosion in organic farms [40]. Moreover, the VineScout project [41] was the follow-up project of VineRobot; their goal is to produce a solution with a score of 9 on the Technological Readiness Level (TRL) scale of technological maturity [42]. The Polytechnic University of Madrid and the Universidad Complutense de Madrid, both involved in the RHEA project [36], have produced work mainly in perception, namely for olive classification [43], air–ground sensor networks for crop monitoring [44], and crop/weed discrimination [45]. The Centro de Automática y Robótica (CAR at CSIC-UPM), also involved in the CROPS project [46], has produced work on perception, namely on the combination of signals for the discrimination of grapevine elements [47]. Spanish companies, such as Robotnik (https://www.robotnik.eu/, accessed on 1 July 2023), involved in projects such as VINBOT [48] and BACCHUS [49], have contributed to various novel robotic platforms, leading to innovation in navigation techniques [50] for viticulture robots.



**Figure 4.** The RHEA robot fleet on a wheat spraying mission. RHEA focused on the development of novel techniques for weed management in agriculture and forestry, mainly through the usage of heterogeneous robotic teams, involving autonomous tractors and Unmanned Aerial Vehicles (UAVs). Reproduced with permission.

### 2.2. Research and Development throughout the Rest of Europe

The Group of Intelligent Robotics at Umeå University, in Sweden, has conducted substantial and pioneering research in this field, having recently participated in the CROPS [46] and SWEEPER projects. This group has worked in several sub-fields of agricultural and forestry robotics, starting with surveying and designing autonomous vehicles for forest operations [51,52] (Figure 5). The work includes perception, e.g., vision-based detection of trees [53] and infrared-based human detection in forestry environments [54]. This group also worked on the decision-making and actuation aspects of the field, specifically by developing a software framework for these agents [55], path-tracking algorithms [52], and navigation and manipulation control for autonomous forwarders [56]. The Department

of Forest Resource Management of the Swedish University of Agricultural Sciences (SLU) has also contributed to the specific field of forestry, namely with a simulation of harvester productivity in the thinning of young forests [57]. The SLU, through the Departments of Forest Resource Management and Forest Biomaterials and Technology, has also been involved in the CROPS project and contributed regularly to the field. These contributions include works on the design of forestry robots [51] (in cooperation with Umeå University) and perception [58].
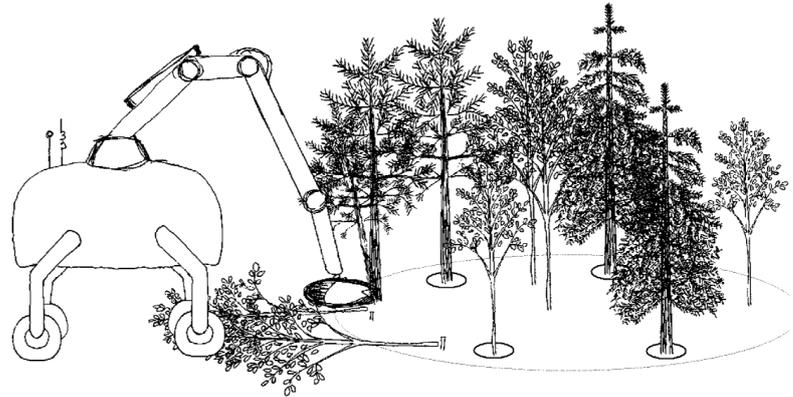


**Figure 5.** An illustration of an autonomous robot performing landscaping on a young forest. The circled trees are the mainstems, and should be kept, while the others are to be cut. Reproduced from [51] with permission.

Wageningen University, in the Netherlands, has also contributed to the field through its Greenhouse Horticulture & Flower Bulbs and Farm Technology groups. These groups have been involved in several high-profile research projects, such as CROPS [46], SWEEPER [59], and SAGA [60], and have worked mainly on the perception abilities of agricultural robots. Contributions include techniques for vision-based localization of sweet-pepper stems [61], obstacle classification for an autonomous sweet-pepper harvester [62], plant classification [63,64], and weed mapping using UAVs [65].

The University of Bonn is part of the Digiforest Horizon Europe project [66]; it produced important work in agriculture and forestry-related perception, namely in crop/weed classification. Noteworthy recent research includes UAVs for field characterization [67], and semi-supervised learning [68] and Convolutional Neural Networks (CNNs) [69,70] for perceptual tasks, such as crop/weed classification.

Research groups in Italy have also contributed to the field. The University of Florence, through the Department for the Management of Agricultural, Food, and Forest Systems (GESAAF), in collaboration with the University of Pisa's Department of Agriculture, Food, and Environment, have participated in the RHEA project [71]. The University of Milan, through the CROPS project, has produced work on vision-based detection of powdery mildew on grapevine leaves [72].

Several groups in France, Belgium, and Switzerland have also led to several related research efforts. The National Institute of Research in Science and Technology for Environment and Agriculture (IRSTEA), a French research institute, has participated in the RHEA project [38], and has produced work on a new method for nitrogen content assessment in wheat [73]. The University of Liège, through its Laboratory of Forest Resources Management, has produced works on the discrimination of deciduous tree species [74] and on the classification of riparian forest species [75], both from data obtained from UAVs, and additional work on estimating the nitrogen content in wheat [76]. ETH Zurich has produced seminal work on precision harvesting by explicitly addressing several key aspects for the operation of fully autonomous forestry excavators, such as localization, mapping, planning, and control [77,78]. The Robotic Systems Lab has an important track record in all-terrain legged robotic platforms and has participated in relevant research projects, such as the EU Horizon Digiforest [66] and EU H2020 THING [79]. Moreover, the Autonomous

Systems Lab at ETH Zurich also performs investigations on related outdoor perception aspects, such as semantic segmentation for tree-like vegetation [80] and weed classification from multispectral images [81].

In the United Kingdom, the Lincoln Centre for Autonomous Systems at the University of Lincoln has been contributing to the development of RASberry, a fleet of robots for horticultural industry [82], and a 3D vision-based crop–weed discrimination for automated weeding operations [83]. The Agriculture Engineering Innovation Centre at the Harper Adams University, also from the United Kingdom, and host to the National Centre for Precision Farming, has been developing a vision-guided weed identification and a robotic gimbal, which can be mounted with a sprayer or laser to eradicate weeds [84].

A few Eastern European groups, namely through the CROPS project, have added their own contributions. The Aleksandras Stulginskis University, in Lithuania, has tackled the problem of UAV-based perception, producing a system for estimating the properties of spring wheat [85]. The Universitatea Transilvania Brasov, in Romania, has contributed to the issue of automating data collection in studies regarding the automation of farming tasks [86]. The Agricultural Institute of Slovenia has contributed to a real-time positioning algorithm for air-assisted sprayers [87].

*2.3. Research and Development throughout the Rest of the World*

Several groups operating outside of Europe have also shown interest in this field. The Northern Robotics Laboratory (NOR) at Laval University in Canada is a research group that specializes in mobile and autonomous systems in harsh outdoor conditions (see project SNOW [88]). Relevant research work includes SLAM in forest environments [89,90] and autonomous navigation in subarctic forests [91]. The Massachusetts Institute of Technology, namely through the Robotic Mobility Group, has produced work on robotic perception for forested environments. This work includes techniques to detect and classify terrain surfaces [92] and to identify tree stems using Light Detection and Ranging (LiDAR) [93]. The Autonomous and Industrial Robotics Research Group (GRAI) from the Technical University of Federico Santa María in Chile has also been very active in both agricultural robotics with work on fleets of N-trailer vehicles for harvesting operations [94] and forestry robotics, with work on UAV multispectral imagery in forested areas [95] and multispectral vegetation features for digital terrain modelling [96].

Similarly, the Ben-Gurion University in Israel was also involved in the CROPS [46] and SWEEPER projects. This group has contributed to an important review that focused on harvesting robots [61] and has recently worked on sweet pepper maturity classification [97] and grapevine detection using thermal imaging [98], paving the way for future efforts in automated harvesting. Another important group is the forestry research group from the Australian Centre for Field Robotics at the University of Sydney; this group has been working on tree detection [99,100] and the use of LiDAR and UAV photogrammetry in forestry resources [101].

Lastly, several research groups are producing emergent developments in this field, despite not having a significant scientific output at this point. An example is the Robot Agriculture group, a community based on the robot operating system (ROS) software framework (http://www.ros.org, accessed on 1 July 2023); the group is looking to apply it to agricultural and forestry operations [102]. Similarly, several recent research projects and respective consortia are also noteworthy, such as SAGA [60], GREENPATROL [103], and SWEEPER [59] (Figure 6).

Table 1 presents an overview of the most relevant research groups involved in this field. We can observe that a substantial amount of work is devoted to applications in agricultural robots, and not forestry *per se*. Indeed, agricultural robots have received significant attention from the research community; see, for example, [104–106]. However, the fundamental research on locomotion, perception, and decision-making should still be largely applicable, as both applications share many of the same challenges: irregular terrains, perception issues introduced by natural conditions, etc. Nonetheless, in general,

the unstructured nature of forests makes robotics for forestry a particularly demanding application area.

**Table 1.** Overview of the most relevant research groups involved in the development of robotic solutions for agriculture and forestry.

| Group * | Country | Main References | Research Projects | Main Focus and Applications |
|---|---|---|---|---|
| FRUC, University of Coimbra | Portugal | [27,32,107,108] | SEMFIRE, CORE, SAFEFOREST | Perception and decision-making for forestry robots. |
| RAISE, Nottingham Trent University | UK | [27,32,107,108] | SEMFIRE, CORE, SAFEFOREST | Perception and decision-making for forestry robots. |
| Carnegie Mellon University | USA | [29–31] | SAFEFOREST | Perception for forestry robots: navigation, mapping, classification and vegetation detection |
| CRIIS, INESTEC | Portugal | [33,35,109,110] | SCORPION, BIOTECFOR | Robotics in industry and intelligent systems for agriculture and forestry environments |
| Centre for Automation and Robotics (CAR at CSIC-UPM) | Spain | [37,44] | RHEA, CROPS | Robot fleets and swarms for agriculture, crop monitoring. |
| Televitis at Universidad de La Rioja | Spain | [40] | VINEROBOT, VineScout | Perception and actuation for agricultural robots for vineyard monitoring. |
| Umeå University | Sweden | [51,52,56] | CROPS, SWEEPER | Perception, manipulation, and decision-making for forestry and agricultural robots, including simulation and literature research. |
| Wageningen University | Netherlands | [61,65] | CROPS, SWEEPER, SAGA | Perception for agricultural robots: crop/weed classification, plant classification, weed mapping. |
| University of Bonn | Germany | [68,69] | Digiforest | Perception for agricultural robots: crop/weed classification using multiple techniques. |
| University of Milan | Italy | [72] | CROPS | Perception for agricultural robots: detection of powdery mildew on grapevine leaves. |
| IRSTEA | France | [73] | RHEA | Perception for crop monitoring: nitrogen content assessment in wheat. |
| University of Liège | Belgium | [74,76] | n/a | Perception for forestry and agricultural robots: discrimination of deciduous tree species and nitrogen content estimation. |
| ETH Zurich | Switzerland | [77,78,80,81] | Digiforest, THING | Fully autonomous forest excavators and vegetation detection and classification. |
| University of Lincoln | UK | [82] | RASberry, BACCHUS | Fleets of robots for horticulture and crop/weed discrimination for automated weeding. |
| Harper Adams University | UK | n/a | L-CAS | Perception and actuation for agricultural robots: vision-guided weed identification and robotic gimbal for spray- or laser-based weeding. |
| Aleksandras Stulginskis University | Lithuania | [85] | n/a | Perception for agricultural robots: UAV-based spring wheat monitoring. |
| Universitatea Transilvania Brasov | Romania | [86] | n/a | Automation of data collection in farm automation operations. |
| Agricultural Institute of Slovenia | Slovenia | [87] | CROPS | Perception for agricultural robots: real-time position of air-assisted sprayers. |
| Laval University | Canada | [88–91] | SNOW | Automated forestry, SLAM and navigation in forest environments. |
| Massachusetts Institute of Technology | USA | [92,93] | n/a | Perception for forestry robots: terrain classification and tree stem identification. |
| Technical University of Federico Santa María | Chile | [94–96] | n/a | Multispectral imagery perception in forests and N-trailers for robotic harvesting. |
| Ben-Gurion University | Israel | [61,97,98] | CROPS, SWEEPER | Agricultural harvesting robots, including literature survey. |
| University of Sydney | Australia | [99–101] | n/a | Tree detection, LiDAR and UAV photogrammetry. |

\* The groups are ordered according to their appearance in the text.

(**a**)                                                                                          (**b**)

**Figure 6.** The Sweeper robot (**a**), a sweet-pepper harvesting robot operating in a greenhouse. (**b**): the output of Sweeper's pepper detection technique. The Sweeper project aims to develop an autonomous harvesting robot, based on the developments of the CROPS project, which can operate in real-world conditions. Reproduced with permission. Source: www.sweeper-robot.eu, accessed on 30 June 2023.

## 3. Sensing Technologies

Perception in forestry robotics faces many challenges imposed at the sensor level by extrinsic environmental conditions (trees and relief, weather conditions, dust, smoke, etc.), among other intrinsic technical challenges brought by robot operations; see Figure 7 for examples. Many different types of sensors have been used in an attempt to improve robustness at the source, including low-cost solutions, such as 2D and 3D, red–green–blue channel (RGB) and near- to far-infrared camera setups, laser technologies, sonars, and higher-end solutions, such as LiDAR and/or Laser Detection and Rangings (LaDARs). Table 2 shows a systematic comparison of the most important sensing technologies used in forestry robotics.

**Table 2.** Comparison of the most important sensor technologies used in forestry robotics.

| Sensor Technology | Sensing Type | Advantages | Disadvantages |
|---|---|---|---|
| RGB camera | imaging sensor | allows for relatively inexpensive high-resolution imaging | does not include depth information |
| RGB-D camera | imaging and ranging sensor | relates images to depth values | generally low-resolution to reduce costs |
| thermal camera | temperature imaging sensor | temperature readings in image format can improve segmentation and help detect animals | generally low-resolution and more expensive than normal cameras |
| hyperspectral sensor | imaging sensor w/ many specialised channels | allows for better segmentation (e.g., using vegetation indices) | expensive and heavy-duty when compared to other imaging techniques |
| multispectral camera | imaging sensor w/ some specialised channels | allows for better segmentation (e.g., using vegetation indices); inexpensive | less powerful than its hyperspectral counterpart |
| sonar | sound-based range sensor | allows for inexpensive obstacle avoidance | limited detection range and resolution |
| LiDAR/LaDAR | laser-based range sensors | allow for precise 3D sensing | relatively expensive and difficult to extract information beyond spatial occupancy |
| electronic compass | orientation sensor | allows for partial pose estimation | may suffer from magnetic interference |
| inertial sensors | motion/vertical orientation sensors | allow for partial pose estimation | suffer from measurement drift |
| GPS/GNSS | absolute positioning sensors | allow for localization and pose estimation | difficult to keep track of satellite signals in remote woodland environments |

To overcome individual drawbacks and take advantage of possible synergies, multisensory setups have been proposed as well. Positioning, orientation, and navigation systems, such as electronic compasses, inertial sensors, and GPS or GNSS, have been used to complement these sensory setups. However, these systems exhibit their own robustness issues, resulting, for example, from drift, the lack of resolution, or occlusions [111].

In [112], technologies, techniques, and strategies are surveyed in the areas of sensor fusion, sensor networks, smart sensing, Geographic Information Systems (GIS), photogrammetry, and other intelligent systems. In these systems, finding optimal solutions to the placement and deployment of multi-modal sensors covering wide areas is important. Technologies, such as radar range sensors, thermal (infrared), visual (optical), and laser range sensors, are reviewed, and sensor integration for cooperative smart sensor networks, sensor hand-off, data fusion, object characterization, and recognition for wide-area sensing are discussed in-depth. In the following subsections, we list and summarily describe technologies that are the most relevant, specifically for forestry robotics.

(**a**) Image homogeneity.

(**b**) Lens flare and overexposure due to direct sunlight.

(**c**) Decreased visibility due to fog or rain.

(**d**) Unexpected motion effects due to wind.

(**e**) Effects of perspective changes.

(**f**) Blurring due to robot motion.

**Figure 7.** Sensing challenges for forestry robotics.

### 3.1. Cameras and Other Imaging Sensors

Plants are part of a complex biological ecosystem, which makes spatiotemporal information crucial for their analysis [113]. A wide range of sensors can be used to monitor the health status of a plant. Khanal et al. [114] describe potential applications of thermal

remote sensing in precision agriculture. The use of Short-Wave Infrared (SWIR) enables scientists to infer the crops' health, crop water stress, soil moisture, plant diseases, crop maturity, etc.; see Figure 8. Such information is required for planning irrigation strategies, soil moisture compensation, water stress monitoring, evapotranspiration, drought stress monitoring, residue cover detection, and crop yield estimation [114].

Hyperspectral imaging has also been used for monitoring the health of growth of plants [115]. A hyperspectral image could have hundreds or thousands of bands. In general, it is produced by an imaging spectrometer. Unlike standard RGB imaging, which captures the wavelengths of approximately 475 nm (blue channel), 520 nm (green channel), and 650 nm (red channel), hyperspectral imaging extends the spectrum, typically from ultraviolet (UV; starting at ~250 nm) up to SWIR (~2500 nm) [115]. The visible and near-infrared ranges are particularly important for plant monitoring and analysis of their features; leaf pigmentation can be inferred from 400 to 700 nm and mesophyll cell structure from 700 to 1300 nm; however, a 1300–2500 nm range is required to measure the water content of a plant [115].

More specifically, the use of hyperspectral imaging is desirable for computing, among others, the following indices [115]:

- Normalised Difference Vegetation Index (NDVI) [680, 800] nm: to improve chlorophyll detection and measure the general health status of crops, the optimal wavelength varies with the type of plant;
- Red edge NDVI [705, 750] nm: to detect changes, in particular, abrupt reflectance increases at the red/near-infrared border (chlorophyll strongly absorbs wavelengths up to around 700 nm);
- Simple Ratio Index (SRI) [680, 800] nm;
- Photochemical Reflectance Index (PRI) [531, 570] nm;
- Plant Senescence Reflectance Index (PSRI) [520, 680, 800] nm;
- Normalised Phaeophytization Index (NPQI) [415,435] nm;
- Structural Independent Pigment Index (SIPI) [445, 680, 800] nm;
- Leaf Rust Disease Severity Index (LRDSI) [455, 605] nm: to allow detection of leaf rust.
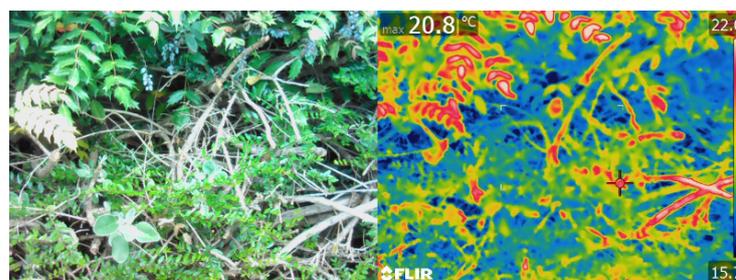


**Figure 8.** RGB (left) and thermal (right) images of bushes and shrubbery captured using a thermal camera. A variation of about 7 °C exists in the heating distribution in the thermal image. Such a temperature variation will have an impact on the overall plant water stress and, therefore, on its health.

The NDVI, in particular, is one of the most used in farming/forestry applications. For example, the red channel of the image is used alongside the Visible-to-Short-Wave-Infrared (VSWIR) (from 750 nm up to about 2500 nm) for calculating a coefficient [−1, 1] and, therefore, infers the water stress of a given plant [113,116]. Also, the NDVI can be used to measure the soil water evapotranspiration. Examples of robotic systems that use hyperspectral imaging include [117–119].

A related type of sensor that is particularly promising for artificial perception in natural environments, which has been increasingly used in precision agriculture and geographic research using unmanned aerial vehicles is the multispectral camera, e.g., [74,120]. For example, this type of sensor significantly improves the discrimination between vegetation and

non-vegetation due to its sensitivity to chlorophyll [121]. More importantly, multispectral cameras have been successfully used to discriminate between different species throughout the so-called vegetation index [122], for instance, to separate weeds from plants [123–125], and to detect diseased crops [126]. Multispectral imaging is often used instead of hyperspectral imaging due to its considerably lower cost. It uses radiometers and generally involves only 3 to 10 bands [127].

*3.2. Range-Based Sensors*

With recent advances in hardware and sensor technology, 3D LiDARs are increasingly being used in field robotics, primarily because they can produce highly accurate and detailed 3D maps of natural and man-made environments and can be used in many contexts due to their robustness to dynamic changes in the environment. In addition, the cost of LiDAR sensors has decreased in recent years, which is an important factor for many application scenarios. Currently, the most popular 3D LiDARs sensors are produced by Velodyne (https://www.velodynelidar.com/, accessed on 1 July 2023), with a very strong presence on the ever-growing market of self-driving cars; see reviews on some of these sensors in [128,129]. This type of sensor has been used in a variety of outdoor applications, such as mapping rough terrain in Disaster City, the world's largest search-and-rescue training facility, as reported by Pellenz et al. [130]. The authors captured a 15 min dataset, integrating Inertial Measurement Unit (IMU) data, GPS data, and camera data. To map the scenario, two 3D techniques were tested: (i) the standard Iterative Closest Point (ICP) approach [131]; and (ii) 6D Simultaneous Localization and Mapping (SLAM) (simultaneous localization and mapping (SLAM) is the computational problem of constructing or updating a map of an unknown environment while simultaneously keeping track of an agent's location within it [132]) software from Nüchter et al. [133], while classifying the terrain online using a principal component analysis (PCA) (a principal component analysis (PCA) is a statistical method that reduces data dimensionality by creating new uncorrelated variables, called principal components, to capture the most significant variance in the data [134])-based approach [135], yielding surprisingly precise results. These advances in 3D LiDARs technology not only represent a significant step for artificial perception for field and service robots, but also in other key areas, such as wearable reality capture systems, as seen in [136], where the Pegasus Backpack from Leica Geosystems is presented as a wearable sensing backpack that combines cameras and LiDARs profilers for real-time mobile mapping.

Despite being used over the last couple of decades for a variety of robotic-perception-related tasks, Laser Range Finders (LRFs) have not been properly explored for vegetation identification, until the seminal work of [137], where a novel approach for detecting low, grass-like vegetation using 2D LiDAR (a SICK LMS291-S05) remission values was presented. The approach relies on a self-supervised learning method for robust terrain classification by means of a support vector machine using a vibration-based terrain classifier to gather training data. With this, the authors were able to obtain a terrain classification accuracy of over 99% in the results presented, and improve the robot navigation in structured outdoor environments.

## 4. Artificial Perception: Approaches, Algorithms, and Systems

Although the right choice of sensors and sensor fusion certainly helps, we believe that the major challenges for artificial perception in outdoor environments will only be addressed by the development of an encompassing perceptual system. In fact, this system should be able to support high-level decision layers by bridging all the appropriate perceptual to action tasks, e.g., active perception, foreground-background segmentation, entity classification in forest scenarios, odometry (in particular visual), trail following and obstacle avoidance for navigation. Additionally, the system should efficiently utilize sensing and computational resources in order to achieve the best possible trade-off between

minimising the energy footprint, to prolong robot operational autonomy and allow for real-time performance of mission-critical sub-systems.

An artificial perception system for robots in our particular application of interest simultaneously faces the following major challenges:

- It should allow the robot to navigate through the site while effectively and safely avoiding obstacles under all expected environmental conditions.
- It should equip the robot with the capacity to ensure the safety of both humans and local fauna.
- It must allow the robot to find, select, and act appropriately, with respect to the diverse vegetation encountered in the target site, according to the designated task and the tree species comprising forest production for that site, namely in distinguishing between what should be protected and what should be removed, as defined by the end user.
- Finally, its outcomes should go beyond reproducing a layman's perspective, and effectively be modulated by the specifics of tasks informed by expert knowledge in forestry operations.

In the following subsections, we will present an overview of the state of the art of artificial perception algorithms, solutions, full-fledged systems, and architectures for forestry applications and related fields.

*4.1. Forest Scene Analysis and Parsing*

To perform tasks in forestry applications, robots must be able to take a scene retrieved by their sensors and parse it in such a way that they are able to distinguish and recognize entities crucial to the task at hand, namely:

- The object of the task, for example, a specific plant and any part of that plant, plants in general of the same species of interest, etc.;
- Any distractors, such as other plants (of other species, or of the same species but not the particular plant to be acted upon, etc.);
- Secondary or ancillary entities to the task, such as humans (co-workers or by-standers), animal wildlife, navigation paths, obstacles to navigation and actuation, geological features (ridges, slopes, and any non-living object), etc.

They also need to be able to tie this semantic parsing of the forest scene to a spatial representation of its surroundings to appropriately perform its operations. There are forestry applications in which specific objects, such as stumps, logs, and standing trees, need to be targeted or avoided. In these cases, simple object detection methods and spatial representations may suffice to perform a gripping task or an evading manoeuvre in closed-loop and real-time fashion. However, in our experience, more complex applications involving intricate operations and long-term mapping and planning in the theatre of operations need a richer, preferably a three-dimensional, integrated perception of the scene. The process of building this perception is called metric-semantic mapping, which includes three stages in its pipeline:

- Semantic segmentation;
- Volumetric mapping;
- Semantic label projection.

Spatial overlap and the challenges of sensing mentioned in Section 3 require a potentially probabilistic approach for the projection process to deal with the uncertainty inherent to perception. Research on approaches solving these three stages and their integration will be described next.

4.1.1. Image Segmentation: Object Detection and Semantic Segmentation

Image segmentation is defined as the partitioning of the image into segments or regions to simplify the parsing of the perceived scene. This generally involves locating and classifying objects by assigning labels to pixels based on shared characteristics. In the artifi-

cial perception of agricultural and forestry robotics, simple and efficient object detection and classification methods have often been used to perform parsing, including deep learning algorithms using bounding box partitioning; for recent examples, see [138–140]. These have the significant advantage of being suited for real-time implementations (Section 4.5 will discuss these issues in more depth). Semantic segmentation, on the other hand, involves assigning class labels to individual atomic parts in a sensory representation (e.g., pixels in an image). In image processing, for example, its goal is to accurately classify and localize objects within the image by grouping pixels into semantically meaningful, task-relevant categories. This allows for more complex, spatially-involved applications, as opposed to bounding box solutions.

Semantic segmentation for vegetation detection in general and plant species discrimination in particular has attracted a lot of interest in the past few years. A considerable amount of work has recently been spurred by the botanical scientific community (e.g., see [141]), leading to the development of several platforms and applications for plant identification for taxonomy, such as PlantSnap (see [142], the PlantSnap site (https://www.plantsnap.com/, accessed on 1 July 2023)) and Pl@ntnet (see [143,144], and the Pl@ntnet site (https://plantnet.org/en/, accessed on 1 July 2023)). Sun et al. [145] recently proposed a plant identification system using deep learning, which is claimed to satisfy both the botanical and computer vision communities. Furthermore, for many years, there has been extensive research on agricultural robotics, particularly in crop–weed discrimination (e.g., see [45,64,69,122,124,125,146–148]). There is also work on vegetation segmentation for robot navigation, as well as studies on path-following and traversability for outdoor robotics in areas outside of urban environments [93,121,149,150]. Finally, additional research efforts have focused on vegetation classification for very specific applications. For instance, Geerling et al. [63] presented a methodology for the classification of floodplain vegetation. Their approach involved the fusion of structural and spectral data obtained from LiDARs and CASI systems.

Over the past decade, semantic segmentation for tree species discrimination, in particular in forestry contexts, has also received considerable attention; however, a substantial portion of this research is focused on the processing of satellite or aerial images (see, for instance, [74,151,152]). These works, while interesting, have only marginal relevance in forestry field robotics, as most robots in this context operate at the ground level. In contrast, a smaller subset of this research has been specifically dedicated to robots in forestry applications or toward processing ground-level images, often referred to as "natural images" (see, for example, [53,153–157]); Figure 9 shows an example of semantic segmentation applied to a specific application in forestry robotics.

Instance segmentation has received much less attention in these contexts. The first recent exception to this rule would be the work by Fortin et al. [158], who focused on instance segmentation for autonomous log grasping in forestry operations. The authors compared three neural network architectures for log detection and segmentation: two region-based methods and one attention-based method. They concluded that the results indicated the potential of attention-based methods for this specific task, as they operated directly at the pixel level, and suggested that such a perception system could be used to assist operators or fully automate log-picking operations in the future. Li et al. [159] tackled a similar problem and proposed a metric learning-based instance segmentation algorithm for log-end face segmentation. The metric learning framework is used to reclassify pixels in the overlapping area, improving the accuracy of instance segmentation by 7% compared to state-of-the-art methods. Additionally, the proposed model demonstrates a faster processing speed and better segmentation of log-end faces of smaller scales. It effectively handles occlusion situations during the shooting process, making it a flexible and intelligent solution for log-end face segmentation in practical production. Another recent work addressing instance segmentation was presented by Grondin et al. [160], who compared two different model architectures, Mask R-CNN and Cascade Mask R-CNN, for tree detection, to segment and estimate the felling cut, diameter, and inclination. For each input image, one of three tested

backbones, the CNN-based ResNext, the transformer-based Swin, or CBNetV2, which uses dual backbones (which can be CNN-based or transformer-based), extracts distinctive features. Then, the first network head predicts the class and bounding box, the second head predicts the segmentation mask, and the third head predicts keypoints. This unified architecture enables end-to-end training without any post-processing. Two densely annotated trunk detection image datasets—the first including 43 K synthetic images and the second 100 real images—were acquired for the bounding box, segmentation mask, and key-point detections to assess the potential of the methods under comparison, together with the dataset by da Silva et al. [139] (see Section 4.1.5) to perform data augmentation (data augmentation and other methods for improving learning are presented in Section 4.1.6) and help measure the extent to which the models generalised to a different woodland environment. The models tested in this work achieved a precision of 90.4% for tree detection, 87.2% for tree segmentation, and centimetre-accurate keypoint estimations.
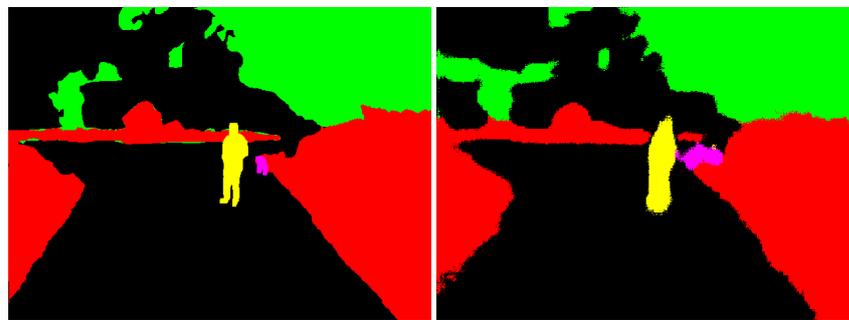


**Figure 9.** Example output of a semantic segmentation model applied to the robotic perception pipeline, designed to perform landscaping in woodlands to reduce the amount of living flammable material (aka "Fuel") for wildfire prevention presented in [157]. The ground-truth image is shown on the left and the corresponding prediction is on the right. The model takes multispectral images as inputs and the classes used for segmentation and respective colour-coding are as follows: "Background" (black), "Fuel" (red), "Canopies" (green), "Trunks" (brown; not present in this example), "Humans" (yellow) and "Animals" (purple). The model consists of an AdapNet++ backbone, an eA-SPP progressive decoder, and fine-tuning trained on Bonnetal, using ImageNet pre-weights for the whole model.

Research on plant part segmentation, on the other hand, has seen a recent rise in interest, going beyond solutions based on more traditional image processing methods of the past, such as the work by Teng et al. [161], who used this type of approach for tree part detection for tree instance segmentation. Examples of such recent research would include the work by Sodhi et al. [162], who proposed a method using 3D imaging for in-field segmentation and the identification of plant structures, work by Barth et al. [163], who proposed a solution using CNNs to improve plant part segmentation by reducing the dependency on large amounts of manually annotated empirical images, and work by [47,62], who used multispectral images to segment plants into leaves, stems, branches, and fruit. Nonetheless, many recent solutions still attempt to solve this problem using traditional approaches; an example would be the work by Anantrasirichai et al. [164], who proposed and compared two methods, watershed and graph–cut-based, respectively, for leaf segmentation in outdoor images.

The presence of humans in the vicinity of forestry machines is also an important aspect that must be perceived autonomously, which is often mission-critical. To this end, ref. [54] presented a technique for segmenting humans in thermal images taken in forestry environments at night. Several techniques, such as kNN, Support Vector Machines (SVMs), and naïve Bayes were tested, yielding 80% precision and 76% recall.

Most of the work reviewed above suffers from a significant reliance on human discretion in choosing the appropriate framing of the visual scene, making it not adaptive enough for use "in the wild". For example, the apps for botanical taxonomy require taking

well-framed, reasonably centred images of very specific parts of the plant, such as its flowers. Some of these works also assume very particular segmentation tasks, such as discriminating plants of a specific species from everything else. Moreover, in general, there seems to be no solution that allows for holistic parsing of forest scenes in order to take into account all or even most of the three major requirements for autonomous robots operating in forestry applications specified at the beginning of this section. This would suggest the lack of all-encompassing perception systems, a subject that will be covered in Section 4.3.

None of these methods explicitly take into account measures of uncertainty, which would be extremely useful in task-relevant decision-making processes or for probabilistic updating in the projection stage of metric-semantic mapping (Section 4.1.3). For example, we are aware of only a few publications on semantic segmentation currently using Bayesian neural network (BNN) for semantic segmentation. Recent work by Dechesne et al. [165] adapted a unimodal U-net architecture with supervised training in Earth observation images, which achieves state-of-the-art accuracy in multiple datasets, while Mukhoti et al. [166] adapted unimodal DeepLabv3+ with supervised learning in the Cityscape dataset, which contains RGB imagery and labelled data in an urban scenario, to achieve comparable results to the original DeepLabv3+ architecture. Finally, Kendall et al. [167] adapted SegNet [168] to achieve a significant improvement over the original model in urban and indoor environments. In all of the above work, implementations of the fully convolutional neural network (FCN) were adapted to incorporate Bayesian inference by adding a Monte Carlo Dropout (MCD) layer at the end of a convolution block that outputs a distribution instead of a fixed value. Overall, BNNs achieved comparable results to current FCNs in modelling uncertainty, which is a valuable feature for a semantic map. Despite the improvements achieved with current BNN implementations, there are still many research opportunities that can be considered, e.g., multimodal BNNs for semantic segmentation.

In recent years, sensors, such as three-dimensional (3D) LiDAR, have begun to be used as alternatives or complementary sensory modalities for object classification and semantic segmentation. This is because these types of sensors are more robust to changes in environmental conditions and are better suited for applications that require depth and position information, such as outdoor robot navigation. The problem with these sensors, however, is that they output large amounts of 3D data in the order of a few million points per second when high-resolution LiDAR is used, making it difficult to process large amounts of 3D data in real time. Moreover, these data are generally unstructured and unordered.

In the past, point cloud processing for artificial perception was based on hand-crafted features [169–172] and was mostly limited to object detection. Hand-crafted features do not require large training data and were rarely used due to insufficient point cloud data; furthermore, this traditional approach was quickly abandoned with the advent of Deep Learning. An overview of techniques based on hand-crafted features can be found in [173].

Due to the increasing performance of deep learning algorithms in computer vision on two-dimensional (2D) image data, researchers have turned to deep learning for processing 3D point data. However, there are many challenges in applying Deep Learning to 3D point cloud data. These challenges include occlusions caused by cluttered scenes or blindsides, noise/outliers that are unintended points, misaligned points, etc. Moreover, CNNs are mainly designed to process ordered, regular, and structured data, so point clouds pose a major challenge. Early approaches overcome these challenges by converting point clouds into a structured intermediate representation.

Several strategies have been proposed for representing 3D LiDAR data before inputting them into a convolutional neural network model. Some of these strategies include rasterising data as 2D images [174] or voxels [175], rasterising data as a series of 2D images acquired from multiple views [176], converting 3D data to grids [177], converting 3D points to the Hough space [178], and converting 3D points to range images [179].

Recently, however, researchers have developed approaches that harness the power of deep learning directly on the raw point cloud without the need for conversion to a structured representation. Recent research shows that it is possible to use 3D data points as

input to a deep neural network. The problem of data irregularity was studied and solved by modelling a symmetric function and designing a transformer network. PointNet [180] and PointNet++ [181] have better accuracy than the best performing state-of-the-art methods, including the rasterization of 3D data into image methods. However, both methods suffer from the problem that they cannot be executed in real-time when there are more than one million points to process. RandLA-Net [182] proposes a solution to this problem that has been shown to perform better than previous methods. The authors use a random selection of points and drop the unselected points before inputting them into a deep neural network. Several other sampling methods can be found in the literature. Some use a heuristic sampling approach, such as farthest point sampling (FPS) [183] and inverse density importance sampling (IDIS) [184], and others are learning-based sampling methods, such as generator-based sampling (GS) [185], Gumbel subset sampling (GSS) [186], and policy gradient-based sampling (PGS) [187]. Despite the good coverage of the entire point cloud and the good representation of the data, the computational costs of these methods are usually high because computationally intensive algorithms or memory-inefficient sampling techniques are used to select the points, limiting their usefulness for real-time applications. Others, such as GS, are simply too difficult to learn. Recently, Mukhandi et al. [188] proposed a method that performs a systematic search and selects points based on a graph-colouring algorithm instead of a random selection of points, and it outperforms RandLA-Net. Figure 10 shows a qualitative evaluation of the semantic segmentation results obtained with this method. For a more detailed review of the literature on deep learning using 3D point clouds, see [189].
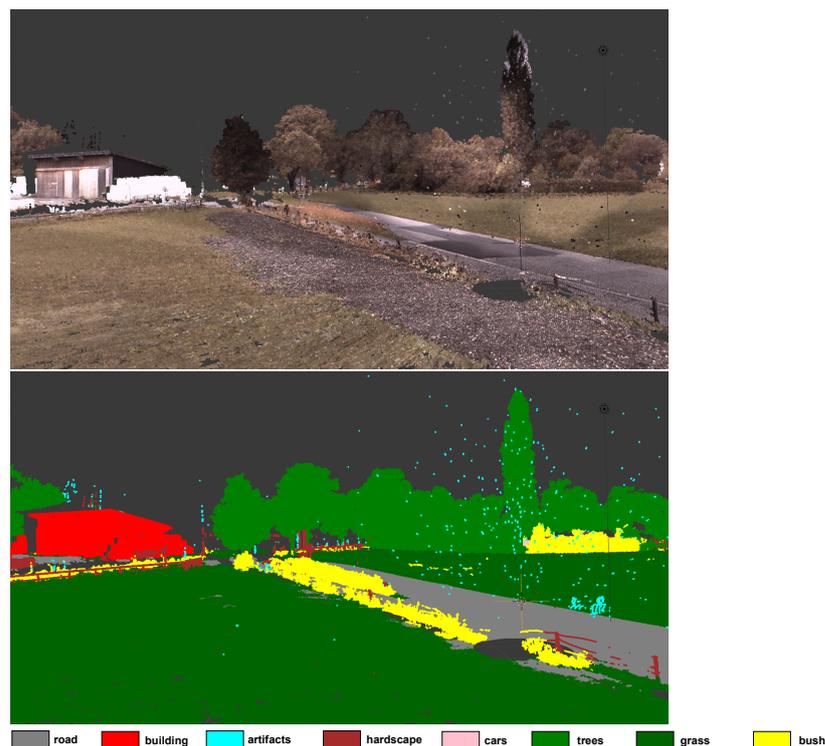


| | road | | building | | artifacts | | hardscape | | cars | | trees | | grass | | bush |

**Figure 10.** Example of the results of semantic segmentation when applied directly to a raw point cloud. The top image shows the original point cloud and the bottom image shows the result of semantic segmentation [188], considering eight different classes (most of which are represented in the example).

In summary, semantic segmentation is crucial for the development of a complete artificial perception system, as it allows the robot to have a better understanding of the scene and its respective context. This semantic context can then be used by the decision-making module for the robot to be able to perform task-relevant operations.

### 4.1.2. Spatial Representations in 3D and Depth Completion Methods

Digital cameras sense the surrounding environment by reading lights that are reflected from the surfaces of objects in the perceptual scenes. A depth image or depth map is an image in which each pixel-wise element stores the distance from its corresponding sensor to the surface from which light is reflected with respect to the camera's referential origin, thereby describing the surrounding geometry from the camera viewpoint and within its field of view (FOV). Depth maps, therefore, provide a 2.5D representation of the camera view. A point cloud, on the other hand, is a collection of points in three-dimensional space commonly representing samples of surfaces in the perceptual scene and often associated with properties, such as surface colour. Point clouds, contrary to depth images, are device-agnostic; they may be derived from depth maps by projecting the 2.5D representations into 3D spaces, but they are also the native representations of sensor readings, from sensors such as LiDARs, which typically generate dense point maps as a result of their scans.

Workaround solutions have been, at times, used to circumvent the lack of depth information in camera images. For instance, recently, Niu et al. [190] proposed a low-cost navigation system tailored for small-sized forest rovers, including a lightweight CNN, to estimate depth images from RGB input images from a low-viewpoint monocular camera. The depth prediction model proposed in this research is designed to infer depth information from low-resolution RGB images using an autoencoder architecture. The goal of the depth prediction model is to provide real-time depth inference on an embedded computer, with sufficient accuracy to facilitate navigation in forest environments. By inferring the depth from RGB images, the model allows for the use of geometric information for navigation while avoiding the high cost of depth sensors. The model's performance was extensively tested and validated in both high-fidelity simulated forests and real-world forest environments under various weather conditions and times of the day. The results demonstrate the model's ability to successfully predict depth from a $16 \times 16$ depth image generated from a $32 \times 32$ monocular RGB image, enabling effective navigation through forest terrains, including several obstacles, such as shrubs, bushes, grass, branches, tree trunks, ditches, mounds, and standing trees.

Dense depth maps provided by sensors, such as RGB-D or stereo cameras, generally range from 10 to 20 m in distance, with low-to-mid accuracy. Point clouds generated by LiDAR sensors, on the other hand, involve distance ranges that easily surpass camera-generated depth maps, and with higher precision; this generally comes at the expense of point density [191]. As an example, whereas a standard 720 p depth image stores 620 K points, a point cloud generated by a typical LiDAR contains only about 15 K points. This accounts for about 3% of the camera's total density. Additionally, LiDAR sensor data are usually mostly geometric-only in nature and as such not as rich as sensor camera data, which also natively includes colour information.

To take advantage of the strengths of both types of sensors, it has become common practice to register and fuse their sensed data into a single 3D representation when the relation between coordinate systems of all sensors is known (e.g., via calibration). However, as mentioned previously, camera systems do not always produce dense depth maps; this has resulted in the recent practice in which point clouds containing high-precision readings taken by LiDAR sensors are used to generate sparse-depth maps by projecting the points onto pixels of property-rich camera images, which are in turn transformed into dense depth maps by estimating the depth of the remaining pixels.

As depicted in Figure 11, depth completion solutions use sparse-depth information together with native camera sensor data (e.g., RGB) to estimate dense depth maps, thus "filling out the gaps". These solutions can be divided into those that use conventional, non-learning techniques and more recent learning- and neural network-based techniques. An example of the former, IPBasic [192], uses computer vision techniques to estimate points with the purpose of being used in systems with limited computational resources. In a similar manner, Zhao et al. [193] achieved precise density by utilising local surface geometry. Both of these methods require sparse-depth information alone as input to

produce their respective estimates for the corresponding dense depth map. Despite their lower complexity and ease of implementation, single-input methods are less accurate due to not taking advantage of the full range of available information. They are also more prone to lower accuracy ratings when processing more complex environments and geometries [194].
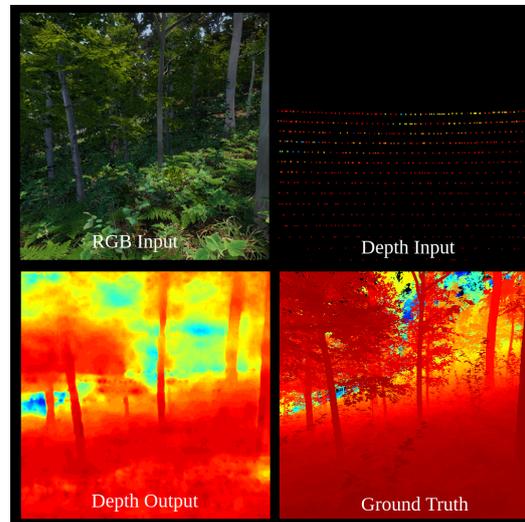


**Figure 11.** Depth completion FCN, called ENet [195], applied to a synthetic forestry dataset [196]. The sparse-depth image shown on the top right is generated by projecting points from a point cloud produced by a (simulated) LiDAR sensor onto the image space of the camera producing the RGB image shown on the top left; since the LiDAR sensor is tilted slightly downwards to prioritise ground-level plants, only the bottom half of the image includes depth information from the point cloud. The depth completion method, which uses both information from the RGB image and the sparse-depth image as inputs to estimate the corresponding dense depth image, produces the output shown on the bottom left, with the ground-truth dense depth image shown on the bottom right for comparison.

Unimodal or multimodal neural networks using all available information (i.e., depth but also colour from cameras) to improve prediction quality have been developed recently (e.g., [197–200]), achieving top-ten results when benchmarked using the well-known Karlsruhe Institute of Technology and Toyota Technological Institute at Chicago (KITTI) standard for depth completion [201]. The benchmark ranking indicates that supervised learning approaches achieve the best performances among published methods with verified reference, and suggests that multimodal neural networks are proven to be more robust than single-input approaches. Examples of these include the dynamic spatial propagation network (DySPN) [197] and the repetitive image-guided network (RigNet) [198].

However, the benchmark's inference rate is ambiguous due to the absence of predetermined computation calculation guidelines. Also, when evaluating depth completion methods, the computational cost should also be taken into consideration. For instance, according to the KITTI benchmark DySPN [197] is five times slower than the new normal network (NNNet) [202] with a difference in RMSE of only 20.5 mm, which can be taken as insignificant.

It is unknown how these methods will perform in woodland scenarios, as they were all trained in urban settings. In fact, we are unaware of any such work for natural environments, which are inherently unstructured. Recent adaptations of depth completion methods using well-established urban models to forestry applications have demonstrated that these neural architectures do not generalize well [203]. An example of how depth completion can be particularly useful in these types of applications can be seen in [107], in which LiDAR point clouds are used to produce dense depth images registered with images produced by a high-resolution multispectral camera that allows for improved semantic segmentation, which in turn allows the generation of complete, camera-referenced,

2.5D representations with semantic information. Representations such as these can then be used to produce feature-rich metric-semantic maps, the significance of which is explained in the following subsection.

### 4.1.3. Metric-Semantic Mapping

Metric-semantic mapping is defined as a map of spatial information about the environment, with assigned confidence values for entities of known classes [23,204], which provides high-level scene understanding. In other words, it constitutes the projection of labels from semantic information (e.g., semantic segmentation images; see Section 4.1.1) onto a spatial map, which can be produced from registered spatial representations such as those presented in Section 4.1.2.

Metric-semantic mapping is a relatively recent research topic, as the integration of semantic information into spatial maps has only become an active area of research in the past few years, driven by advances in deep learning techniques and the increasing demand for robots to navigate and interact with complex environments. As mentioned in Section 4.1.1, as an extension of semantic segmentation, metric-semantic maps provide high-level scene-understanding capabilities to the robot. These, in turn, enable the proper undertaking of current tasks, safe navigation in a given area and allow the avoidance of dynamic entities such as humans and animals, by taking into account not only geometrical but also operation-relevant semantic information. If the metric-semantic maps are probabilistic, they provide the additional capability of tackling uncertainty explicitly. This is potentially important for forestry applications with their many unknown factors, high uncertainty and perceptual challenges [27].

A relevant example of this type of method would be TUPPer-Map (temporal and unified panoptic perception for 3D metric-semantic mapping), proposed by Yang and Liu [205]. The authors propose a new framework that combines both temporal and unified panoptic perception (panoptic perception in this context is a comprehensive and unified understanding of the environment obtained by using the metric-semantic framework). Temporal perception is achieved by using a recurrent neural network (RNN) to incorporate past observations and improve the accuracy of the semantic segmentation. Unified panoptic perception is achieved by jointly processing depth, RGB, and semantic segmentation information. Chang et al. [206] proposed another solution for metric-semantic mapping named Kimera-Multi. This system is based on the Kimera framework and allows multiple robots to collaboratively build a 3D map of an environment while also generating semantic annotations. It was evaluated on both simulated and real-world experiments, demonstrating its ability to generate accurate and consistent 3D maps across multiple robots. Li et al. [207] in turn proposed a semantic-based loop-closure method for LiDAR SLAM called semantic scan context (SSC). The proposed method uses semantic information from the environment to improve the accuracy and robustness of loop-closure detection. Specifically, SSC uses semantic segmentation to extract semantic features from LiDAR scans, which are then used to construct a semantic scan context graph. The graph is used to detect loop closures based on the similarity of semantic features between scans, which reduces the reliance on geometric features and increases the likelihood of correctly detecting loop closures. Another interesting approach was presented by Gan et al. [208], who devised a flexible multitask, multilayer Bayesian mapping framework with readily extendable attribute layers. The proposed framework enables the efficient inference of multi-layered maps with both geometric and semantic information by jointly optimising multiple tasks, including map reconstruction, semantic segmentation, and sensor fusion. The paper describes a detailed implementation of the proposed framework, including its architecture, loss functions, and optimization strategy.

There are only two methods that we found that address forestry applications, the first of which was proposed by Liu et al. [202] and Liu and Jung [209], and the second by Russell et al. [31]. The first group of authors present a comprehensive framework for large-scale autonomous flight with real-time semantic SLAM under dense forest canopies. The frame-

work includes a novel semantic-based approach that integrates semantic segmentation into visual place recognition and allows multiple aerial robots to collaboratively build a 3D map of the environment. The proposed framework is evaluated on large-scale mixed urban-woodland datasets, and the results demonstrate its effectiveness in enabling autonomous flight and accurate mapping under challenging conditions. The second group of authors presents an adaptation of an existing efficient 3D mapping solution, the OctoMap [210], adding metric-semantic information to the original framework and testing it in scenarios including dense forest canopies. The adaptation consists of embedding a probability distribution of the predicted classes in each voxel, multiplying each new observation by the current distribution, and then re-normalising. The voxel is then labelled with the most likely class. This method was evaluated in a dense forestry environment with results demonstrating accurate semantic information and mapping.

We listed the work on metric-semantic mapping we surveyed in Table 3. We analysed the implementation aspect of each method in terms of the environments used, and the geometry and data structure for its spatial representation, as there are few works that have been tested or compared under similar conditions. These particular features are important since they provide an indication of how each technique would perform (particularly in terms of the usage of computational resources) in unstructured, large, outdoor environments, such as woodlands (note that only the active metric-semantic mapping technique by Liu et al. [202], Liu and Jung [209] and the Semantic OctoMap solution by Russell et al. [31] are specifically reported to have been used in this type of environment).

**Table 3.** Comparison of metric-semantic mapping techniques in the context of artificial perception for forestry robotics. Of all these techniques, only TUPPer-Map [205] is reported to not work online.

| Method | Input | Environment | Geometry | Data Structure/Framework |
| --- | --- | --- | --- | --- |
| TUPPer-Map [205] | RGB-D | Urban | Mesh | Truncated Signed Distance Field (TSDF) |
| Kimera-Multi [206] | RGB-D/PCL | Urban | Mesh | TSDF |
| SSC [207] | PCL | Urban | 3D Points | Point Cloud |
| MultiLayerMapping [208] | RGB-D | Urban | Voxels | Multi-Layered BGKOctoMap [211] |
| Semantic OctoMap [31] | RGB-D/PCL | Forest/Urban | Voxels | OctoMap [210] |
| Active MS Mapping [202,209] | Stereo/PCL | Forest/Urban | 3D Models | Factor-Graph |

### 4.1.4. Traversability Analysis for Navigation

An important subtopic of the scene understanding of woodland scenarios (which shares links with metric-semantic mapping, described in the previous section) involves traversability analysis for safe navigation. This involves methods that interpret the environment representation with appropriate fidelity in order to distinguish between traversable and non-traversable areas, considering the platform's configuration footprint and locomotion. Typically, these works are coupled with path-planning techniques, which then allow the robot to navigate through the devised traversable space in order to reach its destination.

A recent and very complete survey on the terrain traversability analysis for autonomous ground vehicles is presented in [212]. The authors define a taxonomy for the different methods by grouping them into vision-based (e.g., [213,214]), LiDAR-based (e.g., [215,216]), alternative sensor-based (e.g., [217–219]), and sensor fusion methods (e.g., [220,221]).

As expected, a clear prevalence of deep learning-based methods with different sensor modalities to tackle terrain classification and image segmentation problems has recently been observed. For instance, in [222], a deep neural network takes an image as input and categorizes every pixel of the image into an assigned class. After a coordinate frame transformation, this can assist the robot's navigation system in traversing the environment. The approach presented in [223] also uses a deep neural network, which outputs navigation commands that need to be post-processed, instead of semantic classes. Additionally, semantic segmentation has also been used for navigation to distinguish between forest trails from less traversable ground [224].

Alongside semantic segmentation methods for traversability analysis, there has been significant progress in the development of target detection techniques for forest navigation recently. Moreover, Sihvo et al. [225] introduced a method for detecting tree locations using on-site terrestrial LiDAR data collected during forest machine operations, employing a planarised ground model and individual tree stem lines to achieve accurate alignment without the need for positioning or orientation systems. More recently, da Silva et al. [140] evaluated several deep learning models for tree trunk detection, achieving high detection accuracy and fast inference times, paving the way for advanced vision systems in forestry robotics. Furthermore, a point-based DNN for accurately classifying tree species using 3D structural features extracted from LiDAR data was proposed in [226], achieving a classification accuracy of 92.5% and outperforming existing algorithms in tree species classification tasks.

Other methods include the discretization of the world into an OctoMap [210] (see Section 4.1.3), which is divided into horizontal layers. Every layer is analysed, considering inter-layer dependency, in order to identify traversable portions of terrain (e.g., levelled areas or slopes) and non-traversable portions of terrain (e.g., cliffs or hills that the robot cannot safely climb) [227]. Another relevant method proposed in [228] uses an efficient 2.5D normal distribution transform (2.5D-NDT) map that stores information about points that are believed to be traversable by ground robots. An analysis was conducted by considering a spherical spatial footprint for the robot, checking whether there were patches inside the sphere to understand if the terrain was suitable for navigation.

Fankhauser et al. [229] proposed a method that incorporates the drift and uncertainties of the robot's state estimation, as well as a noise model of the distance-measuring sensor in order to create a probabilistic terrain estimate as a grid-based elevation map with upper and lower confidence bounds. Adjacent cells are then analysed with respect to some metrics, e.g., the height difference between them, and a traversability map is generated. Moreover, Ruetz et al. [230] conservatively represent the visible spatial surroundings of the robot as a watertight 3D mesh that aims at providing reliable 3D information for path planning while fulfilling real-time constraints. It provides a trade-off between representation accuracy and computational efficiency. A method is then applied to infer the continuity of the surface and its respective roughness value, choosing an appropriate, and path taking these properties into consideration.

A particularly relevant work was described in [231], which presents a motion planner that computes trajectories in the full 6D space of robot poses based on a terrain assessment module that plans directly on 3D point cloud maps. It identifies the geometry and traversability of surfaces represented as point clouds without explicit surface reconstruction or artificial discretization of maps or trajectories. Terrain geometry and traversability are assessed during motion planning, by fitting robot-sized planar patches to the map and analysing the local distribution of map points. The performance of the approach is demonstrated on autonomous navigation tests in three different complex environments: with 3D terrain (rough outdoor terrain, a two-level parking garage, and a dynamic environment).

A recent method proposed by Carvalho et al. [108] focuses on the concept of mechanical effort to solve the problem of terrain traversability and path planning in 3D forest environments. The technique processes 3D point cloud maps to generate terrain gradient information. It then categorizes terrain according to the effort required to traverse it, while identifying key evident obstacles. This allows the generation of efficient paths that avoid obstacles and major hills when more conservative paths are available, potentially minimising fuel consumption and reducing the wear of the equipment and the associated risks.

Finally, it is worth mentioning key works that specifically address methods for fusing information from multiple sources to assist traversability and navigation in outdoor environments. Milella et al. [220] presented a multi-sensor approach using heterogeneous sensors (stereo cameras, VIS–NIR sensor, a thermal camera, and an IMU) to generate a multi-layer map of the ground environment, enabling accurate soil mapping for highly automated agricultural vehicles and potential integration with farm management systems,

while Vulpi et al. [221] explored the use of proprioceptive sensors (wheel encoders and IMU) and DNNs to accurately estimate terrain types for autonomous rough-terrain vehicles, achieving comparable or better performance than standard machine learning methods and eliminating the need for visual or depth information. In a very relevant and recent work, Hera et al. [56] employed exteroceptive sensors, including a dual-antenna GNSS connected to the Swepos Network-RTK for precise positioning in Cartesian coordinates, and a stereo camera (Stereolabs Zed2) that captures 2D RGB images, 2D depth maps, and additional sensor data, such as IMU, barometer, and magnetometer data, to map the ground and facilitate traversability analysis for an autonomous machine operation in a forest environment.

### 4.1.5. Datasets and Learning

For proper benchmarking of the artificial perception methods mentioned previously, learning and testing datasets should allow repeatability and replicability. Datasets concerning outdoor environments are divided into 2D, 2.5D, and 3D data, annotated according to task-relevant classes, the most representative of which are presented in Table 4. In autonomous driving, for example, there are usually classes of people, cars, and cyclists [232]. These datasets are particularly important in the case of supervised learning, where the model is given an expected output that is compared to the predicted output. A loss function is used to calculate the difference between the expected output and the actual output. This is then used to update the weights and parameters of each layer in the model until a loss threshold is met or a final number of iterations is reached.

Two-dimensional datasets include only sensors that contain a 2D representation, such as RGB and thermal images. For example, Tokyo multi-spectral [233] includes RGB and thermal images of an urban street scene; RANUS [234] contains RGB and Near Infrared (NIR) images; it shows a high contrast between natural and artificial objects by using the non-visible light spectrum to its urban street dataset. However, all of these data are for urban environments. The only 2D datasets for forestry research we are aware of are the FinnDataset [235], which uses 4 RGB cameras to map a forest in Finland in summer and winter, and the ForTrunkDet [139,236], which is comprised of just under 3000 images with single-class tree trunk labelling taken from three different Portuguese woodland sites for training and testing detection and segmentation models. There are also a few datasets that are tailored to very specific tasks and conditions in forestry and include a small number of training samples, such as TimberSeg 1.0 [158], which is composed of 220 images showing wood logs in various environments and conditions in Canada.

**Table 4.** Comparison of outdoor environment training/testing datasets in the context of artificial perception for forestry robotics.

| Name | Type | Sensors | Environment | No Frames/Scans | Labelled |
|---|---|---|---|---|---|
| FinnDataset [235] | 2D | RGB | Real/Forest | 360 k | — |
| ForTrunkDet [237] | 2D | RGB/Thermal | Real/Forest | 3 k | 3 k |
| RANUS [234] | 2D | RGB/NIR | Real/Urban | 40 k | 4 k |
| Cityscapes [238] | 2.5D | RGB-D | Real/Urban | 25 k | 25 k |
| LVFDD [239] | 2.5D | RGB-D/GPS/IMU | Real/Forest | 135 k | – |
| Freiburg [156] | 2.5D | RGB-D/NIR | Real/Forest | 15 k | 1 k |
| SynthTree43K [240] | 2.5D | RGB-D | Synthetic/Forest | 43 k | 43 k |
| SynPhoRest [196] | 2.5D | RGB-D | Synthetic/Forest | 3 k | 3 k |
| KITTI [241] | 3D | RGB/LiDAR/IMU | Real/Urban | 216 k/- | 400 |
| nuScenes [242] | 3D | RGB-D/LiDAR/IMU | Real/Urban | 1.4 M/390 k | 93 k |
| SEMFIRE [243] | 3D | NGR/LiDAR/IMU | Real/Forest | 1.7 k/1.7 k | 1.7 k |
| TartanAir [244] | 3D | RGB-D/LiDAR/IMU | Synthetic/Mixed | 1 M/- | 1 M |
| QuintaReiFMD [236] | 3D | RGB-D/LiDAR | Real/Forest | 1.5 k/3 k | — |

To enhance 2D datasets, 2.5D datasets also include corresponding depth images. For outdoor environments, datasets such as SYNTHIA [245] and Cityscapes [238] are popular benchmarks for developing autonomous vehicle perception, as both include high-

quality RGB and depth images in their data. However, there is a lack of datasets focusing on natural environments; therefore, suitable 2.5D forest data are hard to find. Currently, we are aware of only a few 2.5D datasets that are either simulated or recorded in a forest environment. These are the Freiburg forest dataset [156], which contains RGB, NIR, and depth images of a German forest in different seasons. LVFDD [239] provides an off-road forest depth map dataset recorded in Southampton, UK. This dataset features low viewpoints and close-up views of obstacles such as fallen tree branches and shrubs. Another dataset is SynthTree43K [160,240], which is a simulated dataset containing 43,000 synthetic RGB and depth images and over 190,000 annotated trees (i.e., single class annotation), including train, test, and validation splits. Finally, Synthetic PhotoRealistic (SynPhoRest) [196,246] provides simulated data of typical Portuguese woodland, including both RGB and depth images.

Finally, 3D datasets add full 3D representations, such as LiDAR sensor data in the form of 3D point clouds (see Section 4.1.2). For autonomous vehicles in outdoor scenarios, there is the aforementioned KITTI dataset [241], which is one of the most popular benchmarks to test the efficiency of semantic segmentation systems. However, due to the lack of data diversity in different lighting conditions, there are better options in the community today. Lyft [247] provides a robust dataset with 1.3 M 3D annotations of the scene and includes both LiDAR point clouds and RGB images. Similarly, nuScenes [242] provides a comprehensive dataset that includes LiDAR point clouds, 5 RADARs, 6 RGB cameras, annotated LiDAR points for 3D semantic segmentation, and 93 k annotated RGB images. Although both datasets represent advances over the previous KITTI benchmark, they are designed for autonomous driving in urban scenarios and are, therefore, not ideal for forestry robotics applications.

For forestry environments, there are very few publicly available 3D datasets. These include QuintaReiFMD [236], the SEMFIRE dataset [243], and TartanAir [244]. These three datasets are a step forward in forestry robotics, although each still has some limitations. The SEMFIRE dataset, while having the advantage of being annotated, is relatively small. QuintaReiFMD does not provide different lighting conditions, and TartanAir is a synthetic dataset that has yet to be shown to allow generalization to real-world scenarios.

Adding to the problem of the lack of annotated data, none of the datasets listed above fully addresses an important issue in computer vision for forestry robotics: class imbalance. This becomes particularly crucial for cases in which examples of instances of mission-critical classes (e.g., humans or animals) are not represented in sufficient proportion in the dataset, as they are relatively rare occurrences in the wild. In summary, although these datasets represent important contributions to machine learning-based perception for forestry robotics, more extensive, diversified and comprehensive datasets are needed.

### 4.1.6. Improving Learning: Data Augmentation and Transfer Learning

The performance of machine learning (ML) models, deep learning models, in particular, depends on the quality, quantity, and relevance of training data. The process of collecting and annotating data is often time-consuming and expensive due to the need for expert knowledge. In fact, as identified in Section 4.1.5, insufficient data is one of the most common challenges in using datasets for training ML models. Another issue also mentioned previously is the difficulty in resolving class imbalance, i.e., skewed class representation within the dataset.

Data augmentation is a set of techniques that artificially increases available training content by generating new data from existing samples. Data augmentation methods, as illustrated in Figure 12, are applied to enlarge datasets by adding modified copies of existing data or realistic synthetic data.
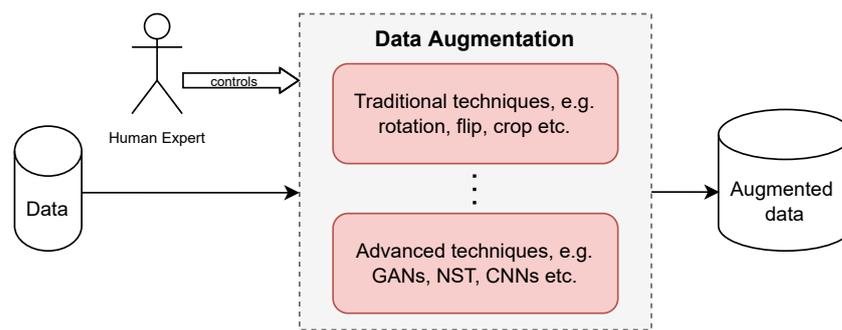
**Figure 12.** Overview diagram of a data augmentation process; from [248]. Data from a specific domain is forwarded into a data augmentation unit, potentially curated by a human expert, which in turn produces an augmented dataset containing the original data and new artificially generated samples.

In case a dataset is small or insufficiently diverse, traditional augmentation methods that create modified copies of images in the dataset, via image manipulation operations such as rotation, mirroring, cropping, darkening, etc., may not be sufficient to satisfactorily increase accuracy. Learning-based image classification performance has been shown to improve when, in addition to real-world data, realistic synthetically rendered images are used for training [249,250]. Synthetic images for this purpose can be obtained in several ways, namely by using renderings taken from realistic simulated 3D environments, by resorting to generative models or by means of a mixed approach.

For forestry applications in particular, synthetic data have already been used in purely synthetic datasets as presented in Section 4.1.5 with promising results, as shown by Nunes et al. [196,246]. Generative Adversarial Networks (GANs) were used by Bittner et al. [251] to generate synthetic NIR multispectral images from fully annotated RGB images for data augmentation purposes. Fully annotated multispectral datasets are difficult to obtain with sufficient training samples when compared to RGB-based datasets; this is because annotation is often time-consuming and expensive due to the need for expert knowledge. This issue is compounded even further by the specific nature of multispectral images [251]. The authors provided proof-of-concept showing that the synthetic images generated by this type of solution yield a level of performance in the semantic segmentation model proposed in [157] comparable to what is obtained using real images for training. This solution can be used with real and/or synthetic (i.e., rendered from simulation) RGB images; see Figures 13 and 14.

A further alternative for improving performance is transfer learning, a technique in which a model is trained and developed for one task and then reused on a second related task. It refers to the situation in which what was learnt in one setting is used to improve optimization in another [252,253]. One of the main benefits of transfer learning is its ability to address imbalanced data. Transfer learning can help to improve classification accuracy and reduce the impact of imbalanced data by transferring knowledge from pre-trained models to new models. In other words, by leveraging pre-trained models, transfer learning can significantly reduce the amount of labelled data required for training, and thus can be an effective way to overcome imbalanced data challenges.

An increasing amount of research has been conducted in applying transfer learning to generic computer vision problems to address the imbalanced class problem [254]. An example includes the work by Liu et al. [255], who proposed a solution that they named the transfer learning classifier (TLC). This solution is composed of an active sampling module, a real-time data augmentation module, and a transfer learning module based on a standard DenseNet network, pre-trained on the ImageNet dataset, and transferred to TLC for relearning, with memory usage adjustment to make it more efficient.
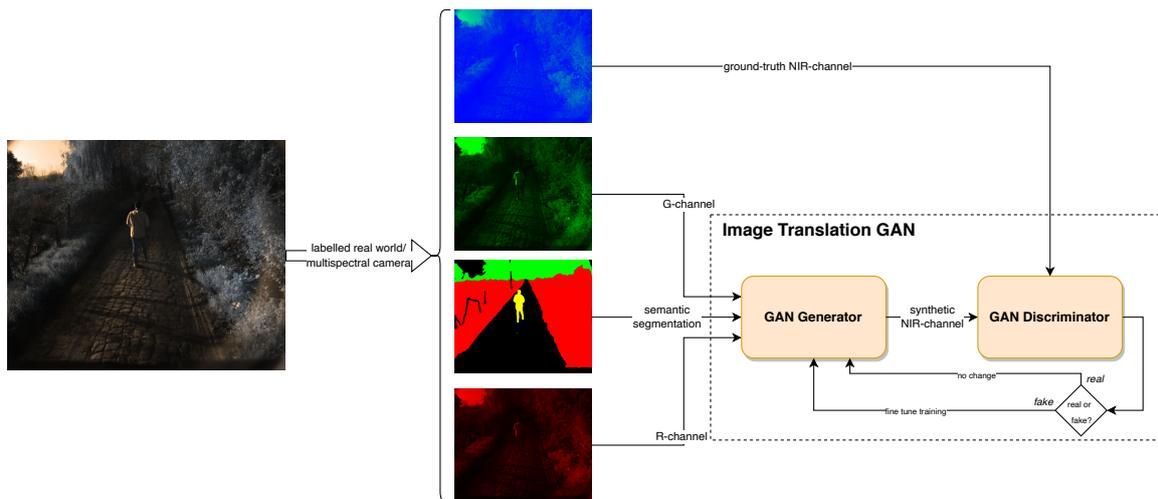
**Figure 13.** GAN image translation training from [251] to generate corresponding NIR channels of multispectral images with an original multispectral image (left image); a model ground truth image, which the model attempts to predict (centre, top image); a green channel image, which is part of the model input image (centre, second image from the top); a semantic segmentation image, where its label values are part of the model input image (centre, third image from the top) and a red channel image, which is part of the model input image (centre, last image).
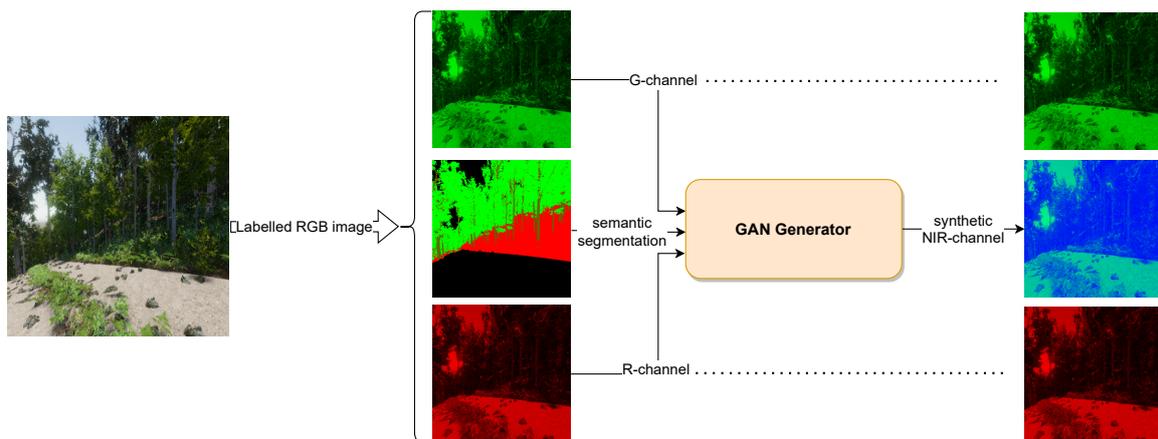


**Figure 14.** GAN image translation generation from [251] of synthetic NIR channel and corresponding final "fake" multispectral image from a fully annotated RGB input image (on the left); a green channel image, which is part of the model input image and is fed forward to be merged after generation with the synthetic NIR channel (centre, top image and right, top image); a semantic segmentation image, where its label values are part of the model input image (centre, second image from the top); a red channel image, which is part of the model input image and is fed forward to be merged after generation with the synthetic NIR channel (centre, last image and right, last image); a synthetic NIR channel image, which the model predicted and is merged afterwards with the real red and green channels as a synthetic multispectral image (right, second image from the top).

Examples of research in applying transfer learning to the more specific context of precision forestry applications, to our knowledge, are relatively few. Niu et al. [190] employed transfer learning to train their model (see Section 4.1.2) on both indoor and outdoor datasets, including the real-world low-viewpoint forest dataset collected by the authors, described in Section 4.1.5. Andrada et al. [157] showed that transfer learning techniques, such as weight initialization with class imbalance (with the "Humans" and "Animals" as the underrepresented classes; see the previous section, Figure 9) and pre-trained weights, improved the overall quality of semantic segmentation.

### 4.2. Localization and Mapping

Visual odometry, along with visual SLAM, have been the objects of extensive research for several decades, as exemplified in the survey work by Younes et al. [256] and the fundamental tutorial article of Scaramuzza and Fraundorfer [257]. Recently, there has been increased interest in applying these techniques to the more challenging natural outdoor scenarios [258]. An account of research conducted during the last ten years on natural outdoor environments would include work by Konolige et al. [259], who explicitly addressed visual odometry for rough terrains, Otsu et al. [260], detected interesting points in untextured terrains by adaptively selecting algorithms for feature tracking, Daftry et al. [261], and introduced a semi-dense depth map estimation in a scale-aware monocular system in cluttered environments, such as natural forests, as well as a very interesting biologically inspired Bayesian approach using convolution neural networks via inference of the sun's direction to reduce drift (by Peretroukhin et al. [262]).

Moreover, for such applications, it would be worthwhile to consider fusing 2D camera fusion with 3D-based devices; recent examples of such research would include work by Giancola et al. [263], Paudel et al. [264], and notable work on SLAM in a forest environment by Pierzchala et al. [89] and Tremblay et al. [90] (see mapping output in Figure 15). Recent developments in visual odometry and navigation for UAVs are also relevant. Research by Giusti et al. [224] is particularly pertinent due to its specificity to forest applications. Contributions also include the work by Smolyanskiy et al. [265], who focused on stable flight and trail-following coupled with obstacle detection in outdoor environments using a DNN, and Mascaro et al. [266], who discussed the fusion of visual–inertial odometry with globally referenced positions for UAV navigation. Moreover, Kocer et al. [267] looked into visual sensing of forest topology. Robust solutions in outdoor applications were explored by Griffith and Pradalier [268], who addressed the accurate alignment of images captured over extensive seasonal variations in natural environments for survey registration, and Naseer et al. [269], who implemented data association exploiting network flows and DNNs to enhance visual localization across seasons. Additionally, Engel et al. [270] proposed the direct sparse odometry (DSO) method, which amalgamates accurate sparse and direct structure and motion estimation with real-time parameter optimization.

One of the pivotal contributions that departed from 2D SLAM in flat indoor environments in 2D to provide a full 3D SLAM system in less structured outdoor settings was introduced by Cole and Newman [271]. The authors employed a 2D scanner that continuously oscillated around a horizontal axis, thus consecutively building 3D scans. This was achieved through a straightforward segmentation algorithm and by executing a "stop-acquire-move" cycle. A scan-matching classification technique was applied for inter-scan registrations, which incorporated an integrity check step. Consequently, they achieved successful 3D probabilistic SLAM in outdoor, uneven terrain; in a related study [272], the system was enhanced with a forward-facing camera. This addition aimed to improve loop closure detection through the use of an appearance-based retrieval approach. Further explorations of loop closure detection using vision in outdoor environments were followed up with in subsequent works, such as [273–275].

Around that period, Thrun and Montemerlo [276] introduced GraphSLAM, a widely recognised offline algorithm. This algorithm extracts a series of soft constraints from the dataset, which are represented using a sparse graph. The map and robot path are subsequently obtained by linearising these constraints and solving the least squares problem through standard optimization techniques. The approach was tested outdoors in large-scale urban structures, employing a bi-directional scanning laser and optionally integrating GPS measurements, achieving satisfactory 3D map representations. Another influential contribution to 3D SLAM is found in [133]. This approach involves a robot equipped with a tiltable SICK LRF in a natural outdoor environment. It relies on 6D ICP scan matching, coupled with a heuristic for closed-loop detection and a global relaxation method, resulting in precise mapping of the environment, aligning closely with an aerial ground truth photograph.
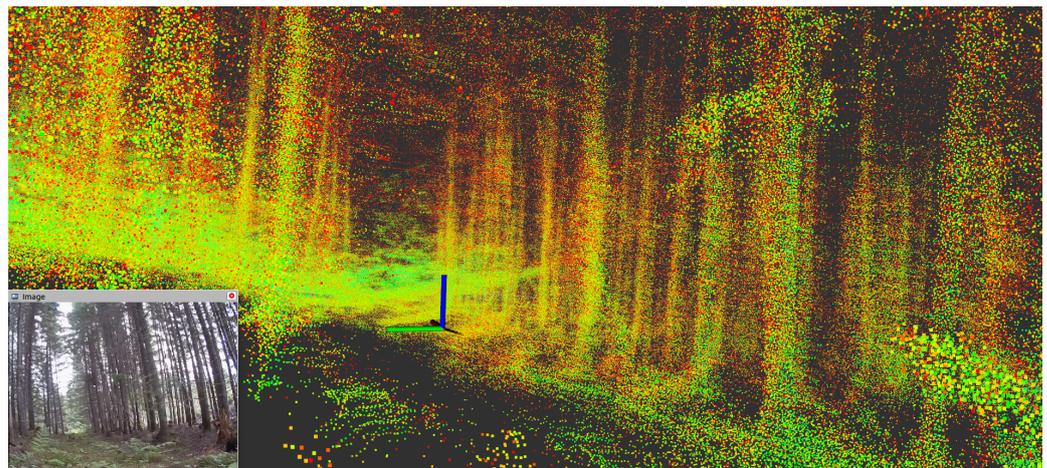
(**a**) Robot perspective.
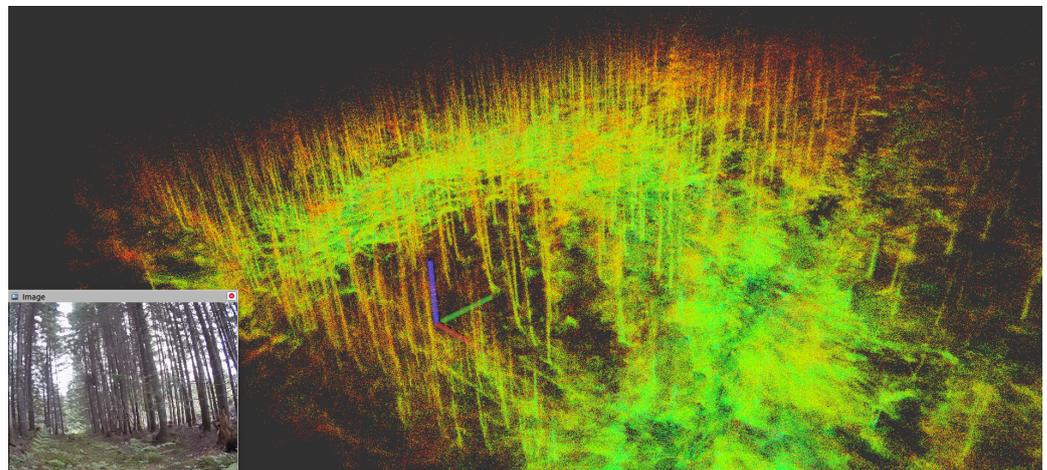


(**b**) Bird's eye view.

**Figure 15.** The 3D point cloud representation of a forest (source: Montmorency dataset [90]). The three axes XYZ at the origin of the robot's coordinate system are represented in red, green, and blue, respectively.

Based on the early work of Singh and Kelly [277] concerning elevation maps for navigating an all-terrain vehicle, Pfaff et al. [278] presented an efficient approach to tackle the 3D SLAM problem. In this approach, individual cells of an elevation map are categorised into four classes, representing different parts of the terrain, as viewed from above, vertical objects, overhanging objects (e.g., branches of trees or bridges), and traversable areas. The ICP registration algorithm then takes this information into consideration, leading to the development of a consistent constraint-based technique for robot pose estimation. For their experimental work, the authors employed a Pioneer II AT robot equipped with a SICK LMS range scanner mounted on a pan/tilt device. The results demonstrated that these techniques yield substantially increased correspondences and alignments in the generated map.

Numerous LiDAR-based SLAM techniques have emerged in the field, with one of the foundational methods being LiDAR Odometry And Mapping, commonly referred to as LOAM [279]. LOAM has been recognised for its capability to generate highly accurate maps; however, it tends to exhibit sub-optimal performance in environments characterised by sparse landmarks, such as lengthy corridors. To address these limitations, LeGO-LOAM was introduced, incorporating two supplementary modules into the LOAM framework: point cloud segmentation and loop closure [280]. These additional components serve to

enhance computational efficiency and mitigate drift over extended distances, although they do not substantially enhance performance in feature-scarce settings.

In the realm of loop closure techniques, while LeGO-LOAM relies on the naive ICP (Iterative Closest Point) algorithm, a more robust approach based on point cloud descriptors is employed in SC-LeGO-LOAM [280,281]. To enhance performance in environments with limited distinguishing features, recent efforts have focused on integrating Inertial Measurement Units (IMUs) into similar systems through tightly coupled approaches [282–284]. This integration has given rise to the term LiDAR Inertial Odometry (LIO).

An exemplar of this tightly coupled approach is the LIO-SAM [285] method, which presents a comprehensive LiDAR framework built upon a factor graph. LIO-SAM incorporates four distinct factors into its framework: IMU preintegration, LiDAR odometry, GPS, and a loop closure factor. This configuration renders it particularly well-suited for multi-sensor fusion and global optimization. Notably, the recent research landscape has witnessed the emergence of numerous methodologies grounded in similar principles [286–293].

An insightful literature review concerning SLAM in outdoor environments up until 2009 can be found in [294]. In this review, the authors separately discuss mapping and localization. They start by comparing different approaches to occupancy grid mapping and then move on to analysing various methods for localization within the map. These approaches are shown to fall into three distinct groups: visual SLAM (using monocular cameras and stereo vision), LiDAR SLAM (comprising scan matching and maximum likelihood estimation techniques using LiDARs), and sensor fusion SLAM (involving techniques integrating various sensors, such as cameras, LRFs, radar, and/or others). A more recent survey addressing localization and mapping specifically for agriculture and forestry has been presented in [23]. Nine works in the context of forestry were analysed. The authors highlight the need for precise localization in autonomous robotic tasks, such as pruning, harvesting, and mowing, and also emphasize the importance of mapping the surroundings of the robot under these scenarios. The prematurity of this research line is made clear, and particularly, 3D full localization is rare in these environments, with notable exceptions, including the work described in [295], involving localization and mapping in woodland scenarios. Also, advanced mapping techniques (common in areas such as topological and semantic mapping) are not yet solidified for agriculture and forestry.

Tian et al. [296] address cooperative mapping with multiple UAVs. This work is particularly relevant due to its application in a forest environment for the GPS-denied search and rescue under the tree canopy. The UAVs perform onboard sensing, estimation, and planning, and transmit compressed submaps to a central station for collaborative SLAM (see Figure 16). The work explicitly addresses the data-association problem, which is even more challenging in forest scenarios due to moving branches and leaves. However, cooperative mapping is limited to 2D due to onboard processing and communication bandwidth limitations. The system is validated in a real-world collaborative exploration mission.

Stereo vision has seen extensive use in outdoor perception studies over the past few decades [297,298]. In [299], an approach employing hierarchical (topological/metric) techniques is introduced for vehicles navigating large-scale outdoor urban environments. This method utilizes an affordable, wide-angle stereo camera and incorporates GPS measurements into low-level visual landmarks metric mapping to enhance vehicle positioning. At a higher level, the approach employs a topological graph-like map that includes vertices representing topological places, each represented by local metric submaps. These vertices are connected by edges containing transformation matrices and uncertainties that describe their relationships. This approach effectively mitigates global errors while adhering to real-time constraints. The authors conducted successful tests of this approach using an autonomous car, traversing a path spanning 3.17 km. Even in cases where GPS signals were unavailable, the method demonstrated only minimal degradation in performance. This was confirmed by comparing the estimated path results with ground truth data obtained from a professional RTK-GPS receiver module mounted on the vehicle.
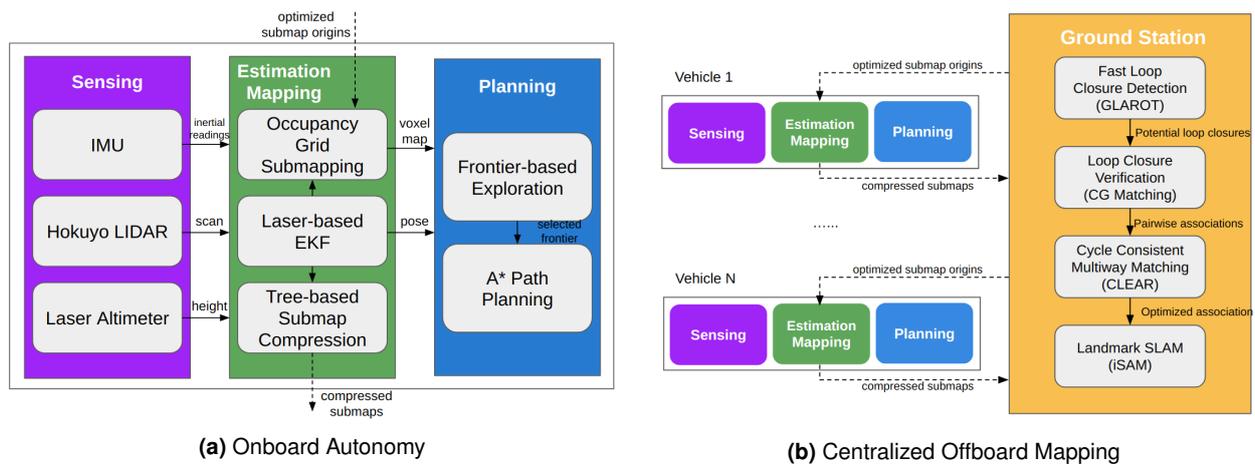
**(a)** Onboard Autonomy



**(b)** Centralized Offboard Mapping

**Figure 16.** System architecture from [296], where multiple UAVs autonomously performed onboard sensing, vehicle state estimation, local mapping, and exploration planning, and a centralised offboard mapping station performs cooperative SLAM, by detecting loop closures and recovering associations observed in multiple submaps, in a forest environment. Reproduced with permission.

In a couple of related studies [300,301], submap matching techniques are applied in the context of stereo-vision outdoor SLAM. Unlike the approach presented in [299], these methods do not presume the existence of flat terrain when matching the 3D local metric maps, and take into account environments where GPS signals are unreliable. The authors introduce an innovative approach to submap matching, relying on robust keypoints derived from local obstacle classification. By characterising distinctive geometric 3D features, this method achieves invariance to changes in the viewpoint and varying lighting conditions, solely relying on an IMU and a pair of cameras mounted on an outdoor robot.

In parallel with the evolution of sensor technologies for 3D perception [302,303] and the ongoing advancements in computational capabilities coupled with increased storage and memory availability [304,305], recent noteworthy contributions have arisen to improve existing solutions. These innovations encompass the refinement of ICP-based scan matching techniques and data registration methodologies [306], the development of more streamlined and efficient approaches to representing 3D data [210,307], the introduction of hierarchical and multiresolution mapping strategies [308], enhanced calibration techniques [309], the introduction of novel probabilistic frameworks [310,311], and the exploration of semantic mapping and terrain modelling [14,312,313]; please also refer to Section 4.1.3.

The feasibility of achieving precise pose and localization estimation with single cameras through the application of advanced SLAM methodologies has been showcased in recent years. Notable examples include RatSLAM [314–316], a SLAM system inspired by biological principles, and ORB SLAM [317,318], a feature-based monocular SLAM method. Furthermore, research has shown the potential for extending individual SLAM approaches to distributed teams of multiple robots (as discussed in Section 4.3). This prospect is particularly pronounced in the context of graphSLAM approaches, which can be leveraged to optimize both the map and 6D pose estimates for all participating robots, as illustrated in [319].

Considering the aforementioned points, a significant body of research has focused on outdoor perception and mapping applications for forest harvesters and precision agriculture [58,89,90,293,320,321]. For example, Miettinen et al. [322] introduced a feature-based approach that combines 2D laser localization and mapping with GPS information to construct global tree maps. This study explores various scan correlation and data association techniques to generate simplified 2D tree maps of the forest. In [323], an extended information filter (EIF) SLAM method for precision agriculture was deployed in an olive grove, relying on stem detection in conjunction with a monocular vision system and a

laser range sensor. In [324], outdoor robot localization for tasks related to steep slope vineyard monitoring was enhanced through the utilization of agricultural wireless sensors. The received signal strength indication (RSSI) of iBeacons (Bluetooth-based sensors) was employed for distance estimation, thus augmenting localization accuracy in environments with unstable GPS signals.

Notwithstanding all the reported advances, however, an examination of this body of work shows that the conclusions drawn by Chahine and Pradalier [258] regarding the many open scientific and technological challenges for visual odometry and SLAM applied to natural outdoor environments still hold. In fact, these challenges, particularly those resulting from the constraints posed by visual sensor limitations listed and described in Section 3, remain relevant today. For example, forestry environments pose many obstacles for sensing, which can lead to corner cases in feature matching and loop closure techniques. Additionally, as can also be inferred from that section, high vibration and other motion biases have a significant effect on proprioceptive sensors and their calibration, resulting in motion blur, hampering optical flow techniques and leading to high variance of image scale, as well as rolling shutter effects on vision sensors [325]. When it comes to LiDARs and IMUs, motion during data acquisition and the lack of stabilization inevitably increase noise in the collected sensing data. The cost aspect also poses a challenge in the development of outdoor-resistant and compliant solutions. Most affordable and commonly used sensors in robotics are ill-suited for use in rainy, foggy/smoky, and windy conditions. Additionally, as noted in [326], the high computational cost of image processing remains a significant technological hurdle that persists today in contemporary visual odometry and visual SLAM solutions.

Finally, SLAM techniques still face difficulties in handling extremely dynamic and harsh environments, especially in the data association step, and there are no fail-safe algorithms that can deal explicitly with metric re-localization or provide proper recovery strategies. Moreover, despite the advances in SLAM in the last decade, there are still three key issues that remain open: (i) the lack of techniques for automatic parameter tuning for out-of-the-box SLAM; (ii) the implementation of an automatic memory-saving solution for offloading parts of the map when not in use and recalling them when needed again; (iii) the use of high-level representations beyond point clouds, meshes, surface models, etc. [327].

*4.3. Cooperative Perception*

Mobile robot teams [328], eventually comprising robots with heterogeneous perception capabilities, have the potential of scaling up robot perception to cover vast areas, thus being an important asset in forestry applications. As sensors of multiple robots, operating simultaneously in different places of the environment, provide data from different locations, distributed robots can perceive simultaneously different spots, i.e., they provide space distribution. Moreover, sensors from different robots, eventually involving diverse sensor modalities, allow distributed robots to provide data from the same place in the environment at different time instants, i.e., they allow for temporal perception. In other words, distributed robots can update percepts at locations previously visited by other robotic teammates, allowing them to adapt to the evolution of dynamic environments over time. Exploiting the complementary features of different sensory modalities through multi-sensor data fusion [329] is important for enabling robots to operate robustly in forestry environments. For instance, the long range and high precision of LiDARs sensors can be complemented and enhanced even further by colour and texture features provided by vision-based sensors (e.g., stereo vision).

Therefore, distributed robot teams potentially allow for persistent, long-term perception in vast areas. However, in order to fulfil this potential of robotic collectives, two fundamental scientific sub-problems of cooperative perception need to be effectively tackled: (i) building and updating a consistent perceptual model shared by multiple robots, eventually over a large time span; (ii) multi-robot coordination to optimise the information gain in active perception. The first sub-problem pertains to the spatial and temporal

distribution of information, involving the fusion of fragmented perceptual data from individual robots into a unified and globally coherent perceptual model that encompasses the collaboratively explored spatial and temporal domains. On the other hand, the second sub-problem revolves around harnessing the partial information contained within the perceptual model to make decisions regarding the next course of action for either an individual robot or the entire robot team. This decision-making process aims to identify areas where new or more recent data can be gathered to enhance or update the existing perceptual model, thus establishing a closed-loop relationship between sensing and action. To maximize collective performance and fully leverage the spatial distribution capabilities offered by the multi-robot system, active perception requires seamless coordination of actions in the robotic team. This coordination typically relies on the exchange of certain coordination data, such as state information, among the robots.

In both of these sub-problems, the development of decentralised collective architectures that do not depend on a central point of failure stands as a critical prerequisite for effectively operating in unstructured and dynamic environments, such as those encountered in forestry settings. Firstly, this is important to scale up perception to large robot teams by exploiting spatial locality to process raw data for perception as close as possible to the place where they are collected, rather than doing it on a remote centralised computational node. This allows for saving communication bandwidth, avoiding transmission latency, and mostly confining the sharing of information to robots located in the same neighbourhood. Then, at a higher level of the architecture, a robot can take the role of perceptual data aggregator within its robotic cluster and share perceptual data with peer representatives of other robotic clusters to build cooperatively a consistent global perceptual model in a decentralised hierarchical way. Secondly, decentralization makes the robotic collective resilient and robust in harsh operational conditions, including individual robots' hardware failures and communication outages. This is achieved at the expense of some loss of optimality provided by centralised schemes because often decisions taken by individual robots need to be made under incomplete and uncertain information. Thus, tackling those problems by using decentralised control schemes tends to significantly increase the complexity of robotic collectives' design and is still an essentially open research problem, especially for many robot systems performing cooperative perception over very wide areas.

Although these two sub-problems have been studied in specific robotics application domains, such as long-term security and care services in man-made environments [330], search and rescue [331,332], environmental monitoring [333–336], or monitoring of atmospheric dispersion of pollutants [337], they have been only partially and sparsely solved and are still essentially open research problems, especially in forestry environments and field robotic applications in general. Only a scarcity of research studies have specifically investigated the utilization of teams of robots in forest-like applications. In particular, the focus has been primarily on UAVs [296,338,339] or UGVs, but without conducting real-world experimentation [340,341]. A few research works have also been devoted to precision agriculture [37,38,342], an application domain having similar requirements to robotic forestry.

The Modular Framework for Distributed Semantic Mapping (MoDSeM) has been proposed and designed in [343] and further detailed in [344] to address the first sub-problem in forestry robotics. MoDSeM aims to systematise artificial perception development by splitting it into three main elements: (i) sensors, producing raw data; (ii) perception modules (PMs), which take sensor data and produce percepts; and (iii) the semantic map (SM), which maintains the percepts produced by PMs in a layered structure. PMs implement particular perception methods that require well-defined sets of inputs and provide a set of outputs. The information from multiple PMs is consolidated into layers within the SM, with each layer corresponding to a specific aspect of the physical world. The system also foresees the sharing of information within teams of robots (see Figure 17). As a consequence, cooperative perception becomes a natural feature of the framework: agents share their

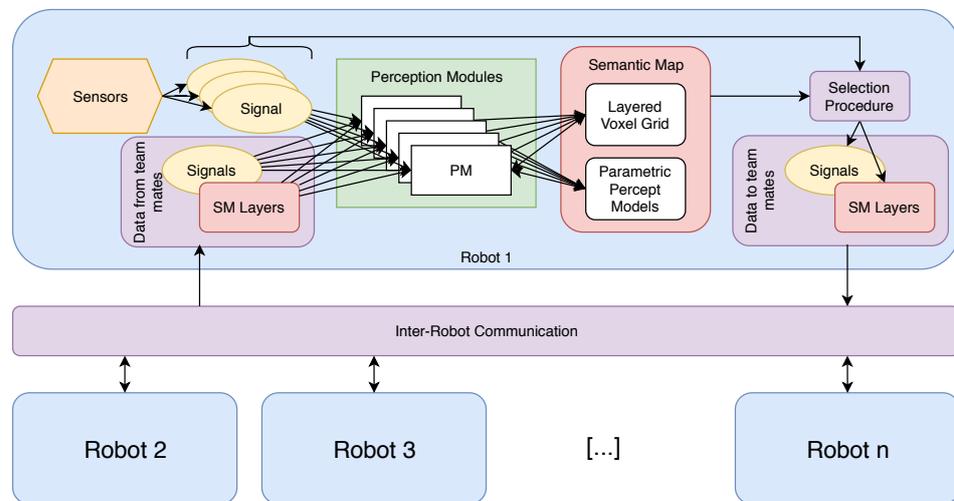semantic map layers, which other agents fuse into theirs and make use of when making decisions.



**Figure 17.** An overview of a robot team operating with the Modular Framework for Distributed Semantic Mapping (MoDSeM) [343,344]. Each team member can have its own sensors, perception modules and semantic map. These can be shared arbitrarily with the rest of the team, as needed. Each robot is also able to receive signals and semantic map layers from other robots, which are used as input by perception modules to achieve a unified semantic map.

Rocha et al. [345] pursued seminal research on cooperative mobile robots based on a distributed control scheme. They demonstrated the potential advantages of this design strategy for multi-robot systems in a 3D cooperative mapping case study, i.e., a specific cooperative perception task. The proposed distributed architecture was based on the reciprocal altruism principle, whereby each robot used an information theoretic criterion, to efficiently share information with its peers when building a 3D map of the environment. Another more recent example of multi-robot system cooperative task allocation and execution is given by Das et al. [346].

Singh et al. [333] studied informative paths to achieve the required sensing coverage by multiple robots in environmental monitoring tasks. A path-planning algorithm to coordinate robots' monitoring actions was presented whereby robots maximize information gain in their visited locations. Information gain was formulated using Gaussian processes and mutual information to measure the reduction of uncertainty at unobserved locations.

More recently, Ma et al. [335] refined the use of this type of information-driven approach to propose a path-planning method for an unmanned underwater vehicle used for long-term ocean monitoring, by considering spatiotemporal variations of ocean phenomena and an information-theoretic component that plans the most informative observation waypoints for reducing the uncertainty of ocean phenomena modelling and prediction. Within a similar line of research, Manjanna and Dudek [336] coined the concept of a multi-scale path to produce a variable resolution map of the spatial field being studied, and proposed an anytime algorithm for active data-driven sampling that provides a trade-off between data sampling resolution and cost (e.g., time spent, distance travelled, or energy spent).

Euler and von Stryk [337] addressed the distributed control of groups of UAVs used for monitoring atmospheric dispersion processes, such as atmospheric dispersion of volcanic ashes or hazardous material in the sequel of industrial or nuclear accidents. By formulating the robotic collective as a hybrid system, a distributed model-predictive control scheme was proposed. Special attention was given to scalability with the team size through the discretization and linearization of time, and limitation of the state variable of each UAV controller.

Emmi and Gonzalez-de-Santos [37] surveyed very few recent works and projects related to the deployment of groups of robots in agriculture tasks, usually possessing

heterogeneous capabilities. In [38], the authors describe the RHEA project (see Section 2), in which robotic collectives, comprising both unmanned ground vehicles and unmanned aerial vehicles with complementary perception and actuation capabilities, are used to detect and eliminate weed patches, thus diminishing the use of pesticides through their selective application in the areas where they are really needed to control pests.

When used in agriculture, these multi-robot systems allow for the required space and time distribution of weed and crop sensors and actuators to cover effectively large areas [342]. Also, it is technologically and economically more interesting to use a collective of simpler and cheaper robots than one big machine possessing all the required perception and actuation capabilities. Although using robotic collectives in cooperative perception and the actuation required by precision agriculture is still a preliminary and pioneering effort, it presents a huge potential that is certainly also important for robotic forestry, which has several intersecting technical requirements with the former. Lightweight robots can cause less topsoil damage even with repeated wheeling, as shown by Calleja-Huerta et al. [347]; therefore, the use of multiple lighter autonomous machines operating as a coordinated team instead of one big and heavier machine can help reduce soil compaction and damage [348], as suggested by Tarapore et al. [341], who propose the use of what they term "sparse swarms" in forestry robotics. However, although some work has been conducted in preparation for this type of solution (e.g., [190,349]), no systematic work has been published that we are aware of. In summary, multi-robot cooperation in forestry can lead to more efficient and sustainable operations, reducing soil damage, improving overall forest health and protecting vegetation that should be preserved in the interventions sought with typical forestry tasks (see discussion in Section 5.1.4).

### 4.4. Perception Systems and Architectures

In their review paper, da Silva et al. [350] presented an analysis of unimodal and multimodal perception systems for forest management. The work made a comparison between existing perception methods and presented a new multimodal dataset, composed of images and laser scanning data, and collected with the platform illustrated in Figure 18. The authors divided the works into vision-based, LiDAR-based and multimodal perception methods, categorising the applications into health and diseases, inventory and structure, navigation, and species classification. The authors concluded that the expected advances in these topics, including fully unmanned navigation, may allow autonomous operations, such as cleaning, pruning, fertilising, and planting by forestry robots.

In general terms, the traditional workflow that perception follows in field robotics usually begins with perception techniques receiving raw signals from sensors, such as images (Section 3), and processing them either linearly or in batches. The results of this process are percepts, i.e., information that encapsulates a meaningful aspect of the signals received by the system, such as the location of trees and humans in the environment, the traversability of the observed locations, or the health status of the observed plants by the system.

In practical terms, complete frameworks that perform all of the above functions are rare, and their development usually constitutes more of a technological advancement than a scientific one [55]. Furthermore, while there exists an open-source architecture for autonomous unmanned systems called FroboMind [351], it does not support multi-robot systems and cooperative perception, and its development website (http://frobomind.org/web/doku.php, accessed on 1 July 2023) does not seem to be maintained any longer; therefore, there is a lack of all-encompassing software frameworks specifically applicable to perception. Given the relative immaturity of the field of precision forestry, we focus on surveying two aspects of the latest literature: works developed in the context of precision forestry applications itself (forestry-specific), and a wider look at unrelated but applicable perception techniques in adjacent fields (forestry-relevant).

**Figure 18.** AgRob V16 mobile platform and its multisensory system for forestry perception. Reproduced from [350] with permission.

Relatively few perception systems are explicitly dedicated to the solution of forestry-related problems, as seen before. Perception systems dedicated to precision forestry tackle mainly problems related to the localization of target plants and the determination of terrain traversability. In contrast, Russell et al. [31] integrated LiDARs, stereo cameras, and IMU into a semantic SLAM system using an OctoMap representation as shown in the diagram in Figure 19 to identify flammable materials from a UAV's point of view. Similarly, Andrada et al. [107] use an unmanned ground vehicle (UGV) with a multispectral camera and LiDARs to identify and locate clusters of flammable materials in a forestry environment in real-time scenarios. A diagram of this pipeline can be seen in Figure 20. Figure 21 provides the full perceptual architecture for decision-making based on this pipeline and the MoDSeM framework introduced in Section 4.3 for the SEMFIRE forest landscaping use case [343,344,352].
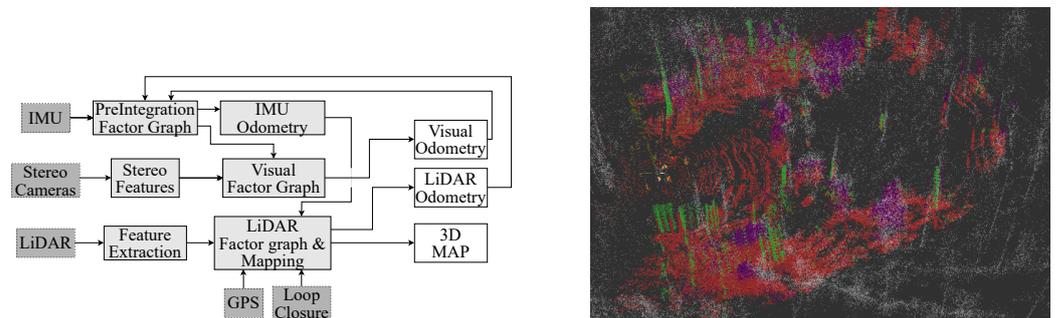


**Figure 19.** Diagram overview of a UAV operating with a perception system developed at Carnegie Mellon's Robotics Institute, which ultimately creates a dense semantic map to identify flammable materials in a forest environment using a full OctoMap representation [31].
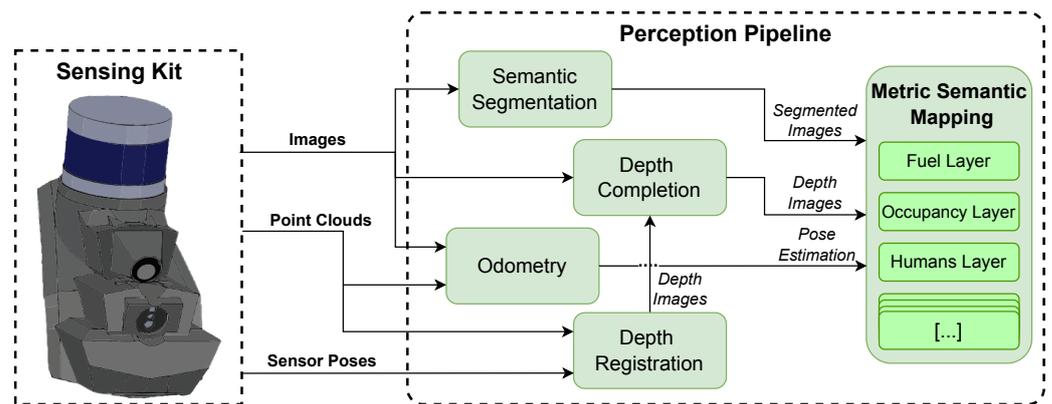
**Figure 20.** An overview of the perceptual pipeline developed by the FRUC group for identifying clusters of flammable material for maintenance using a UGV in forestry environments with a multispectral camera and LiDAR in real-time scenarios [107].
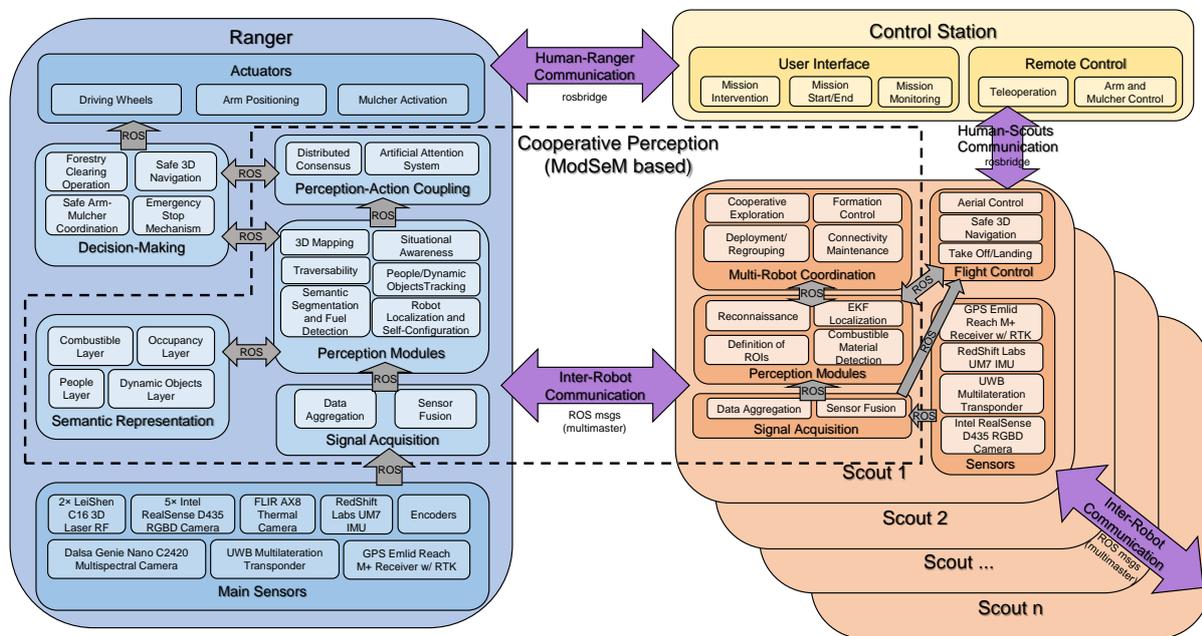


**Figure 21.** SEMFIRE distributed system architecture based on the perceptual pipeline of Figure 20 and the Modular Framework for Distributed Semantic Mapping (MoDSeM) of Figure 17. Please refer to [107,343,344,352] for more details.

Michez et al. [75], aiming to segment riparian (riverside) trees from the landscape using images taken from UAVs, have implemented their developed perception modules using both RGB and NIR cameras. Using a random forest classification approach, the technique was able to achieve an accuracy of 80% in species classification, and 90% in binary health status classification. Bradley et al. [121] presented a technique that, among other functionalities, estimated the ground plane and complemented that information with traversal costs and an estimate of the rigidity of obstacles. The technique employs not only LiDARs but also RGB, NIR, and NDVI data obtained from satellite imagery. More recently, Giusti et al. [224] presented a technique that aims to determine the direction that an off-road trail is taking. This technique implements RGB images collected with a UAV, which are processed using a neural network and output the direction of any trail visible in the image.

Despite not properly tackling precision forestry, many perception systems in field robotics can be applicable to the problems of precision forestry, warranting their discussion in this section. These problems include common autonomous agriculture problems, such as

ground coverage, weed/crop discrimination or plant part segmentation. These techniques are also summarised in Table 5 under the "agriculture" application.

In order to allow a robotic system to position itself correctly for the task at hand, in [87], the research group at the Agricultural Institute of Slovenia presented a technique through which the robot becomes able to localize, with respect to itself, three spraying arms, with a total of eight degrees of freedom. This technique implements the flow using a laser rangefinder as the main sensor, and outputs the kinematic configuration of the robot as a percept.

Agricultural robots employ crop–weed detection techniques to select and distinguish between vegetation at the work site, usually implement the flow with RGB images, and output the classification of each pixel as crop or weed. In the RHEA project, García-Santillán and Pajares [45] developed a solution, focusing on crop–weed detection in initial growth stages, which rules out simpler techniques, such as height filtering. The authors use a Unmanned Ground Vehicle (UGV) to collect RGB images, which are processed by several classification techniques, demonstrating that their approach achieves around 90% accuracy for all of them. Researchers at the University of Bonn presented techniques in [69,70] that used a UGV with a perceptual system comprised of CNN-based models to detect weeds present in RGB and NIR images, achieving 95% precision in detecting both weeds and crops. UAVs have also been used to carry cameras to detect weeds, such as in [68], where the authors employed random forests for classification. In precision forestry, the distinction between plant species is a central problem, particularly if the goal of the mission is to select and intervene in areas affected by invading species, or species that need to be removed or protected in a cleaning operation. In order to distinguish between these, crop–weed detection-based techniques could be used to classify the environment into desirable and undesirable portions.

As seen in Section 4.1, segmenting observable plants and determining the spatial distribution of their parts is a key issue. Bac et al. [62] presented a technique for segmenting plants into five parts: stem, top of a leaf, bottom of a leaf, fruit and petiole, using multi-spectral images from a rig aboard a UGV as input. The authors achieved a recognition rate ranging from 40% to 70% for several different plant segments.

The health status of plants is also a central issue in agricultural robotics: automating information collection and classifying health status is a boon for human users. This represents a broad issue, as the spectrum of possible diseases and health indicators is vast. Mozgeris et al. [85] presented a technique, developed at the Aleksandras Stulginskis University, for estimating the chlorophyll content of spring wheat. This work implemented RGB cameras to provide input and used several algorithms to compare them. In the context of precision forestry, these techniques can provide important information for the selection of areas of intervention, as unhealthy plants may result in higher risks of forest fires.

Table 5 presents a comparison summary of the perception systems reviewed in this section.

**Table 5.** Summary of the perception systems under survey.

| Ref. | Year | Application | Platform [a] | Input [b] | Percepts [c] | Algorithms [d] |
|------|------|-------------|--------------|-----------|--------------|----------------|
| [65] | 2017 | Agriculture | UAV Swarm | Position of agent, position of other agents, detected weed density, confidence | Weed density map | Model fitting |
| [62] | 2013 | Agriculture | UGV | Multispectral Images | Detected hard and soft obstacles, segmented plant parts | CART |
| [68] | 2017 | Agriculture | UAV | RGB | Segmented crop and weed sections | Random Forests |
| [69] | 2018 | Agriculture | UGV | RGB | Segmented crop and weed sections | CNN |
| [70] | 2018 | Agriculture | UGV | RGB, NIR | Segmented crop and weed sections | CNN |

**Table 5.** *Cont.*

| Ref. | Year | Application | Platform [a] | Input [b] | Percepts [c] | Algorithms [d] |
|------|------|-------------|--------------|-----------|--------------|----------------|
| [45] | 2018 | Agriculture | UGV | RGB, crop row model | Segmented crop and weed sections | SVM, LVQ, AES, ODMD |
| [85] | 2018 | Agriculture | MAV | Hyperspectral, RGB | Chlorophyll values in given area | Several |
| [87] | 2013 | Agriculture | UGV | LRF | Geometry of sprayed plant | Model fitting |
| [75] | 2016 | Forestry | UAV | RGB, NIR | Tree species distribution, health status | RF |
| [107] | 2022 | Forestry | UGV | LRF, Multispectral Images | Depth registered image, Segmented Image, 3D point clouds with live flammable material | CNN, IPBasic |
| [121] | 2007 | Forestry | UGV | LRF, RGB, NDVI, NIR | 2D Grid of traversal costs, ground plane estimate, rigid and soft obstacle classification | Linear Maximum Entropy Classifiers |
| [224] | 2016 | Forestry | UAV | RGB | Trail direction | DNN |
| [296] | 2020 | Forestry | Two UAVs | 2D LRF, IMU, Altimeter | 2D collaborative map of explored forest area, 3D voxel grid with tree positions | Frontier-based exploration, CSLAM |
| [31] | 2022 | Forestry | UAV | IMU, Stereo Cameras, LRF | 3D Semantic Map, Traversability indexes | Multi-sensor Factor Graph SLAM, SegFormer CNN |

[a] The type of robot used to test the system. UGV; UAV. [b] The signals used as input for the perception system, such as images, point clouds, etc. LRF RGB; NIR; NDVI. [c] The percepts output by the system, such as maps, plant types, localization, etc. [d] The algorithms employed by the system. CNN; Nearest neighbours (NN); SVM DNN; RF.

### 4.5. Computational Resource Management and Real-Time Operation Considerations

Artificial perception systems often involve sensing and processing on the edge, and generally need to adhere to strict execution time requirements. In fact, manufacturers, such as Xilinx (available online: https://www.xilinx.com/applications/megatrends/machine-learning.html, last accessed 25 June 2023), NVIDIA (available online: https://www.nvidia.com/en-gb/autonomous-machines/embedded-systems/jetson-tx2/, last accessed 25 June 2023), and the Raspberry Pi (RPI) foundation (available online: https://forums.raspberrypi.com//viewtopic.php?t=239812, last accessed 25 June 2023) supply are relevant computing systems for this purpose. Frequently, however, to keep costs low, or due to power considerations, perception systems are implemented using limited computational resources. Autonomy is particularly critical in forestry applications—much more than in agriculture, in fact—and as such, requires specific strategies to be put in place. For example, Niu et al. [190,349] used a Raspberry Pi, chosen due to its compact size, low cost, and low power requirements, for the deployment of their depth prediction model (see Section 4.1.2) and navigation algorithm, which are executed successfully in real time.

However, in more complicated frameworks, as seen in Section 4.3, multiple different modules need to be deployed, executed and coordinated across different systems. This has a direct impact on performance and very important implications in terms of the strategy for the deployment of each of these components on the available computational resources. In this scenario, most likely tens of gigabytes per second of sensor data will be generated which would be intractable to store locally or communicate between systems in bulk. These data must, therefore, be processed at least in part at the edge, forwarding only relevant information or higher-level constructs and representations to other systems. Finally, in many instances, real-time performance and immediate reaction times are absolutely essential for mission-critical tasks, such as avoiding obstacles and protecting surrounding fauna or human beings.

Unfortunately, despite the crucial nature of these issues, we are not aware of works in forestry robotics that explicitly address them in a systematic fashion. A notable exception

would be the research effort by Machado et al. [17] within the context of the SEMFIRE project. In this project, as described in Section 2, a heterogeneous team of robots was used, composed of the *Ranger* UGV platform and the supporting UAV platforms, called *Scouts*, and these robots were fitted with the following computational resources:

- The distributed computing system of the *Ranger* consists of two main components. The first component is the Sundance VCS-1, which is a PC/104 Linux stack comprising two main parts. The EMC2 board serves as a PCIe/104 OneBank carrier for a Trenz-compatible SoC, and the AMD-Xilinx UltraScale+ ZU4EV Multi-Processor System-On-Chip. The ZU4EV includes a quad-core ARM Cortex-A53 MPCore processor, an ARM Cortex-R5 PSU, an ARM MALI 400 GPU, and PL. The VCS-1 provides 2 GB of onboard DDR4 memory that is shared with the processor system and the programmable logic unit. The second component is an Intel I7-8700 CPU with 16GB of DDR4 memory and an NVIDIA GeForce RTX 600. This distributed system is specifically designed for accelerating state-of-the-art AI algorithms using Vitis-AI, CUDA, and CuDNN, implemented with AI frameworks like TensorFlow, Caffe, and PyTorch. Both computational devices are powered by the Robot Operating System (ROS).
- The *Scouts* computing system is based on an Intel NUC i7 with 4 GB of DDR memory, also powered by ROS. The Scouts are responsible for processing all sensory data locally using *Scout* modules. Only a minimal set of localization and post-processed data is exchanged with the *Ranger* through a wireless connection.

As explained in Section 2, the *Ranger* is a heavy-duty robot equipped with a hazardous mulcher system that may pose serious risks to both humans and animals. Therefore, to ensure safety during its operation, the robot includes both critical and non-critical sensor networks. The critical network encompasses all sensors requiring near-zero latency guarantee (i.e., LiDARs, encoders, and depth and IMU sensors) and critical localization information collected by the *Scout* platforms used to improve safety. Conversely, the non-critical network includes non-critical, low-priority sensors (i.e., GPS, encoders, multilateration transponder, and thermal and multispectral cameras). The SEMFIRE computational resource architecture (including these networks) is presented in Figure 22.
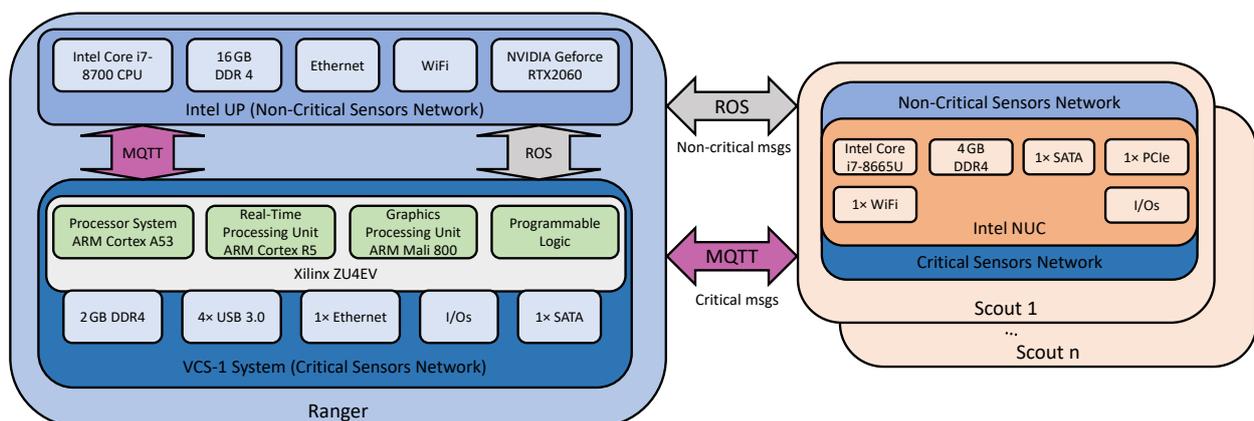


**Figure 22.** SEMFIRE computational resource architecture [17].

## 5. Discussion and Conclusions

In the following text, and given what has been presented and argued in previous sections, the current scientific and technological landscape will be discussed, including an analysis of open research questions. We will finish by drawing our final conclusions and proposing a tentative roadmap for the future.

### 5.1. Current Scientific and Technological Landscape: Open Questions and Opportunities

Despite many advances in field robotics and perceptions of outdoor robotics in particular, the current scientific and technological landscape still allows for substantial future

work in the context of robotics in forestry applications. In the remainder of this subsection, we will present several specific issues in which we believe the state of the art could be improved with further work.

### 5.1.1. Lack of Attention to Precision Forestry

In general terms, the volume of work in agricultural robotics far outweighs that of automated forestry, as seen in Table 5, resulting in the scientific under-development of the latter sub-field. This naturally presents an issue for the development of forestry-specific systems, which in effect lie at a lower TRL. However, the scientific overlap between the fields raises an interesting opportunity: since both sub-fields share many of their specific problems, the transfer of knowledge from one field to the other is expected not to be difficult. In practical terms, while the sweet pepper localization system of [61] may not be directly applicable to forestry, perhaps the underlying computer vision techniques could be used to detect certain kinds of trees instead. Similarly, while the CNN used by Barth et al. [163] to segment plants into their constituents might not find direct application in forestry, the same concept and parts of its formulation can almost certainly be applied to segment trees into, for instance, parts that should be kept and others that should be pruned.

The reasons behind the under-development of precision forestry include a generalised lack of interest by roboticists since forestry work is hard and difficult to execute, as it often involves heavy-duty machines and complex logistics. Additionally, only a relatively low amount of forest machinery is sold each year when compared to automotive, agriculture, or construction equipment industries, which implies a large percentage of R&D costs on each machine and, in turn, reduces the total effort in the field, both in-house development and collaborations with academia. Moreover, work is often published in "system papers", which are not as attractive to the scientific community. Efforts should be made to drive the industry to collaborate in spite of the lack of an obvious return on investment. Opportunities include attracting young scientists and developers by highlighting the potential positive impact of precision forestry on the environment and operational health and safety, and thereby "selling" forestry robotics in a "greener" and more ethical light, as opposed to allowing it to be construed as an invasive technology that will contribute to unemployment.

### 5.1.2. Lack of Available Data

A clear drawback for artificial perception systems in forestry, and specifically machine learning systems in this context, is the lack of annotated data and its impact on the development of advanced perception systems in forest environments. Data quality and quantity are central to the successful performance of learning-based perception systems. Thus, collecting and sharing massive amounts of data from forest scenarios is essential to train models and improve current perception systems.

Recently, with the proliferation of machine learning application techniques for solving complex problems, we witnessed a growing tendency in the scientific community to share datasets. However, current efforts are still insufficient and the lack of data delays the potential advances in this area. Open questions remain, such as: How should datasets be handled? Should editors make dataset sharing mandatory for the publication of learning-based systems in scientific journals? How can we make public sharing of datasets attractive to scholars? How can we consistently share data in a uniform way? Although challenging to put into practice, the community should address these issues sooner rather than later. Other important and related opportunities include assessing the extent to which data augmentation can help mitigate this problem (e.g., see [243]) and developing mechanisms to collect data on the fly when machines are under operation to increase the learning database. Finally, alternatives to fully supervised learning, such as semi-supervised learning, should be explored to further improve model performance.

### 5.1.3. Lack of All-Encompassing Open Software Frameworks for Perception

Software solutions, namely those devoted to artificial perception, tend to operate on disparate standards; while one solution may need images as input and output a point cloud, another might need laser scans and return 3D points, and so on. From a technological point of view, this is a natural consequence of independent development; different groups will employ different technological resources, and this will lead to different solutions. However, in order to promote the development of robust, reusable and impactful solutions, these should be integrated into a common standard framework that can be easily used in different robots and operating conditions. A generally well-regarded effort in this aspect is the ROS, which standardises the development of software modules, including perception modules, intended for use in robots. It promotes the development of decoupled techniques that can be easily reused between robots, have well-defined dependencies, and are generally better engineered. For these reasons, ROS is, in general, the go-to choice for academic researchers. However, it faces some struggles in the industry, as manufacturers often choose to use closed source systems, in-house frameworks or fully distributed middleware frameworks that offer aspects such as security, real-time support, availability on multiple operating systems, reliability in constrained networks and others. For that reason, version 2 of ROS—ROS 2—has recently been released to address these issues, even though adoption by the community proceeds at a very slow pace.

However, when applying artificial perception to robotics in forestry applications, it is important that various techniques contribute to a unified vision of the world. Thus, there is space for a novel methodology that unifies the output of these techniques and enables them to work in tandem seamlessly, while still maintaining their decoupling and portability. This would allow, for instance, the traversal analysis technique of [121] to be combined with the trail detection of [224] to obtain a combined map featuring the terrain's traversability for both human and robotic agents.

### 5.1.4. Lack of Solutions for Robot Swarms and Teams of Robots

Very few state-of-the-art solutions focus on cooperative teams of field robots, with the notable exceptions of the SEMFIRE/SAFEFOREST [27], RHEA [38], and RASberry projects. These projects make use of teams of robots, sometimes heterogeneous (i.e., involving various kinds of platforms), which tend to not be numerous, and usually involve less than 10 robots. Heterogeneous teams are of particular interest since, for instance, they are able to combine the actuation abilities of large robots with the perceptive abilities of UAVs, as seen in Section 4.3.

Swarm robotics, on the other hand, operates on the principle of employing large teams of small, relatively simple and cost-effective robots to perform a certain task. As mentioned in Section 4.3, a sparsely distributed swarm of rovers holds the potential to aid in forest monitoring. The swarm could collect spatiotemporal information, including census data on the growth of healthy tree saplings, or visually examine bark and leaves to detect signs of devastating invasive diseases. A collective swarm could collaboratively identify forest areas susceptible to wildfires, enabling targeted preventive measures. Additionally, multi-agent cooperation holds immense potential for executing distributed tasks, including reforestation, harvesting, thinning, and forwarding. Naturally, the individual rovers comprising the swarm must be compact in size (portable) to minimise their impact on the environment, such as soil compaction [348]. Furthermore, the robots need to be cost-effective to allow their widespread deployment as a swarm. Unfortunately, we could not find significant work exploring the perceptual abilities of swarms of small robots for field operations, which constitutes an important scientific and technological gap. This potential, which has been largely overlooked in previous research in forestry and agricultural robotics, should motivate further research on cooperative robotics in these application domains.

### 5.1.5. Lack of End-User Involvement and Specific Use Case Scenarios

Work on automated forestry, e.g., for harvesting, landscape maintenance, tree inventory, detection of flammable material, etc., can benefit from close contact with the end users, who have the know-how regarding field operations, strategies for increased task effectiveness, and valuable insight into the terrain and the overall forest environment. Nevertheless, the involvement of end users in the development of technological solutions in this area from an early stage is unfortunately a research gap.

Oftentimes, research groups tend to develop a "one-kill-all" system to address multiple issues, which may only work with the desired performance in specific portions of time, due to the significant number of challenges that are tackled simultaneously. However, interaction with end-users shows that robotic systems should be developed for particularly well-defined tasks. As such, having a reliable solution that works in most cases for a specific task can be more impactful than a multipurpose robotic system that is not designed for the task at hand, which will inevitably fail under certain circumstances. Paradoxically, researchers end up spending most of their time with technology issues, instead of scientific issues when dealing with complex multipurpose systems (see also Section 5.1.1). Opportunities can be withdrawn by simplifying use cases for well-defined forestry scenarios, itemising clearly what are the potential applications for end-users, and involving them in the design and requirements stage [352] as well as in the decision on which metrics to use to measure system performance.

### 5.1.6. Lack of Computational Resource Planning and Management to Satisfy Real-Time Operation Constraints

As explained in Section 4.5, there is a lack of a systematic approach to computational resource planning and management to satisfy real-time performance constraints by dealing with the following requirements:

- Power consumption and cost;
- The potentially distributed nature of the available computational resources (both in a single, specific robotic platform and between several members in a swarm or heterogeneous team) and its consequences on module deployment and execution.

These considerations are crucial in order to create mature solutions that reach higher TRLs.

### 5.1.7. Lack of Field Testing

From a scientific perspective, a solution can be validated in a laboratory environment while operating solely on pre-acquired data. This methodology allows for techniques to be tuned and refined offline, avoiding the repetition of potentially costly data collection operations. Nonetheless, technological readiness demands that techniques be validated in the operational scenario and operate in their target application under the supervision of end users, e.g., overcoming hardware failures and actuator degradation. However, many techniques, particularly software solutions, are tested in this scenario only for demonstration purposes, or not at all, since their scientific value does not depend on it.

For example, the STRANDS project [330] provided interesting scientific and technical contributions to the long-term autonomy of mobile robots in everyday environments, i.e., man-made indoor environments, where mapping, localization, and perception software components used in robots were designed to run persistently, consistently, and resiliently over a large time span of continuous operation (e.g., several weeks) in a dynamic, always changing environment, to provide security or care services. The lessons learned from the project need to be extended and adapted to the more demanding and specific challenges of field robotic applications, including robotic forestry.

This exposes a lack of technological maturity in the field of outdoor robotics, which has been recently tackled by projects such as VineScout, which aim to produce marketable and thoroughly tested solutions. Particularly in forestry applications; there are still wide opportunities for thorough field testing.

*5.2. Conclusions*

In this article, we have conducted a comprehensive survey of the state of sensing and artificial perception systems for robots in precision forestry. Our analysis has shed light on the challenges and opportunities, considering the current scientific and technological landscape. Below, we summarize the key takeaways from our review and draw a tentative roadmap for the future in this evolving field.

### 5.2.1. Key Findings

- **Challenges in artificial perception**: We identified significant challenges in the development of artificial perception systems for precision forestry, including ensuring safe operations (for both the robot and other living beings), enabling multiscale sensing and perception, and addressing specialised, expert-informed tasks. Despite remarkable technical advances, these challenges persist and remain relevant today.
- **Importance of precision forestry techniques**: Our survey highlights the paramount importance of precision forestry-specific artificial perception techniques, enabling robots to navigate and perform localised precision tasks. While actuation aspects such as flying, locomotion, or manipulation have seen significant progress, perception remains the most challenging component, requiring ongoing attention.
- **Robust and integrated perception**: Achieving full autonomy in forestry robotics requires the integration of a robust, integrated, and comprehensive set of perceptual functionalities. While progress has been made, fully reliable multipurpose autonomy remains an aspiration, with ongoing discussions about autonomy at task-specific levels, and whether operations with no human in the loop are ever needed, given the safety concerns involved.
- **Software frameworks and interoperability**: The absence of comprehensive software frameworks specific to perception has led to a fragmented landscape of technological resources. Addressing this issue is essential to advance the field, along with promoting interoperability among existing techniques, which tend to have disparate requirements and outputs.

### 5.2.2. Tentative Roadmap for the Future

Building upon our findings, we outline a tentative roadmap for the future of research and development in precision forestry robotics:

- **Advancements in sensing technologies**: We have seen a consolidated use of popular sensors in recent decades, and advancements in sensing and edge-processing technologies can be anticipated, which are likely to impact the next generation of artificial perception and decision-making systems for automated forestry.
- **Multi-robot systems for cooperative perception**: Future research should focus on multi-robot systems, both homogeneous and heterogeneous, to enhance cooperative perception. This approach can help reduce soil damage and improve overall efficiency in forestry operations.
- **Addressing rural abandonment**: Given the increasing issue of rural abandonment in many developed countries, governments should improve the attractiveness and consider investing in automated solutions to maintain and protect forested areas, mitigating risks, such as wildfires in poorly maintained or abandoned forest areas.
- **Clarifying requirements and use cases**: Stakeholders should work together to specify clear requirements and use cases that align machines with the tasks at hand, increasing the academic drive and pushing the industry towards the introduction of robots in forestry.
- **Benchmarks and standards**: The community should collaborate to develop benchmarks and standard methods for measuring success and sharing useful datasets.
- **Integrated co-robot-human teams**: As we address current challenges, we envision a future where autonomous robotic swarms or multiple robots engage in human–robot

interaction (HRI) to assist human co-worker experts, thereby forming an integrated co-robot–human team for joint forestry operations.

In conclusion, in this text, we have shown that the introduction of robotics to precision forestry imposes very significant scientific and technological problems in artificial sensing and perception, making this a particularly challenging field with an impact on economics, society, technology, and standards. However, the potential benefits of forestry robotics also make it an exciting field of research and development. By collectively addressing the identified issues and following our proposed roadmap, we can advance the impact of robotics in precision forestry, contributing more definitively to a sustainable environment and the attainment of the United Nations' sustainable development goals mentioned in Section 1.1.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| 2D | two dimensions/two-dimensional |
| 2.5D | two dimensions plus depth |
| 3D | three dimensions/three-dimensional |
| 6D | six dimensions/six-dimensional |
| BNN | Bayesian neural network |
| CNN | convolutional neural network |
| CORE | Centre of Operations for Rethinking Engineering |
| DNN | deep neural network |
| FCN | fully convolutional neural network |
| GAN | generative adversarial network |
| GDP | gross domestic product |
| GIS | geographic information systems |
| GNSS | global navigation satellite system |
| GPS | global positioning system |

| | |
|---|---|
| HRI | human–robot interaction |
| ICP | iterative closest point |
| IMU | inertial measurement unit |
| LADAR | laser detection and ranging |
| LiDAR | light detection and ranging |
| LRDSI | leaf rust disease severity index |
| LRF | laser range finder |
| MCD | Monte Carlo dropout |
| ML | machine learning |
| MoDSeM | modular framework for distributed semantic mapping |
| NDVI | normalised difference vegetation index |
| NIR | near-infrared |
| NN | nearest neighbour |
| NPQI | normalised phaeophytization index |
| NTFP | non-timber forest product |
| NTU | Nottingham Trent University |
| PCA | principal component analysis |
| PRI | photochemical reflectance index |
| PSRI | plant senescence reflectance index |
| RAISE | robotics and artificial intelligence initiative for a sustainable environment |
| RF | random forest |
| RGB | red–green–blue |
| RGB-D | red–green–blue–depth |
| ROS | Robot Operating System |
| SAFEFOREST | semi-autonomous robotic system for forest cleaning and fire prevention |
| SDG | UN sustainable development goal |
| SEMFIRE | safety, exploration, and maintenance of forests with ecological robotics |
| SIPI | structural independent pigment index |
| SLAM | simultaneous localization and mapping |
| SRI | simple ratio index |
| SVM | support vector machine |
| SWIR | short-wave infrared |
| TRL | technological readiness level |
| TSDF | truncated signed distance field |
| UAV | unmanned aerial vehicle |
| UGV | unmanned ground vehicle |
| VIS-NIR | visible and near-infrared |
| VSWIR | visible-to-short-wave-infrared |
| WWUI | wildland and wildland–urban interface |

## References

1. Agrawal, A.; Cashore, B.; Hardin, R.; Shepherd, G.; Benson, C.; Miller, D. Economic Contributions of Forests. *Backgr. Pap.* **2013**, *1*, 1–132.
2. Vaughan, R.C.; Munsell, J.F.; Chamberlain, J.L. Opportunities for Enhancing Nontimber Forest Products Management in the United States. *J. For.* **2013**, *111*, 26–33. [CrossRef]
3. Hansen, K.; Malmaeus, M. Ecosystem Services in Swedish Forests. *Scand. J. For. Res.* **2016**, *31*, 626–640. [CrossRef]
4. Karsenty, A.; Blanco, C.; Dufour, T. *Forests and Climate Change—Instruments Related to the United Nations Framework Convention on Climate Change and Their Potential for Sustainable Forest Management in Africa*; Forests and Climate Change Working Paper; FAO: Rome, Italy, 2003. Available online: https://www.fao.org/documents/card/en/c/a2e6e6ef-baee-5922-9bc4-c3b2bf5cdb80/ (accessed on 1 July 2023 ).
5. Ringdahl, O. Automation in Forestry: Development of Unmanned Forwarders. Ph.D. Thesis, Institutionen för Datavetenskap, Umeå Universitet, Umeå, Sweden, 2011.
6. Silversides, C.R. *Broadaxe to Flying Shear: The Mechanization of Forest Harvesting East of the Rockies*; Technical Report; National museum of Science and Technology: Ottawa, ON, Canada, 1997; ISBN 0-660-15980-5.
7. UN. Report of the Open Working Group of the General Assembly on Sustainable Development Goals. 2014. Available online: https://digitallibrary.un.org/record/778970 (accessed on 1 July 2023).

8.  Guenat, S.; Purnell, P.; Davies, Z.G.; Nawrath, M.; Stringer, L.C.; Babu, G.R.; Balasubramanian, M.; Ballantyne, E.E.F.; Bylappa, B.K.; Chen, B.; et al. Meeting Sustainable Development Goals via Robotics and Autonomous Systems. *Nat. Commun.* **2022**, *13*, 3559. [CrossRef] [PubMed]

9.  Choudhry, H.; O'Kelly, G. Precision Forestry: A Revolution in the Woods. 2018. Available online: https://www.mckinsey.com/industries/paper-and-forest-products/our-insights/precision-forestry-a-revolution-in-the-woods (accessed on 1 July 2023).

10. San-Miguel-Ayanz, J.; Schulte, E.; Schmuck, G.; Camia, A.; Strobl, P.; Liberta, G.; Giovando, C.; Boca, R.; Sedano, F.; Kempeneers, P.; et al. *Comprehensive Monitoring of Wildfires in Europe: The European Forest Fire Information System (EFFIS)*; IntechOpen: London, UK, 2012. [CrossRef]

11. Moreira, F.; Pe'er, G. Agricultural Policy Can Reduce Wildfires. *Science* **2018**, *359*, 1001. [CrossRef] [PubMed]

12. Gómez-González, S.; Ojeda, F.; Fernandes, P.M. Portugal and Chile: Longing for Sustainable Forestry While Rising from the Ashes. *Environ. Sci. Policy* **2018**, *81*, 104–107. [CrossRef]

13. Ribeiro, C.; Valente, S.; Coelho, C.; Figueiredo, E. A Look at Forest Fires in Portugal: Technical, Institutional, and Social Perceptions. *Scand. J. For. Res.* **2015**, *30*, 317–325. [CrossRef]

14. Suger, B.; Steder, B.; Burgard, W. Traversability Analysis for Mobile Robots in Outdoor Environments: A Semi-Supervised Learning Approach Based on 3D-lidar Data. In Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA), Washington State Convention Center, Seattle, WA, USA, 25–30 May 2015; pp. 3941–3946. [CrossRef]

15. Siegwart, R.; Lamon, P.; Estier, T.; Lauria, M.; Piguet, R. Innovative Design for Wheeled Locomotion in Rough Terrain. *Robot. Auton. Syst.* **2002**, *40*, 151–162. [CrossRef]

16. Habib, M.K.; Baudoin, Y. Robot-Assisted Risky Intervention, Search, Rescue and Environmental Surveillance. *Int. J. Adv. Robot. Syst.* **2010**, *7*, 10. [CrossRef]

17. Machado, P.; Bonnell, J.; Brandenburgh, S.; Ferreira, J.F.; Portugal, D.; Couceiro, M. Robotics Use Case Scenarios. In *Towards Ubiquitous Low-Power Image Processing Platforms*; Jahre, M., Göhringer, D., Millet, P., Eds.; Springer: Cham, Switzerland, 2021; pp. 151–172. [CrossRef]

18. Panzieri, S.; Pascucci, F.; Ulivi, G. An Outdoor Navigation System Using GPS and Inertial Platform. *IEEE/ASME Trans. Mech.* **2002**, *7*, 134–142. [CrossRef]

19. Gougeon, F.A.; Kourtz, P.H.; Strome, M. Preliminary Research on Robotic Vision in a Regenerating Forest Environment. In Proceedings of the International Symposium on Intelligent Robotic Systems, Grenoble, France, 11–15 July 1994; Volume 94, pp. 11–15. Available online: http://cfs.nrcan.gc.ca/publications?id=4582 (accessed on 1 July 2023).

20. Thorpe, C.; Durrant-Whyte, H. Field Robots. In Proceedings of the 10th International Symposium of Robotics Research (ISRR'01), Lorne, Australia, 9–12 November 2001.

21. Kelly, A.; Stentz, A.; Amidi, O.; Bode, M.; Bradley, D.; Diaz-Calderon, A.; Happold, M.; Herman, H.; Mandelbaum, R.; Pilarski, T.; et al. Toward Reliable off Road Autonomous Vehicles Operating in Challenging Environments. *Int. J. Robot. Res.* **2006**, *25*, 449–483. [CrossRef]

22. Lowry, S.; Milford, M.J. Supervised and Unsupervised Linear Learning Techniques for Visual Place Recognition in Changing Environments. *IEEE Trans. Robot.* **2016**, *32*, 600–613. [CrossRef]

23. Aguiar, A.S.; dos Santos, F.N.; Cunha, J.B.; Sobreira, H.; Sousa, A.J. Localization and mapping for robots in agriculture and forestry: A survey. *Robotics* **2020**, *9*, 97. [CrossRef]

24. Oliveira, L.F.P.; Moreira, A.P.; Silva, M.F. Advances in Forest Robotics: A State-of-the-Art Survey. *Robotics* **2021**, *10*, 53. [CrossRef]

25. SEMFIRE. Safety, Exploration and Maintenance of Forests with Ecological Robotics (SEMFIRE, Ref. CENTRO-01-0247-FEDER-03269). 2023. Available online: https://semfire.ingeniarius.pt (accessed on 1 July 2023).

26. CORE. Centre of Operations for Rethinking Engineering (CORE, Ref. CENTRO-01-0247-FEDER-037082). 2023. Available online: https://core.ingeniarius.pt (accessed on 1 July 2023).

27. Couceiro, M.; Portugal, D.; Ferreira, J.F.; Rocha, R.P. SEMFIRE: Towards a New Generation of Forestry Maintenance Multi-Robot Systems. In Proceedings of the IEEE/SICE International Symposium on System Integration, Sorbone University, Paris, France, 14–16 January 2019.

28. SAFEFOREST. Semi-Autonomous Robotic System for Forest Cleaning and Fire Prevention (SafeForest, Ref. CENTRO-01-0247-FEDER-045931). 2023. Available online: https://safeforest.ingeniarius.pt (accessed on 1 July 2023).

29. Fairfield, N.; Wettergreen, D.; Kantor, G. Segmented SLAM in three-dimensional environments. *J. Field Robot.* **2010**, *27*, 85–103. [CrossRef]

30. Silwal, A.; Parhar, T.; Yandun, F.; Baweja, H.; Kantor, G. A robust illumination-invariant camera system for agricultural applications. In Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Prague, Czech Republic, 27 September–1 October 2021; pp. 3292–3298.

31. Russell, D.J.; Arevalo-Ramirez, T.; Garg, C.; Kuang, W.; Yandun, F.; Wettergreen, D.; Kantor, G. UAV Mapping with Semantic and Traversability Metrics for Forest Fire Mitigation. In Proceedings of the ICRA 2022 Workshop in Innovation in Forestry Robotics: Research and Industry Adoption, Philadelphia, PA, USA, 23–27 May 2022.

32. Portugal, D.; Andrada, M.E.; Araújo, A.G.; Couceiro, M.S.; Ferreira, J.F. ROS Integration of an Instrumented Bobcat T190 for the SEMFIRE Project. In *Robot Operating System (ROS)*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 87–119.

33. Reis, R.; dos Santos, F.N.; Santos, L. Forest Robot and Datasets for Biomass Collection. In *Proceedings of the Robot 2019: Fourth Iberian Robotics Conference: Advances in Robotics*; Springer: Berlin/Heidelberg, Germany, 2020; Volume 1, pp. 152–163.

34. SCORPION. Scorpion-H2020—Cost Effective Robots for Smart Precision Spraying. 2023. Available online: https://scorpion-h2 020.eu/ (accessed on 1 July 2023).

35. Aguiar, A.S.; Dos Santos, F.N.; Sobreira, H.; Boaventura-Cunha, J.; Sousa, A.J. Localization and Mapping on Agriculture Based on Point-Feature Extraction and Semiplanes Segmentation From 3D LiDAR Data. *Front. Robot. AI* **2022**, *9*, 832165. [CrossRef]

36. RHEA. Robot Fleets for Highly Effective Agriculture and Forestry Management | Projects | FP7-NMP. 2018. Available online: https://cordis.europa.eu/project/rcn/95055_en.html (accessed on 1 July 2023).

37. Emmi, L.; Gonzalez-de-Santos, P. Mobile Robotics in Arable Lands: Current State and Future Trends. In Proceedings of the 2017 European Conference on Mobile Robots (ECMR), Paris, France, 6–8 September 2017; pp. 1–6. [CrossRef]

38. Gonzalez-de-Santos, P.; Ribeiro, A.; Fernandez-Quintanilla, C.; Lopez-Granados, F.; Brandstoetter, M.; Tomic, S.; Pedrazzi, S.; Peruzzi, A.; Pajares, G.; Kaplanis, G.; et al. Fleets of Robots for Environmentally-Safe Pest Control in Agriculture. *Precis. Agric.* **2017**, *18*, 574–614. [CrossRef]

39. VINEROBOT. VINEyardROBOT | Projects | FP7-ICT. 2018. Available online: https://cordis.europa.eu/project/rcn/111031_en. html (accessed on 1 July 2023).

40. Costantini, E.A.C.; Castaldini, M.; Diago, M.P.; Giffard, B.; Lagomarsino, A.; Schroers, H.J.; Priori, S.; Valboa, G.; Agnelli, A.E.; Akça, E.; et al. Effects of Soil Erosion on Agro-Ecosystem Services and Soil Functions: A Multidisciplinary Study in Nineteen Organically Farmed European and Turkish Vineyards. *J. Environ. Manag.* **2018**, *223*, 614–624. [CrossRef]

41. VineScout. News & Gallery | VineScout. 2023. Available online: http://vinescout.eu/web/newsgallery-2 (accessed on 1 July 2023).

42. Héder, M. From NASA to EU: The Evolution of the TRL Scale in Public Sector Innovation. *Innov. J.* **2017**, *22*, 1–23. Available online: https://innovation.cc/document/2017-22-2-3-from-nasa-to-eu-the-evolution-of-the-trl-scale-in-public-sector-innovation/ (accessed on 1 July 2023).

43. Riquelme, M.T.; Barreiro, P.; Ruiz-Altisent, M.; Valero, C. Olive Classification According to External Damage Using Image Analysis. *J. Food Eng.* **2008**, *87*, 371–379. [CrossRef]

44. Valente, J.; Sanz, D.; Barrientos, A.; del Cerro, J.; Ribeiro, Á.; Rossi, C.; Valente, J.; Sanz, D.; Barrientos, A.; del Cerro, J.; et al. An Air-Ground Wireless Sensor Network for Crop Monitoring. *Sensors* **2011**, *11*, 6088–6108. [CrossRef]

45. García-Santillán, I.D.; Pajares, G. On-Line Crop/Weed Discrimination through the Mahalanobis Distance from Images in Maize Fields. *Biosyst. Eng.* **2018**, *166*, 28–43. [CrossRef]

46. CROPS. Intelligent Sensing and Manipulation for Sustainable Production and Harvesting of High Value Crops, Clever Robots for Crops | Projects | FP7-NMP. 2018. Available online: https://cordis.europa.eu/project/rcn/96216_en.html (accessed on 1 July 2023).

47. Fernández, R.; Montes, H.; Salinas, C.; Sarria, J.; Armada, M. Combination of RGB and Multispectral Imagery for Discrimination of Cabernet Sauvignon Grapevine Elements. *Sensors* **2013**, *13*, 7838–7859. [CrossRef]

48. VINBOT. Autonomous Cloud-Computing Vineyard Robot to Optimise Yield Management and Wine Quality | Projects | FP7-SME. 2018. Available online: https://cordis.europa.eu/project/rcn/111459_en.html (accessed on 1 July 2023).

49. BACCHUS. BACCHUS EU Project. 2023. Available online: https://bacchus-project.eu/ (accessed on 1 July 2023).

50. Guzmán, R.; Ariño, J.; Navarro, R.; Lopes, C.M.; Graça, J.; Reyes, M.; Barriguinha, A.; Braga, R. Autonomous Hybrid GPS/Reactive Navigation of an Unmanned Ground Vehicle for Precision Viticulture-VINBOT. In Proceedings of the 62nd German Winegrowers Conference, Stuttgart, Germany, 27–30 November 2016; pp. 1–12.

51. Vestlund, K.; Hellström, T. Requirements and System Design for a Robot Performing Selective Cleaning in Young Forest Stands. *J. Terramech.* **2006**, *43*, 505–525. [CrossRef]

52. Hellström, T.; Lärkeryd, P.; Nordfjell, T.; Ringdahl, O. Autonomous Forest Vehicles: Historic, Envisioned, and State-of-the-Art. *Int. J. For. Eng.* **2009**, *20*, 31–38. [CrossRef]

53. Hellström, T.; Ostovar, A. Detection of Trees Based on Quality Guided Image Segmentation. In Proceedings of the Second International Conference on Robotics and Associated High-technologies and Equipment for Agriculture and Forestry (RHEA-2014), Madrid, Spain, 21–23 May 2014; pp. 531–540. Available online: https://www.researchgate.net/publication/266556537_ Detection_of_Trees_Based_on_Quality_Guided_Image_Segmentation (accessed on 1 July 2023).

54. Ostovar, A.; Hellström, T.; Ringdahl, O. Human Detection Based on Infrared Images in Forestry Environments. In *Image Analysis and Recognition*; Campilho, A., Karray, F., Eds.; Number 9730 in Lecture Notes in Computer Science; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 175–182. [CrossRef]

55. Hellström, T.; Ringdahl, O. A Software Framework for Agricultural and Forestry Robotics. In *Proceedings of the DIVA*; Pisa University Press: Pisa, Italy, 2012; pp. 171–176. Available online: http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-60154 (accessed on 1 July 2023).

56. Hera, P.M.L.; Trejo, O.M.; Lindroos, O.; Lideskog, H.; Lindbä, T.; Latif, S.; Li, S.; Karlberg, M. Exploring the Feasibility of Autonomous Forestry Operations: Results from the First Experimental Unmanned Machine. *Authorea* **2023**. [CrossRef]

57. Sängstuvall, L.; Bergström, D.; Lämås, T.; Nordfjell, T. Simulation of Harvester Productivity in Selective and Boom-Corridor Thinning of Young Forests. *Scand. J. For. Res.* **2012**, *27*, 56–73. [CrossRef]

58. Lindroos, O.; Ringdahl, O.; La Hera, P.; Hohnloser, P.; Hellström, T.H. Estimating the Position of the Harvester Head—A Key Step towards the Precision Forestry of the Future? *Croat. J. For. Eng.* **2015**, *36*, 147–164.

59. SWEEPER. Sweeper Homepage. 2018. Available online: http://www.sweeper-robot.eu/ (accessed on 1 July 2023).

60. SAGA. SAGA—Swarm Robotics for Agricultural Applications. 2018. Available online: http://laral.istc.cnr.it/saga/ (accessed on 1 July 2023).

61. Bac, C.W.; Hemming, J.; van Henten, E.J. Stem Localization of Sweet-Pepper Plants Using the Support Wire as a Visual Cue. *Comput. Electron. Agric.* **2014**, *105*, 111–120. [CrossRef]

62. Bac, C.W.; Hemming, J.; van Henten, E.J. Robust Pixel-Based Classification of Obstacles for Robotic Harvesting of Sweet-Pepper. *Comput. Electron. Agric.* **2013**, *96*, 148–162. [CrossRef]

63. Geerling, G.W.; Labrador-Garcia, M.; Clevers, J.G.P.W.; Ragas, A.M.J.; Smits, A.J.M. Classification of Floodplain Vegetation by Data Fusion of Spectral (CASI) and LiDAR Data. *Int. J. Remote Sens.* **2007**, *28*, 4263–4284. [CrossRef]

64. Hemming, J.; Rath, T. Computer-Vision-based Weed Identfication under Field Conditions Using Controlled Lighting. *J. Agric. Eng. Res.* **2001**, *78*, 233–243. [CrossRef]

65. Albani, D.; Nardi, D.; Trianni, V. Field Coverage and Weed Mapping by UAV Swarms. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 4319–4325. [CrossRef]

66. Digiforest. Digiforest. 2023. Available online: https://digiforest.eu (accessed on 1 July 2023).

67. Lottes, P.; Stachniss, C. Semi-Supervised Online Visual Crop and Weed Classification in Precision Farming Exploiting Plant Arrangement. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 5155–5161. [CrossRef]

68. Lottes, P.; Khanna, R.; Pfeifer, J.; Siegwart, R.; Stachniss, C. UAV-based Crop and Weed Classification for Smart Farming. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; pp. 3024–3031. [CrossRef]

69. Milioto, A.; Stachniss, C. Bonnet: An Open-Source Training and Deployment Framework for Semantic Segmentation in Robotics Using CNNs. *arXiv* **2018**, arXiv:1802.08960.

70. Lottes, P.; Behley, J.; Milioto, A.; Stachniss, C. Fully Convolutional Networks with Sequential Information for Robust Crop and Weed Detection in Precision Farming. *IEEE Robot. Autom. Lett.* **2018**, *3*, 2870–2877. [CrossRef]

71. Vieri, M.; Sarri, D.; Rimediotti, M.; Lisci, R.; Peruzzi, A.; Raffaelli, M.; Fontanelli, M.; Frasconi, C.; Martelloni, L. RHEA Project Achievement: An Innovative Spray Concept for Pesticide Application to Tree Crops Equipping a Fleet of Autonomous Robots. In Proceedings of the International Conference of Agricultural Engineering. CIGR-AgEng2012 —Valencia Conference Center, Valencia, Spain, 8–12 July 2012; p. 9.

72. Oberti, R.; Marchi, M.; Tirelli, P.; Calcante, A.; Iriti, M.; Borghese, A.N. Automatic Detection of Powdery Mildew on Grapevine Leaves by Image Analysis: Optimal View-Angle Range to Increase the Sensitivity. *Comput. Electron. Agric.* **2014**, *104*, 1–8. [CrossRef]

73. Rabatel, G.; Makdessi, N.A.; Ecarnot, M.; Roumet, P. A Spectral Correction Method for Multi-Scattering Effects in Close Range Hyperspectral Imagery of Vegetation Scenes: Application to Nitrogen Content Assessment in Wheat. *Adv. Anim. Biosci.* **2017**, *8*, 353–358. [CrossRef]

74. Lisein, J.; Michez, A.; Claessens, H.; Lejeune, P. Discrimination of Deciduous Tree Species from Time Series of Unmanned Aerial System Imagery. *PLoS ONE* **2015**, *10*, e0141006. [CrossRef]

75. Michez, A.; Piégay, H.; Lisein, J.; Claessens, H.; Lejeune, P. Classification of Riparian Forest Species and Health Condition Using Multi-Temporal and Hyperspatial Imagery from Unmanned Aerial System. *Environ. Monit. Assess.* **2016**, *188*, 146. [CrossRef]

76. Leemans, V.; Marlier, G.; Destain, M.F.; Dumont, B.; Mercatoris, B. Estimation of Leaf Nitrogen Concentration on Winter Wheat by Multispectral Imaging. In Proceedings of the SPIE Commercial + Scientific Sensing and Imaging, Anaheim, CA, USA, 28 April 2017; Bannon, D.P., Ed.; p. 102130. [CrossRef]

77. Jelavic, E.; Berdou, Y.; Jud, D.; Kerscher, S.; Hutter, M. Terrain-adaptive planning and control of complex motions for walking excavators. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25–29 October 2020; pp. 2684–2691.

78. Jelavic, E.; Jud, D.; Egli, P.; Hutter, M. Robotic Precision Harvesting: Mapping, Localization, Planning and Control for a Legged Tree Harvester. *Field Robot.* **2022**, *2*, 1386–1431. [CrossRef]

79. THING. THING—SubTerranean Haptic INvestiGator. 2023. Available online: https://thing.put.poznan.pl/ (accessed on 1 July 2023).

80. Digumarti, S.T.; Nieto, J.; Cadena, C.; Siegwart, R.; Beardsley, P. Automatic segmentation of tree structure from point cloud data. *IEEE Robot. Autom. Lett.* **2018**, *3*, 3043–3050. [CrossRef]

81. Sa, I.; Chen, Z.; Popović, M.; Khanna, R.; Liebisch, F.; Nieto, J.; Siegwart, R. weedNet: Dense Semantic Weed Classification Using Multispectral Images and MAV for Smart Farming. *IEEE Robot. Autom. Lett.* **2018**, *3*, 588–595. [CrossRef]

82. Johan From, P.; Grimstad, L.; Hanheide, M.; Pearson, S.; Cielniak, G. Rasberry-robotic and autonomous systems for berry production. *Mech. Eng.* **2018**, *140*, S14–S18. [CrossRef]

83. L-CAS. Lincoln Centre for Autonomous Systems Projects. 2012. Available online: https://lcas.lincoln.ac.uk/wp/projects/ (accessed on 1 July 2023).

84. L-CAS. Research—Hyperweeding | Harper Adams University. 2023. Available online: https://www.harper-adams.ac.uk/research/project.cfm?id=187 (accessed on 1 July 2023).

85. Mozgeris, G.; Jonikavičius, D.; Jovarauskas, D.; Zinkevičius, R.; Petkevičius, S.; Steponavičius, D. Imaging from Manned Ultra-Light and Unmanned Aerial Vehicles for Estimating Properties of Spring Wheat. *Precis. Agric.* **2018**, *19*, 876–894. [CrossRef]

86. Borz, S.A.; Talagai, N.; Cheța, M.; Montoya, A.G.; Vizuete, D.D.C. Automating Data Collection in Motor-manual Time and Motion Studies Implemented in a Willow Short Rotation Coppice. *BioResources* **2018**, *13*, 3236–3249. [CrossRef]

87. Osterman, A.; Godeša, T.; Hočevar, M.; Širok, B.; Stopar, M. Real-Time Positioning Algorithm for Variable-Geometry Air-Assisted Orchard Sprayer. *Comput. Electron. Agric.* **2013**, *98*, 175–182. [CrossRef]

88. SNOW. Project SNOW • Northern Robotics Laboratory. 2023. Available online: https://norlab.ulaval.ca/research/snow/ (accessed on 1 July 2023).

89. Pierzchala, M.; Giguère, P.; Astrup, R. Mapping Forests Using an Unmanned Ground Vehicle with 3D LiDAR and Graph-SLAM. *Comput. Electron. Agric.* **2018**, *145*, 217–225. [CrossRef]

90. Tremblay, J.F.; Béland, M.; Pomerleau, F.; Gagnon, R.; Giguère, P. Automatic 3D Mapping for Tree Diameter Measurements in Inventory Operations. *J. Field Robot.* **2020**, *37*, 1328–1346. [CrossRef]

91. Baril, D.; Deschênes, S.P.; Gamache, O.; Vaidis, M.; LaRocque, D.; Laconte, J.; Kubelka, V.; Giguère, P.; Pomerleau, F. Kilometer-scale autonomous navigation in subarctic forests: Challenges and lessons learned. *Field Robot.* **2021**, *2*, 1628–1600. [CrossRef]

92. Zhou, S.; Xi, J.; McDaniel, M.W.; Nishihata, T.; Salesses, P.; Iagnemma, K. Self-Supervised Learning to Visually Detect Terrain Surfaces for Autonomous Robots Operating in Forested Terrain. *J. Field Robot.* **2012**, *29*, 277–297. [CrossRef]

93. McDaniel, M.W.; Nishihata, T.; Brooks, C.A.; Salesses, P.; Iagnemma, K. Terrain Classification and Identification of Tree Stems Using Ground-Based LiDAR. *J. Field Robot.* **2012**, *29*, 891–910. [CrossRef]

94. Guevara, L.; Rocha, R.P.; Cheein, F.A. Improving the manual harvesting operation efficiency by coordinating a fleet of N-trailer vehicles. *Comput. Electron. Agric.* **2021**, *185*, 106103. [CrossRef]

95. Villacrés, J.; Cheein, F.A.A. Construction of 3D maps of vegetation indices retrieved from UAV multispectral imagery in forested areas. *Biosyst. Eng.* **2022**, *213*, 76–88. [CrossRef]

96. Arevalo-Ramirez, T.; Guevara, J.; Rivera, R.G.; Villacrés, J.; Menéndez, O.; Fuentes, A.; Cheein, F.A. Assessment of Multispectral Vegetation Features for Digital Terrain Modeling in Forested Regions. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 4405509. [CrossRef]

97. van Essen, R.; Harel, B.; Kootstra, G.; Edan, Y. Dynamic Viewpoint Selection for Sweet Pepper Maturity Classification Using Online Economic Decisions. *Appl. Sci.* **2022**, *12*, 4414. [CrossRef]

98. Cohen, B.; Edan, Y.; Levi, A.; Alchanatis, V. Early Detection of Grapevine (*Vitis vinifera*) Downy Mildew (*Peronospora*) and Diurnal Variations Using Thermal Imaging. *Sensors* **2022**, *22*, 3585. [CrossRef] [PubMed]

99. Windrim, L.; Bryson, M. Forest tree detection and segmentation using high resolution airborne LiDAR. In Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), The Venetian Macao, Macau, 4–8 November 2019; pp. 3898–3904.

100. Westling, F.; Underwood, J.; Bryson, M. Graph-based methods for analyzing orchard tree structure using noisy point cloud data. *Comput. Electron. Agric.* **2021**, *187*, 106270. [CrossRef]

101. Windrim, L.; Bryson, M.; McLean, M.; Randle, J.; Stone, C. Automated mapping of woody debris over harvested forest plantations using UAVs, high-resolution imagery, and machine learning. *Remote Sens.* **2019**, *11*, 733. [CrossRef]

102. ROS Agriculture. Robot Agriculture. Available online: https://github.com/ros-agriculture (accessed on 1 July 2023).

103. GREENPATROL. Galileo Enhanced Solution for Pest Detection and Control in Greenhouse Fields with Autonomous Service Robots | Projects | H2020. 2018. Available online: https://cordis.europa.eu/project/rcn/212439_en.html (accessed on 1 July 2023).

104. Tiozzo Fasiolo, D.; Scalera, L.; Maset, E.; Gasparetto, A. Recent Trends in Mobile Robotics for 3D Mapping in Agriculture. In *Proceedings of the Advances in Service and Industrial Robotics*; Mechanisms and Machine Science; Müller, A., Brandstötter, M., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 428–435. [CrossRef]

105. Fountas, S.; Mylonas, N.; Malounas, I.; Rodias, E.; Hellmann Santos, C.; Pekkeriet, E. Agricultural Robotics for Field Operations. *Sensors* **2020**, *20*, 2672. [CrossRef]

106. Ding, H.; Zhang, B.; Zhou, J.; Yan, Y.; Tian, G.; Gu, B. Recent Developments and Applications of Simultaneous Localization and Mapping in Agriculture. *J. Field Robot.* **2022**, *39*, 956–983. [CrossRef]

107. Andrada, M.E.; Ferreira, J.F.; Portugal, D.; Couceiro, M.S. Integration of an Artificial Perception System for Identification of Live Flammable Material in Forestry Robotics. In Proceedings of the 2022 IEEE/SICE International Symposium on System Integration (SII), Online, 9–12 January 2022; pp. 103–108.

108. Carvalho, A.E.; Ferreira, J.F.; Portugal, D. 3D Traversability Analysis in Forest Environments based on Mechanical Effort. In Proceedings of the 17th International Conference on Intelligent Autonomous Systems (IAS-17), Zagreb, Croatia, 13–16 June 2022; pp. 457–468.

109. Mendes, J.; Pinho, T.M.; Neves dos Santos, F.; Sousa, J.J.; Peres, E.; Boaventura-Cunha, J.; Cunha, M.; Morais, R. Smartphone Applications Targeting Precision Agriculture Practices—A Systematic Review. *Agronomy* **2020**, *10*, 855. [CrossRef]

110. Oliveira, L.F.; Moreira, A.P.; Silva, M.F. Advances in agriculture robotics: A state-of-the-art review and challenges ahead. *Robotics* **2021**, *10*, 52. [CrossRef]

111. Rovira-Más, F.; Chatterjee, I.; Sáiz-Rubio, V. The Role of GNSS in the Navigation Strategies of Cost-Effective Agricultural Robots. *Comput. Electron. Agric.* **2015**, *112*, 172–183. [CrossRef]

112. Abidi, B.R.; Aragam, N.R.; Yao, Y.; Abidi, M.A. Survey and Analysis of Multimodal Sensor Planning and Integration for Wide Area Surveillance. *ACM Comput. Surv.* **2009**, *41*, 7:1–7:36. [CrossRef]
113. Asner, G.P.; Martin, R.E.; Anderson, C.B.; Knapp, D.E. Quantifying Forest Canopy Traits: Imaging Spectroscopy versus Field Survey. *Remote Sens. Environ.* **2015**, *158*, 15–27. [CrossRef]
114. Khanal, S.; Fulton, J.; Shearer, S. An Overview of Current and Potential Applications of Thermal Remote Sensing in Precision Agriculture. *Comput. Electron. Agric.* **2017**, *139*, 22–32. [CrossRef]
115. Lowe, A.; Harrison, N.; French, A.P. Hyperspectral Image Analysis Techniques for the Detection and Classification of the Early Onset of Plant Disease and Stress. *Plant Methods* **2017**, *13*, 80. [CrossRef] [PubMed]
116. Rapaport, T.; Hochberg, U.; Shoshany, M.; Karnieli, A.; Rachmilevitch, S. Combining Leaf Physiology, Hyperspectral Imaging and Partial Least Squares-Regression (PLS-R) for Grapevine Water Status Assessment. *ISPRS J. Photogramm. Remote Sens.* **2015**, *109*, 88–97. [CrossRef]
117. Ristorto, G.; Gallo, R.; Gasparetto, A.; Scalera, L.; Vidoni, R.; Mazzetto, F. A Mobile Laboratory for Orchard Health Status Monitoring in Precision Farming. *Chem. Eng. Trans.* **2017**, *58*, 661–666. [CrossRef]
118. Cubero, S.; Marco-Noales, E.; Aleixos, N.; Barbé, S.; Blasco, J. RobHortic: A Field Robot to Detect Pests and Diseases in Horticultural Crops by Proximal Sensing. *Agriculture* **2020**, *10*, 276. [CrossRef]
119. Clamens, T.; Alexakis, G.; Duverne, R.; Seulin, R.; Fauvet, E.; Fofi, D. Real-Time Multispectral Image Processing and Registration on 3D Point Cloud for Vineyard Analysis. In Proceedings of the 16th International Conference on Computer Vision Theory and Applications, Online, 8–10 February 2021; pp. 388–398. Available online: https://www.scitepress.org/Link.aspx?doi=10.5220/0010266203880398 (accessed on 1 July 2023).
120. Halounová, L.; Junek, P.; Petruchová, J. Vegetation Indices–Tools for the Development Evaluation in Reclaimed Areas. In *Proceedings of the Global Developments in Environmental Earth Observation from Space: Proceedings of the 25th Annual Symposium of the European Association of Remote Sensing Laboratories*; IOS Press Inc.: Porto, Portugal, 2005; p. 339.
121. Bradley, D.M.; Unnikrishnan, R.; Bagnell, J. Vegetation Detection for Driving in Complex Environments. In Proceedings of the 2007 IEEE International Conference on Robotics and Automation, Rome, Italy, 10–14 April 2007; pp. 503–508. [CrossRef]
122. Meyer, G.E. Machine Vision Identification of Plants. In *Recent Trends for Enhancing the Diversity and Quality of Soybean Products*; Krezhova, D., Ed.; InTech Europe: Rijeka, Croatia, 2011; pp. 401–420. Available online: https://www.intechopen.com/chapters/22613 (accessed on 1 July 2023).
123. Symonds, P.; Paap, A.; Alameh, K.; Rowe, J.; Miller, C. A Real-Time Plant Discrimination System Utilising Discrete Reflectance Spectroscopy. *Comput. Electron. Agric.* **2015**, *117*, 57–69. [CrossRef]
124. Noble, S.D.; Brown, R.B. Plant Species Discrimination Using Spectral/Spatial Descriptive Statistics. In Proceedings of the 1st International Workshop on Computer Image Analysis in Agriculture, Potsdam, Germany, 27–28 August 2009; pp. 27–28.
125. Feyaerts, F.; van Gool, L. Multi-Spectral Vision System for Weed Detection. *Pattern Recognit. Lett.* **2001**, *22*, 667–674. [CrossRef]
126. Di Gennaro, S.F.; Battiston, E.; Di Marco, S.; Facini, O.; Matese, A.; Nocentini, M.; Palliotti, A.; Mugnai, L. Unmanned Aerial Vehicle (UAV)-Based Remote Sensing to Monitor Grapevine Leaf Stripe Disease within a Vineyard Affected by Esca Complex. *Phytopathol. Mediterr.* **2016**, *55*, 262–275. [CrossRef]
127. Hagen, N.A.; Kudenov, M.W. Review of Snapshot Spectral Imaging Technologies. *Opt. Eng.* **2013**, *52*, 090901. [CrossRef]
128. Ross, P.E. Velodyne Unveils Monster Lidar with 128 Laser Beams. 2017. Available online: https://spectrum.ieee.org/cars-that-think/transportation/sensors/velodyne-unveils-monster-lidar-with-128-laser-beams (accessed on 1 July 2023).
129. Schwarz, B. Mapping the World in 3D. *Nat. Photonics* **2010**, *4*, 429–430. [CrossRef]
130. Pellenz, J.; Lang, D.; Neuhaus, F.; Paulus, D. Real-Time 3D Mapping of Rough Terrain: A Field Report from Disaster City. In Proceedings of the 2010 IEEE Safety Security and Rescue Robotics, Bremen, Germany, 26–30 July 2010; pp. 1–6. [CrossRef]
131. Besl, P.; McKay, N.D. A Method for Registration of 3-D Shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 239–256. [CrossRef]
132. Durrant-Whyte, H.; Rye, D.; Nebot, E. Localization of autonomous guided vehicles. In Proceedings of the Robotics Research: The Seventh International Symposium, Munich, Germany, 21–24 October 1995; Springer: Berlin/Heidelberg, Germany, 1996; pp. 613–625.
133. Nüchter, A.; Lingemann, K.; Hertzberg, J.; Surmann, H. 6D SLAM—3D Mapping Outdoor Environments. *J. Field Robot.* **2007**, *24*, 699–722. [CrossRef]
134. Pearson, K. LIII. On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1901**, *2*, 559–572. [CrossRef]
135. Neuhaus, F.; Dillenberger, D.; Pellenz, J.; Paulus, D. Terrain Drivability Analysis in 3D Laser Range Data for Autonomous Robot Navigation in Unstructured Environments. In Proceedings of the 2009 IEEE Conference on Emerging Technologies & Factory Automation, Palma de Mallorca, Spain, 22–25 September 2009; pp. 1–4. [CrossRef]
136. Woods, S. Laser Scanning on the Go. *GIM Int.* **2016**, *29* –31. Available online: https://www.gim-international.com/content/article/laser-scanning-on-the-go (accessed on 1 July 2023).
137. Wurm, K.M.; Kümmerle, R.; Stachniss, C.; Burgard, W. Improving Robot Navigation in Structured Outdoor Environments by Identifying Vegetation from Laser Data. In Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, USA, 11–15 October 2009; pp. 1217–1222. [CrossRef]

138. dos Santos, A.A.; Marcato Junior, J.; Araújo, M.S.; Di Martini, D.R.; Tetila, E.C.; Siqueira, H.L.; Aoki, C.; Eltner, A.; Matsubara, E.T.; Pistori, H.; et al. Assessment of CNN-Based Methods for Individual Tree Detection on Images Captured by RGB Cameras Attached to UAVs. *Sensors* **2019**, *19*, 3595. [CrossRef]

139. da Silva, D.Q.; dos Santos, F.N.; Sousa, A.J.; Filipe, V. Visible and Thermal Image-Based Trunk Detection with Deep Learning for Forestry Mobile Robotics. *J. Imaging* **2021**, *7*, 176. [CrossRef]

140. da Silva, D.Q.; dos Santos, F.N.; Filipe, V.; Sousa, A.J.; Oliveira, P.M. Edge AI-Based Tree Trunk Detection for Forestry Monitoring Robotics. *Robotics* **2022**, *11*, 136. [CrossRef]

141. Goeau, H.; Bonnet, P.; Joly, A. Plant Identification Based on Noisy Web Data: The Amazing Performance of Deep Learning (LifeCLEF 2017). In Proceedings of the CLEF 2017—Conference and Labs of the Evaluation Forum, Dublin, Ireland, 11–14 September 2017; pp. 1–13. Available online: https://hal.archives-ouvertes.fr/hal-01629183 (accessed on 1 July 2023).

142. Kumar, N.; Belhumeur, P.N.; Biswas, A.; Jacobs, D.W.; Kress, W.J.; Lopez, I.C.; Soares, J.V.B. Leafsnap: A Computer Vision System for Automatic Plant Species Identification. In *Computer Vision—ECCV 2012*; Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C., Eds.; Number 7573 in Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germnay, 2012; pp. 502–516. [CrossRef]

143. Affouard, A.; Goëau, H.; Bonnet, P.; Lombardo, J.C.; Joly, A. Pl@Ntnet App in the Era of Deep Learning. In Proceedings of the ICLR 2017 Workshop Track—5th International Conference on Learning Representations, Toulon, France, 24–26 April 2017.

144. Goëau, H.; Joly, A.; Yahiaoui, I.; Bakić, V.; Verroust-Blondet, A.; Bonnet, P.; Barthélémy, D.; Boujemaa, N.; Molino, J.F. Plantnet Participation at Lifeclef2014 Plant Identification Task. In Proceedings of the CLEF2014 Working Notes Working Notes for CLEF 2014 Conference CEUR-WS, Sheffield, UK, 15–18 September 2014; pp. 724–737.

145. Sun, Y.; Liu, Y.; Wang, G.; Zhang, H. Deep Learning for Plant Identification in Natural Environment. *Comput. Intell. Neurosci.* **2017**, *2017*, 7361042. [CrossRef] [PubMed]

146. Borregaard, T.; Nielsen, H.; Nørgaard, L.; Have, H. Crop–Weed Discrimination by Line Imaging Spectroscopy. *J. Agric. Eng. Res.* **2000**, *75*, 389–400. [CrossRef]

147. Piron, A.; Leemans, V.; Kleynen, O.; Lebeau, F.; Destain, M.F. Selection of the Most Efficient Wavelength Bands for Discriminating Weeds from Crop. *Comput. Electron. Agric.* **2008**, *62*, 141–148. [CrossRef]

148. Weiss, U.; Biber, P.; Laible, S.; Bohlmann, K.; Zell, A. Plant Species Classification Using a 3D LIDAR Sensor and Machine Learning. In Proceedings of the IEEE Ninth International Conference on Machine Learning and Applications (ICMLA'10), Washington, DC, USA, 12–14 December 2010; pp. 339–345.

149. Bradley, D.; Thayer, S.; Stentz, A.; Rander, P. Vegetation Detection for Mobile Robot Navigation. Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-04-12. 2004. Available online: http://www.ri.cmu.edu/pub_files/pub4/bradley_david_2004_2/bradley_david_2004_2.pdf (accessed on 1 July 2023).

150. Brunner, A.; Gizachew, B. Rapid Detection of Stand Density, Tree Positions, and Tree Diameter with a 2D Terrestrial Laser Scanner. *Eur. J. For. Res.* **2014**, *133*, 819–831. [CrossRef]

151. Fiel, S.; Sablatnig, R. *Automated Identification of Tree Species from Images of the Bark, Leaves and Needles*; Technical Report CVL-TR-3; TU Wien, Faculty of Informatics, Computer Vision Lab: Vienna, Austria, 2010. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.379.1376&rep=rep1&type=pdf#page=67 (accessed on 1 July 2023).

152. Dechesne, C.; Mallet, C.; Le Bris, A.; Gouet-Brunet, V. Semantic Segmentation of Forest Stands of Pure Species as a Global Optimization Problem. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *4*, 141–148. [CrossRef]

153. Li, H.; Zhang, X.; Jaeger, M.; Constant, T. Segmentation of Forest Terrain Laser Scan Data. In Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry, Seoul, Republic of Korea, 12–13 December 2010; pp. 47–54.

154. Cerutti, G.; Tougne, L.; Mille, J.; Vacavant, A.; Coquin, D. Understanding Leaves in Natural Images—A Model-Based Approach for Tree Species Identification. *Comput. Vis. Image Underst.* **2013**, *117*, 1482–1501. [CrossRef]

155. Carpentier, M.; Giguère, P.; Gaudreault, J. Tree Species Identification from Bark Images Using Convolutional Neural Networks. *arXiv* **2018**, arXiv:1803.00949.

156. Valada, A.; Mohan, R.; Burgard, W. Self-Supervised Model Adaptation for Multimodal Semantic Segmentation. *Int. J. Comput. Vis.* **2020**, *128*, 1239–1285. [CrossRef]

157. Andrada, M.E.; Ferreira, J.; Portugal, D.; Couceiro, M. Testing Different CNN Architectures for Semantic Segmentation for Landscaping with Forestry Robotics. In Proceedings of the Workshop on Perception, Planning and Mobility in Forestry Robotics, Virtual Workshop, 29 October 2020.

158. Fortin, J.M.; Gamache, O.; Grondin, V.; Pomerleau, F.; Giguère, P. Instance Segmentation for Autonomous Log Grasping in Forestry Operations. In Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Kyoto, Japan, 23–27 October 2022; pp. 6064–6071. [CrossRef]

159. Li, H.; Liu, J.; Wang, D. A Fast Instance Segmentation Technique for Log End Faces Based on Metric Learning. *Forests* **2023**, *14*, 795. [CrossRef]

160. Grondin, V.; Fortin, J.M.; Pomerleau, F.; Giguère, P. Tree Detection and Diameter Estimation Based on Deep Learning. *For. Int. J. For. Res.* **2023**, *96*, 264–276. [CrossRef]

161. Teng, C.H.; Chen, Y.S.; Hsu, W.H. Tree Segmentation from an Image. In Proceedings of the 9th IAPR Conference on Machine Vision Applications (MVA), Tsukuba Science City, Japan, 16–18 May 2005; pp. 59–63.

162. Sodhi, P.; Vijayarangan, S.; Wettergreen, D. In-Field Segmentation and Identification of Plant Structures Using 3D Imaging. In Proceedings of the Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference, Vancouver, BC, Canada, 24–28 September 2017; pp. 5180–5187.

163. Barth, R.; Hemming, J.; van Henten, E.J. Improved Part Segmentation Performance by Optimising Realism of Synthetic Images Using Cycle Generative Adversarial Networks. *arXiv* **2018**, arXiv:1803.06301.

164. Anantrasirichai, N.; Hannuna, S.; Canagarajah, N. Automatic Leaf Extraction from Outdoor Images. *arXiv* **2017**, arXiv:1709.06437.

165. Dechesne, C.; Lassalle, P.; Lefèvre, S. Bayesian U-Net: Estimating Uncertainty in Semantic Segmentation of Earth Observation Images. *Remote Sens.* **2021**, *13*, 3836. [CrossRef]

166. Mukhoti, J.; Gal, Y. Evaluating Bayesian Deep Learning Methods for Semantic Segmentation. *arXiv* **2018**, arXiv:1811.12709

167. Kendall, A.; Badrinarayanan, V.; Cipolla, R. Bayesian Segnet: Model Uncertainty in Deep Convolutional Encoder-Decoder Architectures for Scene Understanding. *arXiv* **2015**, arXiv:1511.02680

168. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]

169. Lo, T.W.; Siebert, J. Local feature extraction and matching on range images: 2.5D SIFT. *Comput. Vis. Image Underst.* **2009**, *113*, 1235–1250. [CrossRef]

170. Knopp, J.; Prasad, M.; Willems, G.; Timofte, R.; Van Gool, L. Hough Transform and 3D SURF for Robust Three Dimensional Classification. In Proceedings of the Computer Vision—ECCV 2010, Heraklion, Crete, 5–11 September 2010; Daniilidis, K., Maragos, P., Paragios, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 589–602. [CrossRef]

171. Aubry, M.; Schlickewei, U.; Cremers, D. The wave kernel signature: A quantum mechanical approach to shape analysis. In Proceedings of the 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), Barcelona, Spain, 6–13 November 2011; IEEE Computer Society: Los Alamitos, CA, USA, 2011; pp. 1626–1633. [CrossRef]

172. Ghrabat, M.; Ma, G.; Maolood, I.; Alresheedi, S.; Abduljabbar, Z. An effective image retrieval based on optimized genetic algorithm utilized a novel SVM-based convolutional neural network classifier. *Hum.-Centric Comput. Inf. Sci.* **2019**, *9*, 31. [CrossRef]

173. Hänsch, R.; Weber, T.; Hellwich, O. Comparison of 3D interest point detectors and descriptors for point cloud fusion. *Isprs Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *II-3*, 57–64. [CrossRef]

174. Li, B.; Zhang, T.; Xia, T. Vehicle Detection from 3D Lidar Using Fully Convolutional Network. In *Proceedings of the Robotics: Science and Systems XII*; Hsu, D., Amato, N.M., Berman, S., Jacobs, S.A., Eds.; University of Michigan: Ann Arbor, MI, USA, 2016. [CrossRef]

175. Graham, B.; Engelcke, M.; Maaten, L.v.d. 3D Semantic Segmentation with Submanifold Sparse Convolutional Networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–22 June 2018; pp. 9224–9232. [CrossRef]

176. Qi, C.R.; Su, H.; NieBner, M.; Dai, A.; Yan, M.; Guibas, L.J. Volumetric and Multi-view CNNs for Object Classification on 3D Data. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; IEEE Computer Society: Los Alamitos, CA, USA, 2016; pp. 5648–5656. [CrossRef]

177. Ben-Shabat, Y.; Lindenbaum, M.; Fischer, A. 3DmFV: Three-Dimensional Point Cloud Classification in Real-Time Using Convolutional Neural Networks. *IEEE Robot. Autom. Lett.* **2018**, *3*, 3145–3152. [CrossRef]

178. Song, W.; Zhang, L.; Tian, Y.; Fong, S.; Liu, J.; Gozho, A. CNN-based 3D Object Classification Using Hough Space of LiDAR Point Clouds. *Hum.-Centric Comput. Inf. Sci.* **2020**, *10*, 19. [CrossRef]

179. Milioto, A.; Vizzo, I.; Behley, J.; Stachniss, C. RangeNet ++: Fast and Accurate LiDAR Semantic Segmentation. In Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), The Venetian Macao, Macau, 4–8 November 2019; pp. 4213–4220. [CrossRef]

180. Charles, R.; Su, H.; Kaichun, M.; Guibas, L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 22–25 July 2017; IEEE Computer Society: Los Alamitos, CA, USA, 2017; pp. 77–85. [CrossRef]

181. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. In Proceedings of the 31st International Conference on Neural Information Processing Systems NIPS'17, Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017; pp. 5105–5114. Available online: https://proceedings.neurips.cc/paper_files/paper/2017/file/d8bf84be3800d12f74d8b05e9b89836f-Paper.pdf (accessed on 1 July 2023).

182. Hu, Q.; Yang, B.; Xie, L.; Rosa, S.; Guo, Y.; Wang, Z.; Trigoni, N.; Markham, A. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Online, 14–19 June 2020; pp. 11105–11114. [CrossRef]

183. Li, Y.; Bu, R.; Sun, M.; Wu, W.; Di, X.; Chen, B. PointCNN: Convolution on x-Transformed Points. In Proceedings of the 32nd International Conference on Neural Information Processing Systems NIPS'18, Montreal, QC, Canada, 2–8 December 2018; Curran Associates Inc.: Red Hook, NY, USA, 2018; pp. 828–838. [CrossRef]

184. Groh, F.; Wieschollek, P.; Lensch, H.P.A. Flex-Convolution (million-scale point-cloud learning beyond grid-worlds). In Proceedings of the Computer Vision—ACCV 2018, Perth, Australia, 2–6 December 2018; Jawahar, C.V., Li, H., Mori, G., Schindler, K., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 105–122. [CrossRef]

185. Dovrat, O.; Lang, I.; Avidan, S. Learning to Sample. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–20 June 2019; IEEE Computer Society: Los Alamitos, CA, USA, 2019; pp. 2755–2764. [CrossRef]

186. Yang, J.; Zhang, Q.; Ni, B.; Li, L.; Liu, J.; Zhou, M.; Tian, Q. Modeling Point Clouds With Self-Attention and Gumbel Subset Sampling. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–20 June 2019; pp. 3318–3327. [CrossRef]

187. Sutton, R.S.; McAllester, D.; Singh, S.; Mansour, Y. Policy Gradient Methods for Reinforcement Learning with Function Approximation. In Proceedings of the Advances in Neural Information Processing Systems 1999, Denver, CO, USA, 29 November–4 December 1999; Solla, S., Leen, T., Müller, K., Eds.; MIT Press: Cambridge, MA, USA; Volume 12.

188. Mukhandi, H.; Ferreira, J.F.; Peixoto, P. Systematic Sampling of Large-Scale LiDAR Point Clouds for Semantic Segmentation in Forestry Robotics. *Electr. Electron. Eng.* 2023, *preprints*.

189. Bello, S.A.; Yu, S.; Wang, C.; Adam, J.M.; Li, J. Review: Deep Learning on 3D Point Clouds. *Remote Sens.* **2020**, *12*, 1729. [CrossRef]

190. Niu, C.; Zauner, K.P.; Tarapore, D. An Embarrassingly Simple Approach for Visual Navigation of Forest Environments. *Front. Robot. AI* **2023**, *10* . [CrossRef]

191. Xie, D.; Chen, L.; Liu, L.; Chen, L.; Wang, H. Actuators and Sensors for Application in Agricultural Robots: A Review. *Machines* **2022**, *10*, 913. [CrossRef]

192. Ku, J.; Harakeh, A.; Waslander, S.L. In Defense of Classical Image Processing: Fast Depth Completion on the CPU. In Proceedings of the 2018 15th Conference on Computer and Robot Vision (CRV), Toronto, ON, Canada, 8–10 May 2018; pp. 16–22.

193. Zhao, Y.; Bai, L.; Zhang, Z.; Huang, X. A Surface Geometry Model for LiDAR Depth Completion. *IEEE Robot. Autom. Lett.* **2021**, *6*, 4457–4464. [CrossRef]

194. Xie, Z.; Yu, X.; Gao, X.; Li, K.; Shen, S. Recent Advances in Conventional and Deep Learning-Based Depth Completion: A Survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, 1–12. [CrossRef]

195. Hu, M.; Wang, S.; Li, B.; Ning, S.; Fan, L.; Gong, X. Towards Precise and Efficient Image Guided Depth Completion. In Proceedings of the 2021 International Conference on Robotics and Automation (ICRA 2021), Xi'an, China, 30 May–5 June 2021.

196. Nunes, R.; Ferreira, J.; Peixoto, P. SynPhoRest—Synthetic Photorealistic Forest Dataset with Depth Information for Machine Learning Model Training. 2022. Available online: https://doi.org/10.5281/zenodo.6369445 (accessed on 1 July 2023).

197. Lin, M.; Cao, L.; Zhang, Y.; Shao, L.; Lin, C.W.; Ji, R. Pruning Networks with Cross-Layer Ranking & k-Reciprocal Nearest Filters. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, 1–10. [CrossRef]

198. Yan, Z.; Wang, K.; Li, X.; Zhang, Z.; Li, J.; Yang, J. RigNet: Repetitive Image Guided Network for Depth Completion. *arXiv* **2021**, arXiv:2107.13802. https://doi.org/10.48550/ARXIV.2107.13802.

199. Wong, A.; Cicek, S.; Soatto, S. Learning Topology from Synthetic Data for Unsupervised Depth Completion. *IEEE Robot. Autom. Lett.* **2021**, *6*, 1495–1502. [CrossRef]

200. Eldesokey, A.; Felsberg, M.; Holmquist, K.; Persson, M. Uncertainty-Aware CNNs for Depth Completion: Uncertainty from Beginning to End. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020), Online, 14–19 June 2020; pp. 12014–12023.

201. KITTI. The KITTI Vision Benchmark Suite. 2023. Available online: https://www.cvlibs.net/datasets/kitti/eval_depth.php?benchmark=depth_completion (accessed on 1 July 2023).

202. Liu, X.; Nardari, G.V.; Ojeda, F.C.; Tao, Y.; Zhou, A.; Donnelly, T.; Qu, C.; Chen, S.W.; Romero, R.A.; Taylor, C.J.; et al. Large-Scale Autonomous Flight with Real-Time Semantic Slam under Dense Forest Canopy. *IEEE Robot. Autom. Lett.* **2022**, *7*, 5512–5519. [CrossRef]

203. Andrada, M.E.; Ferreira, J.F.; Kantor, G.; Portugal, D.; Antunes, C.H. Model Pruning in Depth Completion CNNs for Forestry Robotics with Simulated Annealing. In Proceedings of the ICRA 2022 Workshop in Innovation in Forestry Robotics: Research and Industry Adoption, Philadelphia, PA, USA, 23 May 2022.

204. Han, X.; Li, S.; Wang, X.; Zhou, W. Semantic Mapping for Mobile Robots in Indoor Scenes: A Survey. *Information* **2021**, *12*, 92. [CrossRef]

205. Yang, Z.; Liu, C. TUPPer-Map: Temporal and Unified Panoptic Perception for 3D Metric-Semantic Mapping. In Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Prague, Czech Republic, 27 September–1 October 2021; pp. 1094–1101.

206. Chang, Y.; Tian, Y.; How, J.P.; Carlone, L. Kimera-Multi: A System for Distributed Multi-Robot Metric-Semantic Simultaneous Localization and Mapping. In Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 30 May–5 June 2021; pp. 11210–11218. [CrossRef]

207. Li, L.; Kong, X.; Zhao, X.; Huang, T.; Liu, Y. Semantic Scan Context: A Novel Semantic-Based Loop-Closure Method for LiDAR SLAM. *Auton. Robots* **2022**, *46*,535–551. [CrossRef]

208. Gan, L.; Kim, Y.; Grizzle, J.W.; Walls, J.M.; Kim, A.; Eustice, R.M.; Ghaffari, M. Multitask Learning for Scalable and Dense Multilayer Bayesian Map Inference. *IEEE Trans. Robot.* **2022**, *39*, 699–717. [CrossRef]

209. Liu, J.; Jung, C. NNNet: New Normal Guided Depth Completion from Sparse LiDAR Data and Single Color Image. *IEEE Access* **2022**, *10*, 114252–114261. [CrossRef]

210. Hornung, A.; Wurm, K.M.; Bennewitz, M.; Stachniss, C.; Burgard, W. OctoMap: An Efficient Probabilistic 3D Mapping Framework Based on Octrees. *Auton. Robot.* **2013**, *34*, 189–206. [CrossRef]

211. Doherty, K.; Shan, T.; Wang, J.; Englot, B. Learning-Aided 3-D Occupancy Mapping With Bayesian Generalized Kernel Inference. *IEEE Trans. Robot.* **2019**, *35*, 953–966. [CrossRef]
212. Borges, P.; Peynot, T.; Liang, S.; Arain, B.; Wildie, M.; Minareci, M.; Lichman, S.; Samvedi, G.; Sa, I.; Hudson, N.; et al. A Survey on Terrain Traversability Analysis for Autonomous Ground Vehicles: Methods, Sensors, and Challenges. *Field Robot.* **2022**, *2*, 1567–1627. [CrossRef]
213. Wu, H.; Liu, B.; Su, W.; Chen, Z.; Zhang, W.; Ren, X.; Sun, J. Optimum pipeline for visual terrain classification using improved bag of visual words and fusion methods. *J. Sens.* **2017**, *2017*, 8513949. [CrossRef]
214. Palazzo, S.; Guastella, D.C.; Cantelli, L.; Spadaro, P.; Rundo, F.; Muscato, G.; Giordano, D.; Spampinato, C. Domain adaptation for outdoor robot traversability estimation from RGB data with safety-preserving Loss. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25–29 October 2020; pp. 10014–10021.
215. Reina, G.; Leanza, A.; Milella, A.; Messina, A. Mind the ground: A power spectral density-based estimator for all-terrain rovers. *Measurement* **2020**, *151*, 107136. [CrossRef]
216. Goodin, C.; Dabbiru, L.; Hudson, C.; Mason, G.; Carruth, D.; Doude, M. Fast terrain traversability estimation with terrestrial lidar in off-road autonomous navigation. In Proceedings of the SPIE Unmanned Systems Technology XXIII, Online, 12–16 April 2021; Volume 11758, pp. 189–199.
217. Rankin, A.L.; Matthies, L.H. Passive sensor evaluation for unmanned ground vehicle mud detection. *J. Field Robot.* **2010**, *27*, 473–490. [CrossRef]
218. Ahtiainen, J.; Peynot, T.; Saarinen, J.; Scheding, S.; Visala, A. Learned ultra-wideband RADAR sensor model for augmented LIDAR-based traversability mapping in vegetated environments. In Proceedings of the 18th International Conference on Information Fusion (Fusion 2015), Washington, DC, USA, 6–9 July 2015; pp. 953–960.
219. Winkens, C.; Sattler, F.; Paulus, D. Hyperspectral Terrain Classification for Ground Vehicles. In Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP—5: VISAPP), Porto, Portugal, 27 February–1 March 2017; pp. 417–424.
220. Milella, A.; Reina, G.; Nielsen, M. A multi-sensor robotic platform for ground mapping and estimation beyond the visible spectrum. *Precis. Agric.* **2019**, *20*, 423–444. [CrossRef]
221. Vulpi, F.; Milella, A.; Marani, R.; Reina, G. Recurrent and convolutional neural networks for deep terrain classification by autonomous robots. *J. Terramech.* **2021**, *96*, 119–131. [CrossRef]
222. Usui, K. Data augmentation using image-to-image translation for detecting forest strip roads based on deep learning. *Int. J. For. Eng.* **2021**, *32*, 57–66. [CrossRef]
223. Tai, L.; Li, S.; Liu, M. A deep-network solution towards model-less obstacle avoidance. In Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Republic of Korea, 9–14 October 2016; pp. 2759–2764.
224. Giusti, A.; Guzzi, J.; Cireşan, D.C.; He, F.L.; Rodríguez, J.P.; Fontana, F.; Faessler, M.; Forster, C.; Schmidhuber, J.; Caro, G.D.; et al. A Machine Learning Approach to Visual Perception of Forest Trails for Mobile Robots. *IEEE Robot. Autom. Lett.* **2016**, *1*, 661–667. [CrossRef]
225. Sihvo, S.; Virjonen, P.; Nevalainen, P.; Heikkonen, J. Tree detection around forest harvester based on onboard LiDAR measurements. In Proceedings of the 2018 Baltic Geodetic Congress (BGC Geomatics), Olsztyn, Poland, 21–23 June 2018; pp. 364–367.
226. Liu, M.; Han, Z.; Chen, Y.; Liu, Z.; Han, Y. Tree species classification of LiDAR data based on 3D deep learning. *Measurement* **2021**, *177*, 109301. [CrossRef]
227. Wang, C.; Wang, J.; Li, C.; Ho, D.; Cheng, J.; Yan, T.; Meng, L.; Meng, M.Q.H. Safe and robust mobile robot navigation in uneven indoor environments. *Sensors* **2019**, *19*, 2993. [CrossRef]
228. Yang, S.; Yang, S.; Yi, X. An efficient spatial representation for path planning of ground robots in 3D environments. *IEEE Access* **2018**, *6*, 41539–41550. [CrossRef]
229. Fankhauser, P.; Bloesch, M.; Hutter, M. Probabilistic terrain mapping for mobile robots with uncertain localization. *IEEE Robot. Autom. Lett.* **2018**, *3*, 3019–3026. [CrossRef]
230. Ruetz, F.; Hernández, E.; Pfeiffer, M.; Oleynikova, H.; Cox, M.; Lowe, T.; Borges, P. Ovpc mesh: 3d free-space representation for local ground vehicle navigation. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; pp. 8648–8654.
231. Krüsi, P.; Furgale, P.; Bosse, M.; Siegwart, R. Driving on point clouds: Motion planning, trajectory optimization, and terrain assessment in generic nonplanar environments. *J. Field Robot.* **2017**, *34*, 940–984. [CrossRef]
232. Ramachandram, D.; Taylor, G.W. Deep Multimodal Learning: A Survey on Recent Advances and Trends. *IEEE Signal Process. Mag.* **2017**, *34*, 96–108. [CrossRef]
233. Bao, Y.; Song, K.; Wang, J.; Huang, L.; Dong, H.; Yan, Y. Visible and Thermal Images Fusion Architecture for Few-Shot Semantic Segmentation. *J. Vis. Commun. Image Represent.* **2021**, *80*, 103306. [CrossRef]
234. Choe, G.; Kim, S.H.; Im, S.; Lee, J.Y.; Narasimhan, S.G.; Kweon, I.S. RANUS: RGB and NIR Urban Scene Dataset for Deep Scene Parsing. *IEEE Robot. Autom. Lett.* **2018**, *3*, 1808–1815. [CrossRef]
235. Ali, I.; Durmush, A.; Suominen, O.; Yli-Hietanen, J.; Peltonen, S.; Collin, J.; Gotchev, A. FinnForest Dataset: A Forest Landscape for Visual SLAM. *Robot. Auton. Syst.* **2020**, *132*, 103610. [CrossRef]
236. da Silva, D.Q.; dos Santos, F.N.; Santos, L.; Aguiar, A. QuintaReiFMD - ROS1.0 Bag Dataset Acquired with AgRob V16 in Portuguese Forest. 2021. Available online: https://doi.org/10.5281/zenodo.5045355 (accessed on 1 July 2023).

237. da Silva, D.Q.; dos Santos, F.N. ForTrunkDet—Forest Dataset of Visible and Thermal Annotated Images for Object Detection. 2021. Available online: https://doi.org/10.5281/zenodo.5213825 (accessed on 1 July 2023).

238. Cordts, M.; Omran, M.; Ramos, S.; Scharwächter, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B. The Cityscapes Dataset. In Proceedings of the CVPR Workshop on the Future of Datasets in Vision, Boston, MA, USA, 8–10 June 2015; Volume 2.

239. Niu, C.; Tarapore, D.; Zauner, K.P. Low-viewpoint forest depth dataset for sparse rover swarms. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25–29 October 2020; pp. 8035–8040.

240. Grondin, V.; Pomerleau, F.; Giguère, P. Training Deep Learning Algorithms on Synthetic Forest Images for Tree Detection. In Proceedings of the ICRA 2022 Workshop in Innovation in Forestry Robotics: Research and Industry Adoption, Philadelphia, PA, USA, 23 May 2022. Available online: https://openreview.net/forum?id=SxWgxLtyW7c (accessed on 1 July 2023).

241. Uhrig, J.; Schneider, N.; Schneider, L.; Franke, U.; Brox, T.; Geiger, A. Sparsity Invariant Cnns. In Proceedings of the 2017 International Conference on 3D Vision (3DV), Qingdao, China, 10–12 October 2017; pp. 11–20.

242. Caesar, H.; Bankiti, V.; Lang, A.H.; Vora, S.; Liong, V.E.; Xu, Q.; Krishnan, A.; Pan, Y.; Baldan, G.; Beijbom, O. nuScenes: A Multimodal Dataset for Autonomous Driving. *arXiv* **2019**, arXiv:1903.11027.

243. Bittner, D.; Andrada, M.E.; Portugal, D.; Ferreira, J.F. SEMFIRE Forest Dataset for Semantic Segmentation and Data Augmentation. 2021. Available online: https://doi.org/10.5281/ZENODO.5819064 (accessed on 1 July 2023).

244. Wang, W.; Zhu, D.; Wang, X.; Hu, Y.; Qiu, Y.; Wang, C.; Hu, Y.; Kapoor, A.; Scherer, S. Tartanair: A Dataset to Push the Limits of Visual Slam. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, Las Vegas, NV, USA, 25–29 October 2020; pp. 4909–4916.

245. Ros, G.; Sellart, L.; Materzynska, J.; Vazquez, D.; Lopez, A.M. The Synthia Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 3234–3243.

246. Nunes, R.; Ferreira, J.; Peixoto, P. Procedural Generation of Synthetic Forest Environments to Train Machine Learning Algorithms. In Proceedings of the ICRA 2022 Workshop in Innovation in Forestry Robotics: Research and Industry Adoption, Philadelphia, PA, USA, 23 May 2022. Available online: https://irep.ntu.ac.uk/id/eprint/46417/ (accessed on 1 July 2023).

247. Kesten, R.; Usman, M.; Houston, J.; Pandya, T.; Nadhamuni, K.; Ferreira, A.; Yuan, M.; Low, B.; Jain, A.; Ondruska, P.; et al. Level 5 Perception Dataset 2020. 2019. Available online: https://apera.io/a/tech/561428/lyft-level-5-dataset#:~:text=The%20Lyft%20Level%205%20Dataset,(including%20lanes%20and%20crosswalks) (accessed on 1 July 2023).

248. Bittner, D. Data Augmentation Solutions for CNN-Based Semantic Segmentation in Forestry Applications. Bachelor's Thesis, Regensburg University of Applied Sciences (OTH), Regensburg, Germany, 2022.

249. Shorten, C.; Khoshgoftaar, T.M. A Survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [CrossRef]

250. Bird, J.J.; Faria, D.R.; Ekárt, A.; Ayrosa, P.P.S. From Simulation to Reality: CNN Transfer Learning for Scene Classification. In Proceedings of the 2020 IEEE 10th International Conference on Intelligent Systems (IS), Varna, Bulgaria, 28–30 August 2020; pp. 619–625. [CrossRef]

251. Bittner, D.; Ferreira, J.F.; Andrada, M.E.; Bird, J.J.; Portugal, D. Generating Synthetic Multispectral Images for Semantic Segmentation in Forestry Applications. In Proceedings of the ICRA 2022 Workshop in Innovation in Forestry Robotics: Research and Industry Adoption, Philadelphia, PA, USA, 23 May 2022. Available online: https://irep.ntu.ac.uk/id/eprint/46416/ (accessed on 1 July 2023).

252. Gao, Y.; Mosalam, K.M. Deep Transfer Learning for Image-Based Structural Damage Recognition. *Comput.-Aided Civ. Infrastruct. Eng.* **2018**, *33*, 748–768. [CrossRef]

253. Hussain, M.; Bird, J.J.; Faria, D.R. A Study on CNN Transfer Learning for Image Classification. In *Advances in Computational Intelligence Systems*; Advances in Intelligent Systems and Computing; Lotfi, A., Bouchachia, H., Gegov, A., Langensiepen, C., McGinnity, M., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 191–202. [CrossRef]

254. Johnson, J.M.; Khoshgoftaar, T.M. Survey on Deep Learning with Class Imbalance. *J. Big Data* **2019**, *6*, 27. [CrossRef]

255. Liu, Y.; Yang, G.; Qiao, S.; Liu, M.; Qu, L.; Han, N.; Wu, T.; Yuan, G.; Wu, T.; Peng, Y. Imbalanced Data Classification: Using Transfer Learning and Active Sampling. *Eng. Appl. Artif. Intell.* **2023**, *117*, 105621. [CrossRef]

256. Younes, G.; Asmar, D.; Shammas, E.; Zelek, J. Keyframe-Based Monocular SLAM: Design, Survey, and Future Directions. *Robot. Auton. Syst.* **2017**, *98*, 67–88. [CrossRef]

257. Scaramuzza, D.; Fraundorfer, F. Visual Odometry [Tutorial]. *IEEE Robot. Autom. Mag.* **2011**, *18*, 80–92. [CrossRef]

258. Chahine, G.; Pradalier, C. Survey of Monocular SLAM Algorithms in Natural Environments. In Proceedings of the 15th Conference on Computer and Robot Vision (CRV), Toronto, ON, Canada, 8–10 May 2018 2018; pp. 345–352. [CrossRef]

259. Konolige, K.; Agrawal, M.; Sola, J. Large-Scale Visual Odometry for Rough Terrain. In *Robotics Research*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 201–212.

260. Otsu, K.; Otsuki, M.; Ishigami, G.; Kubota, T. Terrain Adaptive Detector Selection for Visual Odometry in Natural Scenes. *Adv. Robot.* **2013**, *27*, 1465–1476. [CrossRef]

261. Daftry, S.; Dey, D.; Sandhawalia, H.; Zeng, S.; Bagnell, J.A.; Hebert, M. Semi-Dense Visual Odometry for Monocular Navigation in Cluttered Environment. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2015), Seattle, WA, USA, 25–30 May 2015.

262. Peretroukhin, V.; Clement, L.; Kelly, J. Reducing Drift in Visual Odometry by Inferring Sun Direction Using a Bayesian Convolutional Neural Network. In Proceedings of the Robotics and Automation (ICRA), 2017 IEEE International Conference, Marina Bay Sands, Singapore, 29 May–3 June 2017; pp. 2035–2042.

263. Giancola, S.; Schneider, J.; Wonka, P.; Ghanem, B.S. Integration of Absolute Orientation Measurements in the KinectFusion Reconstruction Pipeline. *arXiv* **2018**, arXiv:1802.03980.

264. Paudel, D.P.; Demonceaux, C.; Habed, A.; Vasseur, P. 2D–3D Synchronous/Asynchronous Camera Fusion for Visual Odometry. *Auton. Robots* **2018**, *43*, 21–35. [CrossRef]

265. Smolyanskiy, N.; Kamenev, A.; Smith, J.; Birchfield, S. Toward low-flying autonomous MAV trail navigation using deep neural networks for environmental awareness. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 4241–4247.

266. Mascaro, R.; Teixeira, L.; Hinzmann, T.; Siegwart, R.; Chli, M. GOMSF: Graph-Optimization Based Multi-Sensor Fusion for Robust UAV Pose Estimation. In Proceedings of the International Conference on Robotics and Automation (ICRA 2018) IEEE, Brisbane, Australia, 21–26 May, 2018.

267. Kocer, B.B.; Ho, B.; Zhu, X.; Zheng, P.; Farinha, A.; Xiao, F.; Stephens, B.; Wiesemüller, F.; Orr, L.; Kovac, M. Forest drones for environmental sensing and nature conservation. In Proceedings of the 2021 IEEE Aerial Robotic Systems Physically Interacting with the Environment (AIRPHARO), Biograd Na Moru, Croatia, 4–5 October 2021; pp. 1–8.

268. Griffith, S.; Pradalier, C. Survey Registration for Long-Term Natural Environment Monitoring. *J. Field Robot.* **2017**, *34*, 188–208. [CrossRef]

269. Naseer, T.; Burgard, W.; Stachniss, C. Robust Visual Localization Across Seasons. *IEEE Trans. Robot.* **2018**, *34*, 289–302. [CrossRef]

270. Engel, J.; Koltun, V.; Cremers, D. Direct Sparse Odometry. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 611–625. [CrossRef]

271. Cole, D.; Newman, P. Using Laser Range Data for 3D SLAM in Outdoor Environments. In Proceedings of the 2006 IEEE International Conference on Robotics and Automation (ICRA 2006), Orlando, FL, USA, 15–19 May 2006; pp. 1556–1563. [CrossRef]

272. Newman, P.; Cole, D.; Ho, K. Outdoor SLAM Using Visual Appearance and Laser Ranging. In Proceedings of the 2006 IEEE International Conference on Robotics and Automation (ICRA 2006), Orlando, FL, USA, 15–19 May 2006; pp. 1180–1187. [CrossRef]

273. Ramos, F.T.; Nieto, J.; Durrant-Whyte, H.F. Recognising and Modelling Landmarks to Close Loops in Outdoor SLAM. In Proceedings of the 2007 IEEE International Conference on Robotics and Automation (ICRA 2007), Rome, Italy, 10–14 April 2007; pp. 2036–2041. [CrossRef]

274. Angeli, A.; Filliat, D.; Doncieux, S.; Meyer, J.A. Fast and Incremental Method for Loop-Closure Detection Using Bags of Visual Words. *IEEE Trans. Robot.* **2008**, *24*, 1027–1037. [CrossRef]

275. Han, L.; Fang, L. MILD: Multi-index Hashing for Appearance Based Loop Closure Detection. In Proceedings of the IEEE International Conference on Multimedia and Expo (ICME 2017), Hong Kong, 10–14 July 2017; pp. 139–144. [CrossRef]

276. Thrun, S.; Montemerlo, M. The Graph SLAM Algorithm with Applications to Large-Scale Mapping of Urban Structures. *Int. J. Robot. Res.* **2006**, *25*, 403–429. [CrossRef]

277. Singh, S.; Kelly, A. Robot Planning in the Space of Feasible Actions: Two Examples. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 1996), Minneapolis, MN, USA, 22–28 April 1996; Volume 4, pp. 3309–3316. [CrossRef]

278. Pfaff, P.; Triebel, R.; Burgard, W. An Efficient Extension to Elevation Maps for Outdoor Terrain Mapping and Loop Closing. *Int. J. Robot. Res.* **2007**, *26*, 217–230. [CrossRef]

279. Zhang, J.; Singh, S. LOAM: Lidar odometry and mapping in real-time. In Proceedings of the Robotics: Science and Systems (RSS 2014), Berkeley, CA, USA, 12–16 July 2014; Volume 2, pp. 1–9.

280. Shan, T.; Englot, B. LeGO-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; pp. 4758–4765.

281. Kim, G.; Kim, A. Scan Context: Egocentric Spatial Descriptor for Place Recognition within 3D Point Cloud Map. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018. [CrossRef]

282. Ye, H.; Chen, Y.; Liu, M. Tightly Coupled 3D Lidar Inertial Odometry and Mapping. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2019), Montreal, QC, Canada, 20–24 May 2019.

283. Xu, W.; Zhang, F. FAST-LIO: A Fast, Robust LiDAR-Inertial Odometry Package by Tightly-Coupled Iterated Kalman Filter. *IEEE Robot. Autom. Lett.* April **2021**, *6*, 3317–3324 . [CrossRef]

284. Xu, W.; Cai, Y.; He, D.; Lin, J.; Zhang, F. FAST-LIO2: Fast Direct LiDAR-inertial Odometry. *arXiv* **2021**, arXiv:2107.06829.

285. Shan, T.; Englot, B.; Meyers, D.; Wang, W.; Ratti, C.; Daniela, R. LIO-SAM: Tightly-coupled Lidar Inertial Odometry via Smoothing and Mapping. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25–29 October 2020; pp. 5135–5142.

286. Reinke, A.; Palieri, M.; Morrell, B.; Chang, Y.; Ebadi, K.; Carlone, L.; Agha-Mohammadi, A.A. LOCUS 2.0: Robust and Computationally Efficient Lidar Odometry for Real-Time 3D Mapping. *IEEE Robot. Autom. Lett.* **2022**, *7*, 9043–9050. [CrossRef]

287. Lin, J.; Zhang, F. R 3 LIVE: A Robust, Real-time, RGB-colored, LiDAR-Inertial-Visual tightly-coupled state Estimation and mapping package. In Proceedings of the 2022 International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 23–27 May 2022; pp. 10672–10678.

288. Yin, H.; Li, S.; Tao, Y.; Guo, J.; Huang, B. Dynam-SLAM: An Accurate, Robust Stereo Visual-Inertial SLAM Method in Dynamic Environments. *IEEE Trans. Robot.* **2022**, *39*, 289–308. [CrossRef]

289. Wang, Y.; Ma, H. mVIL-Fusion: Monocular Visual-Inertial-LiDAR Simultaneous Localization and Mapping in Challenging Environments. *IEEE Robot. Autom. Lett.* **2022**, *8*, 504–511. [CrossRef]

290. Yuan, Z.; Wang, Q.; Cheng, K.; Hao, T.; Yang, X. SDV-LOAM: Semi-Direct Visual-LiDAR Odometry and Mapping. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 11203–11220. [CrossRef] [PubMed]

291. He, D.; Xu, W.; Chen, N.; Kong, F.; Yuan, C.; Zhang, F. Point-LIO: Robust High-Bandwidth Light Detection and Ranging Inertial Odometry. *Adv. Intell. Syst.* **2023**, *5*, 2200459. [CrossRef]

292. Vizzo, I.; Guadagnino, T.; Mersch, B.; Wiesmann, L.; Behley, J.; Stachniss, C. KISS-ICP: In Defense of Point-to-Point ICP Simple, Accurate, and Robust Registration If Done the Right Way. *IEEE Robot. Autom. Lett.* **2023**, *8*, 1029–1036. [CrossRef]

293. Karfakis, P.T.; Couceiro, M.S.; Portugal, D. NR5G-SAM: A SLAM Framework for Field Robot Applications Based on 5G New Radio. *Sensors* **2023**, *23*, 5354. [CrossRef]

294. Lu, Z.; Hu, Z.; Uchimura, K. SLAM Estimation in Dynamic Outdoor Environments: A Review. In *Intelligent Robotics and Applications*; Lecture Notes in Computer Science; Xie, M., Xiong, Y., Xiong, C., Liu, H., Hu, Z., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 255–267. [CrossRef]

295. Cristóvão, M.P.; Portugal, D.; Carvalho, A.E.; Ferreira, J.F. A LiDAR-Camera-Inertial-GNSS Apparatus for 3D Multimodal Dataset Collection in Woodland Scenarios. *Sensors* **2023**, *23*, 6676. [CrossRef]

296. Tian, Y.; Liu, K.; Ok, K.; Tran, L.; Allen, D.; Roy, N.; How, J.P. Search and rescue under the forest canopy using multiple UAVs. *Int. J. Robot. Res.* **2020**, *39*, 1201–1221. [CrossRef]

297. Agrawal, M.; Konolige, K. Real-Time Localization in Outdoor Environments Using Stereo Vision and Inexpensive GPS. In Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, 20–24 August 2006; Volume 3, pp. 1063–1068. [CrossRef]

298. Konolige, K.; Agrawal, M.; Bolles, R.C.; Cowan, C.; Fischler, M.; Gerkey, B. Outdoor Mapping and Navigation Using Stereo Vision. In Proceedings of the Experimental Robotics: The 10th International Symposium on Experimental Robotics (ISER 2006), Rio de Janeiro, Brazil, 6–12 July 2006; Khatib, O., Kumar, V., Rus, D., Eds.; Springer Tracts in Advanced Robotics; Springer: Berlin/Heidelberg, Germany, 2008; pp. 179–190. [CrossRef]

299. Schleicher, D.; Bergasa, L.M.; Ocana, M.; Barea, R.; Lopez, M.E. Real-Time Hierarchical Outdoor SLAM Based on Stereovision and GPS Fusion. *IEEE Trans. Intell. Transp. Syst.* **2009**, *10*, 440–452. [CrossRef]

300. Brand, C.; Schuster, M.J.; Hirschmüller, H.; Suppa, M. Stereo-Vision Based Obstacle Mapping for Indoor/Outdoor SLAM. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014), Chicago, IL, USA, 14–18 September 2014; pp. 1846–1853. [CrossRef]

301. Brand, C.; Schuster, M.J.; Hirschmüller, H.; Suppa, M. Submap Matching for Stereo-Vision Based Indoor/Outdoor SLAM. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015), Hamburg, Germany, 28 September–2 October, 2015; pp. 5670–5677. [CrossRef]

302. Moosmann, F.; Stiller, C. Velodyne SLAM. In Proceedings of the 2011 IEEE Intelligent Vehicles Symposium (IV), Baden-Baden, Germany, 5–9 June 2011; pp. 393–398. [CrossRef]

303. Abbas, S.M.; Muhammad, A. Outdoor RGB-D SLAM Performance in Slow Mine Detection. In Proceedings of the 7th German Conference on Robotics (ROBOTIK 2012), Munich, Germany, 21–22 May 2012; pp. 1–6.

304. Portugal, D.; Gouveia, B.D.; Marques, L. A Distributed and Multithreaded SLAM Architecture for Robotic Clusters and Wireless Sensor Networks. In *Cooperative Robots and Sensor Networks 2015*; Koubâa, A., Martínez-de Dios, J., Eds.; Studies in Computational Intelligence; Springer International Publishing: Cham, Switzerland, 2015; pp. 121–141. [CrossRef]

305. Sakai, T.; Koide, K.; Miura, J.; Oishi, S. Large-Scale 3D Outdoor Mapping and on-Line Localization Using 3D-2D Matching. In Proceedings of the 2017 IEEE/SICE International Symposium on System Integration (SII), Taipei, Taiwan, 11–14 December 2017; pp. 829–834. [CrossRef]

306. Lee, Y.J.; Song, J.B.; Choi, J.H. Performance Improvement of Iterative Closest Point-Based Outdoor SLAM by Rotation Invariant Descriptors of Salient Regions. *J. Intell. Robot. Syst.* **2013**, *71*, 349–360. [CrossRef]

307. Suzuki, T.; Kitamura, M.; Amano, Y.; Hashizume, T. 6-DOF Localization for a Mobile Robot Using Outdoor 3D Voxel Maps. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 18–22 October 2010; pp. 5737–5743. [CrossRef]

308. Droeschel, D.; Behnke, S. Efficient Continuous-Time SLAM for 3D Lidar-Based Online Mapping. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2018), Brisbane, Australia, 21–26 May 2018; pp. 5000–5007. [CrossRef]

309. Harrison, A.; Newman, P. High Quality 3D Laser Ranging under General Vehicle Motion. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2008), Pasadena, CA, USA, 19–23 May 2008; pp. 7–12. [CrossRef]

310. Levinson, J.; Thrun, S. Robust Vehicle Localization in Urban Environments Using Probabilistic Maps. In Proceedings of the 2010 IEEE International Conference on Robotics and Automation (ICRA 2010), Anchorage, AK, USA, 3–7 May 2010; pp. 4372–4378. [CrossRef]

311. Simanek, J.; Reinstein, M.; Kubelka, V. Evaluation of the EKF-Based Estimation Architectures for Data Fusion in Mobile Robots. *IEEE/ASME Trans. Mech.* **2015**, *20*, 985–990. [CrossRef]

312. Bernuy, F.; Ruiz Del Solar, J. Semantic Mapping of Large-Scale Outdoor Scenes for Autonomous Off-Road Driving. In Proceedings of the IEEE International Conference on Computer Vision Workshop (ICCVW 2015), Santiago, Chile, 7–13 December 2015; pp. 124–130. [CrossRef]

313. Boularias, A.; Duvallet, F.; Oh, J.; Stentz, A. Grounding Spatial Relations for Outdoor Robot Navigation. In Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA 2015), Seattle, WA, USA, 26–30 May 2015; pp. 1976–1982. [CrossRef]

314. Milford, M.J.; Wyeth, G.F. Mapping a Suburb with a Single Camera Using a Biologically Inspired SLAM System. *IEEE Trans. Robot.* **2008**, *24*, 1038–1053. [CrossRef]

315. Glover, A.J.; Maddern, W.P.; Milford, M.J.; Wyeth, G.F. FAB-MAP + RatSLAM: Appearance-based SLAM for Multiple Times of Day. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2010), Anchorage, AK, USA, 3–7 May 2010; pp. 3507–3512. [CrossRef]

316. Milford, M.; George, A. Featureless Visual Processing for SLAM in Changing Outdoor Environments. In *Field and Service Robotics: Results of the 8th International Conference [Springer Tracts in Advanced Robotics, Volume 92]*; Yoshida, K., Tadokoro, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 569–583. Available online: https://link.springer.com/chapter/10.1007/978-3-642-40686-7_38 (accessed on 1 July 2023).

317. Mur-Artal, R.; Montiel, J.M.M.; Tardós, J.D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Trans. Robot.* **2015**, *31*, 1147–1163. [CrossRef]

318. Mur-Artal, R.; Tardós, J.D. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. *IEEE Trans. Robot.* **2017**, *33*, 1255–1262. [CrossRef]

319. Schuster, M.J.; Brand, C.; Hirschmüller, H.; Suppa, M.; Beetz, M. Multi-Robot 6D Graph SLAM Connecting Decoupled Local Reference Filters. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015), Hamburg, Germany, 28 September–2 October 2015; pp. 5093–5100. [CrossRef]

320. Rossmann, J.; Schluse, M.; Schlette, C.; Buecken, A.; Krahwinkler, P.; Emde, M. Realization of a Highly Accurate Mobile Robot System for Multi Purpose Precision Forestry Applications. In Proceedings of the International Conference on Advanced Robotics (ICAR 2009), Munich, Germany, 22–26 June 2009; pp. 1–6.

321. Post, M.A.; Bianco, A.; Yan, X.T. Autonomous Navigation with ROS for a Mobile Robot in Agricultural Fields. In Proceedings of the 14th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2017), Madrid, Spain, 26–28 July 2017. Available online: https://strathprints.strath.ac.uk/61247/ (accessed on 1 July 2023).

322. Miettinen, M.; Ohman, M.; Visala, A.; Forsman, P. Simultaneous Localization and Mapping for Forest Harvesters. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2007), Rome, Italy, 10–14 April 2007; pp. 517–522. [CrossRef]

323. Auat Cheein, F.; Steiner, G.; Perez Paina, G.; Carelli, R. Optimized EIF-SLAM Algorithm for Precision Agriculture Mapping Based on Stems Detection. *Comput. Electron. Agric.* **2011**, *78*, 195–207. [CrossRef]

324. Duarte, M.; dos Santos, F.N.; Sousa, A.; Morais, R. Agricultural Wireless Sensor Mapping for Robot Localization. In Proceedings of the Robot 2015: Second Iberian Robotics Conference, Advances in Intelligent Systems and Computing, Lisbon, Portugal, 19–21 November 2015; Reis, L.P., Moreira, A.P., Lima, P.U., Montano, L., Muñoz-Martinez, V., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 359–370. [CrossRef]

325. Yang, N.; Wang, R.; Gao, X.; Cremers, D. Challenges in Monocular Visual Odometry: Photometric Calibration, Motion Bias, and Rolling Shutter Effect. *IEEE Robot. Autom. Lett.* **2018**, *3*, 2878–2885. [CrossRef]

326. Aqel, M.O.A.; Marhaban, M.H.; Saripan, M.I.; Ismail, N.B. Review of Visual Odometry: Types, Approaches, Challenges, and Applications. *SpringerPlus* **2016**, *5*, 1897. [CrossRef] [PubMed]

327. Cadena, C.; Carlone, L.; Carrillo, H.; Latif, Y.; Scaramuzza, D.; Neira, J.; Reid, I.; Leonard, J.J. Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. *IEEE Trans. Robot.* **2016**, *32*, 1309–1332. [CrossRef]

328. Arkin, R.C.; Balch, T. Cooperative Multiagent Robotic Systems. In *Artificial Intelligence and Mobile Robots: Case Studies of Successful Robot Systems*; MIT Press: Cambridge, MA, USA, 1998; pp. 277–296.

329. Zhang, J.; Singh, S. Laser-visual-inertial Odometry and Mapping with High Robustness and Low Drift. *J. Field Robot.* **2018**, *35*, 1242—1264. [CrossRef]

330. Hawes, N.; Burbridge, C.; Jovan, F.; Kunze, L.; Lacerda, B.; Mudrova, L.; Young, J.; Wyatt, J.; Hebesberger, D.; Kortner, T.; et al. The STRANDS Project: Long-Term Autonomy in Everyday Environments. *IEEE Robot. Autom. Mag.* **2017**, *24*, 146–156. [CrossRef]

331. Rocha, R.P.; Portugal, D.; Couceiro, M.; Araújo, F.; Menezes, P.; Lobo, J. The CHOPIN project: Cooperation between human and rObotic teams in catastrophic incidents. In Proceedings of the 2013 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR 2013), Linkoping, Sweden, 21–26 October 2013; pp. 1–4. [CrossRef]

332. Kruijff-Korbayová, I.; Colas, F.; Gianni, M.; Pirri, F.; de Greeff, J.; Hindriks, K.; Neerincx, M.; Ögren, P.; Svoboda, T.; Worst, R. TRADR Project: Long-Term Human-Robot Teaming for Robot Assisted Disaster Response. *Künstliche Intell.* **2015**, *29*, 193–201. [CrossRef]

333. Singh, A.; Krause, A.R.; Guestrin, C.; Kaiser, W.J.; Batalin, M.A. Efficient Planning of Informative Paths for Multiple Robots. In Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007), Hyderabad, India, 6–12 January 2007; Volume 7, pp. 2204–2211. Available online: https://openreview.net/forum?id=ryVLY4G_ZS (accessed on 1 July 2023).

334. La, H.M.; Sheng, W.; Chen, J. Cooperative and Active Sensing in Mobile Sensor Networks for Scalar Field Mapping. *IEEE Trans. Syst. Man Cybern. Syst.* **2015**, *45*, 1–12. [CrossRef]

335. Ma, K.C.; Liu, L.; Sukhatme, G.S. An Information-Driven and Disturbance-Aware Planning Method for Long-Term Ocean Monitoring. In Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Republic of Korea, 9–14 October 2016; pp. 2102–2108. [CrossRef]

336. Manjanna, S.; Dudek, G. Data-Driven Selective Sampling for Marine Vehicles Using Multi-Scale Paths. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 6111–6117. [CrossRef]

337. Euler, J.; von Stryk, O. Optimized Vehicle-Specific Trajectories for Cooperative Process Estimation by Sensor-Equipped UAVs. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2017), Marina Bay Sands, Singapore, 29 May–3 June 2017; pp. 3397–3403. [CrossRef]

338. Merino, L.; Caballero, F.; Martínez-de Dios, J.R.; Maza, I.; Ollero, A. An unmanned aircraft system for automatic forest fire monitoring and measurement. *J. Intell. Robot. Syst.* **2012**, *65*, 533–548. [CrossRef]

339. Ahmad, A.; Walter, V.; Petráček, P.; Petrlík, M.; Báča, T.; Žaitlík, D.; Saska, M. Autonomous aerial swarming in gnss-denied environments with high obstacle density. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2021), Xi'an, China, 30 May–5 June 2021; pp. 570–576.

340. Couceiro, M.S.; Portugal, D. Swarming in forestry environments: Collective exploration and network deployment. *Swarm Intell. Princ. Curr. Algoritm. Methods* **2018**, *119*, 323–344.

341. Tarapore, D.; Groß, R.; Zauner, K.P. Sparse robot swarms: Moving swarms to real-world applications. *Front. Robot. AI* **2020**, *7*, 83. [CrossRef]

342. Ju, C.; Kim, J.; Seol, J.; Son, H.I. A review on multirobot systems in agriculture. *Comput. Electron. Agric.* **2022**, *202*, 107336. [CrossRef]

343. Martins, G.S.; Ferreira, J.F.; Portugal, D.; Couceiro, M.S. MoDSeM: Modular Framework for Distributed Semantic Mapping. In Proceedings of the UK-RAS Robotics and Autonomous Systems Conference: "Embedded Intelligence: Enabling and Supporting RAS Technologies", Loughborough University, Loughborough, UK, 24 January 2019; pp. 12–15. [CrossRef]

344. Martins, G.S.; Ferreira, J.F.; Portugal, D.; Couceiro, M.S. MoDSeM: Towards Semantic Mapping with Distributed Robots. In Proceedings of the 20th Towards Autonomous Robotic Systems Conference, London, UK, 3–5 July 2019, pp. 131–142. [CrossRef]

345. Rocha, R.; Dias, J.; Carvalho, A. Cooperative Multi-Robot Systems: A Study of Vision-Based 3-D Mapping Using Information Theory. *Robot. Auton. Syst.* **2005**, *53*, 282–311. [CrossRef]

346. Das, G.P.; McGinnity, T.M.; Coleman, S.A.; Behera, L. A Fast Distributed Auction and Consensus Process Using Parallel Task Allocation and Execution. In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2011), San Francisco, CA, USA, 25–30 September 2011; pp. 4716–4721. [CrossRef]

347. Calleja-Huerta, A.; Lamandé, M.; Green, O.; Munkholm, L.J. Impacts of Load and Repeated Wheeling from a Lightweight Autonomous Field Robot on the Physical Properties of a Loamy Sand Soil. *Soil Tillage Res.* **2023**, *233*, 105791. [CrossRef]

348. Batey, T. Soil compaction and soil management—A review. *Soil Use Manag.* **2009**, *25*, 335–345. [CrossRef]

349. Niu, C.; Zauner, K.P.; Tarapore, D. End-to-End Learning for Visual Navigation of Forest Environments. *Forests* **2023**, *14*, 268. [CrossRef]

350. da Silva, D.Q.; dos Santos, F.N.; Sousa, A.J.; Filipe, V.; Boaventura-Cunha, J. Unimodal and Multimodal Perception for Forest Management: Review and Dataset. *Computation* **2021**, *9*, 127. [CrossRef]

351. Jensen, K.; Larsen, M.; Nielsen, S.; Larsen, L.; Olsen, K.; Jørgensen, R. Towards an Open Software Platform for Field Robots in Precision Agriculture. *Robotics* **2014**, *3*, 207–234. [CrossRef]

352. Portugal, D.; Ferreira, J.F.; Couceiro, M.S. Requirements specification and integration architecture for perception in a cooperative team of forestry robots. In Proceedings of the Annual Conference towards Autonomous Robotic Systems, Online, 16 September 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 329–344.