remote sensing

Article

Ethical Framework to Assess and Quantify the Trustworthiness of Artificial Intelligence Techniques: Application Case in Remote Sensing

Marina Paolanti [1], Simona Tiribelli [1], Benedetta Giovanola [1], Adriano Mancini [2], Emanuele Frontoni [1] and Roberto Pierdicca [3,*]

[1] Department of Political Sciences, Communication and International Relations, University of Macerata, 62100 Macerata, Italy; marina.paolanti@unimc.it (M.P.); simona.tiribelli@unimc.it (S.T.); benedetta.giovanola@unimc.it (B.G.); emanuele.frontoni@unimc.it (E.F.)

[2] Department of Information Engineering (DII), Università Politecnica delle Marche, 60131 Ancona, Italy; a.mancini@staff.univpm.it

[3] Department of Civil Engineering Building and Architecture (DICEA), Università Politecnica delle Marche, 60131 Ancona, Italy

* Correspondence: r.pierdicca@staff.univpm.it

Abstract: In the rapidly evolving field of remote sensing, Deep Learning (DL) techniques have become pivotal in interpreting and processing complex datasets. However, the increasing reliance on these algorithms necessitates a robust ethical framework to evaluate their trustworthiness. This paper introduces a comprehensive ethical framework designed to assess and quantify the trustworthiness of DL techniques in the context of remote sensing. We first define trustworthiness in DL as a multidimensional construct encompassing accuracy, reliability, transparency and explainability, fairness, and accountability. Our framework then operationalizes these dimensions through a set of quantifiable metrics, allowing for the systematic evaluation of DL models. To illustrate the applicability of our framework, we selected an existing case study in remote sensing, wherein we apply our ethical assessment to a DL model used for classification. Our results demonstrate the model's performance across different trustworthiness metrics, highlighting areas for ethical improvement. This paper not only contributes a novel framework for ethical analysis in the field of DL, but also provides a practical tool for developers and practitioners in remote sensing to ensure the responsible deployment of DL technologies. Through a dual approach that combines top-down international standards with bottom-up, context-specific considerations, our framework serves as a practical tool for ensuring responsible AI applications in remote sensing. Its application through a case study highlights its potential to influence policy-making and guide ethical AI development in this domain.

Keywords: remote sensing; AI ethics; guidelines; ethical framework; GeoAI

Citation: Paolanti, M.; Tiribelli, S.; Giovanola, B.; Mancini, A.; Frontoni, E.; Pierdicca, R. Ethical Framework to Assess and Quantify the Trustworthiness of Artificial Intelligence Techniques: Application Case in Remote Sensing. *Remote Sens.* **2024**, *16*, 4529. https://doi.org/10.3390/rs16234529

Academic Editors: Zhe Wang, Chao Fan, Sanaz Salati, Marshall (Xiaogang) Ma, Xiang Que and Hui Wang

Received: 9 October 2024
Revised: 27 November 2024
Accepted: 2 December 2024
Published: 3 December 2024

1. Introduction

Geospatial data are a foundation of modern society, and play a vital role in a wide range of essential activities, spanning both social and political spheres [1]. Their applications range from the production of detailed topographic maps of urban areas, the updating of real-time traffic information, and the digital preservation of cultural and archaeological sites, to the classification of urban and rural landscapes for environmental risk monitoring [2]. In this regard, the Committee on Earth Observation Satellite [3] noted that using Artificial Intelligence (AI) models in geomatics and geography, resulting in the definition of a new concept named Geospatial Artificial Intelligence (GeoAI) [4], it is fundamental for states to measure and achieve global sustainable development goals (SDGs) [5]. In the discipline of geomatics, remote sensing emerges as a key case study that illustrates the broadening scope of the field [6]. Remote sensing has given geomatics a powerful tool to address challenges ranging from the impact of climate change to sustainable land management and urban planning. The integration of DL techniques in remote sensing has further elevated its capabilities, enabling more sophisticated analysis and interpretation of satellite imagery [7]. These advancements have significantly enhanced tasks such as land cover classification, change detection, and environmental monitoring, which are crucial for sustainable development and disaster management [8–10].

Despite these technological advances, the field of remote sensing faces a critical gap: the lack of an ethical framework for the application of DL techniques, given the processes automation leading to a strong reduction in human intervention. This gap raises significant concerns, particularly given the socio-political implications of decisions based on remote sensing data. Without ethical guidelines, there is a risk of misinterpretation, misuse, or even manipulation of data, which could lead to adverse consequences. With a few exceptions [11], data ethics and the ethical implications in this area have not been explored so far [12–14]. This lack of an ethical framework for the evaluation of DL techniques applied to remote sensing makes it difficult for non-experts to understand which technique should be used. For example, at a policy level, *what could be the best technical DL for the management and prevention of environmental disasters? Or which one would be more appropriate for the classification of urban areas?* In addition, there has been a huge increase in the availability of data from Earth Observation (EO) programs. All around the world, these programs are experiencing exponential growth (such as China's GAOFEN, South America's Perusat, USA's NASA, and Europe's Copernicus, to name a few), offering vast opportunities while also presenting significant risks that need to be addressed [15]. From a mere legal point of view, strict rules have been defined, but the same cannot be said from an ethical point of view. In the field of AI, the latest achievements demonstrate that multi-task learning can be adopted for building segmentation [16], settlement mapping in urban and rural areas [17,18], and multitemporal flooding detection and forecasting [19]. After the AI-based task is performed, a series of actions can be undertaken by decision-makers for each of these subdomains. Additional to the standard metrics developed to evaluate the performances of algorithms, there is a surge to determine the ethical implications that such decision might bring.

Our paper addresses this urgent need by introducing a comprehensive ethical framework tailored to assess and quantify the trustworthiness of DL techniques in remote sensing. This framework aims to ensure that the application of DL in remote sensing is not only technologically robust, but also ethically responsible. It takes into account the unique characteristics and challenges of remote sensing data, such as spatial resolution, temporal frequency, and privacy issues. It ensures that the ethical use of DL techniques is consistent with the complexities of the domain. The motivation for developing such a framework arises from broader concerns about the ethics of AI in general. Several international bodies, including the United Nations, UNESCO, the European Union, and the World Health Organization, have established guidelines for the development of trustworthy AI. However, there remains a gap in how these high-level ethical standards are applied and interpreted in specific domains, particularly in remote sensing. Our framework, therefore, aims to bridge this gap by aligning global ethical guidelines with the unique challenges and applications of remote sensing. Our aim is not to highlight issues unique to remote sensing. Instead, we aim to unpack today's benchmark AI ethics guidelines in a manner deeply relevant and useful specifically for the application of AI in remote sensing, rather than for its general use. Given the lack of attention to this crucial domain, without claiming to be exhaustive, the framework aims to pave the way for the development of the first practical tool in this field. This tool is designed to complement, rather than replace, qualitative (human) assessment. Drawing on benchmark ethical guidelines and frameworks, we propose: (a) a structured evaluation framework for assessing fairness in remote sensing AI systems through bias detection, prevention, and mitigation strategies; (b) a framework for assessing transparency and explainability by dimension and stakeholder level; and (c) quantifiable metrics for assessing ethical compliance across these dimensions. These efforts aim to support the development and implementation of trustworthy AI systems tailored to the unique challenges and opportunities of remote sensing.

The main contributions of the paper can be summarized as follows: (i) *Introduction of an ethical framework for ai in remote sensing*: it addresses the gap in ethical considerations in the field of remote sensing, which is critical given the increasing reliance on AI for environmental monitoring and disaster management. (ii) *Quantified metrics for ethical compliance*: development of quantified metrics that allow for the objective assessment of AI algorithms' compliance with ethical standards. These metrics provide tangible and measurable criteria, facilitating a structured and objective approach in evaluating the ethical integrity of DL techniques in remote sensing. (iii) *Alignment with global ethical standards*: As stated before, the paper aligns its ethical framework with the guidelines set by leading international bodies. This alignment ensures that the framework is grounded in globally recognized principles of ethical AI. (iv) *Context-sensitive, dual approach*: It adopts a dual approach: a top-down perspective adhering to international ethical standards, coupled with a bottom-up, context-sensitive approach. This approach takes into account the specific challenges and complexities of remote sensing applications, making the framework both globally relevant and locally applicable. (v) *Practical tool for ethical AI application*: by introducing this framework and its associated metrics, the paper provides a practical tool for assessing and ensuring the ethical application of AI in remote sensing, where AI-driven decisions have far-reaching implications for environmental monitoring and policy-making.

The paper is structured as follows: the 'Related Works' provides the background, and also identifies the gaps our paper seeks to address; 'Methodology' presents the core of our paper: the development of the ethical framework. This section details the creation of quantified metrics for assessing AI's ethical compliance in remote sensing and how these metrics align with global ethical standards. 'Results' and 'Discussions' illustrates the practical application of our framework in remote sensing scenarios, showcasing its relevance and utility in real-world situations. Finally, 'Conclusions and Future Works' concludes with a summary of our findings, contributions, and suggestions for future research directions.

2. Related Works

In this section, we provide a thorough examination of the current status of remote sensing and AI, with a particular focus on ethical considerations. This section is divided into two distinct but linked parts: "Technical State of the Art" and "Ethical State of the Art".

2.1. Technical State of the Art

DL, a subset of machine learning, has become an important tool in remote sensing [20,21]. Over time, Deep Neural Networks (DNNs) effectively transform data into forms that are well suited for specific tasks, including image pre-processing, object detection, and pixel-level classification. Many proposed applications of DL methods with satellite data span fields such as astronomy, planetary science, and Earth observation [22,23]. Due to their layered learning approach, DL models are able to accurately mimic complex non-linear interactions between environmental factors. This ability is critical for remote sensing tasks such as retrieval, fusion, and downscaling in order to identify possible relationships between different environmental elements. In addition, DL is particularly effective in the extraction of features at multiple scales and levels from remotely sensed imagery, and in the integration of these features from the most basic to the most advanced levels. This contributes significantly to improved performance in image processing and classification tasks. As a result, DL models have demonstrated superior performance to traditional models. This has led to significant advances in the monitoring of the Earth's environment using remote sensing data [24]. Applications of DL to remote sensing images differ from those to common images. Remotely sensed images often have more complex and varied patterns, as well as rich spatial, temporal, and spectral details, requiring more sophisticated processing techniques. Due to its robust feature representation capability, DL has been adopted in the field of environmental remote sensing and has been applied in several areas. These include land cover mapping, environmental parameter retrieval, data fusion, and downscaling, as well as information construction and prediction [25]. DL is a promising approach to environmental parameter retrieval. First, DL can either replicate or streamline the physical models used for environmental parameter retrieval. Physical models often require complex computations, and DL, with its significant simulation capabilities, can be used to perform partial or complete forward simulations of these models. This simplifies the process of environmental parameter retrieval. Furthermore, DL is able to establish statistical correlations between remote sensing observations and in situ environmental parameters due to its ability to approximate complex relationships [26,27]. DL has been applied to land cover mapping, providing optimal results due to its excellence in extracting features across multiple scales and levels. Land cover mapping from remote sensing imagery is fundamentally dependent on image classification. Traditional classification methods sort images based on different spatial units such as pixels, moving windows, objects, and scenes. However, accurately identifying complex land structures or patterns with a limited set of rules can be challenging, as traditional methods typically use only basic spectral and spatial features for classification [28,29]. DL is capable of identifying aggregate features in remote sensing observations and discovering potential links between different observations through its multi-layered learning approach. As a result, DL can thoroughly encapsulate the complex relationships required for data fusion and downscaling. Furthermore, DL constructs these relationships by extracting abstract features from data samples that are less affected by observational properties such as sensor type and spatial scale. This ability allows DL models to establish more stable and reliable relationships [30]. The use of generative adversarial networks (GANs) in remote sensing applications is attracting increasing attention [31]. These networks are highly adapted to manage complex, high-dimensional data and can perform effectively even when there is scarce or no labeled training data available [32].

2.2. Ethical State of the Art

As mentioned in the introduction, although GeoAI has a notable impact on decision-making processes, to date the ethical implications resulting from the use of DL techniques in geomatics have been poorly investigated. In this sense, the work of Gevaert et al. [13] is important, as it highlights how in the context of Earth Observation (EO) the concept of explainability is considered differently at a regulatory and technical level. First, the regulatory framework considers the entire AI model design and use, from data collection to the establishment of the AI algorithm itself and how that algorithm is used. This approach is broader than the simple consideration of algorithmic inputs and outputs that occur at the technical level of EO. Secondly, the audiences addressed by the two levels are different. While the constituent steps of the models must be understandable for technicians to be applicable in different domains, it is necessary for society to understand the conformity of the outcomes of the models with socio-political purposes. This difference in technical and regulatory understanding necessarily refers to different concepts of explainability [33]. Despite this, the literature relating to EO is dedicated exclusively to the study of the concept of explainability from a technical point of view [13]. This paper, by proposing an ethical evaluation of DL models on remote sensing useful for the decision-makers, offers an analysis of the concept of explainability focused on the regulatory level. To build an ethical framework useful for evaluating DL techniques for the use of remote sensing data, this paper is based on three main sources, all three of which have a pivotal role at a global level. The first is the Assessment List for Trustworthy Artificial Intelligence (ALTAI), drawn up by the European Commission High-Level Group on Artificial Intelligence [34]. Specifically, ALTAI identifies seven requirements to achieve trustworthy AI: (1) human agency and oversight; (2) technical robustness and safety; (3) privacy and data governance; (4) transparency and explainability; (5) diversity, non-discrimination, and fairness; (6) societal and environmental well-being; and (7) accountability. The role of these ethical requirements is regulative and not legally binding. It has only the function of providing guidance for the development of trustworthy technology. The second document is UNESCO's Recommendation on The Ethics of Artificial Intelligence [35]. These recommendations provide systematic regulatory guidance, based on a global and multicultural approach aimed at guiding companies to responsibly address the impact of AI on humans and society. To this end, these recommendations highlight the importance of addressing digital and knowledge gaps between countries throughout the lifecycle of AI models. In fact, they define precisely those values that should guide the responsible development and use of DL systems. As with the EU guidelines, UNESCO also identifies 'Transparency and explainability' as one of the key principles for trustworthy AI. According to such principles people have the right to know when a political decision is supported by an AI system and to understand its value. Transparency is necessary for public understanding of the possible social and political role of DL-based system outcomes and, therefore, to guarantee equity and inclusiveness. Explainability refers to understanding the behavior of various algorithm mechanisms and how they contribute to the transformation of inputs into the final outcome. The third document used in this work is the United Nations Integrated Geospatial Information Network (UN-IGIF), published in 2020 by the Committee of Experts on Global Geospatial Information Management of the United Nations [36]. This document is specific for the management of geospatial data aimed at meeting the sustainable development goals (SDGs) outlined in 2015 by the United Nations in the 2030 Agenda [37]. UN-IGIF was developed to outline a guide for understanding and exploiting the potential of geospatial data. This document presents three sections: the overarching strategic framework, the implementation guide, and the country-level action plan. Although this document lacks a specific ethical dimension, it is based on seven underpinning principles that define the best use of geospatial data: (1) strategic enabling; (2) transparent and accountable; (3) reliable, accessible, and easily used; (4) collaboration and cooperation; (5) integrative solutions; (6) sustainable and valued; (7) leadership and commitment. As shown by Calzati and van Loenen [5], ALTAI requirements and UN-IGIF principles have many common points. UN-IGIF principles 1–6 recall ALTAI requirements 3–6, as they refer to the public role of governments and societies in requiring that the development of AI models is respectful of sustainability criteria. Similarly, principles 2 and 3 of the UN-IGIF document appeal to the ethical principles of transparency, accountability, trustworthiness, fairness, and accessibility, which are in line with almost all ALTAI requirements.

3. Methodology

Embedding AI ethics principles into the development and deployment of AI systems for remote sensing requires (i) *explaining* what benchmark AI ethics principles entail, and (ii) *unpacking* them into categories and criteria to enable their enforcement and understanding by engineers, policy makers, and other stakeholders involved in evaluating or approving such systems. To achieve this, we draw on the "Ethics Guidelines for Trustworthy AI" developed by the High-Level Expert Group on AI established by the European Commission (EC) in 2019, and on the conceptual tool (*Assessment List for Trustworthy AI (ALTAI)*) proposed to support AI developers and users in operationalizing these principles in specific domains, such as remote sensing. The ethical requirements proposed by the EC for trustworthy AI aim to ensure respect for five core ethical principles that are widely recognized in AI ethics research [38]:

1. Benevolence;
2. Nonmaleficence;
3. Autonomy;
4. Justice and Fairness;
5. Explicability.

These principles define the ethical foundations for the responsible use of AI in remote sensing, ensuring that systems are designed and deployed to (1) benefit society, (2) minimize harm, (3) respect human autonomy, (4) promote justice and fairness by mitigating bias, and (5) ensure transparency through explicability and interpretability. In this section, we present a detailed approach to developing and applying an ethical framework that assesses the trustworthiness of DL techniques in remote sensing. The comprehensive approach of our study begins with bias identification, recognizing that addressing bias is a fundamental step in ensuring ethical AI systems. This process involves the detection and mitigation of potential biases inherent in remote sensing data and models, such as geographic, demographic, temporal, and label biases, which can significantly impact the fairness and reliability of AI outcomes. Following the bias identification, the study reviews guidelines from key regulatory frameworks, including the Assessment List for Trustworthy Artificial Intelligence (ALTAI), the UNESCO Recommendation on the Ethics of Artificial Intelligence, and the United Nations Integrated Geospatial Information Network (UN-IGIF). Following this review, the study selects a specific remote sensing method for evaluation, applies an ethical evaluation framework with defined metrics derived from these guidelines, and concludes with a detailed analysis and quantification of the ethical compliance scores. This study is methodically divided into three integral components: "Ethical Framework", "Evaluation Questions", and "Quantification and Metrics", each of which plays a pivotal role in our research. Figure 1 schematically depicts the ethical evaluation process for remote sensing methodologies. Finally, we present the Ethical Framework for Explainability in Remote Sensing, which assesses the intelligibility and interpretability of AI systems. Inspired by the work of Tiribelli et al. in the retail sector [39], this framework adapts their methods to the unique challenges of remote sensing. It focuses on multi-dimensional explainability, including computational, justificatory and cautionary dimensions, and proposes practical tools such as visualizations, confidence scores and plain-language summaries to make AI outputs more accessible and actionable for different stakeholders.

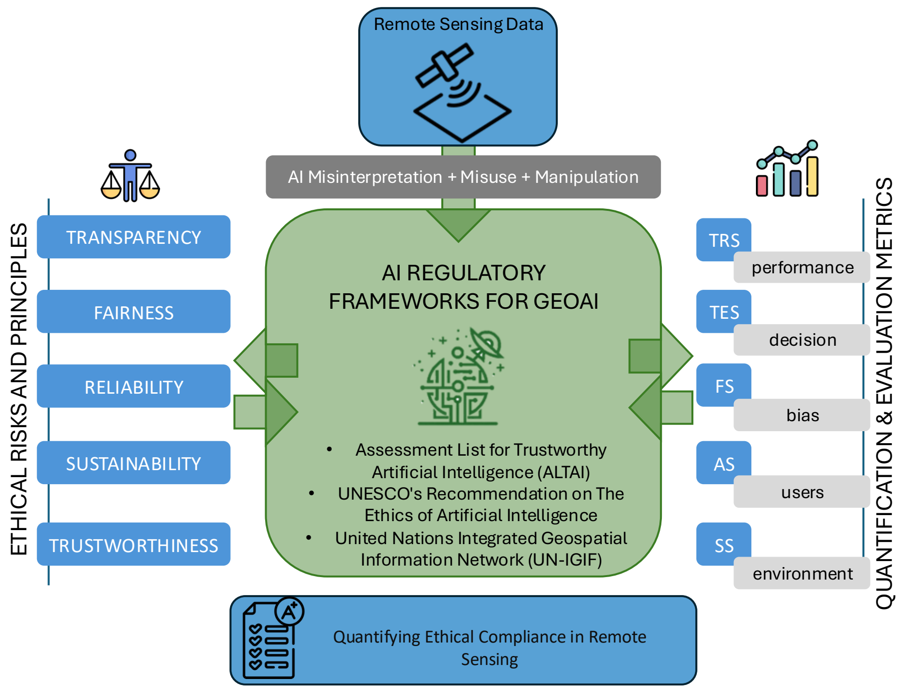

Figure 1. Ethical evaluation process for remote sensing methodologies. The approach starts from the selection of the remote sensing methodology, applying the ethical evaluation framework with defined metrics (please refer to Section 3.5), and concluding with the analysis and quantification of ethical compliance scores.

3.1. Ethical Framework for Bias Detection in Remote Sensing

Addressing ethical challenges in remote sensing requires a focused examination of bias in AI systems, particularly in detecting and mitigating unintentional biases that may arise during data processing or model training. These biases are particularly critical in remote sensing applications where AI-driven systems influence decisions related to land use, environmental monitoring, and disaster response. Left unchecked, such biases could lead to inequitable resource allocation, misclassification of vulnerable regions, or environmental damage. Our framework implements the principle of fairness, as outlined in the EU High Level Expert Group's Ethics Guidelines for Trustworthy AI, by addressing biases specific to remote sensing data and workflows [40–44]. This includes developing structured methods to identify and address bias, and to ensure that AI systems produce equitable outcomes across different geographies, populations and ecosystems.

Table 1 summarizes the main types of bias commonly found in remote sensing data, their descriptions and recommended mitigation strategies. For example, population target bias occurs when datasets do not adequately cover the intended target populations, leading to inaccurate predictions in rural or under-represented areas. Similarly, temporal bias—common in remote sensing datasets—occurs when models rely on outdated imagery, perpetuating historical inequalities. Addressing such biases in the early stages of system design will ensure more fair and accurate outcomes, fostering trust in AI applications for remote sensing.

Table 1. Bias types in data input for enacting fairness in remote sensing.

Type of Bias	Description in Remote Sensing	Recommended Action (RA)
Population Target Bias	Occurs when datasets fail to represent the actual target regions or populations, such as overrepresenting urban areas while underrepresenting rural or remote areas.	Clearly define target regions and ensure datasets include diverse and representative samples of all geographical and demographic groups.
Missing Data Bias	Results from gaps in satellite imagery or missing metadata, leading to incomplete analysis and reduced accuracy.	Conduct quality checks and use imputation techniques or data fusion to address spatial and temporal gaps.
Minority Bias	Arises from underrepresentation of minority geographical features (e.g., wetlands, indigenous lands), reducing accuracy in classification or monitoring.	Prioritize the inclusion of underrepresented features during data collection and augment with synthetic data if necessary.
Informativeness Bias	Occurs when features, such as low-resolution imagery, fail to provide adequate detail for accurate analysis (e.g., distinguishing vegetation types).	Enhance dataset informativeness with high-resolution imagery or multispectral/hyperspectral data.
Temporal Bias	Results from using outdated satellite imagery, failing to capture current land-use or environmental conditions.	Regularly update datasets and conduct temporal consistency checks to align analyses with recent conditions.

Table 1. *Cont.*

Type of Bias	Description in Remote Sensing	Recommended Action (RA)
Socio-Behavioral Context Bias	Stems from variations in socio-economic and behavioral patterns across regions, such as informal settlements versus urban planning.	Incorporate socio-economic data and design models to handle regional differences effectively.
Self-Selection Bias	Arises when ground truth data are collected predominantly from accessible regions, excluding remote or hard-to-reach areas.	Ensure data collection efforts cover both accessible and remote regions, leveraging local partnerships if needed.
Historical Bias	Occurs when training datasets reflect outdated societal norms or land-use practices, perpetuating inaccuracies in predictions.	Audit datasets to identify and correct correlations reflecting historical inequalities or outdated land-use patterns.
Label Bias	Results from inconsistent or inaccurate labeling during annotation, leading to misclassification in remote sensing imagery.	Ensure consistent and accurate labeling by involving diverse and skilled annotators and conducting consensus validation.
Omitted Variable Bias	Arises when key contextual variables, such as elevation or soil type, are excluded, reducing model accuracy.	Include relevant variables by consulting domain experts and conducting comprehensive feature engineering.
Aggregation Bias	Results from assumptions based on aggregated data, which may obscure critical variations within the dataset.	Use granular data wherever possible and design models to account for intra-class variations.

By using these structured guidelines, developers and decision-makers can systematically identify specific biases relevant to remote sensing projects and implement measures to mitigate their impact. The tables serve as a conceptual guide for building more equitable AI systems that ensure inclusivity and accountability.

3.2. Ethical Principles for Trustworthy AI in Remote Sensing

The ethical framework proposed in this paper extends the work of Calzati and Van Loenen [5], who introduced the Geoinformation Ecosystem Ethics Assessment List. This is a qualitative approach to the evaluation of geospatial information within the broader framework of the International Geospatial Information Framework (IGIF). However, while [5] provides a refined general framework, it falls short in providing a quantitative measure specifically for DL techniques. Our paper addresses this gap by focusing on the ethical analysis of DL systems, specifically for remote sensing. DL in remote sensing requires a framework that goes beyond qualitative analysis to provide quantitative measures, as it presents unique ethical challenges. Therefore, our approach focuses on providing a quantitative ethical assessment. The aim is to transform this ethical assessment into a practical operationalization tool, potentially automated or application-based, making ethical assessment accessible and actionable.

In our ethical framework, we identify and address four primary ethical risks that are associated with DL systems for remote sensing related to: (1) *transparency and explainability*, (2) *fairness*, (3) *reliability*, and (4) *sustainability*. These risks summarize the key areas where DL applications in remote sensing could possibly violate benchmark ethical principles and values acknowledged in the core frameworks and guidelines at the global level [34].

Transparency and explainability focus on ensuring that the processes and results of AI models are understandable and interpretable, especially for non-technical stakeholders. This is critical in remote sensing, where decision makers often rely on AI-driven insights to make high-stakes decisions. A lack of transparency can lead to mistrust or misuse of the system's results. For example, if a DL model classifies urban and rural areas without a clear rationale, policymakers may misinterpret the results, leading to inappropriate policy decisions. Our framework emphasizes the need for models to provide interpretable results, documentation, and detailed justifications to enable accountability and trust.

Fairness refers to the balanced treatment of all societal groups, ensuring that DL systems do not propagate or exacerbate existing biases. In remote sensing, fairness is particularly relevant when analyzing data from underrepresented or disadvantaged regions. For example, DL models trained on datasets with limited representation of rural or low-income areas may produce inaccurate or inequitable results, affecting resource allocation or disaster response strategies. Our framework highlights mechanisms to identify, mitigate and monitor biases, thereby promoting outcomes that are equitable and inclusive.

Reliability addresses the technical reliability of DL models in remote sensing, with a focus on data quality, accuracy, and completeness. Remote sensing data, often obtained from satellites or drones, can suffer from gaps, noise, or inconsistencies that, if not properly handled, can compromise model performance. Reliable DL systems need to be robust to missing or erroneous data and provide consistent results across different scenarios. Our framework evaluates these aspects to ensure that AI systems deliver reliable and reproducible results.

Sustainability emphasizes the alignment of DL technologies with environmentally and socially responsible practices. Remote sensing applications, such as land use monitoring or environmental risk detection, have significant implications for sustainability. For example, the high computational demands of DL models can contribute to carbon emissions, while their outputs influence decisions that affect ecosystems and communities. Our framework advocates minimizing the environmental footprint and integrating DL systems into broader sustainability goals to ensure long-term societal and environmental benefits.

In addition to the four ethical risks, our work explores the broader concept of **trustworthiness**, which we define in terms of two dimensions:

- **Technical robustness**: The ability of the system to function reliably, accurately and consistently under varying conditions. This includes error handling, adaptability to diverse data sets, and resilience to adversarial scenarios.
- **Social robustness**: The ability of the system to operate ethically in different societal contexts, taking into account cultural, economic and political variations. This includes assessing whether the system integrates with societal norms, respects privacy, and complies with local governance frameworks.

Detailed ethical considerations for each risk are outlined in Table 2.

Table 2. Ethical evaluation framework for deep learning techniques in remote sensing.

	Fairness	**Reliability**	**Sustainability**	**Transparency and Explainability**	**Technical and Social Robustness**
Ethical Principle	DL techniques in remote sensing must ensure equitable treatment for all societal groups, particularly the disadvantaged.	The quality and completeness of datasets are critical for the reliability of AI-driven remote sensing applications.	DL techniques in remote sensing should align with sustainable development principles, aiding in management and disaster risk assessment.	AI models in remote sensing should be understandable and interpretable by non-technical stakeholders.	The system should exhibit robustness both technically and socially, functioning effectively and ethically across diverse contexts.

Table 2. *Cont.*

	Fairness	Reliability	Sustainability	Transparency and Explainability	Technical and Social Robustness
Challenges	- *Minority bias*: algorithms may favor majority groups. - *Training-serving issues*: discrepancies between training data and real-world scenarios. - *Cohort bias*: over- or under-representation of certain groups. - *Label bias*: inaccurate labeling affecting minority groups.	- *Missing data bias*: lack of data from underrepresented regions. - *Label bias*: incorrect labeling due to ethnic and racial diversity not being captured. - *Automation bias*: overreliance on AI results without local expertise.	*Social and environmental costs*: environmental and societal impact of deploying AI systems, including energy consumption and community impacts.	*Lack of audit*: difficulties in auditing AI algorithms due to complexity or proprietary nature.	- *System adaptability*: ability of the system to adapt to diverse social settings. - *Ethical responsiveness*: responsiveness of the system to ethical concerns across different cultures and societies.
Technical Risks	*Training sample*: selection and composition of training data are crucial for fairness.	*Training sample and change detection*: comprehensive and diverse datasets are vital to avoid skewed outcomes.	*Hardware and software issues*: balancing computational requirements with sustainable practices.	- *Image registration*: aligning images from different sources. - *Image complexity*: handling complex image data. - *Change detection*: assessing changes over time or between datasets.	- *Technical resilience*: ability of the system to maintain functionality in diverse and changing conditions. - *Social integration*: integrating social considerations into the system's functionality.
Detailed Explanation	Fairness is challenged when DL models perpetuate biases or fail to provide equitable analysis due to biased training data.	Reliability is at risk when algorithms are trained on incomplete datasets, leading to inaccurate analysis and predictions.	Sustainability is compromised when AI in remote sensing negatively impacts sustainable development goals.	Transparency is undermined when AI models are too complex for non-technical users.	Technical and Social Robustness is achieved when the system operates effectively and ethically.

3.3. Evaluation Questions

We have rethought general AI ethics guidelines [34–36] through the lens of a few frameworks available for AI in remote sensing to assess how to reframe general questions in a context-specific way (e.g., the majority of frameworks are developed for social media data or health data, but not considering peculiarities of for remote sensing data), and what questions are missing. We reserved a limited space for ethical norms, as we referred to ethical documents that expand such ethical norms/concepts from a theoretical point of view (HLEGAI 2019 (https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai, accessed on 1 December 2024); ALTAI 2022 (https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment, accessed on 1 December 2024); UNESCO 2023 (https://www.unesco.org/en/artificial-intelligence/recommendation-ethics, accessed on 1 December 2024)). Reading these documents together shows how the questions proposed are specifically reframed considering specificities to the field of geomatics and remote sensing. The following ethical assessment questions are derived from benchmark ethical documents and have been contextualized through the analysis of three application scenarios (including the one extended in the paper), which allowed us to further refine them and test their relevance. These questions are based on both global ethical standards and domain-specific sensitivities. They include technical robustness, transparency and explainability, fairness, accessibility, and sustainability.

1. **Technical and social robustness**:
    - Under diverse and changing environmental conditions, how does DL technique maintain consistent performance?

- What strategies are implemented to ensure data integrity and mitigate the risk of corruption/loss?
- How does the model deal with unexpected or anomalous data inputs?
- What safeguards are in place to prevent system failures, and how are these failures managed?
- How adaptable is the DL system to changes in the technology landscape and new data trends?

2. **Transparency and explainability**:
    - Are the DL system's decision-making processes clearly documented and understandable to stakeholders?
    - Can the system provide an understandable explanation for its output, especially in critical decision scenarios?
    - How does the model allow for external auditing and reviewing by third parties?
    - Are there any features that allow users to query or interact with the system in order to have a better understanding of its functionality?
    - What measures are taken to ensure that the operation of the model is transparent and without black box ambiguities?
3. **Fairness**:
    - How does the system ensure that the data collected and processed remains unbiased, particularly regarding underrepresented groups?
    - What steps are taken to identify and correct training bias?
    - How does the model ensure that all demographic groups (e.g., from the economic point of view) are treated in an equal and fair manner?
    - Are there mechanisms in place to review and update the model on a regular basis to prevent bias creep?
    - How are potential biases monitored and managed over time?
4. **Accessibility**:
    - Is the technology user-friendly and accessible to non-experts?
    - How does the system accommodate users with different technical skills or knowledge?
    - What provisions are in place to ensure that the technology is affordable and accessible from an economic point of view?
    - Are there any features or support systems in place that will make the technology accessible to people with disabilities?
    - How does the technology consider and respect the cultural and linguistic diversity of those who use it?
5. **Sustainability**:
    - What are the long-term environmental impacts of the use of this technology, and how are they mitigated?
    - How does DL contribute to sustainable practices in the remote sensing field?
    - Are measures taken to ensure that it uses energy and resources efficiently?
    - In what ways does the technology promote socio-environmental benefits in the communities in which it is used?
    - How is the system designed and operated to reflect a commitment to sustainability and ethical environmental practices?

These comprehensive questions guide the evaluation of DL techniques in remote sensing, ensuring that their deployment aligns with critical ethical principles. By rigorously addressing each of these areas, developers and users can ensure the ethical integrity and societal responsibility of their AI applications.

3.4. Ethical Framework for Explainability in Remote Sensing

Explainability is a cornerstone of ethical AI, especially in remote sensing where stakeholders often include policy makers, scientists and local communities with varying levels of technical expertise. A robust explainability framework ensures that AI models are transparent, interpretable, and trustworthy, enabling informed decision-making and fostering public trust.

Our proposed explainability framework for remote sensing is based on multidimensional principles adapted from established methodologies [45,46]. These principles include:

1. **Computational explainability:** understanding the algorithmic processes by which outputs are produced, such as how satellite imagery is processed to detect deforestation or land-use change.
2. **Rational explainability:** explaining why certain outputs, such as flood risk predictions, are accurate and relevant.
3. **Informative accountability:** communicating the practical implications of outputs, e.g., advising evacuation plans based on identified flood zones.
4. **Cautionary explainability:** highlighting the uncertainties and limitations of AI predictions, such as areas with insufficient data coverage.

The framework further considers explainability at three operational levels:

- **Global explainability:** provides an overall understanding of the behavior and logic of the AI model, which is critical for system audits and regulatory compliance.
- **Local explainability:** provides case-specific insights, such as explaining the classification of a particular region as vulnerable to deforestation.
- **Semi-local accountability:** bridging global and local insights to provide comprehensive explanations tailored to stakeholder needs.

Table 3 provides detailed definitions and examples specific to remote sensing that illustrate how explainability can be integrated into the development and deployment of AI systems. For example, in land cover classification, computational explainability could involve detailing how vegetation indices are used to make classifications, while justificatory explainability could explain why certain areas are classified as forests or urban regions.

By embedding these dimensions and levels of accountability into the workflow, developers can create systems that are not only technically robust, but also ethically aligned, fostering accountability and inclusivity in remote sensing applications.

Table 3. AI ethics framework for explainability in remote sensing.

Multidimensional Explainability		
Dimension	**Definition**	**Example in Remote Sensing**
Computational	How the algorithm generates outputs.	Detecting deforestation by analyzing vegetation index changes over time.
Mechanistic	Why the algorithm produced the output.	Deforestation identified due to vegetation indices dropping below a defined threshold.
Justificatory	Why the output is correct.	Matches ground truth showing recent logging activities in the area.
Informative	What the output means.	Indicates potential illegal logging or land-use change.
Cautionary	Confidence or uncertainty of the output.	Reduced detection accuracy in areas with frequent cloud cover or data gaps.
Explainability by Level		
Level	**Definition**	**Details/Example in Remote Sensing**
Global	Overall model logic.	Explains how multispectral and elevation data classify land use across regions.
Local	Case-specific insights.	Explains why a pixel is classified as water instead of bare soil.
Semi-Local	Combines global and local insights.	Highlights urban classification patterns and pixel-specific misclassifications.

3.5. Quantification and Metrics

We translate the theoretical principles of our ethical framework and the insights gained from the evaluation questions into concrete, measurable metrics. This crucial step involves the careful development of quantifiable metrics that embody the ethical principles identified earlier. These metrics are specifically tailored to evaluate DL techniques in remote sensing. These metrics are implemented through well-defined indicators, such as data anonymization methods for privacy, or diversity in training datasets for fairness, allowing an objective assessment of AI applications. To facilitate the understanding of these evaluations, a scoring system is introduced that allows a structured interpretation of the results. This system not only aggregates individual scores to provide an overall ethical assessment, but also provides clear guidelines for interpreting these results, thus ensuring that the use of AI techniques in remote sensing is not only technologically advanced, but also ethically sound and in line with established standards. To simplify the approach, we can define a set of overall metrics that reflect the key ethical principles (technical and social robustness, transparency and explainability, fairness, accessibility, and sustainability). Each metric is evaluated based on the answers to the corresponding set of pre-formulated evaluation questions. In the methodology, we used two focus groups composed of diverse stakeholders (ethicists, engineers, geomatics, a privacy expert, members of the society at large, and two public decision-makers at the city level in the EU and US), and elected a board of 10 members for a consensus vote on the questions and the scoring used. The questions have been revised and perfected after consulting a pool of experts. Scores were calculated as the average of the individual scores of each panel member. This approach ensured a balanced assessment, minimizing individual bias. Where there were significant discrepancies in scores between panel members, a consensus discussion was held to agree on a final score. Each member of the focus groups has been provided with a metric from 0 to 10, expressing the increasing severity of each ethical aspect/risk/related question, as well as provided with a handbook explaining the ethical issue/risk at stake in the considered context (with a blank space for qualitative observations). To maintain homogeneity and avoid background bias, we used the same weight for each member in average rate elaboration. Indeed, as we know, any AI ethics framework is never static, especially if context-sensitive, but needs to be flexible to continuous input from a growing number of stakeholders in the field considered.

The **technical and social robustness score (TRS)** measures the reliability, error handling, adaptability, and performance consistency of the AI system. It is based on the average of the scores from all relevant questions about the system's performance in various conditions, data integrity, error management, and adaptability to changing technologies.

The **transparency and explainability score (TES)** assesses the clarity of the AI system's decision-making processes and its ability to provide understandable outputs. It is determined by the average score from the system's documentation clarity, audibility, user interaction, and rationale for its outputs.

The **fairness score (FS)** evaluates the AI system's ability to provide unbiased processing and equitable outcomes. It derives the average from the responses to the fairness-related questions, based on the system's handling of data biases, equitable treatment, and mechanisms for bias monitoring and management.

The **accessibility score (AS)** measures how user-friendly and economically accessible the AI system is for a diverse range of users. It determines the system's user-friendliness, adaptability to various technical skill levels, economic accessibility, and inclusivity.

The **sustainability score (SS)** assesses the AI system's environmental impact and its contribution to sustainable practices. It is calculated on the average score from the sustainability questions, based on the system's environmental footprint, resource efficiency, and alignment with long-term sustainability goals.

Table 4 provides an overview of the metrics used for assessing the explainability of AI systems, as introduced in [39] for retail domain. Designed to assess critical aspects such as trust, performance and user understanding, these metrics enable a comprehensive and transparent analysis of the system's ethical compliance.

Table 4. Explainability metrics and evaluation scale [39].

Category	**Indicator**	**Scale (1–10)**
(a) Goodness and Satisfaction		
Understanding	Does the explanation help the user understand how the system works?	1–10
Satisfaction	Is the explanation satisfying?	1–10
Detail	Is the explanation sufficiently detailed?	1–10
Completeness	Is the explanation complete?	1–10
Actionability	Is the explanation actionable (i.e., helps the user handle the system)?	1–10
Accuracy	Does the explanation convey how accurate or reliable the system is?	1–10
Trustworthiness	Does the explanation indicate the trustworthiness of the system?	1–10
(b) Curiosity		
What Happened?	Does the user want to know what the system did?	1–10
What is Next?	Does the user want to understand what the system will do next?	1–10
Alternative Decisions	Does the user want to know why the system did not make another decision?	1–10
Counterfactuals	Does the user want to know what the system would have done if conditions were different?	1–10
(c) Trust		
Confidence	Is the user confident that the system works well?	1–10
Predictability	Are the outputs of the system predictable?	1–10
Reliability	Can users rely on the system to be consistently correct?	1–10
Safety	Does the user feel safe relying on the system?	1–10
Efficiency	Is the system efficient in its operations?	1–10
Skepticism	Is the user wary of the system?	1–10
Comparison	Does the system outperform a novice human user?	1–10
(d) Performance		
Improvement	Will user performance improve with satisfying explanations?	1–10
Epistemic Trust	Does user performance depend on their level of trust?	1–10
Competence Exploration	Is user performance better when exploring the system's competence envelope?	1–10

4. Results

The application of our ethical evaluation framework to DL techniques in remote sensing has yielded a broad spectrum of insights. Based on the Technical State of the Art section, we conducted an audit of these systems to assess their compliance with our ethical principles and to recommend improvements for accountability. This section discusses these results in a general context, reflecting on the potential ethical compliance of AI systems across the key areas of technical and social robustness, transparency and explainability, fairness, accessibility, and sustainability.

We apply our ethical evaluation framework to the methodology presented by [47] for habitat mapping in the Mediterranean Special Area of Conservation "Gola di Frasassi", which yielded insightful findings. In the work, a new methodology has been presented for generating habitat maps using remotely sensed time-series data. The methodology is based on supervised classification supported by functional data analysis. The training data involved 308 plots with 11 different target habitat classes. By using temporal vegetation indices and Functional Principal Component Analysis, an high overall accuracy of was achieved in habitat classification. This region, characterized by its rich biodiversity and complex ecosystem, presents unique challenges, making it an ideal case for assessing the ethical implications of using DL in habitat monitoring. Our framework assesses how these innovative techniques fit with ethical standards, ensuring that they meet environmental, regulatory and societal needs. This assessment is crucial, not only for understanding the wider implications of technological advances in remote sensing, but also for guiding future applications towards more responsible and sustainable environmental practices. By reviewing this methodology, we are contributing to the development of ethically sound practices in remote sensing, which are essential for informed decision-making in conservation policy and practice.

We quantified the compliance of their system across five key metrics: TRS, TES, FS, AS, and SS.

1. **TRS**: given its strong technical performance but less clear social applicability, it receives a score of 8/10.
    - Performance in various conditions: the system showed excellent adaptability and reliability across different environmental scenarios in the Mediterranean region, a key strength in remote sensing applications.
    - Performance consistency: consistent performance was observed in the processing and analysis of time-series data, essential for reliable habitat monitoring.
    - Effectiveness in diverse social contexts: while the system performed well technically, its effectiveness in diverse social contexts (like stakeholder engagement and local community considerations) was not explicitly detailed.
2. **TES**: considering these factors, a score of 8.5/10 reflects the system's transparency and explainability strengths.
    - Decision-making clarity: the methodology provided clear insights into habitat characteristics, aiding decision-making in conservation efforts.
    - Understandability of outputs: while technically sound, the system's outputs and processes might be challenging for non-technical stakeholders to fully grasp.
    - Documentation and auditability: the study's documentation was thorough, facilitating auditability and reproducibility.
    - User interaction: the extent to which users can interact with and query the system was not explicitly covered, suggesting an area for potential enhancement.
3. **FS**: given these considerations, a score of 7/10 is assigned, acknowledging the efforts to address fairness with room for enhanced bias management.
    - Handling of data biases: there is a potential risk of bias in data collection and annotation, a common challenge in remote sensing.
    - Equitable treatment and outcomes: the methodology aims to provide equitable insights across different habitat types, but the degree of success in this regard could vary.
    - Bias monitoring and management: the extent to which biases are systematically monitored and managed was not detailed, suggesting an area for further development.
4. **AS**: reflecting these aspects, a score of 6/10 is appropriate, indicating the need for improvements in making the system more accessible and inclusive.
    - User-friendliness for non-experts: the system's specialized nature might limit its accessibility to non-expert users in the field of remote sensing and conservation.
    - Adaptability to various technical skill levels: the methodology appears to require a certain level of technical expertise, potentially limiting its wider use.
    - Economic accessibility: the economic accessibility of the system for diverse user groups, particularly in resource-limited settings, was not detailed.
    - Cultural and linguistic inclusivity: considerations for cultural and linguistic diversity in the system's use were not explicitly addressed.
5. **SS**: considering its contribution to sustainability in remote sensing, a score of 8/10 is given, acknowledging its positive role with some areas not fully explored.

- Environmental impact and resource efficiency: The methodology's focus on habitat conservation inherently supports sustainability goals. However, the specifics of its environmental impact and resource efficiency were not detailed.
- Contribution to sustainable practices: the system contributes to sustainable habitat monitoring practices, which is a key aspect of environmental conservation.

Final overall score: The average of these scores provides a detailed understanding of the system's ethical compliance. The final score of 7.5/10 indicates strong performance across several ethical dimensions. There is scope for further improvement in areas such as fairness, accessibility, and detailed sustainability assessment.

This expanded evaluation will help guide future efforts to improve ethical compliance in remote sensing methodologies by providing a more comprehensive view of the system's ethical strengths and areas for improvement.

To assess the explainability of the system, we evaluated several dimensions that measure how well the AI methodology supports user understanding, trust, and usability. These dimensions were grouped into categories: Goodness and Satisfaction, Curiosity, Trust, and Performance. Each category reflects specific aspects of explainability relevant to remote sensing applications. Table 5 provides a summary of the evaluation, including the mean scores for each category on a scale of 1 to 10. This evaluation highlights the areas where the system excels and where there is room for improvement, providing a comprehensive view of its strengths and limitations in promoting explainability.

Table 5. Evaluation of explainability metrics.

Explainability Category	Mean Score (1–10)
Goodness and Satisfaction	7.5
Curiosity	8.0
Trust	7.0
Performance	6.8

5. Discussions

Based on these results, it is clear that while the system has strong technical capabilities, there is significant potential to improve its explainability to better meet the needs of different stakeholders. For example, the relatively high scores in categories such as **Goodness** and **Satisfaction** and **Curiosity** indicate that the system provides valuable insights and stimulates user engagement. However, moderate scores in Trust and Performance highlight areas where the system could be improved to inspire greater trust and usability.

To address these gaps, targeted enhancements such as interactive visualization tools, user-friendly dashboards and tailored documentation can bridge the gap between technical complexity and user accessibility. These enhancements will not only make the system more transparent, but will also enable decision makers to better understand, trust and act on its output. By focusing on these improvements, the system can evolve into a more ethically robust tool that meets the needs of both technical and non-technical stakeholders, ensuring its effective application in habitat mapping and conservation efforts. These refinements also set the stage for the development of generalizable explainability practices that can extend beyond this specific use case, benefiting other remote sensing applications and fostering wider adoption of AI technologies in environmental monitoring.

6. Conclusions and Future Works

Our ethical evaluation of [47], based on a framework of quantified metrics, has provided profound insights into the trustworthiness of such advanced technologies. The application of our framework, which assesses technical and social robustness, transparency and accountability, fairness, accessibility, and sustainability, revealed a rich picture of the system's ethical standing. Our quantification process, which resulted in an overall trustworthiness score of 7.5 out of 10, highlights the system's strong performance in certain areas, such as technical robustness and sustainability, while also identifying critical areas for improvement, particularly in fairness and accessibility. This comprehensive assessment underscores the importance of integrating ethical considerations into the development and deployment of remote sensing technologies to ensure that they are not only effective, but also in line with ethical and societal standards. There are promising directions for future research. Key among these is the refinement and expansion of our ethical metrics to capture a broader range of ethical considerations specific to remote sensing technologies. This initiative requires the application and validation of our ethical framework in different ecological and geographical contexts, thereby testing its generalizability and robustness. Integrating these ethical considerations into the development processes of remote sensing technologies emerges as a critical focus, ensuring that ethical standards are embedded from the outset of these technologies.

Author Contributions: Conceptualization, M.P. and R.P.; methodology, M.P. and R.P.; formal analysis, M.P. and R.P.; writing—original draft preparation, M.P. and R.P.; writing—review and editing, M.P. and R.P.; visualization, B.G., S.T. and A.M.; supervision, E.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Union–NextGenerationEU under the Italian Ministry of University and Research (MIUR), National Innovation Ecosystem grant ECS00000041-VITALITY-CUP D83C22000710005.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Workman, S.; Rafique, M.U.; Blanton, H.; Jacobs, N. Revisiting near/remote sensing with geospatial attention. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1778–1787.
2. Yang, F.; Ma, C. Sparse and Complete Latent Organization for Geospatial Semantic Segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1809–1818.
3. Paganini, M.; Petiteville, I.; Ward, S.; Dyke, G.; Steventon, M.; Harry, J.; Kerblat, F. Satellite earth observations in support of the sustainable development goals. In *The CEOS Earth Observation Handbook*; ESA: Paris, France, 2018.
4. Pierdicca, R.; Paolanti, M. GeoAI: A review of artificial intelligence approaches for the interpretation of complex geomatics data. *Geosci. Instrum. Methods Data Syst.* **2022**, *11*, 195–218. [CrossRef]
5. Calzati, S.; van Loenen, B. An ethics assessment list for geoinformation ecosystems: Revisiting the integrated geospatial information framework of the United Nations. *Int. J. Digit. Earth* **2023**, *6*, 1418–1438. [CrossRef]
6. Robinson, C.; Ortiz, A.; Park, H.; Lozano, N.; Kaw, J.K.; Sederholm, T.; Dodhia, R.; Ferres, J.M.L. Fast building segmentation from satellite imagery and few local labels. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1463–1471.
7. Marí, R.; Facciolo, G.; Ehret, T. Sat-nerf: Learning multi-view satellite photogrammetry with transient objects and shadow modeling using rpc cameras. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1311–1321.
8. Huang, X.; Ren, L.; Liu, C.; Wang, Y.; Yu, H.; Schmitt, M.; Hänsch, R.; Sun, X.; Huang, H.; Mayer, H. Urban Building Classification (UBC)—A Dataset for Individual Building Detection and Classification From Satellite Imagery. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1413–1421.
9. Gominski, D.; Gouet-Brunet, V.; Chen, L. Cross-dataset Learning for Generalizable Land Use Scene Classification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1382–1391.
10. Chappuis, C.; Zermatten, V.; Lobry, S.; Le Saux, B.; Tuia, D. Prompt-RSVQA: Prompting visual context to a language model for remote sensing visual question answering. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1372–1381.
11. Petrocchi, E.; Tiribelli, S.; Paolanti, M.; Giovanola, B.; Frontoni, E.; Pierdicca, R. GeomEthics: Ethical Considerations About Using Artificial Intelligence in Geomatics. In Proceedings of the International Conference on Image Analysis and Processing, Udine, Italy, 11–15 September 2023; pp. 282–293.
12. Gevaert, C.M.; Carman, M.; Rosman, B.; Georgiadou, Y.; Soden, R. Fairness and accountability of AI in disaster risk management: Opportunities and challenges. *Patterns* **2021**, *2*, 100363. [CrossRef] [PubMed]

13. Gevaert, C.M. Finding biases in geospatial datasets in the Global South–are we missing vulnerable populations? In Proceedings of the 41st EARSeL Symposium 2022: Earth Observation for Environmental Monitoring, Paphos, Cyprus, 13–16 September 2022.
14. Micheli, M.; Gevaert, C.M.; Carman, M.; Craglia, M.; Daemen, E.; Ibrahim, R.E.; Kotsev, A.; Mohamed-Ghouse, Z.; Schade, S.; Schneider, I.; et al. AI ethics and data governance in the geospatial domain of Digital Earth. *Big Data Soc.* **2022**, *9*, 20539517221138767. [CrossRef]
15. Bennett, M.M. Chinese sociotechnical imaginaries of Earth observation: From sight to foresight. *Big Data Soc.* **2023**, *10*, 20539517231191527. [CrossRef]
16. Bischke, B.; Helber, P.; Folz, J.; Borth, D.; Dengel, A. Multi-task learning for segmentation of building footprints with deep neural networks. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; pp. 1480–1484.
17. Bischke, B.; Helber, P.; Borth, D.; Dengel, A. Segmentation of imbalanced classes in satellite imagery using adaptive uncertainty weighted class loss. In Proceedings of the IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 6191–6194.
18. Gram-Hansen, B.J.; Helber, P.; Varatharajan, I.; Azam, F.; Coca-Castro, A.; Kopackova, V.; Bilinski, P. Mapping informal settlements in developing countries using machine learning and low resolution multi-spectral data. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, Honolulu, HI, USA, 27–28 January 2019; pp. 361–368.
19. Rudner, T.G.; Rußwurm, M.; Fil, J.; Pelich, R.; Bischke, B.; Kopačková, V.; Biliński, P. Multi3net: Segmenting flooded buildings via fusion of multiresolution, multisensor, and multitemporal satellite imagery. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; Volume 33, pp. 702–709.
20. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [CrossRef]
21. Zhang, L.; Zhang, L. Artificial intelligence for remote sensing data analysis: A review of challenges and opportunities. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 270–294. [CrossRef]
22. Ishikawa, S.N.; Matsumura, H.; Uchiyama, Y.; Glesener, L. Automatic detection of occulted hard X-ray flares using deep-learning methods. *Sol. Phys.* **2021**, *296*, 39. [CrossRef]
23. Cheng, G.; Xie, X.; Han, J.; Guo, L.; Xia, G.S. Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 3735–3756. [CrossRef]
24. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat, F. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204. [CrossRef]
25. Yuan, Q.; Shen, H.; Li, T.; Li, Z.; Li, S.; Jiang, Y.; Xu, H.; Tan, W.; Yang, Q.; Wang, J.; et al. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sens. Environ.* **2020**, *241*, 111716. [CrossRef]
26. Li, Z.; Yuan, Q.; Zhang, L. Geo-Intelligent Retrieval Framework Based on Machine Learning in the Cloud Environment: A Case Study of Soil Moisture Retrieval. *IEEE Trans. Geosci. Remote. Sens.* **2023**, *61*, 4502615. [CrossRef]
27. Tian, S.; Guo, H.; Xu, W.; Zhu, X.; Wang, B.; Zeng, Q.; Mai, Y.; Huang, J.J. Remote sensing retrieval of inland water quality parameters using Sentinel-2 and multiple machine learning algorithms. *Environ. Sci. Pollut. Res.* **2023**, *30*, 18617–18630. [CrossRef] [PubMed]
28. Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atkinson, P.M. Joint Deep Learning for land cover and land use classification. *Remote Sens. Environ.* **2019**, *221*, 173–187. [CrossRef]
29. Bergamasco, L.; Bovolo, F.; Bruzzone, L. A dual-branch deep learning architecture for multisensor and multitemporal remote sensing semantic segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 2147–2162. [CrossRef]
30. Nadeem, A.A.; Zha, Y.; Shi, L.; Ali, S.; Wang, X.; Zafar, Z.; Afzal, Z.; Tariq, M.A.U.R. Spatial downscaling and gap-filling of SMAP soil moisture to high resolution using MODIS surface variables and machine learning approaches over ShanDian River Basin, China. *Remote Sens.* **2023**, *15*, 812. [CrossRef]
31. Jozdani, S.; Chen, D.; Pouliot, D.; Johnson, B.A. A review and meta-analysis of generative adversarial networks and their applications in remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *108*, 102734. [CrossRef]
32. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [CrossRef]
33. Raz, A.; Heinrichs, B.; Avnoon, N.; Eyal, G.; Inbar, Y. Prediction and explainability in AI: Striking a new balance? *Big Data Soc.* **2024**, *11*, 20539517241235871. [CrossRef]
34. Ala-Pietilä, P.; Bonnet, Y.; Bergmann, U.; Bielikova, M.; Bonefeld-Dahl, C.; Bauer, W.; Bouarfa, L.; Chatila, R.; Coeckelbergh, M.; Dignum, V.; et al. *The Assessment List for Trustworthy Artificial Intelligence (ALTAI)*; European Commission: Brussels, Belgium, 2020.
35. UNESCO. *Preliminary Report on the First Draft of the Recommendation on the Ethics of Artificial Intelligence*; United Nations Educational, Scientific and Cultural Organization: Paris, France, 2021.
36. UN-GGIM. Integrated Geospatial Information Framework (IGIF). 2020. Available online: https://www.google.com.hk/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.efgs.info/wp-content/uploads/2020/10/EFGS_2020_session2_20.10.2020_Greg-Scott.pdf&ved=2ahUKEwjltf71uoqKAxU3avUHHfgYHWwQFnoECBsQAQ&usg=AOvVaw1FM6Wl28qjtGAA5t1RY33d (accessed on 1 December 2024).
37. Nations, U. *Transforming Our World: The 2030 Agenda for Sustainable Development*; United Nations, Department of Economic and Social Affairs: New York, NY, USA, 2015.

38. Jobin, A.; Ienca, M.; Vayena, E. The global landscape of AI ethics guidelines. *Nat. Mach. Intell.* **2019**, *1*, 389–399. [CrossRef]
39. Tiribelli, S.; Giovanola, B.; Pietrini, R.; Frontoni, E.; Paolanti, M. Embedding AI ethics into the design and use of computer vision technology for consumer's behaviour understanding. *Comput. Vis. Image Underst.* **2024**, *248*, 104142. [CrossRef]
40. Giovanola, B.; Tiribelli, S. Beyond bias and discrimination: Redefining the AI ethics principle of fairness in healthcare machine-learning algorithms. *AI Soc.* **2023**, *38*, 549–563. [CrossRef] [PubMed]
41. Migliorelli, L.; Tiribelli, S.; Cacciatore, A.; Giovanola, B.; Frontoni, E.; Moccia, S. Accountable deep-learning-based vision systems for preterm infant monitoring. *Computer* **2023**, *56*, 84–93. [CrossRef]
42. Mehrabi, N.; Morstatter, F.; Saxena, N.; Lerman, K.; Galstyan, A. A survey on bias and fairness in machine learning. *ACM Comput. Surv. (CSUR)* **2021**, *54*, 1–35. [CrossRef]
43. Suresh, H.; Guttag, J. A framework for understanding sources of harm throughout the machine learning life cycle. In Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization, Virtually, 5–9 October 2021; pp. 1–9.
44. Olteanu, A.; Castillo, C.; Diaz, F.; Kıcıman, E. Social data: Biases, methodological pitfalls, and ethical boundaries. *Front. Big Data* **2019**, *2*, 13. [CrossRef]
45. Cabitza, F.; Campagner, A.; Malgieri, G.; Natali, C.; Schneeberger, D.; Stoeger, K.; Holzinger, A. Quod erat demonstrandum?—Towards a typology of the concept of explanation for the design of explainable AI. *Expert Syst. Appl.* **2023**, *213*, 118888. [CrossRef]
46. Ding, W.; Abdel-Basset, M.; Hawash, H.; Ali, A.M. Explainability of artificial intelligence methods, applications and challenges: A comprehensive survey. *Inf. Sci.* **2022**, *615*, 238–292. [CrossRef]
47. Pesaresi, S.; Mancini, A.; Quattrini, G.; Casavecchia, S. Functional analysis for habitat mapping in a special area of conservation using sentinel-2 time-series data. *Remote Sens.* **2022**, *14*, 1179. [CrossRef]