

Deep Learning and Earth Observation to Support the Sustainable Development Goals

Claudio Persello, Jan Dirk Wegner, Ronny Hänsch, Devis Tuia, Pedram Ghamisi, Mila Koeva and Gustau Camps-Valls

This is the pre-acceptance version of the paper. The final version will appear in the IEEE Geoscience and Remote Sensing Magazine.

The synergistic combination of deep learning models and Earth observation promises significant advances to support the sustainable development goals (SDGs). New developments and a plethora of applications are already changing the way humanity will face the living planet challenges. This paper reviews current deep learning approaches for Earth observation data, along with their application towards monitoring and achieving the SDGs most impacted by the rapid development of deep learning in Earth observation. We systematically review case studies to 1) achieve zero hunger, 2) sustainable cities, 3) deliver tenure security, 4) mitigate and adapt to climate change, and 5) preserve biodiversity. Important societal, economic and environmental implications are concerned. Exciting times ahead are coming where algorithms and Earth data can help in our endeavor to address the climate crisis and support more sustainable development.

I. INTRODUCTION

Machine learning has played a fundamental role in the analysis of Earth observation (EO) data for over three decades, and its importance has been continuously growing. From the early investigations in artificial neural networks and statistical techniques [1], [2], the EO community has been striving for effective algorithms to automate the extraction of information from various sources of remotely sensed images, in situ data and models. The developments in sensor technologies and the increasing availability of voluminous data go hand-in-hand with the demand for more accurate and scalable information extraction methods and tools. This demand is spurred by many geospatial applications and the growing awareness of the necessity to monitor system Earth for the multiple threats to our natural environment, climate, and the sustainable development of human societies.

After a long period when neural networks fell out of fashion, the deep learning (DL) revolution started about a decade ago, and brought back the attention to these powerful learning algorithms [3], [4]. Thanks to the development of specialized hardware, i.e., graphical processing units (GPUs), and the availability of large benchmark data sets, DL networks became more and more popular. They revealed extremely versatile learning machines able to learn virtually any task in data and image analysis [5]. Neural networks can be seen

as trainable data-processing graphs, where the input data is gradually transformed through a sequence of layers that extract intermediate features and are finally used to predict the target output. In a supervised setting, the network is trained with a set of input-output instances, exemplifying the functional relationship between the explanatory covariates and the target variable to predict. Although simplistic, this view depicts the flexibility of neural networks for data analysis purposes. A variety of architectures have been developed so far, the most popular being multilayer perceptrons (MLPs) [6], convolutional neural networks (CNNs) [7], recurrent neural networks (RNNs) [8], autoencoders [9], and generative adversarial networks (GANs) [10]. Moreover, as the number of consecutive layers increases, i.e., the network becomes deeper, the algorithm tends to improve its ability to learn informative features capturing intricate structures within input variables and their relation with the target output.

In the context of EO applications, deep networks can address a large variety of analysis tasks, from image classification and segmentation to data fusion, change detection, object detection and delineation. Deep networks can be designed according to the characteristics of the remotely sensed data and possibly fuse different sensor data types and information layers. One of the main advantages of DL is the ability to learn abstract hierarchical representations of the data, allowing networks to bring spatial, spectral, and temporal patterns hidden in the data to the surface. This results in state-of-the-art performance and enables researchers and engineers to streamline the information extraction processing chain, potentially integrating multi-modal data fusion, feature extraction, and inference tasks into one, holistic, end-to-end learning framework. The combination of DL with powerful computing infrastructure and massive EO data sets opens up tremendous opportunities for geospatial applications. State of the art DL methods are closing the gap between the performance of automated workflows and the need for accurate and reliable information imposed by real applications. Departing from research laboratories, EO and DL have nowadays the opportunity to contribute to some of the most pressing global societal challenges, such as those identified by the 2030 Agenda for Sustainable Development [11]. The United Nations (UN) has defined a set of sev-

“Earth observations can generate data for monitoring a number of SDG targets and indicators. Deep Learning contributes extracting meaningful and consistent information.”

seventeen sustainable development goals (SDGs) as a plan of action to reach peace and prosperity for all people on our planet by 2030. The goals are related to social, economic, and environmental challenges and provide a blueprint for shared action. It is recognised that eradicating poverty in all its forms and dimensions is the greatest global challenge and an indispensable requirement for sustainable development. Each of the seventeen goals has a set of targets and indicators to measure, monitor, and report the progress of each country. The global framework established by the UN is designed around 169 targets and 232 indicators, representing the first truly data-driven framework in which countries can engage with evidence-based decision making and policy development [12]. The 2030 agenda recognizes that *if you can't measure it, you can't manage it*, thus emphasizing the importance of objective, accurate and trustworthy information for decision making. This approach requires using multiple types of data such as traditional national accounts, household surveys and routine administrative data, as well as new sources such as EO data for the extraction of updated geospatial information.

The role of EO in support of the SDGs has been recognized and facilitated by international organizations such as the group on Earth observation (GEO), the committee on Earth observation satellites (CEOS), and the European space agency (ESA) [12]–[14]. EO can provide continuous temporal information over the globe, capturing the sustainability of the developments underpinning the SDG framework. Satellite, airborne, and unmanned aerial vehicle (UAV) acquisitions provide data at multiple scales for monitoring the state of natural ecosystems, natural resources, oceans, coasts, land, built infrastructure and their change over time. EO data are spatially and temporally consistent, allowing for effective comparison of the results among different countries and in different years. EO data are also complementary with traditional statistical methods, offering a source of information to cross-check the validity of in-situ data measurements (such as survey and inventory data) that are commonly collected by national statistical offices. Moreover, EO can significantly reduce the cost of monitoring SDG targets and indicators with respect to traditional data collection methods. According to the “Compendium of EO contributions to the SDG targets and indicators” recently released by ESA [14], 34 SDG indicators can be either directly (17 indicators) or indirectly (17 indicators) informed with space-based EO data across 29 targets and 11 goals. Table I summarizes where EO data can contribute to SDG targets and indicators, providing examples of EO applications in support of monitoring the progress and achieving the goals. The ESA analysis [14] also recognizes the role of the technical infrastructure for storing and processing big EO data, and in particular the relevance of cloud computing, parallel processing systems, and data cubes. However, the contribution of machine learning and DL towards the SDGs is not equally emphasized for their ability to extract meaningful and consistent information from EO data.

This paper aims to analyze the role and opportunities of DL in EO to support the 2030 agenda for sustainable development (Fig. 1). Section II provides a review of the recent developments in DL for EO data analysis, placing them in the

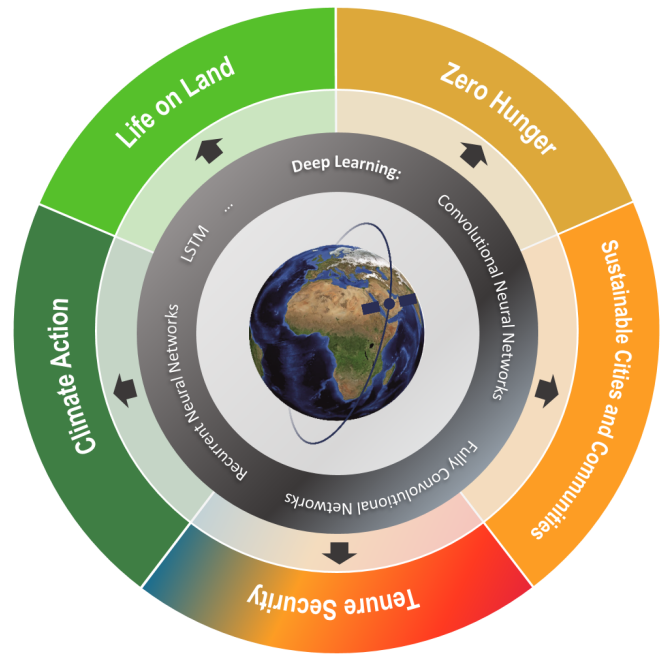


Fig. 1. Framework of the study showing the considered application domains.

context of relevant geospatial applications. Section III gives an overview of applications where DL methods applied to EO data contribute towards key SDGs most impacted by the new technological developments. Section IV reflects on the current achievements, challenges, and future perspectives. Section V draws the conclusion.

II. DEEP LEARNING FOR EARTH OBSERVATION DATA

Advancements in DL, often based on computer vision research, had a large influence in EO image analysis, resulting in the adoption of DL for a variety of data types and geospatial applications [15]–[18].

A. Very high resolution images

The analysis of very high resolution (VHR) images has been the first to benefit from DL. Given the large amount of spatial information and context contained in VHR images, the extraction of features has always been an active field of investigation [19]. With DL, it became possible to learn large dictionaries of convolutional filters directly from data. The appearance of publicly available large-scale data sets issued from competitions enabled the appearance of deep networks specific for scene classification [20], [21] and semantic segmentation [22]–[24] of VHR data. The authors in [20] introduced an explicit metric learning regularization term in the loss function to learn more discriminative features. A wide number of works appeared to process these data sets for classifying land cover at single pixels level: in [25], the authors proposed a hybrid

¹Our list differs slightly from those reported in [12]–[14]. We focus here on target and indicators that can be more directly supported by EO applications and derived products and is not limited to satellite EO, but also considers other EO platforms.

TABLE I
SDG TARGETS AND INDICATORS THAT CAN BE SUPPORTED BY EO-DERIVED PRODUCTS¹.

SDG	Targets	Indicators	Earth observation application in support of SDG targets and indicators
 1 NO POVERTY	1.4	1.4.1, 1.4.2	Extraction of visible cadastral boundaries and information in support of fit-for-purpose land administration systems
	1.5		Risk assessment of natural and climate-induced disasters, early warning and post-event damage assessment
 2 ZERO HUNGER	2.3	2.3.1	Spatial distribution of cropland and smallholder farms; estimation of agricultural productivity
	2.4	2.4.1	Assessment of vulnerability to climate change, extreme weather, drought, flooding and other disasters
 3 GOOD HEALTH AND WELL-BEING	3.3	3.9.1, 3.9.2	Early warning system for vector-borne disease
	3.6		Extraction of road maps and assessment of road conditions (paved/unpaved)
	3.9		Mapping of hazardous chemicals and pollutants in air, water and soil
	3.d		Geospatial information in support of assessing health risks
 5 GENDER EQUALITY	5.a	5.a.1	Extraction of visible cadastral boundaries and geospatial information in support of assessing ownership and secure rights over agricultural land
 6 CLEAN WATER AND SANITATION	6.1, 6.3, 6.4	6.3.2	Mapping of water quality and pollutant concentrations
	6.5		Geospatial data for runoff modeling; Global rainfall data
	6.6	6.6.1	Mapping of water-related ecosystems; Change in the extent of water-related ecosystems over time
 7 AFFORDABLE AND CLEAN ENERGY	7.1	7.1.1	Mapping human presence and availability of electricity (e.g., using night-time images)
	7.2		Geospatial information in support of renewable energies
 9 INDUSTRY, INNOVATION AND INFRASTRUCTURE	9.1	9.1.1	Road and transportation network information in rural areas to support assessing accessibility to all-season roads
 11 SUSTAINABLE CITIES AND COMMUNITIES	11.1, 11.3	11.1.1, 11.3.1	Mapping of slum distributions and extent, housing quality, density and socio-economic conditions of slum dwellers
	11.2	11.2.1	Road network information for assessing the accessibility to public transport
	11.4		Geospatial Mapping and monitoring of cultural and natural heritage sites
	11.5		Risk assessment and early warning of vulnerable urban areas and disaster-induced damage assessment
	11.6	11.6.2	Air quality maps (PM 2.5 and PM 10 concentration); Mapping waste sites
	11.7	11.7.1	Maps of urban green and public open spaces
	11.b, 11.c		Geospatial information for the development of resilient cities in developing countries
 12 RESPONSIBLE CONSUMPTION AND PRODUCTION	12.2		Maps of natural resources
	12.4		Information about waste and pollutant released in air, water and soil
 13 CLIMATE ACTION	13.1		Risks and damages associated with climate-related hazards and natural disasters
	13.2		Environmental variables for climate change models
 14 LIFE BELOW WATER	14.1	14.1.1	Coastal eutrophication and floating plastic debris density
	14.2		Maps of marine and coastal ecosystems
	14.3	14.3.1	Marine acidity (pH)
	14.4	14.4.1	Geochemical (chlorophyll concentration) and geophysical analysis (sea surface temperature and ocean currents) and forecast for global and regional seas
 15 LIFE ON LAND	15.1	15.1.1	Forest maps
	15.2	15.2.1	Forest inventories, deforestation/afforestation maps, wildfire risk assessment
	15.3	15.3.1	Maps of deserts and degraded land, prediction of drought and floods
	15.4	15.4.1;15.4.2	Mountain biodiversity maps
	15.5		Biodiversity maps
	15.7		Wildlife detection to support actions to end poaching and trafficking of protected species

system based on both CNN and traditional descriptor features and then used them in a random forest. In a subsequent paper [26], the author trained two CNN models, one for

color and one for height data. In both cases, predictions were provided at the patch level (i.e., a single label is predicted for the whole patch), and a conditional random field was used to

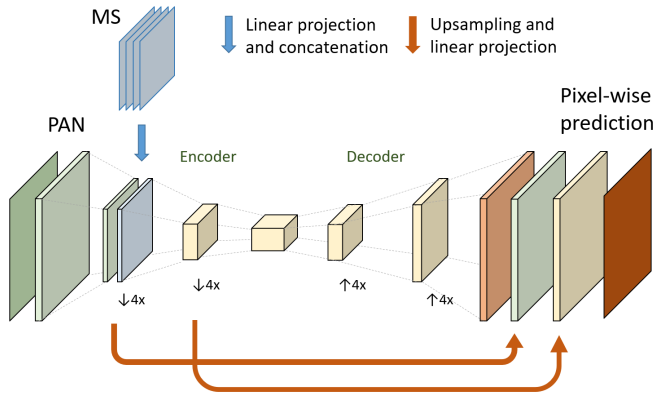


Fig. 2. Architecture of the PAN-MS fusion network (FuseNet) for the pixel-wise classification of VHR multiresolution images [37], [40].

smooth the results.

After these first efforts, papers started to appear applying fully convolutional networks (FCNs) [27], providing predictions for each pixel of the patch in one go and greatly reducing the computational cost at inference time [28]–[32]. Since these pioneering works, a large number of papers tackling semantic segmentation have been published in the field and pushed the boundaries of performance on these data sets. Of notable interest are papers that tackled issues such as the integration of prior knowledge [33], edge information [34], invariances [35], and the explicit inclusion of spatial reasoning [36].

Other works looked at methods to fuse the multiresolution bands acquired by most VHR satellite sensors, such as panchromatic (PAN) and multispectral (MS) images [37]–[39]. Bergado et al. [37], introduces a multiresolution FCN, called FuseNet, to perform an end-to-end image fusion and land-cover classification (Fig. 2). This architecture, tailored to VHR satellite data characteristics, resulted in higher performance than methods based on pansharpening. Contextual label information is included in ReuseNet [40], a fully-convolutional recurrent network able to learn the contextual label-to-label dependencies that are commonly captured by techniques based on conditional random fields.

Going beyond the human design of CNN architecture, Wang et al. investigate a neural architecture search (NAS) approach [41], [42] to automatically design the CNN for the classification of VHR images [43]. Unlike other NAS methods based on reinforcement learning or evolutionary algorithms over a discrete and non-differentiable search space, their framework uses a gradient-based method to optimize both architecture and model parameters [44]. A switchable module allows for addressing both image classification and semantic segmentation.

Another promising research line investigates the extension of DL models to the direct prediction of regularized vector outcomes, i.e., outputs that can be immediately ingested in geographic information system (GIS) environments [45]–[48] (Fig. 3). These developments are expected to have many practical uses for building outline delineation, road network extraction, and more in general, for urban planning and

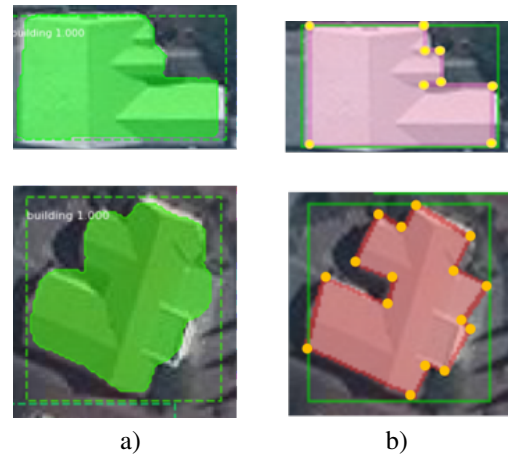


Fig. 3. a) Instance segmentation algorithm with raster output (Mask R-CNN) [59]. b) Regularized outline extraction with polygonal output [48].

monitoring applications in the context of SDG 11.

B. Image time series

Recurrent Neural Networks (RNN) are a powerful method for modeling sequential data leading to much progress especially in language processing [49]–[51]. Their capability of learning long-range patterns over time make RNNs promising tools for a variety of tasks in remote sensing, too. One important example in the context of SDGs is food security and assessment of famine risk (SDG 2), which calls for large-scale mapping of agricultural activities. RNNs allow learning temporal patterns specific to different kinds of agricultural land use like demonstrated in [52]–[55] (see more details in Section III-A1). A cell is the basic building block of an RNN. It combines the data of the current time step in the sequence and the unit’s output from the previous time step as two inputs. While RNNs can, in principle, handle sequences of arbitrary and varying length, they are (in their basic form) challenged by long-term dependencies since learning those would require the propagation of gradients over many time steps. Gated architectures like Long Short-Term Memory (LSTM) cells [56] and Gated Recurrent Units (GRU) [57] aim at mitigating this problem. They use gating mechanisms to store and propagate information over longer time intervals to reduce the vanishing gradient problem. In general, abstract features are often represented better by deeper architectures [58]. In the same way that multiple hidden layers can be stacked in traditional feed-forward networks, multiple recurrent cells can also be stacked on top of each other, i.e., the output (or the hidden state) of the lower cell is connected to the input of the next-higher cell, allowing for different dynamics.

C. Hyperspectral images

Hyperspectral images (HSIs) have intensively contributed to SDGs, in particular, SDG 2 [60], SDG 6 [61], SDG 14 [62], and SDG 15 [63]. CNNs might be the most widely used deep architecture for feature extraction and classification due to the utilization of shared weights and local connections, which

substantially decrease the number of trainable parameters in such networks compared to its fully connected alternatives. In the literature, 1D- [64], 2D- [65], and 3D-CNNs [66] have been employed to extract spectral, spatial, and spectral-spatial information, respectively, from HSI images. The high dimensionality of HSIs, which leads to a higher number of trainable parameters compared to gray scale or multispectral images, along with the availability of only a limited number of training samples, make the training stage of such data extremely challenging. To address these issues, some basic strategies such as dropout and weight decay can be used. In addition, four sets of strategies have been investigated to properly train such high-dimensional data with only a limited number of training samples such as dimensionality reduction [67], data augmentation [68], transfer learning [69], and semisupervised or even unsupervised learning [70].

RNNs have also been applied to HSI image analysis. By considering the spectral signature of each pixel vector as sequential data, RNN can be applied to a single hyperspectral image for classification [71]. In this context, for each pixel vector, the spectral values are usually fed into RNN from the first band to the last one (this can also be done in a bidirectional way [72]), and the output of the hidden layer at the last band is the extracted spectral feature. In real applications, the sequences' length can be very long (equal to the number of bands), which leads to training difficulties such as gradient vanishing or explosion. To address this issue, a possible solution is to group the spectral bands into shorter sequences [73] or use LSTM [74] and GRU [75].

D. Synthetic aperture radar

Synthetic Aperture Radar (SAR) emits coherent microwave pulses and records the amplitude and phase of their backscattered echo. As an active sensor, it is independent of daylight, and due to the used frequencies, it can penetrate clouds, dust, and to some degree vegetation, soil, ice, and other materials. Applications range from estimating surface characteristics such as roughness and moisture and using polarimetric SAR for land cover/use classification, interferometric SAR for the generation of digital elevation models, and tomographic SAR to estimate height profiles over forests or urban areas. However, SAR has not yet seen the same attention of DL as optical sensors (for a detailed review of DL in SAR we refer the reader to [76]). The reasons for this are manifold. First, the imaging geometry differs greatly from optical cameras causing effects unknown in optical imagery such as layover or displacement of moving objects. Furthermore, objects' appearance is strongly view dependent (e.g., certain types of backscatter happen only for certain geometric arrangements between sensor and object). Second, SAR records the amplitude and phase of the received backscatter of a coherent pulse and is therefore complex-valued. While the absolute phase of a single-channel SAR image has no direct meaning, the relative phase between two polarimetric channels or two SAR acquisitions is highly important. Since most machine learning methods and frameworks are designed for real-valued data, early approaches to apply DL to SAR data relied on the

extraction of real-valued (and mostly hand-crafted) features used as input for the network (e.g. [77]). To address this issue, complex-valued CNNs and FCNs that directly work on the complex-valued data by using complex-valued convolutions, activation functions, as well as loss functions have been introduced in [78] and [79], respectively. A NAS approach is proposed in [80] to automate CNN architecture design for SAR data and applied to LC/LU classification. CNNs for scene classification are studied in [81], and RNN architectures for object detection are investigated in [82].

The phase of multiple SAR images plays a particular role in interferometric SAR (InSAR) as it relates to changes in height. It is used to generate digital elevation models, monitor earthquakes and volcanoes or general land subsidence. Corresponding networks need to be invariant to constant phase offsets as well as take the cyclic nature of the phase angle into account. CNNs have been used to enhance the quality of measured interferograms [83], to directly estimate the interferometric phase and coherence [84], as well as performing phase unwrapping [85], i.e. the conversion of the cyclic phase into an absolute phase field to estimate topographic heights or deformations. One particular problem in InSAR processing is decorrelation of the two SAR measurements which can be due to several causes including temporal changes as well as volumetric scattering. The latter often indicates vegetation and is used in [86] as input to a U-Net [87] to derive forests maps.

Another effect that has hindered the direct application of methods designed for optical images to SAR data is speckle: a chaotic fluctuation inherent to all measurements based on coherent waves caused by the interference of multiple backscattering in one resolution cell. Speckle reduction has greatly benefited from DL, by using supervised CNN-based denoising [88], [89], a multistream complex-valued FCN [90] or by exploiting approaches such as noise2noise (e.g., in [91], [92]) that do not require clean data.

E. Big Geodata fusion

A sharp increase in the amount of data captured by sensing devices has led to the *big data* deluge, the creation of the new field of data science and the popularization of DL algorithms to deal with such data [93]. In a similar manner, the field of remote sensing (RS) has been influenced by an ever-growing number of spaceborne, airborne, and proximate sensing devices such as UAVs to acquire multiscale data from a particular scene. The increase in the number, quality and volume of passive sensing devices has been coupled with the growth in the number of alternative modes of measurement such as airborne light detection and ranging (LiDAR), which generates point clouds representing elevation [94], and SAR sensors [95]. Furthermore, the new sources of ancillary data (e.g., data from crowd-sourcing and social media [96]) have been used along with RS devices for a variety of applications in the context of smart cities and smart environment, hazards and disaster identification, or tracking.

Multi-source data fusion aims to integrate the data of different types, distributions, and sources (can be from a single sensor or from different sensors) by leveraging modality-

specific information to improve the performance of the processing approaches compared to a single modality. CNN and its variants have significantly contributed to a wide range of multisource data fusion scenarios such as 1) spatio-spectral fusion to produce a fine-spectral fine-spatial resolution image [97], [98]; 2) spatio-temporal fusion to create a fine-spatio-temporal resolution image [99], [100]; 3) active (e.g., SAR and LiDAR) and passive (e.g., multispectral and hyperspectral) data fusion mainly to improve classification performance or data matching [101], [102]; 4) RS and social media fusion [93], [103].

III. DL APPLICATIONS CONTRIBUTING TO THE SDGs

This section provides an overview of DL and EO applications contributing to the monitoring and achievement of selected SDGs. We focus on applications in the context of 1) zero hunger (SDG 2), 2) sustainable cities and communities (SDG 11), 3) tenure security (multiple SDGs), 4) climate actions (SDG 13), and 5) life on land (SDG 15).

A. Zero Hunger

Monitoring agricultural land use and production is essential to achieve zero hunger (SDG 2). It is of high importance for food production, biodiversity, and forestry [104]. An increasing world population, climate change, and changes in food consumption habits put yet uncultivated areas under pressure, while leading to intensification in existing agricultural areas [105]. Cropland expansion and intensive use of agricultural areas are often connected with negative ecological impacts like deforestation and biodiversity loss, but also degradation of ecosystem services like ground and surface water quality [106], [107]. Therefore, dense, accurate monitoring of agricultural lands plays an essential role for their optimal and sustainable management.

Some of the communities most vulnerable to hunger are smallholder farmers, who dominate agriculture in sub-Saharan Africa with an estimated 51 million farms predominantly characterized by rain-fed production for household consumption [108]. African smallholder farmers often live in poverty in areas prone to natural hazards, where climate change is exacerbating risks of hunger and breakdown of food systems. The large population growth in these areas urgently requires increased production, resilience to natural disasters, improvements in financial services and the governance of food production systems. These improvements are fundamental for defeating hunger and malnutrition, realizing the SDG 2, and in particular target 2.3, which aims to double the agricultural productivity and the incomes of small-scale food producers by 2030. Sustainably increasing the productivity of these agricultural systems and thus improving food security and the livelihoods of smallholder families is a challenge, partly due to a lack of information about these systems.

Knowledge of crop areas and certain land uses is of importance for many political programs that aim to reduce and alleviate the environmental impacts of intensive agriculture too [104]. Policy-driven incentives, for instance, foster a particular share of a farm's area to remain extensively used grassland to

promote biodiversity, or give subsidies to promote a certain crop mix in the rotation [109]. Collecting information is traditionally based on farmer self-reporting and spot checking by the authorities in the field, which is laborious, costly, and prone to errors. Modern machine learning methods in combination with publicly available satellite imagery provides new possibilities for more accurate spatially dense monitoring of agricultural sites at high temporal resolution and low cost. One particularly promising recent sensor is Sentinel-2, due to its low ground sampling distance (10 m) at a revisit rate of 3-5 days. In general, the spectral signal of the vegetation as captured by the satellite has specific characteristics as a function of (i) soil structure and composition (e.g., soil brightness, soil water content, soil type, etc.), (ii) vegetation structure (e.g., canopy cover, Leaf Area Index (LAI), plant height, leaf angle, etc.) and (iii) leaf biochemistry (e.g., chlorophyll, water content, nitrogen content, etc.) [110]. Not only does each plant species have its own spectral signature, but spectral characteristics are also highly dependent on the phenological stage of the plant [111], [112]. Instead of merely analysing images at a single point in time, time-series (sequences) of satellite images thus provide significant additional evidence about crop species, and time-series analysis is a standard practice in agricultural RS.

1) *Crop type mapping*: Crop classification from satellite data has been widely studied in RS. Traditional machine learning approaches with handcrafted features [113], [114] predominantly rely on vegetation indices like the Normalized Difference Vegetation Index (NDVI) [115], [116]. Different strategies have been explored to include the temporal evolution as further evidence for classification, such as temporal windows [117], hidden Markov models and dynamic time warping [118], [119], and conditional random fields [120]. These traditional machine learning models have in common that they struggle to represent the complex spatio-temporal dynamics of spectral features.

Recent DL models no longer rely on hand-engineered features to encode spectral, spatial, and temporal patterns. They can learn very complex, highly non-linear relationships if given sufficient

labeled training data and computational resources. The authors in [122] propose to use a CNN that combines cross-entropy and regression losses for simultaneously mapping and counting oil palm, coconut palm, and olive trees at a country-scale. In [52], the authors use an RNN with LSTM to encode temporal dependencies in the data, while in [123] the results on the same data set are improved by encoding both temporal and spatial dependencies via convolutional LSTM. In [124], satellite images are first processed individually with a CNN to obtain per-image features; then temporal dependencies between these features are modeled with a separate RNN. Further options are temporal CNNs that combine features also across time with convolutions [125], or models that use the attention principle [126] to aggregate information across time [53]. The work in [54] combines pixel-set encoder

“Machine learning in combination with open EO data provides new possibilities for monitoring agricultural sites at low cost”

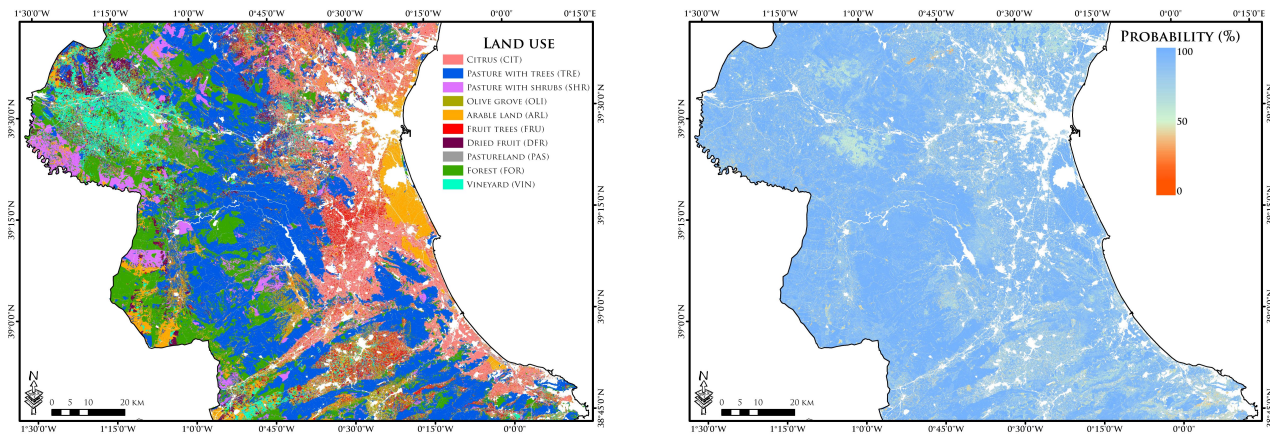


Fig. 4. Example of spatio-temporal deep learning for crop monitoring from high resolution images. A bi-LSTM deep network model was used to exploit the spatial regularities and the temporal dimension of a sequence of Sentinel-2 time series to predict the land use over a number of classes of interest (left) and the pixel probability map (right). Figure adapted from [121].

and transformer [126] and shows improved performance over RNN-based approaches. The authors in [55] build a deep RNN with a new cell structure termed STAR that trains better than LSTM- and GRU-type models and is more parameter-efficient. This makes it possible to train deeper models, which translates to improved performance across a range of sequence modeling tasks. A recent alternative to RNN approaches for crop mapping are neural ordinary differential equations (NODE) that can interpolate in case of missing data [127] due cloud coverage, for example. Finally, recent approaches have considered spatio-temporal bi-LSTM architectures to fully exploit information of long time series of high-resolution Sentinel-2 data to classify different crop types (rice, fallow, barley, oat, wheat, sunflower, and triticale) [121], see Fig. 4.

2) *Delineation of field boundaries*: Field boundaries are essential for digital agricultural services enabling the estimation of cropland areas, to aggregate and record specific information in a spatial database such as crop grown, soil type, yield, application of pesticide and fertilizer. Moreover, they facilitate the extraction of land tenure boundaries for recording land rights in cadastral systems (see Section III-C). Early research on field boundary delineation from EO data has focused on unsupervised techniques based on edge detection or segmentation [128]–[130]. These approaches are typically applied to areas characterized by intensive agriculture with large plots using medium resolution images. However, small area fields (<2 hectare) represent 40% of the fields worldwide and make up 70% of the cropland in Asia and Africa [131]. The delineation of such fields is extremely challenging since plots are small, irregularly shaped often with indistinct boundaries. In these circumstances, standard techniques fail in achieving the required accuracy. To this end, DL-based strategies have resulted in significantly higher performance [132], [133]. An approach based on SegNet [134] and combinatorial grouping was proposed in [132] (Fig. 5). The FCN is trained to detect field contours discarding irrelevant edges. The detected sparse edges are then used as input to the oriented watershed transform algorithm to extract a hierarchy of closed segments and iteratively merging adjacent regions based on the strength

of their common boundary [135]. The final segmentation is obtained by applying the single-scale combinatorial grouping algorithm that explores the segmentation hierarchy to generate accurate field segments [136]. Promising results are obtained in two study areas in Nigeria and Mali. Marvaniya et al. present a multi-stage approach that uses a combination of DL for edge detection and a sequence of post-processing steps for improving the results [133].

Other recent DL-based solutions include [137]–[140]. A method based on U-Net and open data from the land parcel identification system of Spain was investigated in [137]. Waldner and Diakogiannis adopted a multitask approach to tackle the problem. They used a ResUNet-a FCN to identify: 1) the extent of fields; 2) the field boundaries; and 3) the distance to the closest boundary [138]. Using a single monthly composite image from Sentinel-2 as input, their model could accurately map field extent and boundaries. Other notable works have investigated a super-resolution mapping approach [139] and the combination of neural networks with graph-based growing contours method to extract agricultural field polygons [140].

B. Sustainable cities and communities

Cities are the economic hubs of modern nations and home of an estimated 55.3% of the world’s population. By 2030, urban areas are projected to house 60% of the people globally, reaching 68% by 2050 [141]. While urbanization creates opportunities for economic developments, it also creates enormous social and environmental challenges. Some of the most pressing issues are the management of natural hazards, pollution, and the surge of socioeconomic inequalities resulting from excluding the poor from the social fabric. According to UN-Habitat, approximately 1 billion people worldwide reside in informal settlements, commonly called slums, living in deprived conditions lacking access to essential services such

“Approximately 1 billion people worldwide reside in informal settlements, living in deprived conditions lacking access to essential services”

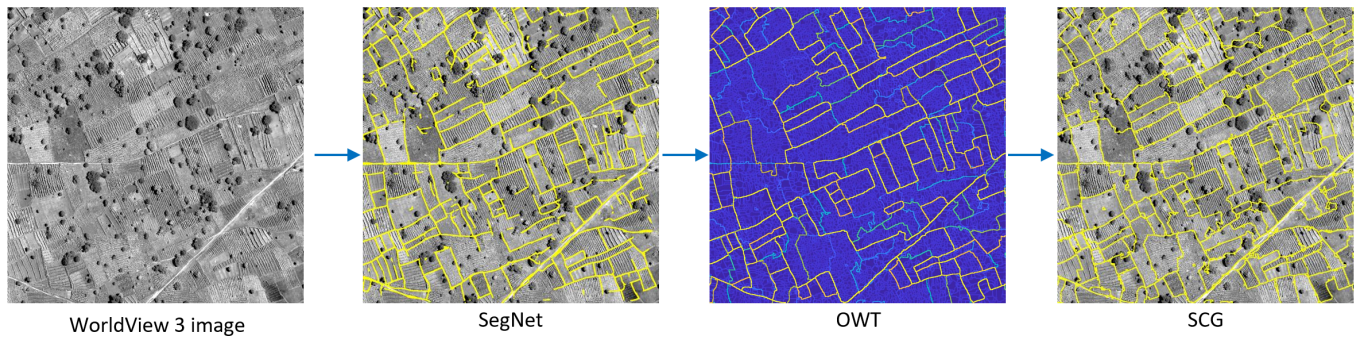


Fig. 5. DL workflow for field boundary delineation in smallholder farms from VHR imagery [132]. SegNet is first applied to extract (fragmented) contours of the agricultural fields. The oriented watershed transform (OWT) is then utilized to extract a hierarchical segmentation. Finally, the single-scale combinatorial grouping (SCG) algorithm globalises local cues using an efficient implementation of normalized cuts and explores the combinatorial space of the segmentation hierarchy to generate regions that are likely to represent complete fields.

as safe water, acceptable sanitation, and durable housing [142].

The rapid urbanization processes in low and middle-income countries contribute to the proliferation of deprived neighborhoods where dwellers live in crowded areas in unhealthy conditions and often without tenure security. In addition to that, these communities are also among the most vulnerable to the effects of climate change and the increasing frequency and intensity of natural disasters such as floods, heatwaves, droughts, landslides, storms, wildfires, and cyclones [143].

1) *Mapping slums and urban poverty*: The 2030 agenda pays particular attention to these global challenges with SDG 11, which aims at inclusive, safe, resilient, and sustainable cities and human settlements. The key indicator 11.1.1 requires monitoring “the proportion of urban population living in slums, informal settlements or inadequate housing.” Current global statistics show a decline in the percentage of urban population living in slums, but an absolute increase of inhabitants living in such areas [144]. Nevertheless, official national statistics are often outdated, inconsistent, or simply inaccurate. Small slum pockets are generally neglected, and population counts based on census data subject to large uncertainties, especially in large metropolitan areas [145]. More accurate and globally consistent methods to gather data on the slum population and their socioeconomic conditions are therefore needed.

Several studies show the ability of remote-sensing techniques to identify informal settlements, providing a relatively consistent mapping approach applicable over large areas and repeatable in time [146]. Maps derived from VHR satellite data can support SDG 11 and the monitoring indicator 11.1.1 in particular. Detailed 2D and 3D geospatial information extracted from UAV data can support the planning and monitoring of urban upgrading projects, thus contributing to targets 11.b and 11.c [147], [148] (Fig. 6). Mapping informal settlements can be performed on the basis of physical and morphological characteristics captured by VHR satellite images. Slums are commonly densely built-up areas characterized by small buildings arranged according to irregular layout patterns and lack of green spaces. Extracting these characteristics automatically from images is, however,

a difficult task. The spectral information alone is insufficient to discriminate between different urban typologies (formal vs. informal). It is necessary to extract contextual features capable of capturing long-range pixel dependency for distinguishing the different spatial patterns. Conventional machine learning approaches resort to the extraction of texture statistics, local binary patterns, oriented gradients, and segment-based features [149], [150]. However, these methods depend on several free parameters, which are difficult to optimize and usually set according to user experience.

The ability of CNNs to automatically learn high-level spatial features results in a streamlined workflow for slum mapping and higher classification accuracy. Mboga et al. apply CNN to detect informal settlements in Dar es Salaam, Tanzania, reporting an accuracy improvement over a support vector machine classifier trained with texture features and local binary patterns [151]. The authors in [32] introduce FCNs for mapping informal settlements from VHR images. To this end, they adopt an FCN architecture with dilated convolutions (named FCN-DK), therefore capturing long-range pixel dependencies while keeping a limited number of network parameters. The best results are obtained by a network with six convolutional layers using increasing dilation factors. Moreover, they report a significant advantage in terms of computational cost at testing time with respect to patch-based CNN. Wurm et al. investigate the transferability of an FCN model pre-trained on VHR images to map slums in coarser resolution Sentinel-2 images and SAR data acquired by TerraSAR-X [152]. They use an FCN-VGG19 architecture adapted from [27]. Their results show that transfer learning can significantly improve the results on Sentinel-2 while not on TerraSAR-X data. Wang et al. investigate a U-Net compound model, including dilated convolution operations, to map deprivation pockets in Bangalore, India [153]. The authors in [154] use an FCN-based approach to study the temporal dynamics of slums, looking in particular at the temporary slum pockets. The study investigates two change detection approaches based on FCN with dilated convolutions. The first approach uses a post-classification change detection, and the second trains FCNs to directly classify the transition in the land cover classes.



a) 2D Orthomosaic



b) 3D photogrammetric point cloud

Fig. 6. UAV image acquired over an informal settlement in Kigali, Rwanda. a) 2D Orthomosaic. b) 3D photogrammetric point cloud [147].

2) *Revealing socioeconomic inequalities*: The papers mentioned above cast the slum mapping problem as a crisp classification, assuming that a boundary can be drawn to separate formal and deprived settlements. Departing from this dichotomy, Ajami et al. adopt a framework conceptualizing the multi-dimensional nature of deprivation including not only the physical (e.g., poor house material) and financial level (e.g., low-income residents), but also human, social, and contextual variables such as accessibility to healthcare, education, and other services or social exclusion factors [155]. The study introduces a data-driven approach to summarize multiple deprivation variables (both categorical and real-valued) into a single real-valued socioeconomic index, named data-driven index of multiple deprivations. A CNN-based transfer-learning method predicts the socioeconomic index values based on VHR images and geographic information system (GIS) features. The results show that an ensemble non-linear regression model, combining the results of the CNN and models based on hand-crafted and GIS features, can explain 75% of the variation in the poverty index obtained from household data.

Other works have applied DL models to nighttime satellite images, street-view, and aerial imagery to infer socioeconomic

conditions. Jean et al. use a CNN-based model to predict economic well-being across five African countries [156]. The CNN model, pre-trained on ImageNet, is fine-tuned to predict nighttime light intensities (used as a proxy for economic activities) corresponding to input daytime satellite imagery. Finally, the CNN-extracted features, along with survey data, are used as input to a ridge regression algorithm to infer the economic well-being. Social, environmental, and health conditions are extracted in [157] by a DL method applied to street-view images for major cities in the UK. Abitbol et al. use a modified EfficientNetB0 CNN architecture [158] to predicting socioeconomic status across France from aerial images and use activation maps to interpret the urban topology [159].

C. Deliver tenure security for all

Secure property rights and efficient registration systems are essential for the modern economy. They give guaranty to individuals and businesses to invest in land, creating the conditions for improving the livelihoods and sustainable management of natural resources, and enabling governments to collect property taxes, which are necessary to finance infrastructure and services to citizens. Unfortunately, a mere 30% of the global population has legally registered rights to their land and homes [160], [161], which means that more than 5 billion of the world's 7.8 billion people do not have documented land rights. Moreover, this percentage drops down to 10% in African countries [162]. The insecurity of land tenure and property rights is often at the root of poverty and inequality [163], [164], leading to legal conflicts, unequal economic systems, locks of assets, challenging effective and democratic governance principles.

The 2030 agenda recognizes the fundamental role of land rights security in several targets and indicators under SDG 1, 2, 5, 11, 15, and 16. The correct registration of land tenure rights directly impacts food security, environmental sustainability, and advancing women's empowerment worldwide.

“The insecurity of land tenure and property rights is often at the root of poverty and inequality”

In many countries, the land is communally owned, but tenure insecurity is often the product of the government's inability to respond to the technical regularization needs [165]. Therefore, significant efforts are needed to formalize land ownership of the poor and vulnerable (target 1.4, indicator 1.4.2). Secure access to land is essential for small-scale agricultural producers to invest in their land and contribute to the market (targets 2.3 and 2.4). It is also fundamental for gender equality, ensuring women's rights to land tenure (SDG 5). The authors in [166] show that secure land tenure for women improves investments in agricultural developments and enhances the chance of women involvement in family food and agricultural productivity. Lack of tenure security also impacts the development of sustainable cities (SDG 11), the management of natural resources (SDG 15), and the synergy between land administration agencies, courts, and legal support services (SDG 16).

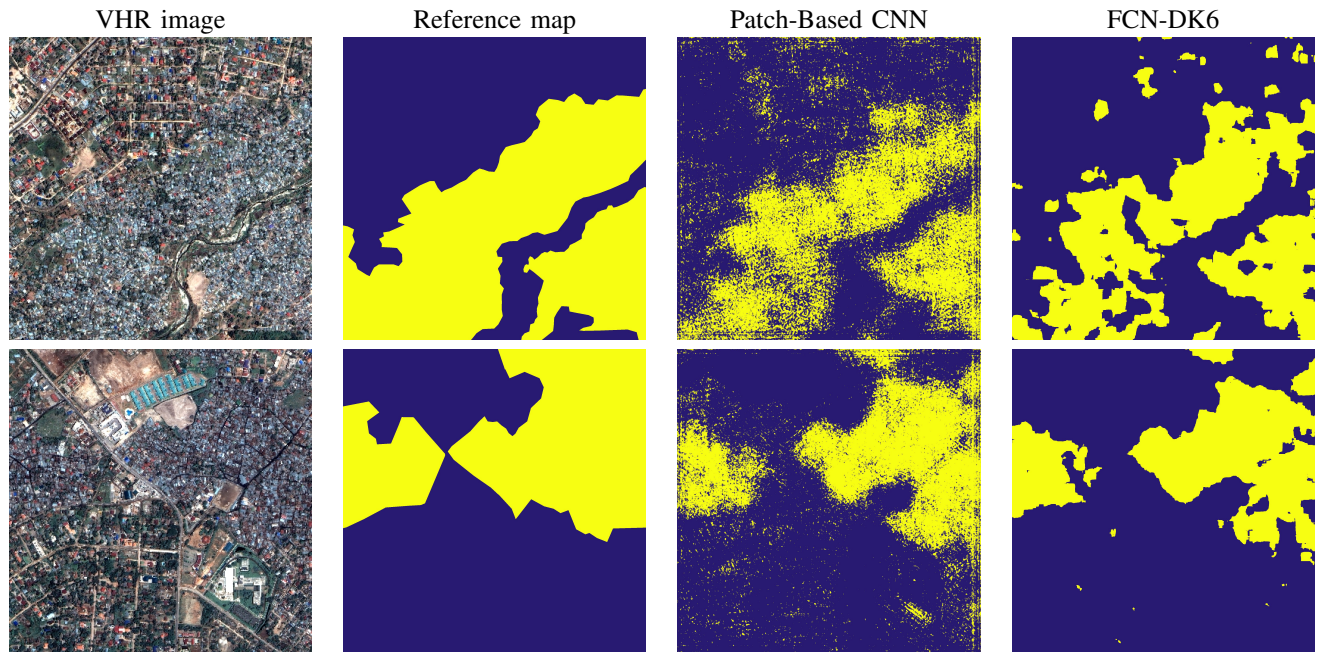


Fig. 7. Informal settlement mapping over two test areas in Dar es Salaam, Tanzania, using patch-based CNN and FCN-DK6 [32]. “informal settlements” in yellow and “rest” in blue.

Strategies to support these goals rely partly on the development of land administration systems (LAS) to formalize land rights and implement land-related policies [167]. There is a clear need for innovation for fast, accurate, and cost-effective cadastral mapping needed for LASs [168]. The traditional surveying methods prove to be quite costly, slow, and labor-intensive. In response, fit-for-purpose (FFP) land administration [169] advises and supports the development of new technologies using remotely sensed data and taking the country context into account. The FFP concept is also included in the recently developed *framework for effective land administration* (FELA) developed by the UN expert group on land administration and management, which is acting as a standard at international level [170]. Spatial land rights recording, i.e., cadastral mapping, is the most expensive part of a land administration system [171]. Automation or semi-automation of cadastral boundary delineation based on satellite or UAV images has been investigated since physical objects often coincide with visible cadastral boundaries and can be detected through image analysis [172]–[176]. The extracted boundaries should be combined with legal information, a procedure known as *adjudication*, and should incorporate local knowledge from human operators.

Initially explored methods for the automated extraction of cadastral boundaries from EO images are based on image segmentation and edge detection [173]. However, the main disadvantages of these methods are sensitivity to intra-parcel variability and dependence of the selected parameters [177], [178]. Better results have been achieved using learning-based contour detectors such as the globalized probability of boundary (gPb) [135], which combines brightness, color, and texture cues into a globalization framework using spectral clustering. Recent studies explored DL methods [179], [180].

The consortium of *its4land*² project developed a suite of open-source solutions for land tenure recording using EO data [168], [181]–[183]. They developed methods based on gPb, SLIC superpixels, and CNN to extract cadastral boundaries and a strategy to assign costs to each line incorporating local user knowledge. This work resulted in an open-source plugin for QGIS providing user-guided delineation functions calculating least-cost paths along the extracted and weighted boundaries. Experiments were conducted using aerial images acquired in Ethiopia and UAV images from Rwanda and Kenya (Fig. 8). Overall, the obtained results based on CNN-derived boundaries achieved a precision of 76%. The use of this semi-automated interactive method leads users to spend 38% less time and 80% fewer clicks compared to manual delineation [179], [184]. Following this research line, Xue et al. [180] explored the potential of FCNs to extract cadastral boundaries in urban and semi-urban areas in Rwanda using UAV data. The authors adopted the FCN-DK architecture, which compared to gPb and multiresolution segmentation, resulted in overall better performance. Nevertheless, the performance of any automated method depends on the presence of visible objects delimiting the boundary of the property (e.g., fences, pathways, walls, roads, land cover transitions). In support of the full land recording process, Chipofya et al. developed an approach incorporating hand-drawn sketch maps with remotely sensed data [185]. The approach converts the raster sketch map into vector automatically, and the hand-drawn symbols are detected and recognized using a CNN. The system performs a stroke-based image segmentation wherein boundaries of sketched objects are drawn and delineated. Finally, the concepts corresponding to the detected symbols are applied to the image segments based on distance and a fixed set of rules specifying

²<https://its4land.com/>

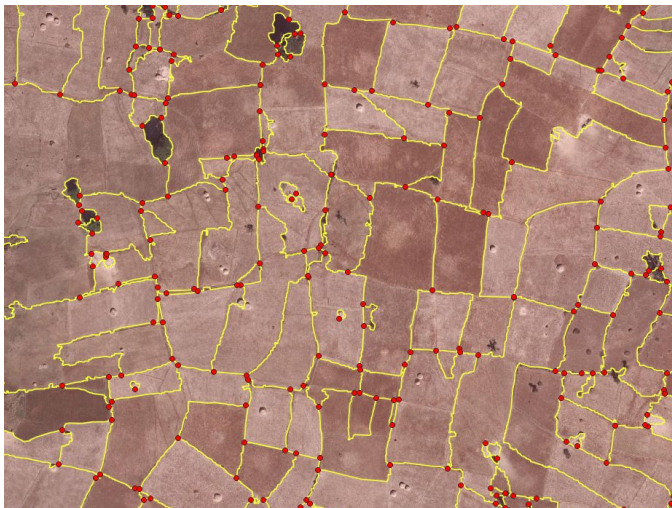


Fig. 8. Cadastral mapping tool (its4land) applied to a rural areas in Ethiopia

spatial constraints on configurations of different types of features.

D. Climate Action

The current scenario of climate change and the projections from climate models call for a definite and urgent action as requested by SDG 13 [186], [187]. Extreme events are more severe, more frequent, and more unexpected in space and time [188]. The Earth system is actually changing globally but also at local and regional scales, with huge implications in ecosystems, biodiversity and agriculture just to name a few. In this scenario, humanity faces the challenge of both mitigation and adaptation to climate change, that is; try to reduce emissions as much as possible while preparing for unavoidable consequences that are no longer a future, but a reality [189]. Machine learning and DL, in particular, can help in the myriad of aspects concerned in both issues. The 2030 agenda focuses mainly on adaptation aspects with target 13.1 demanding to strengthen resilience and adaptive capacity to climate-related hazards and natural disasters in all countries.

Mitigation of greenhouse gas (GHG) emissions requires important changes to electricity systems, transportation, buildings, industry, and land use. Adaptation requires planning for resilience and disaster management, given an understanding of climate and extreme events, see [189] for a organized collective effort to synthesize both the methods and challenges.

1) *DL for climate change mitigation:* Reducing emissions can be achieved with machine learning and DL models. For instance, several DL models have been used to forecast electricity supply and demand, e.g., create short and medium-term forecasts of solar power [190] and wind power [191], [192] or even to use deep networks to produce demand forecasts that optimize for electricity scheduling costs rather than forecast accuracy [193]. DL in combination with RS satellite imagery has been also used to generate size and location data for rooftop solar panels [194], [195], and there are some deep networks that estimate the state of the system [196], [197]. As electricity gets transported from generators to consumers,

some of it gets lost as resistive heat on electricity lines. Prior work has performed predictive maintenance using LSTMs [198] and neural network-plus-clustering techniques [199] on electric grid data. Another important field of action is that of transportation. Decarbonizing transport is essential to a low-carbon society, and there are numerous applications where machine learning can make an impact. For instance, vehicles can be detected in VHR images accurately [200]–[202] and image counts can serve to estimate average vehicle traffic [203]. Neural networks have also been used for analyzing preferences of customers traveling by high-speed trains. Many critical systems inside buildings can be made radically more efficient. Deep autoencoders can be used to simplify information about machine operation so that deep neural networks can then more easily predict multiple kinds of faults [204]. Occupancy detection in buildings can help identify energy demands, a problem where deep neural networks have been also applied [205]. DL can also help to monitor and optimize the operations in smart buildings [206]. Machine learning may be able to help with many aspects of CO₂ sequestration. While still in its infancy, we have seen recent approaches on the use of convolutional image-to-image regression techniques for uncertainty quantification in a global carbon storage simulation study [207]. Such models can help in the development of novel strategies to monitor and develop underground carbon sequestration techniques.

2) *DL for climate change adaptation:* We use general circulation and Earth system models to anticipate climate scenarios on our planet, and to inform local and national governments for decision making. Models have become very precise in projecting scenarios, but still they disagree in some particular cases and are very computationally expensive to run. Machine learning in general and DL in particular can help to mitigate both aspects. The largest part of the uncertainty comes from the parameterization of clouds and aerosols in the models, which have clear implications as bright clouds block sunlight and cool the Earth. Deep neural networks have been used to emulate the behavior of high-resolution cloud resolving simulations at a fraction of the computational cost [208]. Improvements are expected by the combination of DL and process understanding in a new form of hybrid modeling approaches that are data-driven while respecting the fundamental laws of physics [4], [209]. Future improvements in climate modelling will necessarily have to account for the proper characterization and modeling of ice sheet dynamics and sea level rise, yet machine learning has not approached such problems yet systematically [209]–[211].

Weather models are optimized to track the rapid, chaotic changes of the atmosphere, and DL has recently impacted the associated problems. For instance, deep networks are now heavily used to make local forecasts from coarse 10–100 km climate or weather model predictions [212], while other researchers try to translate high-resolution climate forecasts into risk scenarios, e.g. of localized flooding patterns from past data [213], which has clear impacts on individuals. Accurately forecasting hazards and their impacts has societal, economical, and environmental implications. DL is now present in initiatives involving preserving ecosystems at risk [214],

monitoring the risk of food insecurity [215] and deployment of a swift, effective disaster response [216]. Yet, humans can also intervene in the system directly. This is the field of geoengineering. For example, neural networks approaches could facilitate the fast release of aerosols in both space and time [217]. Modeling impacts is also of high relevance; the authors in [218] use deep neural networks to estimate the effects of aerosols on human health, while Crane-Droesch et al. use them to estimate the effects of solar geoengineering on agriculture [219]. Finally, we should note that geoengineering raises many ethical questions, where explainable, accountable AI and fair learning should be part of the discussion.

E. Life on Land

SDG 15 aims to protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, halt and reverse land degradation, and stop the loss of biodiversity. Achieving this goal has far reaching consequences that are closely interlinked with many other SDGs. We review here several application domains where DL and EO play a central role.

1) *Sustainable forest management*: The relevance of sustainable management of forests is linked directly to SDG 15, and target 15.2 in particular, but goes well beyond that. Indeed, the “State of the World’s Forests 2018” report of the Food and Agriculture Organization (FAO) [220] identifies that forests and trees are relevant for 28 targets from ten different SDGs. On the one hand, forests are a key variable to mitigate the effects of climate change, they protect soil and water, and contain more than 75% of the world’s terrestrial biodiversity. On the other hand, forests provide products and services such as food, medicine, and fuel that are of high socio-economic importance in particular in rural areas. The combination of RS and DL has been extensively used to monitor forests, e.g. by producing global forests maps [86], delineating individual tree crowns in aerial imagery [221], [222], or performing damage assessment after storms [223].

Despite their relevance, the loss of the world’s forests through deforestation and forest degradation is an increasing issue destroying natural habitats, limiting resources for the world’s poorest and in the long term worsening climate change by significantly contributing to CO₂ emissions. Reasons for deforestation include tree logging for materials, mining, and farming. Agriculture producing palm oil, beef, soy, pulp and paper is responsible for nearly three quarter of tropical deforestation [220].

“Remote sensing is the ideal tool to monitor large forests that are difficult to access”

Deforestation is mostly happening in rural areas and often performed in secrecy, which requires monitoring large regions that are difficult to access. Thus, RS is the ideal tool

to map and monitor forest that inspired the publication of open data (e.g., in the context of the “Understanding the Amazon from Space” challenge organized by Planet [224]) as well as the usage of DL approaches. Several works focus on experimental comparison between different network models

and shallow learners [225], [226] or apply ensembles of different CNN architectures (e.g., [227] - a participant in the aforementioned challenge). In [228], deforestation mapping is modelled via spatio-temporal deep CNNs by taking several domain-specific components (e.g. handling of clouds) into account. Modern approaches go beyond a mere mapping of forest areas or directly deforestation and instead aim for identifying possible reasons for forest loss. ForestNet [229] not only proposes a deep convolutional network to characterize the processes leading to deforestation but also provides a data set based on Landsat 8 imagery of forest loss events annotated by expert interpreters.

2) *Wildfire risk*: Wildfires, as one of the major factors contributing to deforestation, are becoming more frequent and destructive due to several reasons including higher temperatures, increased droughts, fuel accumulation and dead vegetation, as well as increased population density in close proximity to forests and wildlands. Traditionally, wildfires are detected by human observers either by chance (and then reported to local emergency numbers) or from dedicated watchtowers. Current works aim to complement or even replace the latter by deploying ground-based camera networks (e.g., HPWREN and Alert Wildfire in the State of California, USA, where in 2018 more than 8,000 wildfires burnt 800,000 hectares of land). The camera feed of these networks can then be automatically analyzed by DL approaches as e.g., in [230] which uses an Inception network to detect smoke. An alternative is the usage of UAVs as proposed in [231], which uses a saliency-based system to generate image region proposals that are then analyzed by a standard CNN for classification. Satellite imagery can be used for wildfire detection as well but comes with its own challenges. Geosynchronous satellites such as GOES 16 or GOES 17 constantly observe large parts of a hemisphere but have a rather coarse resolution of several square kilometers which makes detection of wildfires in their early stages difficult. Nevertheless, their image data has been used in combination with DL for wildfire detection (e.g., in [232]). Orbiting satellites such as MODIS, VIIRS, Landsat, and Sentinel-1/2, on the other hand, have a much finer spatial resolution but revisit times of several hours to days. In particular, SAR sensors offer unique benefits as they are able to penetrate clouds and smoke and are independent of daylight. Sentinel-1 time series data and DL have for example been used in [233] to provide a near real-time progression monitoring of wildfires.

Beyond detection and monitoring of wildfires, forecasting their future burns and spread is another important application area [234], [235]. Bergado et al. [235], use a big geodata set to predict wildfire burns. They design FCNs for predicting daily maps of the probability of a wildfire burn over the next 7 days utilizing an extensive set of wildfire related input variables taken from various data sources. 29 quantitative features are selected as input to the models. These features encode factors associated to wildfire burn such as topography (elevation, slope, and aspect), weather (temperature, humidity, solar radiation, rainfall, wind speed and direction, and lightning flash density), proximity to anthropogenic interfaces (distance to roads, distance to power lines) and fuel characteristics

(fuel type, fuel moisture, emissivity). Historical wildfire burn records for Victoria, Australia, collected over the period of 2006–2017 are used for training and testing the DL models. DL and RS have also been used for post-event analysis e.g., for damage assessment [236] or to analyse the impact of wildfires on tree species [237].

3) *Bio-physical parameter estimation*: The problem of retrieving bio-geophysical parameters spans a wide variety of applications and has found a direct impact on achieving the SDGs. The related goals require creating spatially explicit and temporally-resolved maps of quantities and essential climate variables to monitor vegetation status and health, agricultural and forest production. Parameters should be estimated in a consistent and standardized manner to improve accountability. The use and abuse of vegetation indices as proxies of vegetation status and health has been challenged recently by machine learning approaches, from nonlinear generalization of indices [238] to more advanced, yet supervised, machine learning models [209]. Many parameters are now estimated using machine learning; for example, surface temperature and moisture are key parameters for weather prediction with great impact on agriculture and environment, in ecology, hydrology, meteorology, and biology, while leaf area index (LAI) and fractional vegetation cover help in assessing the vegetation cover and dynamics, with implications on crop production. Several seminal works relevant for DL parameter retrieval have been published with focus on Earth Sciences [4], [209], environmental applications [239], and remote sensing [240].

Land parameter retrieval often concerns bio-chemical parameters but can also include physical parameters such as land surface temperature (LST), which was retrieved from microwave radiometer data with DL in [241], which was tested on reference data from both ground stations and other optical satellite data with good results. Leaf-Area-Index (LAI) and Leaf-Chlorophyll-Content (LCC) have been retrieved with optical sensors and using neural networks, yet mostly using shallow architectures [242].

Retrieving parameters can be often hampered by the scarcity in measurements and observations to spatialize them with machine learning. This is the case of relevant parameters for monitoring the land and vegetation such as canopy water content (CWC). For such a case, one can resort to radiative transfer models to generate a look-up-table of expressive simulations to learn from and upscale it in the Google Earth Engine globally [243]. There are some other cases where samples are available in big databases but not sufficiently complete, with many missing attributes or uncertainty in the wild. This was the case of important leaf and plant traits, like phosphorus or nitrogen concentrations, that were not upscaled until the exploitation of the TRY database along with multi-sensor fusion and machine learning [244]. Upscaling carbon, heat and energy fluxes from eddy covariance data has been recently tackled with all kind of machine learning models and neural networks in particular. The key parameters for sensing the health and sensitivity of our warming planet are gross primary production (GPP) and net ecosystem exchange (NEE). Their estimation with neural networks and ensemble methods allow us to quantify global land-atmosphere interactions and

benchmark land surface model simulations [245], [246].

Research in farming applications also relates to biological parameter retrieval. Often though, the goal is not to use predictions as parameters in models, but as proxies for the health condition of crops in so-called smart farming applications. By monitoring and optimising these vegetation indices the goal is to increase crop yield. The variables of interest such as crop type, crop yield, soil moisture and weather variables, can also be used to model and understand the ecosystems that farming effects [247]. Most often though, it is applied to data sets covering only smaller regions of agricultural areas. As opposed to biological parameter retrieval applications, DL is frequently used in farming applications. Some country level work on agriculture has been done for e.g. corn crop yield, [248] and wheat [249], but little work exists on larger scale studies where predictions could be used in models. The authors in [250] provide a comparison of several artificial intelligence methods on a case study in mid-western USA.

Forest cover, biomass and vegetation height are other types of biological parameters which are of high importance to understand and monitor the Earth, with obvious societal and economical implications. DL has also been applied to this problem although mostly on continental level scale, e.g. in [251] used the LSTM networks, while the authors in [252] modeled forest dynamics over a 28-year period by stacking time-series and formulating the task as a change-classification problem, and [253] predicted above ground forest biomass from Lidar and Landsat 8 data with Stacked Sparse Autoencoders (SSAE). The authors in [254] map vegetation height densely at 10-meter resolution from stacks of Sentinel-2 multi-spectral optical satellite imagery at country-scale using CNNs with a regression loss.

4) *Wildlife conservation*: Global loss of biodiversity is observed at all levels [255], and mammals are no exception, with one fifth of them at risk of extinction [256]. Conservation relies heavily on monitoring to estimate biodiversity, as well as resources to sustain life and risks related to human activities (hunting, poaching, expanding agriculture, etc). Despite the urgency of protecting animal populations, the population monitoring is more often done locally in reserves, by experts on foot, and hardly meets the scaling and update requirements to monitor fauna effectively. Due to their larger field of view and the relatively high revisit time potential, satellites [257] and – more importantly – drones [258], are considered more and more for surveying by wildlife ecologists [259], [260]. Drones open perspectives for monitoring on demand, safe detection of poachers and estimation of grazing potential. To process the sheer amount of data collected by drones, researchers are starting to resort to DL massively, to detect animals in the wild with object detection pipelines [261]–[264].

These efforts go hand in hand with computer vision-based community efforts aiming at processing wide archives of camera traps images, i.e. static cameras placed at strategic locations in reserves [265]. To support research in deep

“Remote sensing and deep learning can accelerate conservation efforts and play a central role in the battle against poaching”

learning-based animal conservation, a number of software suites are being proposed, including AIDE [266], which allows ecologists to upload their camera traps or aerial survey and to deploy pre-trained (or own) models in the cloud on Microsoft Azure. Examples of deployments of AIDE in camera traps image classification and single animal detection are reported in Figure 9. The questions of accuracy with respect to animal size and image resolution, or when related to the ratio of background vs animal occupancy in the images (the animals only occupy a fraction of the data collected) are central in these papers. Detection and tracking of poachers is also on the rise, with approaches using thermal images at night [267] or based on deep reinforcement learning [268].

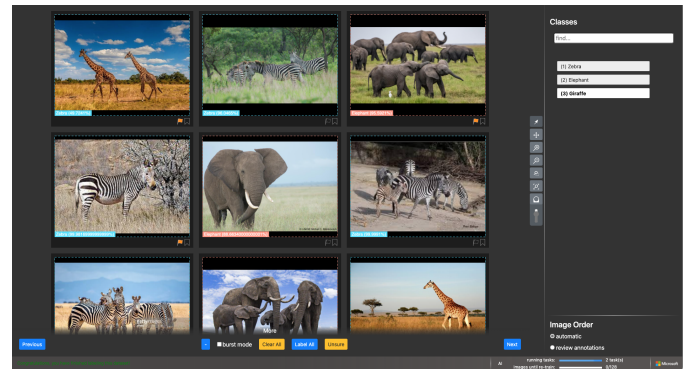
In the context of animal censuses, a question of wide interest is the time efficiency versus the number of animals these algorithms miss. Precise counts are of prime importance and a low recall would force rangers to go through the entire image collection for verification, which would negate the benefits of the DL detection pipeline. Recent research compared DL and citizen science counting methods [269] and concluded that both led to similar accuracy, with a significant speedup achieved when using DL. Also, an active research field is the joining of these two worlds via active learning [270] algorithms: by allowing interactive back and forth between the annotators and the DL models, significant speedups, as well as increased generalization to new campaigns and transfer to new reserves, has been achieved [271]. Finally, these interactive pipelines are nowadays made accessible to the large ecological community, for instance via web-based platforms enabling interactive annotation guided by DL models classifying or detecting in the background [266].

IV. CHALLENGES AND FUTURE OPPORTUNITIES

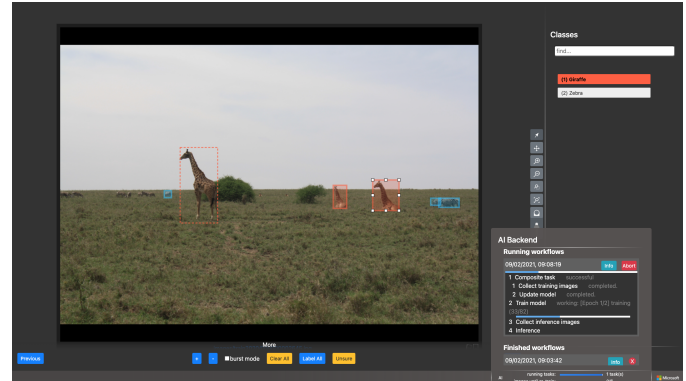
The previous section shows several examples of geospatial applications where DL and EO allow a systematic investigation of global phenomena, providing continuous and spatially consistent information supporting evidence-based decision-making and local interventions. We expect that the coming decade will see a surge of research in this direction, with innovative methodological developments and with an increase in the number and scope of applications in support of the SDGs. However, several questions remain to be addressed: some are purely scientific, others are at the interface between scientific communities, stakeholders and decision-makers. Will the EO scientific community succeed in producing accurate, reliable, consistent and up-to-date geospatial information? Moreover, are these results trustworthy for governmental authorities, stakeholders and local communities? In other words, are DL models trusted by non-experts, who are in charge of decision-making and policy development? The success of evidence-based decision making largely depends on the trust that people have in the data. Transparent data analysis methods and clear communication are fundamental to set proper expectations and build trust between data providers and decision-makers.

A. Open challenges

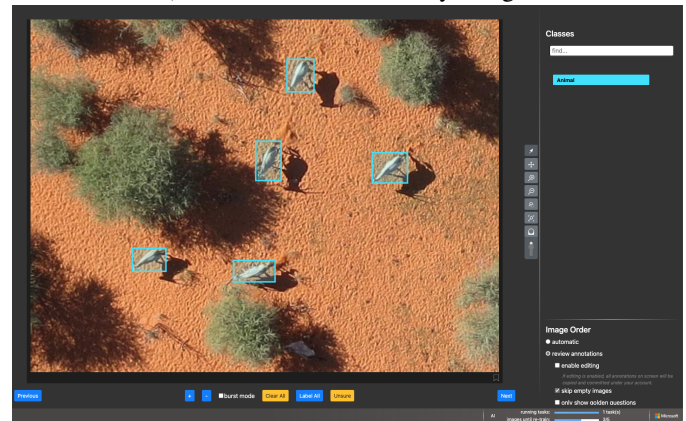
1) *Uncertainty quantification*: To be of true value and gain trust by people, DL models need to provide an indication of



a) Classification of camera traps images



b) Detection in community images



c) Detection in UAV data

Fig. 9. Three deployment cases of the AIDE platform [266] in camera traps images, on (a) camera traps image classification; (b) on animal detection from tourists and photographers pictures acquired during a safari (using the ‘Great Zebra and Giraffe Count’ (GZGC) campaign (<http://lila.science/datasets/great-zebra-giraffe-id>)) and (c) detecting wildlife in UAV images in the Kuzikus reserve in Namibia. In the GZGC case, a typical deployment is shown: a user is editing the detections, while a model is training in the cloud (bottom right box); the predictions of the current model (dashed lines) are also used as a guidance.

the reliability of model predictions. Assigning well-calibrated uncertainties to model outputs plays a critical role in many real-world applications. A significant additional benefit of uncertainty estimates assigned to each data point of the model output is that this creates a practical interface to more traditional post-processing steps using Bayesian models at their core. We can define uncertainty within DL in a twofold way: inherent to all models is epistemic uncertainty, inherent

to all data is aleatoric uncertainty. The former captures the dissimilarity of unseen data compared to what our model has been trained on, i.e. samples that lie within the training distribution have low epistemic uncertainty, samples that are out-of-distribution (OOD) have high epistemic uncertainty. Aleatoric uncertainty results from noise inherent in the observations, such as sensor noise [272]. In addition, there are uncertainties in the spatial domain that arise, for instance, when variables are aggregated over spatial units (e.g., districts or administrative units), resulting in the so-called modifiable areal unit problem (MAUP). It is thus important to consider all sources of uncertainties and their propagation through the whole processing pipeline that affect the quality of the final product.

2) *Data quality quantification for decision making*: It is essential to realize that to support evidence-based policy-making and promote data-driven policy and decision-making, the quality of data products must be carefully assessed and carefully communicated so that non-experts can understand. On the one hand, we encourage the scientific community to pay more attention to how data quality is assessed and communicated. On the other hand, we promote policy developers to incorporate data uncertainties in the decision-making process explicitly. In this respect, we recognize the importance of defining standard data quality measures. Thus, we promote a tighter collaboration between the scientific community and policy makers to define standards on quality measures to quantify the SDG indicators.

3) *Model explainability*: In addition to uncertainty quantification, there is a growing interest in making machine learning and DL models more interpretable and understandable, aiming at neural networks that provide human-understandable justifications for their output, leading to insights about the inner workings [273], [274]. In EO, explainable artificial intelligence is a relatively new field but is quickly becoming important due to the implications that trustable black-box models can have on the usage of DL in societal applications. In the context of agricultural EO, Campos-Taberner et al. investigated how to deepen the understanding of a recurrent neural network for land use classification based on Sentinel-2 time series [121]. In [275], authors studied how land use can be used to explain automatic prediction of the landscapes scenic value (a form of cultural ecosystem service). To do so, they used semantic bottlenecks [276] as intermediate layers of a regression network predicting landscape beauty from Sentinel-2 images. Forcing the network to choose among human-interpretable solutions, then recombined linearly, the model allows understanding why (in terms of land use) the model predicts a given score.

4) *Model transferability*: One of the greatest challenges of DL in EO is the often limited model transferability. For example, a slum mapping DL model trained in Dar es Salaam is unlikely to produce accurate results in Bangalore or São Paulo. This happens not only because the RS images may be affected by different acquisition and radiometric conditions but also because cities in various parts of the world have different characteristics and definition of what constitutes a slum. Despite several studies in domain adaptation and transfer learning [277], model transferability remains a challenge to

ensure spatial and national consistency of indicators derived from DL models. Moreover, non-experts might be unaware of this problem. It is therefore essential that DL model developers provide clear guidance to users regarding the domain where the model is expected to produce valid results.

5) *Interdisciplinary approach*: Addressing global societal problems requires a vast palette of expertise ranging from RS, DL algorithms development, advanced computational skills, as well as domain knowledge in fields such as agriculture, forestry, ecology, urban management and planning, social sciences, land administration, animal conservation, etc. It requires researchers to collaborate and co-design solutions together with other scientists and engage with stakeholders, industrial partners, local communities, governmental and non-governmental organizations. Barriers between different scientific (and non-scientific) communities are often a challenge for an effective interdisciplinary approach.

B. Future opportunities

In this paper, we recognize the importance of the availability of data and computational facilities for the success of DL models. Developments in this direction are offering new opportunities to the EO community. For half a century, Earth has been under continuous observation by satellites to monitor and understand environmental processes. However, historically, RS data was foremost available to those governmental agencies, research institutes, and commercial companies that had direct access to the corresponding sensors. A mixture of different political, organizational, and legal reasons made a free distribution of acquired data difficult to impossible, resulting in a limited number of mostly small data sets [278]. The traditional approach to developing and testing new methods on small and local data sets prevails until today. In particular, in the context of DL, this is problematic as approaches are evaluated on data sets (often consisting of only a single, small image as, e.g., the HSI data set Indian Pines) that do not provide a sufficient amount of independent test samples.

1) *Big and open geodata*: Fortunately, during the last years, large parts of the community have been moving away from closed data and embraced open science principles such as FAIR (findable, accessible, interoperable and reusable [279]) and FOSS (free and open-source software [280]). These developments enable transparent and reproducible scientific research, allow the distribution and reuse of data and methods, and lead to more efficient creation of new data products. This led to open code libraries (such as Open RS <http://openremotesensing.net>) or the IEEE RS Code Library <http://www.grss-ieee.org/publication-category/rscl>), public evaluation servers (such as the IEEE GRSS Data and Algorithm Standard Evaluation website <http://dase.grss-ieee.org>), as well as modern benchmark data sets for a multitude of combinations of EO sensors and tasks which are rapidly replacing older small-scale data sets. These are more in line with the actual situation in RS as current EO data has passed the petabyte-scale and poses common big data challenges regarding volume (i.e., the amount of data), velocity (i.e., the temporal pace with which new data is acquired), and variety (i.e., heterogeneity

regarding image acquisition such as sensors types and modes, environmental factors) [281].

2) *Cloud computing infrastructures*: New opportunities are also arising from the availability of cloud computing infrastructures that allow to visualize and analyze large-scale data (e.g., Landsat [282] and Sentinel data, <https://scihub.copernicus.eu/>) directly in the cloud without the need for local download, storing, and processing. Examples include Digital Earth Australia [283], Earth System Data Lab (ESDL) [284], the Swiss Data Cube [285], the Copernicus Data and Information Access Services (DIAS) and the Google Earth Engine [286]. The availability of large and open data sets in combination with powerful computing infrastructures set the premise for researchers to work more cohesively on addressing the environmental and societal challenges of our time.

3) *A global picture of global phenomena*: Finally, we want to remark that the combination of DL and EO offers the opportunity to obtain a truly global picture of environmental and societal phenomena that go beyond national boundaries as opposed to the data typically collected by national statistical agencies. As discussed in [287], the national-level reporting structure of the SDGs limits the ability to capture environmental phenomena that cross national borders. Moreover, differences in data collection practices of the national offices often result in inconsistent data. EO can provide spatially and temporally consistent data, while DL offers the tools to extract semantic information in an objective and reproducible manner. We thus advocate the use of DL and EO to monitor the progress towards the SDGs and encourage the geoscience and remote sensing community to play an active role in the discussion with stakeholders and policymakers.

V. CONCLUSION

We have reviewed the latest developments in the context of DL for EO and a large number of applications that contribute to the UN agenda for sustainable development. The combination of DL and EO appears to be a strategic asset that can play an essential role in addressing many of the challenges raised by the UN agenda, and beyond that, some of the most urgent demands of human societies. Understanding the role of DL in EO for extracting nationwide geospatial statistical data has far-reaching societal implications for policy development and decision making. Going beyond the SDG agenda, DL and EO can play a significant role in other international agendas such as the New Urban Agenda (NUA) (<https://habitat3.org/the-new-urban-agenda>) or the Sendai framework for disaster risk reduction.

REFERENCES

- [1] J. A. Benediktsson, P. H. Swain, and O. K. Ersoy, "Neural Network Approaches Versus Statistical Methods in Classification of Multisource Remote Sensing Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 4, pp. 540–552, 1990.
- [2] J. Lee, R. C. Weger, S. K. Sengupta, and R. M. Welch, "A Neural Network Approach to Cloud Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 5, pp. 846–855, 1990.
- [3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances In Neural Information Processing Systems*, 2012, pp. 1–9.

- [4] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, and N. C. and, "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, 2019.
- [5] Y. LeCun, B. Yoshua, and H. Geoffrey, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [6] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 10 1986.
- [7] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 12 1989.
- [8] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, p. 1735–1780, 11 1997.
- [9] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science (New York, N.Y.)*, vol. 313, no. July, pp. 504–507, 7 2006.
- [10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," in *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2014, pp. 2672–2680.
- [11] United Nations General Assembly, "Transforming our world: the 2030 Agenda for Sustainable Development," 2015. [Online]. Available: <https://sdgs.un.org/2030agenda>
- [12] European Space Agency (ESA), "Satellite Earth Observations in support of the Sustainable Development Goals," *The CEOS Earth Observation Handbook. Special 2018 Edition*, pp. 1–107, 2018.
- [13] GEO, "Earth Observations in support of the 2030 Agenda for Sustainable Development," Group on Earth Observation (GEO), Tech. Rep., 2017.
- [14] European Space Agency (ESA), "Earth Observation for SDG: Compendium of Earth Observation contributions to the SDG Targets and Indicators," 2020.
- [15] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 8–36, 12 2017.
- [16] Q. Yuan, H. Shen, T. Li, Z. Li, S. Li, Y. Jiang, H. Xu, W. Tan, Q. Yang, J. Wang, J. Gao, and L. Zhang, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sensing of Environment*, vol. 241, no. January, p. 111716, 2020.
- [17] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 152, pp. 166–177, 6 2019.
- [18] T. Hofer and C. Kuenzer, "Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review-Part I: Evolution and Recent Trends," *Remote Sensing*, vol. 12, no. 10, p. 1667, 5 2020. [Online]. Available: <https://www.mdpi.com/2072-4292/12/10/1667>
- [19] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, 2013.
- [20] G. Cheng, C. Yang, X. Yao, L. Guo, and J. Han, "When Deep Learning Meets Metric Learning: Remote Sensing Image Scene Classification via Learning Discriminative CNNs," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 5, pp. 2811–2821, 2018.
- [21] G. Cheng, X. Xie, J. Han, S. Member, and L. Guo, "Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 3735–3756, 2020.
- [22] F. Rottensteiner, G. Sohn, M. Gerke, J. D. Wegner, U. Breitkopf, and J. Jung, "Results of the isprs benchmark on urban object detection and 3d building reconstruction," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, no. 0, pp. 256–271, 2014.
- [23] M. Campos-Taberner, A. Romero-Soriano, C. Gatta, G. Camps-Valls, A. Lagrange, B. L. Saux, A. Beaupère, A. Boulch, A. Chan-Hon-Tong, S. Herbin, H. Randrianarivo, M. Ferecatu, M. Shimoni, G. Moser, and D. Tuia, "Processing of extremely high resolution LiDAR and RGB data: Outcome of the 2015 IEEE GRSS Data Fusion Contest. Part A: 2D contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 12, pp. 5547–5559, 2016.
- [24] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Can semantic labeling methods generalize to any city? the inria aerial image labeling

- benchmark,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2017.
- [25] S. Paisitkriangkrai, J. Sherrah, P. Janney, and A. Van Den Hengel, “Semantic labeling of aerial and satellite imagery,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 7, pp. 2868–2881, 2016.
- [26] J. Sherrah, “Fully convolutional networks for dense semantic labelling of high-resolution aerial imagery,” *arXiv preprint arXiv:1606.02585*, 2016.
- [27] J. Long, E. Shelhamer, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” in *Computer Vision and Pattern Recognition*. IEEE, 6 2015, pp. 3431–3440.
- [28] N. Audebert, B. L. Saux, and S. Lefèvre, “Semantic segmentation of earth observation data using multimodal and multi-scale deep networks,” in *Asian Conference on Computer Vision (ACCV)*, 2016.
- [29] M. Volpi and D. Tuia, “Dense semantic labeling of subdecimeter resolution images with convolutional neural networks,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 881–893, 2017.
- [30] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, “Convolutional neural networks for large-scale remote-sensing image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 645–657, 2017.
- [31] —, “High-resolution aerial image labeling with convolutional neural networks,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7092–7103, 2017.
- [32] C. Persello and A. Stein, “Deep Fully Convolutional Networks for the Detection of Informal Settlements in VHR Images,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2325–2329, 12 2017.
- [33] D. Tuia, M. Volpi, and G. Moser, “Decision fusion with multiple spatial supports by conditional random fields,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3277–3289, 2018.
- [34] D. Marmanis, K. Schindler, J. D. Wegner, S. Galliani, M. Datcu, and U. Stilla, “Classification with an edge: Improving semantic image segmentation with boundary detection,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 135, pp. 158–172, 2018.
- [35] D. Marcos, M. Volpi, B. Kellenberger, and D. Tuia, “Land cover mapping at very high resolution with rotation equivariant CNNs: towards small yet accurate models,” *ISPRS J. Int. Soc. Photo. Remote Sens.*, vol. 145, no. A, pp. 96–107, 2018.
- [36] L. Mou, Y. Hua, and X. X. Zhu, “Relation matters: Relational context-aware fully convolutional network for semantic segmentation of high-resolution aerial images,” *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–13, 2020.
- [37] J. R. Bergado, C. Persello, and A. Stein, “FuseNet: End-to-end multispectral VHR image fusion and classification,” in *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2018-July. IEEE, 2018, pp. 2091–2094.
- [38] Z. Shao and J. Cai, “Remote Sensing Image Fusion with Deep Convolutional Neural Network,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 5, pp. 1656–1669, 2018.
- [39] R. Gaetano, D. Ienco, K. Ose, and R. Cresson, “A Two-Branch CNN Architecture for Land Cover Classification of PAN and MS Imagery,” *Remote Sensing*, vol. 10, no. 11, p. 1746, 11 2018.
- [40] J. Bergado, C. Persello, and A. Stein, “Recurrent Multiresolution Convolutional Networks for VHR Image Classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 11, pp. 6361–6374, 2018.
- [41] B. Zoph and Q. V. Le, “Neural architecture search with reinforcement learning,” in *International Conference on Learning Representations*, 2017, pp. 1–16.
- [42] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, “Learning Transferable Architectures for Scalable Image Recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2018, pp. 8697–8710.
- [43] J. Wang, Y. Zhong, Z. Zheng, A. Ma, and L. Zhang, “RSNet: The Search for Remote Sensing Deep Neural Networks in Recognition Tasks,” *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–15, 2020.
- [44] H. Liu, K. Simonyan, and Y. Yang, “DARTS: Differentiable architecture search,” in *International Conference on Learning Representations (ICLR)*, 2019.
- [45] D. Marcos, D. Tuia, B. Kellenberger, L. Zhang, M. Bai, R. Liao, and R. Urtaşun, “Learning deep structure active contours end-to-end,” in *Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [46] Z. Li, J. D. Wegner, and A. Lucchi, “Topological Map Extraction from Overhead Images,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, p. 715–724.
- [47] N. Girard, D. Smirnov, J. Solomon, and Y. Tarabalka, “Polygonal Building Segmentation by Frame Field Learning,” *arXiv*, 4 2020. [Online]. Available: <http://arxiv.org/abs/2004.14875>
- [48] W. Zhao, C. Persello, and A. Stein, “Building outline delineation: From aerial images to polygons with an improved end-to-end learning framework,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 175, pp. 119–131, 5 2021.
- [49] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning with neural networks,” in *Advances in neural information processing systems*, 2014.
- [50] A. Graves, A.-r. Mohamed, and G. Hinton, “Speech recognition with deep recurrent neural networks,” in *ICASSP*, 2013.
- [51] O. Vinyals and Q. Le, “A neural conversational model,” *arXiv preprint arXiv:1506.05869*, 2015.
- [52] M. Rußwurm and M. Körner, “Temporal vegetation modelling using long short-term memory networks for crop identification from medium-resolution multi-spectral satellite images,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017.
- [53] M. Rußwurm, S. Lefèvre, and M. Körner, “Breizhcrocs: A satellite time series dataset for crop type identification,” in *ICML Workshop*, 2019.
- [54] V. S. F. Garnot, L. Landrieu, S. Giordano, and N. Chehata, “Satellite image time series classification with pixel-set encoders and temporal self-attention,” *arXiv preprint arXiv:1911.07757*, 2019.
- [55] M. O. Turkoglu, S. D’Aronco, J. D. Wegner, and K. Schindler, “Gating revisited: Deep multi-layer rnns that can be trained,” *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2021.
- [56] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [57] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” in *NIPS Workshop*, 2014.
- [58] Y. Bengio, “Learning deep architectures for AI,” *Foundations and trends® in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009.
- [59] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2980–2988.
- [60] T. Hank, K. Berger, H. Bach, J. Clevers, A. Gitelson, P. Zarco-Tejada, and W. Mauser, “Spaceborne imaging spectroscopy for sustainable agriculture: Contributions and challenges,” *Surveys in Geophysics*, vol. 40, pp. 515–551, 2018.
- [61] A. I. Flores-Anderson, R. Griffin, M. Dix, C. S. Romero-Oliva, G. Ochaeta, J. Skinner-Alvarado, M. V. Ramirez Moran, B. Hernandez, E. Cherrington, B. Page, and F. Barreno, “Hyperspectral satellite remote sensing of water quality in lake atitlan, guatemala,” *Frontiers in Environmental Science*, vol. 8, p. 7, 2020.
- [62] J. P. Ryan, C. O. Davis, N. B. Tufflaro, R. M. Kudela, and B.-C. Gao, “Application of the hyperspectral imager for the coastal ocean to phytoplankton ecology studies in monterey bay, ca, usa,” *Remote Sensing*, vol. 6, no. 2, pp. 1007–1025, 2014.
- [63] M. E. Fagan, R. S. DeFries, S. E. Sesnie, J. P. Arroyo-Mora, C. Soto, A. Singh, P. A. Townsend, and R. L. Chazdon, “Mapping species composition of forests and tree plantations in northeastern costa rica with an integration of hyperspectral and multitemporal landsat imagery,” *Remote Sensing*, vol. 7, no. 5, pp. 5660–5696, 2015.
- [64] P. Ghamisi, J. Plaza, Y. Chen, J. Li, and A. J. Plaza, “Advanced spectral classifiers for hyperspectral images: A review,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 1, pp. 8–32, 2017.
- [65] W. Zhao and S. Du, “Spectral-spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4544–4554, 2016.
- [66] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, “Deep feature extraction and classification of hyperspectral images based on convolutional neural networks,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, 2016.
- [67] W. Song, S. Li, L. Fang, and T. Lu, “Hyperspectral image classification with deep feature fusion network,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3173–3184, 2018.
- [68] Y. Kong, X. Wang, and Y. Cheng, “Spectral-spatial feature extraction for hsi classification based on supervised hypergraph and sample expanded cnn,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 11, pp. 4128–4140, Nov 2018.
- [69] X. He, Y. Chen, and P. Ghamisi, “Heterogeneous transfer learning for hyperspectral image classification based on convolutional neural

- network,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3246–3263, 2020.
- [70] L. Mou, P. Ghamisi, and X. X. Zhu, “Unsupervised spectral–spatial feature learning via deep residual conv–deconv network for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 1, pp. 391–406, 2018.
- [71] F. Zhou, R. Hang, Q. Liu, and X. Yuan, “Hyperspectral image classification using spectral-spatial lstms,” *Neurocomputing*, vol. 328, pp. 39–47, 2019.
- [72] Q. Liu, F. Zhou, R. Hang, and X. Yuan, “Bidirectional-convolutional lstm based spectral-spatial feature learning for hyperspectral image classification,” *Remote Sens.*, vol. 9, no. 12, p. 1330, 2017.
- [73] R. Hang, Q. Liu, D. Hong, and P. Ghamisi, “Cascaded recurrent neural networks for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5384–5394, 2019.
- [74] Y. Xu, L. Zhang, B. Du, and F. Zhang, “Spectral-spatial unified networks for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5893–5909, 2018.
- [75] L. Mou, P. Ghamisi, and X. X. Zhu, “Deep recurrent neural networks for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3639–3655, 2017.
- [76] X. Zhu, S. Montazeri, M. Ali, Y. Hua, Y. Wang, L. Mou, Y. Shi, F. Xu, and R. Bamler, “Deep Learning Meets SAR: Concepts, Models, Pitfalls, and Perspectives,” *IEEE Geoscience and Remote Sensing Magazine*, 2021.
- [77] J. Geng, J. Fan, H. Wang, X. Ma, B. Li, and F. Chen, “High-resolution sar image classification via deep convolutional autoencoders,” *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 11, pp. 2351–2355, 2015.
- [78] Z. Zhang, H. Wang, F. Xu, and Y. Jin, “Complex-valued convolutional neural network and its application in polarimetric sar image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 12, pp. 7177–7188, 2017.
- [79] A. G. Mullissa, C. Persello, and A. Stein, “Polsarnet: A deep fully convolutional network for polarimetric sar image classification,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 12, pp. 5300–5309, 2019.
- [80] H. Dong, B. Zou, L. Zhang, and S. Zhang, “Automatic design of cnns via differentiable neural architecture search for polsar image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 9, pp. 6362–6375, 2020.
- [81] Z. Huang, M. Datcu, P. Pan, and B. Lei, “Deep sar-net: Learning objects from signals,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 161, pp. 179 – 193, 2020.
- [82] S. Kazemi, B. Yonel, and B. Yazici, “Deep learning for direct automatic target recognition from sar data,” in *2019 IEEE Radar Conference (RadarConf)*, 2019, pp. 1–6.
- [83] K. Ichikawa and A. Hirose, “Singular unit restoration in insar using complex-valued neural networks in the spectral domain,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 3, pp. 1717–1723, 2017.
- [84] F. Sica, G. Gobbi, P. Rizzoli, and L. Bruzzone, “ ϕ -net: Deep residual learning for insar parameters estimation,” *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–25, 2020.
- [85] F. Sica, F. Calvanese, G. Scarpa, and P. Rizzoli, “A cnn-based coherence-driven approach for insar phase unwrapping,” *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2020.
- [86] A. Mazza, F. Sica, P. Rizzoli, and G. Scarpa, “Tandem-x forest mapping using convolutional neural networks,” *Remote Sensing*, vol. 11, no. 24, p. 2980, Dec 2019.
- [87] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” *Lecture Notes in Computer Science*, vol. 9351, pp. 234–241, 2015.
- [88] P. Wang, H. Zhang, and V. M. Patel, “Sar image despeckling using a convolutional neural network,” *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1763–1767, 2017.
- [89] G. Chierchia, D. Cozzolino, G. Poggi, and L. Verdoliva, “Sar image despeckling through convolutional neural networks,” in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 5438–5441.
- [90] A. G. Mullissa, C. Persello, and J. Reiche, “Despeckling Polarimetric SAR Data Using a Multistream Complex-Valued Fully Convolutional Network,” *IEEE Geoscience and Remote Sensing Letters*, 2021.
- [91] T. Davis, V. Jain, A. Ley, O. D’Hondt, S. Valade, and O. Hellwich, “Reference-free despeckling of synthetic aperture radar images using a deep convolutional network,” in *2020 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2020.
- [92] X. Ma, C. Wang, Z. Yin, and P. Wu, “Sar image despeckling by noisy reference-based deep learning method,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 12, pp. 8807–8818, 2020.
- [93] P. Ghamisi, B. Rasti, N. Yokoya, Q. Wang, B. Hofle, L. Bruzzone, F. Bovolo, M. Chi, K. Anders, R. Gloaguen, P. M. Atkinson, and J. A. Benediktsson, “Multisource and multitemporal data fusion in remote sensing: A comprehensive review of the state of the art,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, no. 1, pp. 6–39, 2019.
- [94] S. A. Bello, S. Yu, C. Wang, J. M. Adam, and J. Li, “Review: Deep learning on 3d point clouds,” *Remote Sensing*, vol. 12, no. 11, p. 1729, May 2020.
- [95] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, “A tutorial on synthetic aperture radar,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 1, pp. 6–43, 2013.
- [96] S. Salcedo-Sanz, P. Ghamisi, M. Piles, M. Werner, L. Cuadra, A. Moreno-Martínez, E. Izquierdo-Verdiguier, J. Muñoz-Marí, A. Mosavi, and G. Camps-Valls, “Machine learning information fusion in earth observation: A comprehensive review of methods, applications and data sources,” *Information Fusion*, vol. 63, pp. 256 – 272, 2020.
- [97] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, “Pansharpening by convolutional neural networks,” *Remote Sensing*, vol. 8, no. 7, p. 594, Jul 2016.
- [98] G. Scarpa, S. Vitale, and D. Cozzolino, “Target-adaptive cnn-based pansharpening,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, pp. 5443–5457, 2018.
- [99] H. Song, Q. Liu, G. Wang, R. Hang, and B. Huang, “Spatiotemporal satellite image fusion using deep convolutional neural networks,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 3, pp. 821–829, 2018.
- [100] M. Belgiu and A. Stein, “Spatiotemporal image fusion in remote sensing,” *Remote Sensing*, vol. 11, no. 7, p. 818, Apr 2019.
- [101] L. H. Hughes, D. Marcos, S. Lobry, D. Tuia, and M. Schmitt, “A deep learning framework for matching of sar and optical imagery,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 169, pp. 166 – 179, 2020.
- [102] R. Hang, Z. Li, P. Ghamisi, D. Hong, G. Xia, and Q. Liu, “Classification of hyperspectral and lidar data using coupled cnns,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 7, pp. 4939–4950, 2020.
- [103] J. Kang, M. Korner, Y. Wang, H. Taubenbock, and X. X. Zhu, “Building instance classification using street view images,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 145, pp. 44 – 59, 2018, deep Learning RS Data.
- [104] C. Gómez, J. C. White, and M. A. Wulder, “Optical remotely sensed time series data for land cover classification: A review,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, pp. 55–72, 2016.
- [105] W. F. Laurance, J. Sayer, and K. G. Cassman, “Agricultural expansion and its impacts on tropical nature,” *Trends in Ecology and Evolution*, vol. 29, no. 2, pp. 107–116, 2014.
- [106] F. Herzog, V. Prasuhn, E. Spiess, and W. Richner, “Environmental cross-compliance mitigates nitrogen and phosphorus pollution from Swiss agriculture,” *Environmental Science and Policy*, vol. 11, no. 7, pp. 655–668, 2008.
- [107] N. B. Dise, M. Ashmore, S. Belyazid, A. Bleeker, R. Bobbink, W. D. Vries, J. W. Erisman, T. Spranger, C. J. Stevens, and L. V. D. Berg, *Nitrogen as a threat to European terrestrial biodiversity*. Cambridge University Press, 2011.
- [108] S. K. Lowder, J. Skoet, and T. Raney, “The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide,” *World Development*, vol. 87, pp. 16–29, 2016.
- [109] R. Finger and N. Lehmann, “Policy reforms to promote efficient and sustainable water use in swiss agriculture,” *Water Policy*, vol. 14, no. 5, pp. 887–901, 2012.
- [110] P. S. Thenkabail, I. Mariotto, M. K. Gumma, E. M. Middleton, D. R. Landis, and K. F. Huemmrich, “Selection of hyperspectral narrowbands (hnbs) and composition of hyperspectral twoband vegetation indices (HVIS) for biophysical characterization and discrimination of crop types using field reflectance and hyperion/EO-1 data,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 427–439, 2013.
- [111] A. Walter, F. Liebisch, and A. Hund, “Plant phenotyping: from bean weighing to image analysis,” *Plant methods*, vol. 11, no. 1, pp. 1–11, 2015.

- [112] J. Anderegg, K. Yu, H. Aasen, A. Walter, F. Liebisch, and A. Hund, "Spectral vegetation indices to track senescence dynamics in diverse wheat germplasm," *Frontiers in Plant Science*, vol. 10, p. 1749, 2020.
- [113] J. Inglada, M. Arias, B. Tardy, O. Hagolle, S. Valero, D. Morin, G. Dedieu, G. Sepulcre, S. Bontemps, P. Defourny *et al.*, "Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery," *Remote Sensing*, vol. 7, no. 9, pp. 12356–12379, 2015.
- [114] F. Vuolo, M. Neuwirth, M. Immitzer, C. Atzberger, and W.-T. Ng, "How much does multi-temporal sentinel-2 data improve crop type classification?" *International journal of applied earth observation and geoinformation*, vol. 72, pp. 122–130, 2018.
- [115] S. Foerster, K. Kaden, M. Foerster, and S. Itzerott, "Crop type mapping using spectral-temporal profiles and phenological information," *Computers and Electronics in Agriculture*, vol. 89, pp. 30–40, 2012.
- [116] M. Ustuner, F. B. Sanli, S. Abdikan, M. Esetlili, and Y. Kurucu, "Crop type classification using vegetation indices of rapideye imagery," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, no. 7, p. 195, 2014.
- [117] C. Conrad, S. Dech, O. Dubovyk, S. Fritsch, D. Klein, F. Löw, G. Schorch, and J. Zeidler, "Derivation of temporal windows for accurate crop discrimination in heterogeneous croplands of uzbekistan using multitemporal rapideye images," *Computers and Electronics in Agriculture*, vol. 103, pp. 63–74, 2014.
- [118] S. Siachalou, G. Mallinis, and M. Tsakiri-Strati, "A hidden markov models approach for crop classification: Linking crop phenology to time series of multi-sensor remote sensing data," *Remote Sensing*, vol. 7, no. 4, pp. 3633–3650, 2015.
- [119] M. Belgiu and O. Csillik, "Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis," *Remote sensing of environment*, vol. 204, pp. 509–523, 2018.
- [120] S. Bailly, S. Giordano, L. Landrieu, and N. Chehata, "Crop-rotation structured classification using multi-source sentinel images and Ipiis for crop type mapping," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2018, pp. 1950–1953.
- [121] M. Campos-Taberner, F. J. García-Haro, B. Martínez, E. Izquierdo-Verdiguier, C. Atzberger, G. Camps-Valls, and M. A. Gilabert, "Understanding deep learning in land use classification based on sentinel-2 time series," *Scientific reports*, vol. 10, no. 1, pp. 1–12, 2020.
- [122] A. Rodriguez and J. Wegner, "Counting the uncountable: deep semantic density estimation from space," in *German Conference on Pattern Recognition, LNCS 11269*, 2018, pp. 351–362.
- [123] M. Rußwurm and M. Körner, "Multi-temporal land cover classification with sequential recurrent encoders," *ISPRS International Journal of Geo-Information*, vol. 7, no. 4, p. 129, 2018.
- [124] R. Rustowicz, R. Cheong, L. Wang, S. Ermon, M. Burke, and D. Lobell, "Semantic segmentation of crop type in africa: A novel dataset and analysis of deep learning methods," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019.
- [125] C. Pelletier, G. I. Webb, and F. Petitjean, "Temporal convolutional neural network for the classification of satellite image time series," *Remote Sensing*, vol. 11, no. 5, p. 523, 2019.
- [126] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in neural information processing systems*, 2017, pp. 5998–6008.
- [127] N. Metzger, M. O. Turkoglu, S. D'Aronco, J. D. Wegner, and K. Schindler, "Crop classification under varying cloud cover with neural ordinary differential equations," *arXiv 2012.02542*, 2020.
- [128] A. Rydberg and G. Borgefors, "Integrated Method for Boundary Delineation of Agricultural Fields in Multispectral Satellite Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 11, pp. 2514–2520, 2001.
- [129] L. Yan and D. P. Roy, "Automated crop field extraction from multi-temporal Web Enabled Landsat Data," *Remote Sensing of Environment*, vol. 144, pp. 42–64, 2014.
- [130] J. Graesser and N. Ramankutty, "Detection of cropland field parcels from Landsat imagery," *Remote Sensing of Environment*, vol. 201, no. August, pp. 165–180, 2017.
- [131] M. Lesiv, J. C. Laso Bayas, L. See, M. Duerauer, D. Dahlia, N. Durando, R. Hazarika, P. Kumar Sahariah, M. Vakolyuk, V. Blyshchyk, A. Bilous, A. Perez-Hoyos, S. Gengler, R. Prestele, S. Bilous, I. u. H. Akhtar, K. Singha, S. B. Choudhury, T. Chetri, Z. Malek, K. Bungnamei, A. Saikia, D. Sahariah, W. Narzary, O. Danylo, T. Sturn, M. Karner, I. McCallum, D. Schepaschenko, E. Moltchanova, D. Fraisl, I. Moorthy, and S. Fritz, "Estimating the global distribution of field size using crowdsourcing," *Global Change Biology*, vol. 25, no. 1, pp. 174–186, 1 2019.
- [132] C. Persello, V. A. Tolpekin, J. R. Bergado, and R. A. de By, "Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping," *Remote Sensing of Environment*, vol. 231, p. 111253, 9 2019.
- [133] S. Marvaniya, U. Devi, J. Hazra, S. Mujumdar, and N. Gupta, "Small, sparse, but substantial: Techniques for segmenting small agricultural fields using sparse ground data," 2020.
- [134] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–14, 2017.
- [135] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 898–916, 2011.
- [136] J. Pont-Tuset, P. Arbelaez, J. T. B. Barron, F. Marques, and J. Malik, "Multiscale combinatorial grouping for image segmentation and object proposal generation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 1, pp. 128–140, 2017.
- [137] A. Garcia-Pedrero, M. Lillo-Saavedra, D. Rodriguez-Esparragon, and C. Gonzalo-Martin, "Deep Learning for Automatic Outlining Agricultural Parcels: Exploiting the Land Parcel Identification System," *IEEE Access*, vol. 7, pp. 158 223–158 236, 2019.
- [138] F. Waldner and F. I. Diakogiannis, "Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network," *Remote Sensing of Environment*, vol. 245, p. 111741, 8 2020.
- [139] K. M. Masoud, C. Persello, and V. A. Tolpekin, "Delineation of Agricultural Field Boundaries from Sentinel-2 Images Using a Novel Super-Resolution Contour Detector Based on Fully Convolutional Networks," *Remote Sensing*, vol. 12, no. 1, p. 59, 12 2019.
- [140] M. P. Wagner and N. Oppelt, "Deep learning and adaptive graph-based growing contours for agricultural field extraction," *Remote Sensing*, vol. 12, no. 12, p. 1990, 2020.
- [141] U. Nations, "World Urbanization Prospects: The 2018 Revision," Tech. Rep., 2018.
- [142] UN-Habitat, "Habitat III Issue papers 22- Informal Settlements," 2015.
- [143] —, "Addressing the most Vulnerable First - Pro-poor Climate Action in Informal Settlements," 2018.
- [144] —, "Slum Almanac 2015 2016," 2015.
- [145] M. Kuffer, J. Wang, M. Nagenborg, K. Pfeffer, D. Kohli, R. Sliuzas, and C. Persello, "The Scope of Earth-Observation to Improve the Consistency of the SDG Slum Indicator," *ISPRS International Journal of Geo-Information*, vol. 7, no. 11, p. 428, 11 2018.
- [146] M. Kuffer, K. Pfeffer, and R. Sliuzas, "Slums from Space—15 Years of Slum Mapping Using Remote Sensing," *Remote Sensing*, vol. 8, no. 6, p. 455, 5 2016.
- [147] C. M. Gevaert, C. Persello, R. Sliuzas, and G. Vosselman, "Informal settlement classification using point-cloud and image-based features from UAV data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 125, pp. 225–236, 2017.
- [148] —, "Monitoring household upgrading in unplanned settlements with unmanned aerial vehicles," *International Journal of Applied Earth Observation and Geoinformation*, vol. 90, p. 102117, 8 2020.
- [149] J. Graesser, A. Cheriyyadat, R. R. Vatsavai, V. Chandola, J. Long, and E. Bright, "Image based characterization of formal and informal neighborhoods in an urban landscape," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 4, pp. 1164–1176, 2012.
- [150] M. Kuffer, K. Pfeffer, R. Sliuzas, and I. Baud, "Extraction of Slum Areas From VHR Imagery Using GLCM Variance," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 5, pp. 1830–1840, 2016.
- [151] N. Mboga, C. Persello, J. Bergado, and A. Stein, "Detection of Informal Settlements from VHR Images Using Convolutional Neural Networks," *Remote Sensing*, vol. 9, no. 11, p. 1106, 10 2017. [Online]. Available: <http://www.mdpi.com/2072-4292/9/11/1106>
- [152] M. Wurm, T. Stark, X. X. Zhu, M. Weigand, and H. Taubenbock, "Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 59–69, 4 2019.
- [153] J. Wang, M. Kuffer, D. Roy, and K. Pfeffer, "Deprivation pockets through the lens of convolutional neural networks," *Remote Sensing of Environment*, vol. 234, p. 111448, 12 2019.

- [154] R. Liu, M. Kuffer, and C. Persello, "The Temporal Dynamics of Slums Employing a CNN-Based Change Detection Approach," *Remote Sensing*, vol. 11, no. 23, p. 2844, 11 2019.
- [155] A. Ajami, M. Kuffer, C. Persello, and K. Pfeffer, "Identifying a Slums' Degree of Deprivation from VHR Images Using Convolutional Neural Networks," *Remote Sensing*, vol. 11, no. 11, p. 1282, 5 2019.
- [156] N. Jean, M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon, "Combining satellite imagery and machine learning to predict poverty," *Science*, vol. 353, no. 6301, pp. 790–794, 2016. [Online]. Available: <http://www.sciencemag.org/cgi/doi/10.1126/science.aaf7894>
- [157] E. Suel, J. W. Polak, J. E. Bennett, and M. Ezzati, "Measuring social, environmental and health inequalities using deep learning and street imagery," *Scientific Reports*, vol. 9, no. 1, pp. 1–10, 2019.
- [158] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *International Conference on Machine Learning*, 2019.
- [159] J. L. Abitbol and M. Karsai, "Socioeconomic correlations of urban patterns inferred from aerial images: interpreting activation maps of Convolutional Neural Networks," *Nature machine intelligence*, vol. 2, p. 684–692, 2020.
- [160] P. Dale and J. McLaughlin, *Land administration*. Oxford University Press, 1999.
- [161] S. Enemark, K. Bell, C. Lemmen, and R. McLaren, *Fit-for-purpose land administration*, S. Enemark, Ed. FIG, 2014.
- [162] C. Boone, "Legal Empowerment of the Poor through Property Rights Reform: Tensions and Trade-offs of Land Registration and Titling in Sub-Saharan Africa," *Journal of Development Studies*, vol. 55, no. 3, pp. 384–400, 3 2019.
- [163] C. Lakner, D. G. Mahler, M. Negre, and E. B. Prydz, "How Much Does Reducing Inequality Matter for Global Poverty?" World Bank Group, Tech. Rep., 2019.
- [164] P. Conforti, *Looking ahead in world food and agriculture: Perspectives to 2050*. FAO, 2011.
- [165] L. Alden Wily, "Collective Land Ownership in the 21st Century: Overview of Global Trends," *Land*, vol. 7, no. 2, p. 68, 5 2018.
- [166] R. Meinzen-Dick, A. Quisumbing, C. Doss, and S. Theis, "Women's land rights as a pathway to poverty reduction: Framework and review of available evidence," *Agricultural Systems*, vol. 172, pp. 72–82, 2019.
- [167] I. Williamson, "The justification of cadastral systems in developing countries," *Geomatica*, vol. 51, no. 1, pp. 21–36, 10 1997.
- [168] M. Koeva, C. Stöcker, S. Crommelinck, M. Chipofya, K. Kundert, A. Schwering, J. Sahib, T. Zein, C. Timm, M. Humayun, J. Crompvoets, E. Tan, C. Casiano, and P. Nkerabigwi, "Innovative geospatial solutions for land tenure mapping," *Rwanda Journal of Engineering, Science, Technology and Environment*, vol. 3, no. 1, 7 2020.
- [169] S. Enemark, R. McLaren, C. Lemmen, D. Antonio, and J. Gitau, *Guiding principles for building fit-for-purpose land administration systems in developing countries*. Nairobi: UN-Habitat, 2016.
- [170] UN-GGIM, "Framework for Effective Land Administration: A reference for developing, reforming, renewing, strengthening or modernizing land administration and management systems Expert Group on Land Administration and Management United Nations Committee of Experts on," 2019.
- [171] I. Williamson, S. Enemark, J. Wallace, and A. Rajabifard, *Land administration for sustainable development*. Redlands: ESRI Press Academic, 2010.
- [172] X. Luo, R. Bennett, M. Koeva, C. Lemmen, and N. Quadros, "Quantifying the Overlap between Cadastral and Visual Boundaries: A Case Study from Vanuatu," *Urban Science*, vol. 1, no. 4, p. 32, 11 2017.
- [173] S. Crommelinck, R. Bennett, M. Gerke, F. Nex, M. Y. Yang, and G. Vosselman, "Review of automatic feature extraction from high-resolution optical sensor data for UAV-based cadastral mapping," 2016.
- [174] M. Manyoky, P. Theiler, D. Steudler, and H. Eisenbeiss, "Unmanned Aerial Vehicle in Cadastral Applications," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVIII-1/, no. September, pp. 57–62, 2012.
- [175] I. Jazayeri, A. Rajabifard, and M. Kalantari, "A geometric and semantic evaluation of 3D data sourcing methods for land and property information," *Land Use Policy*, vol. 36, pp. 219–230, 1 2014.
- [176] M. Koeva, M. Muneza, C. Gevaert, M. Gerke, and F. Nex, "Using UAVs for map creation and updating. A case study in Rwanda," *Survey Review*, vol. 50, no. 361, pp. 1–14, 2016.
- [177] A. García-Pedrero, C. Gonzalo-Martín, and M. Lillo-Saavedra, "A machine learning approach for agricultural parcel delineation through agglomerative segmentation," *International Journal of Remote Sensing*, vol. 38, no. 7, pp. 1809–1819, 4 2017.
- [178] O. E. Dragomir, R. Gouriveau, F. Dragomir, E. Minca, and N. Zerhouni, "Review of prognostic problem in condition-based maintenance," in *2009 European Control Conference (ECC)*, 2009, pp. 1587–1592.
- [179] S. Crommelinck, M. Koeva, M. Y. Yang, and G. Vosselman, "Application of deep learning for delineation of visible cadastral boundaries from remote sensing imagery," *Remote Sensing*, vol. 11, no. 21, pp. 1–22, 2019.
- [180] X. Xia, C. Persello, and M. Koeva, "Deep Fully Convolutional Networks for Cadastral Boundary Detection from UAV Images," *Remote Sensing*, vol. 11, no. 14, p. 1725, 7 2019.
- [181] M. Koeva, S. Crommelinck, C. Stöcker, and J. Crompvoets, "Its4land - Challenges and Opportunities in Developing Innovative Geospatial Tools for Fit-For-Purpose Land Rights Mapping," in *FIG Congress 2018*. Copenhagen: International Federation of Surveyors (FIG), 5 2018, pp. 1–17.
- [182] M. Koeva, C. Stöcker, S. Crommelinck, S. Ho, M. Chipofya, J. Sahib, R. Bennett, J. Zevenbergen, G. Vosselman, C. Lemmen, J. Crompvoets, I. Buntinx, G. Wayumba, R. Wayumba, P. O. Odwe, G. T. Osewe, B. Chika, and V. Pattyn, "Innovative remote sensing methodologies for Kenyan land tenure mapping," *Remote Sensing*, vol. 12, no. 2, pp. 1–27, 1 2020.
- [183] M. N. Koeva, R. M. Bennett, M. Gerke, S. C. Crommelinck, E. C. Stöcker, J. Crompvoets, S. Ho, A. Schwering, M. Chipofya, C. Schultz, T. Zein, M. Biraro, B. K. Alemie, R. Wayumba, and K. Kundert, "Towards innovative geospatial tools for fit-for-purpose land rights mapping," Wuhan, pp. 37–43, 2017.
- [184] S. Crommelinck, B. Höfle, M. Koeva, M. Yang, and G. Vosselman, "Interactive cadastral boundary delineation from UAV data," in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, no. 2, 2018.
- [185] M. Chipofya, M. Karamesouti, C. Schultz, and A. Schwering, "Local Domain Models for Land Tenure Documentation and their Interpretation into the LADM," *Land Use Policy*, vol. 99, p. 105005, 12 2020.
- [186] J. Romm, *Climate Change: What Everyone Needs to Know*. Oxford University Press, 2018.
- [187] D. Archer and S. Rahmstorf, *The climate crisis: An introductory guide to climate change*. Cambridge University Press, 2010.
- [188] C. B. Field, V. Barros, T. F. Stocker, and Q. Dahe, *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*. Cambridge University Press, 2012.
- [189] D. Rolnick, P. L. Donti, L. H. Kaack, K. Kochanski, A. Lacoste, K. Sankaran, A. S. Ross, N. Milojevic-Dupont, N. Jaques, A. Waldman-Brown *et al.*, "Tackling climate change with machine learning," *arXiv preprint arXiv:1906.05433*, 2019.
- [190] J. Mathe, N. Miolane, N. Sebastien, and J. Lequeux, "Pvnet: A lrcn architecture for spatio-temporal photovoltaic power forecasting from numerical weather prediction," 2020.
- [191] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation," *Renewable Energy*, vol. 37, no. 1, pp. 1–8, 2012.
- [192] [Online]. Available: <https://deepmind.com/blog/machine-learning-can-boost-value-wind-energy/>
- [193] P. Donti, B. Amos, and J. Z. Kolter, "Task-based end-to-end model learning in stochastic optimization," in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30. Curran Associates, Inc., 2017, pp. 5484–5494.
- [194] J. M. Malof, K. Bradbury, L. M. Collins, and R. G. Newell, "Automatic detection of solar photovoltaic arrays in high resolution aerial imagery," *Applied energy*, vol. 183, pp. 229–240, 2016.
- [195] J. Yu, Z. Wang, A. Majumdar, and R. Rajagopal, "DeepSolar: A machine learning framework to efficiently construct a solar deployment database in the United States," *Joule*, vol. 2, no. 12, pp. 2605–2617, 2018.
- [196] H. Jiang and Y. Zhang, "Short-term distribution system state forecast based on optimal synchrophasor sensor placement and extreme learning machine," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1–5.
- [197] M. Pertl, K. Heussen, O. Gehrke, and M. Rezkalla, "Voltage estimation in active distribution grids using neural networks," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1–5.
- [198] B. Bhattacharya and A. Sinha, "Deep fault analysis and subset selection in solar power grids," 2017, *Preprint*, archivePrefix = arXiv, eprint = 1711.02810.

- [199] V. N. Nguyen, R. Jenssen, and D. Roverso, "Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning," *International Journal of Electrical Power & Energy Systems*, vol. 99, pp. 107–120, 2018.
- [200] Q. Jiang, L. Cao, M. Cheng, C. Wang, and J. Li, "Deep neural networks-based vehicle detection in satellite images," in *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*, 2015, pp. 184–187.
- [201] T. N. Mundhenk, G. Konjevod, W. A. Sakla, and K. Boakye, "A large contextual dataset for classification," in *European Conference on Computer Vision*. Springer: detection and counting of cars with deep learning. In, 2016, pp. 785–800.
- [202] Z. Deng, H. Sun, S. Zhou, J. Zhao, and H. Zou, *Toward fast and accurate vehicle detection in aerial images using coupled region-based convolutional neural networks*. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017.
- [203] L. H. Kaack, G. H. Chen, and M. G. Morgan, "Truck traffic monitoring with satellite images," in *Proceedings of the 2Nd ACM SIGCAS Conference on Computing and Sustainable Societies*. New York, NY, USA ACM: COMPASS '19, 2019, pp. 155–164.
- [204] F. Jia, Y. Lei, J. Lin, X. Zhou, and A. Na Lu, "Deep neural networks: Promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," *Mechanical Systems and Signal Processing*, vol. 72, pp. 303–315, 2016.
- [205] H. Zou, Y. Zhou, J. Yang, and C. J. Spanos, "Towards occupant activity driven smart buildings via wifi-enabled IoT devices and deep learning," *Energy and Buildings*, vol. 177, pp. 12–22, 2018.
- [206] M. Ateeq, F. Ishmanov, M. K. Afzal, and M. Naeem, "Multi-parametric analysis of reliability and energy consumption in IoT: A deep learning approach," *Sensors*, vol. 19, p. 2, 2019.
- [207] S. Mo, Y. Zhu, N. Zabarax, X. Shi, and J. Wu., "Deep convolutional encoder-decoder networks for uncertainty quantification of dynamic multiphase flow in heterogeneous media," *Water Resources Research*, vol. 55, no. 1, pp. 703–728, 2019.
- [208] P. Gentine, M. Pritchard, S. Rasp, G. Reinaudi, and G. Yacalis, "Could machine learning break the convection parameterization deadlock?" *Geophysical Research Letters*, vol. 45, pp. 5742–5751, 2018.
- [209] G. Camps-Valls, D. Tuia, X. X. Zhu, and M. R. (Editors), *Deep learning for the Earth Sciences: A comprehensive approach to remote sensing, climate science and geosciences*. Wiley & Sons, 2021.
- [210] J. Bolibar, A. Rabatel, I. Gouttevin, C. Galiez, T. Condom, and E. Sauquet, "Deep learning applied to glacier evolution modelling," *The Cryosphere*, vol. 14, no. 2, pp. 565–584, 2020.
- [211] V. Nieves, C. Radin, and G. Camps-Valls, "Predicting regional coastal sea level changes with machine learning," *Sci. Rep. (submitted)*, 2021.
- [212] W. Li, L. Ni, Z. liang Li, S. bo Duan, and H. Wu., "Evaluation of machine learning algorithms in spatial downscaling of MODIS land surface temperature," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, pp. 2299–2307, 2019.
- [213] M. C. Perignon, P. Passalacqua, T. M. Jarriel, J. M. Adams, and I. Overeem, *Patterns of geomorphic processes across deltas using image analysis and machine learning*. In AGU Fall Meeting Abstracts, 2018.
- [214] L. Bragilevsky and I. V. Bajić, "Deep learning for Amazon satellite image analysis," in *2017 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*. IEEE, 2017, pp. 1–5.
- [215] A. X. Wang, C. Tran, N. Desai, D. Lobell, and S. Ermon, "Deep transfer learning for crop yield prediction with remote sensing data," in *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*. page 50. ACM, 2018.
- [216] S. Voigt, T. Kemper, T. Riedlinger, R. Kiefl, K. Scholte, and H. Mehl, "Satellite image analysis for disaster and crisis-management support," *IEEE transactions on geoscience and remote sensing*, vol. 45, no. 6, pp. 1520–1528, 2007.
- [217] C. S. de Witt and T. Hornigold, "Stratospheric aerosol injection as a deep reinforcement learning problem," 2019, *Preprint*, archivePrefix = arXiv, eprint = 1905.07366.
- [218] Q. Di, I. Kloog, P. Koutrakis, A. Lyapustin, Y. Wang, and J. Schwartz, "Assessing PM_{2.5} exposures with high spatiotemporal resolution across the continental United States," *Environmental science & technology*, vol. 50, no. 9, pp. 4712–4721, 2016.
- [219] A. Crane-Droesch, B. Kravitz, and J. T. Abatzoglou, *Using deep learning to model potential impacts of geoeengineering via solar radiation management on US agriculture*. In AGU Fall Meeting Abstracts, 2018.
- [220] "The state of the world's forests 2018 - forest pathways to sustainable development," <http://www.fao.org/policy-support/tools-and-publications/resources-details/en/c/1144279/>, accessed: 2020-12-05.
- [221] B. G. Weinstein, S. Marconi, S. Bohlman, A. Zare, and E. White, "Individual tree-crown detection in rgb imagery using semi-supervised deep learning neural networks," *Remote Sensing*, vol. 11, no. 11, p. 1309, Jun 2019.
- [222] T. Dong, Y. Shen, J. Zhang, Y. Ye, and J. Fan, "Progressive cascaded convolutional neural networks for single tree detection with google earth imagery," *Remote Sensing*, vol. 11, no. 15, p. 1786, Jul 2019.
- [223] Z. M. Hamdi, M. Brandmeier, and C. Straub, "Forest damage assessment using deep learning on high resolution remote sensing data," *Remote Sensing*, vol. 11, no. 17, p. 1976, Aug 2019.
- [224] "Planet: Understanding the amazon from space (2017)," <https://www.kaggle.com/c/planet-understanding-the-amazon-from-space/>, accessed: 2020-12-05.
- [225] S.-H. Lee, K.-J. Han, K. Lee, K.-J. Lee, K.-Y. Oh, and M.-J. Lee, "Classification of landscape affected by deforestation using high-resolution remote sensing data and deep-learning techniques," *Remote Sensing*, vol. 12, p. 3372, Oct 2020.
- [226] P. de Bem, O. de Carvalho Junior, R. Fontes Guimarães, and R. Trancoso Gomes, "Change detection of deforestation in the brazilian amazon using landsat data and convolutional neural networks," *Remote Sensing*, vol. 12, no. 6, p. 901, Mar 2020.
- [227] L. Bragilevsky and I. V. Bajić, "Deep learning for amazon satellite image analysis," in *2017 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, 2017, pp. 1–5.
- [228] R. V. Maretto, L. M. G. Fonseca, N. Jacobs, T. S. Körting, H. N. Bendini, and L. L. Parente, "Spatio-temporal deep learning approach to map deforestation in amazon rainforest," *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2020.
- [229] J. Irvin, H. Sheng, N. Ramachandran, S. Johnson-Yu, S. Zhou, K. Story, R. Rustowicz, C. Elsworth, K. Austin, and A. Y. Ng, "Forestnet: Classifying drivers of deforestation in indonesia using deep learning on satellite imagery," in *NeurIPS 2020 Workshop Tackling Climate Change with Machine Learning*, 12 2020.
- [230] K. Govil, M. L. Welch, J. T. Ball, and C. R. Pennypacker, "Preliminary results from a wildfire detection system using deep learning on remote camera images," *Remote Sensing*, vol. 12, no. 1, p. 166, Jan 2020.
- [231] Y. Zhao, J. Ma, X. Li, and J. Zhang, "Saliency detection and deep learning-based wildfire identification in uav imagery," *Sensors*, vol. 18, no. 3, p. 712, Feb 2018.
- [232] N. T. Toan, P. Thanh Cong, N. Q. Viet Hung, and J. Jo, "A deep learning approach for early wildfire detection from hyperspectral satellite images," in *2019 7th International Conference on Robot Intelligence Technology and Applications (RiTA)*, 2019, pp. 38–45.
- [233] Y. Ban, P. Zhang, A. Nascetti, A. R. Bevington, and M. A. Wulder, "Near real-time wildfire progression monitoring with sentinel-1 sar time series and deep learning," *Scientific Reports*, vol. 10, no. 1, p. 1322, Jan 2020.
- [234] D. Radke, A. Hessler, and D. Ellsworth, "Firecast: Leveraging deep learning to predict wildfire spread," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization, 7 2019, pp. 4575–4581.
- [235] J. R. Bergado, C. Persello, K. Reinke, and A. Stein, "Predicting wildfire burns from big geodata using deep learning," *Safety Science*, vol. 140, p. 105276, 8 2021.
- [236] A. Farasin, L. Colomba, and P. Garza, "Double-step u-net: A deep learning-based approach for the estimation of wildfire damage severity through sentinel-2 satellite data," *Applied Sciences*, vol. 10, no. 12, p. 4332, Jun 2020.
- [237] W. Zhou and L. Klein, "Monitoring the impact of wildfires on tree species with deep learning," in *NeurIPS 2020 Workshop Tackling Climate Change with Machine Learning*, 12 2020.
- [238] G. Camps-Valls, M. Campos-Taberner, A. Moreno-Martin, S. Walther, G. Duveiller, A. Cescatti, M. Mahecha, J. M. noz Marí, F. J. García-Haro, L. Guanter, J. Gamon, M. Jung, and S. W. R. Markus Reichstein, "A unified vegetation index for quantifying the terrestrial biosphere," *Science Advances*, vol. 7, p. eabc7447, 2021.
- [239] Q. Yuan, H. Shen, T. Li, Z. Li, S. Li, Y. Jiang, H. Xu, W. Tan, Q. Yang, J. Wang *et al.*, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sensing of Environment*, vol. 241, p. 111716, 2020.
- [240] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 8–36, 2017.

- [241] J. Tan, N. NourEldeen, K. Mao, J. Shi, Z. Li, T. Xu, and Z. Yuan, "Deep learning convolutional neural network for the retrieval of land surface temperature from amsr2 data in china," *Sensors*, vol. 19, no. 13, p. 2987, 2019.
- [242] J. Verrelst, G. Camps-Valls, J. Muñoz-Marí, J. P. Rivera, F. Veroustraete, J. G. Clevers, and J. Moreno, "Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties—a review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 108, pp. 273–290, 2015.
- [243] F. J. García-Haro, M. Campos-Taberner, A. Moreno, H. T. Tagesson, F. Camacho, B. Martínez, S. Sánchez, M. Piles, G. Camps-Valls, M. Yebra, and M. A. Gilabert, "A global Canopy Water Content product from AVHRR/Metop," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 161, no. 7, pp. 1–18, 2020.
- [244] A. Moreno-Martínez, G. Camps-Valls, J. Kattge, N. Robinson, M. Reichstein, P. van Bodegom, K. Kramer, J. Cornelissen, P. Reich, M. Bahn, U. Niinemets, J. Penuelas, J. Craine, B. Cerabolini, V. Minden, D. Laughlin, L. Sack, B. Allred, C. Baraloto, C. Byun, N. Soudzilovskaia, and S. Running, "A methodology to derive global maps of leaf traits using remote sensing and climate data," *Remote Sensing of Environment*, vol. 218, no. 12, pp. 69–88, 2018.
- [245] M. Jung, S. Koirala, U. Weber, K. Ichii, F. Gans, G. Camps-Valls, D. Papale, C. Schwalm, G. Tramontana, and M. Reichstein, "The FLUXCOM ensemble of global land-atmosphere energy fluxes," *Scientific Data*, 2019.
- [246] M. Jung, C. Schwalm, M. Migliavacca, S. Walther, G. Camps-Valls, S. Koirala, P. Anthoni, S. Besnard, P. Bodesheim, N. Carvalhais, F. Chevallier, F. Gans, D. S. Groll, V. Haverd, K. Ichii, A. K. Jain, J. Liu, D. Lombardozzi, J. E. M. S. Nabel, J. A. Nelson, M. Pallandt, D. Papale, W. Peters, J. Pongratz, C. Rödenbeck, S. Sitch, G. Tramontana, U. Weber, M. Reichstein, P. Koehler, M. O'Sullivan, and A. Walker, "Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the FLUXCOM approach," *Biogeosciences*, vol. 17, pp. 1343–1365, 2020.
- [247] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.
- [248] K. Kuwata and R. Shibasaki, "Estimating corn yield in the united states with modis evi and machine learning methods," *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 3, no. 8, 2016.
- [249] A. Wolanin, G. Mateo-Garcia, G. Camps-Valls, L. Gomez-Chova, M. Meroni, G. Duveiller, Y. Liangzhi, and L. Guanter, "Estimating and understanding crop yields with explainable deep learning in the indian wheat belt," *Environmental Research Letters*, vol. 15, no. 2, pp. 1–12, 2020.
- [250] S. K. Kim, S. Ames, J. Lee, C. Zhang, A. C. Wilson, and D. Williams, "Massive scale deep learning for detecting extreme climate events," *Climate Informatics*, 2017.
- [251] L. Ye, L. Gao, R. Marcos-Martínez, D. Mallants, and B. A. Bryan, "Projecting australia's forest cover dynamics and exploring influential factors using deep learning," *Environmental Modelling & Software*, vol. 119, pp. 407–417, 2019.
- [252] S. H. Khan, X. He, F. Porikli, and M. Bennamoun, "Forest change detection in incomplete satellite images with deep neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 9, pp. 5407–5423, 2017.
- [253] L. Zhang, Z. Shao, J. Liu, and Q. Cheng, "Deep learning based retrieval of forest aboveground biomass from combined lidar and landsat 8 data," *Remote Sensing*, vol. 11, no. 12, p. 1459, 2019.
- [254] N. Lang, K. Schindler, and J. D. Wegner, "Country-wide high-resolution vegetation height mapping with Sentinel-2," *Remote Sensing of Environment*, vol. 233, 2019.
- [255] G. Ceballos, P. R. Ehrlich, and P. H. Raven, "Vertebrates on the brink as indicators of biological annihilation and the sixth mass extinction," *Proc. Natl. Acad. Sci.*, vol. 117, no. 24, pp. 13 596–13 602, 2020.
- [256] H. S. Young, D. J. McCauley, Galetti, M., and R. Dirzo, "Patterns, causes, and consequences of anthropocene defaunation," *Annu. Rev. Ecol. Evol. Syst.*, vol. 47, pp. 333–358, 2016.
- [257] I. Duporge, O. Isupova, S. Reece, D. W. Macdonald, and T. Wang, "Using very-high-resolution satellite imagery and deep learning to detect and count african elephants in heterogeneous landscapes," *Remote Sensing in Ecology and Conservation*, 2020.
- [258] N. Rey, M. Volpi, S. Joost, and D. Tuia, "Detecting animals in african savanna with UAVs and the crowds," *Remote Sensing of Environment*, vol. 200, pp. 341–351, 2017.
- [259] J. Linchant, J. Lisein, J. Semeki, P. Lejeune, and C. Vermeulen, "Are unmanned aircraft systems (UASs) the future of wildlife monitoring? a review of accomplishments and challenges," *Mammal review*, vol. 45, pp. 239–252, 2015.
- [260] P. Fust and Loos, "Development perspectives for the application of autonomous, unmanned aerial systems (uass) in wildlife conservation," *J. Biol. Conserv.*, vol. 241, p. 108380, 2020.
- [261] B. Kellenberger, D. Marcos, and D. Tuia, "Detecting mammals in UAV images: Best practices to address a substantially imbalanced dataset with deep learning," *Remote Sens. Environ.*, vol. 216, pp. 139–153, 2018.
- [262] M. S. Norouzzadeh, A. Nguyen, M. Kosmala, A. Swanson, M. S. Palmer, C. Packer, and J. Clune, "Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning," *Proceedings of the National Academy of Sciences*, vol. 115, no. 25, pp. E5716–E5725, 2018.
- [263] J. A. J. Eikelboom, J. Wind, E. van de Ven, L. M. Kenana, B. Schroder, H. J. de Knegt, F. van Langevelde, and H. H. T. Prins, "Improving the precision and accuracy of animal population estimates with aerial image object detection," *Methods Ecol. Evol.*, vol. 10, no. 11, pp. 1875–1887, 2019.
- [264] J. Peng, D. Wang, X. Liao, Q. Shao, Z. Sun, H. Yue, and H. Ye, "Wild animal survey using UAS imagery and deep learning: modified faster R-CNN for kiang detection in tibetan plateau," *ISPRS J. Photo. Remote Sens.*, vol. 169, pp. 364–376, 2020.
- [265] S. Beery, D. Morris, and S. Yang, "Efficient pipeline for camera trap image review," *arXiv preprint arXiv:1907.06772*, 2019.
- [266] B. Kellenberger, D. Tuia, and D. Morris, "AIDE: Accelerating image-based ecological surveys with artificial intelligence," *Methods Ecol. Evol.*, vol. 11, no. 12, pp. 1716–1727, 2020.
- [267] L. D. Brown and D. Schormann, "Poacher detection and wildlife counting system," in *Proc. SATNAC*, 2019.
- [268] Y. Wang, Z. R. Shi, L. Yu, R. Singh, L. Joppa, and F. Fang, "Deep reinforcement learning for green security games with real-time information," in *Proc. AAAI*, 2019.
- [269] C. J. Torney, D. Lloyd-Jones, M. Chevallier, D. C. Mozer, H. T. Maliti, M. Mwitwa, E. M. Kohi, and G. C. Hopcraft, "A comparison of deep learning and citizen science techniques for counting wildlife in aerial survey images," *Meth. Ecol. Evol.*, vol. 10, no. 6, pp. 779–787, 2019.
- [270] B. Settles, *Active Learning*. Morgan and Claypool Publishers, 2012.
- [271] B. Kellenberger, D. Marcos, S. Lobry, and D. Tuia, "Half a percent of labels is enough: Efficient animal detection in UAV imagery using deep CNNs and active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9524–9533, 2019.
- [272] A. Kendall and Y. Gal, "What uncertainties do we need in bayesian deep learning for computer vision?" in *Advances in neural information processing systems*, 2017, pp. 5574–5584.
- [273] S. Chakraborty, R. Tomsett, R. Raghavendra, D. Harborne, M. Alzantot, F. Cerutti, M. Srivastava, A. Preece, S. Julier, R. M. Rao, T. D. Kelley, D. Braines, M. Sensoy, C. J. Willis, and P. Gurrarn, "Interpretability of deep learning models: A survey of results," in *2017 IEEE SmartWorld Ubiquitous Intelligence and Computing, Advanced and Trusted Computed, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation, SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI 2017*. IEEE, 2018, pp. 1–6.
- [274] W. Samek, G. Montavon, S. Lapuschkin, C. J. Anders, and K. R. Müller, "Toward interpretable machine learning: Transparent deep neural networks and beyond," *arXiv*, pp. 1–24, 2020.
- [275] A. Levering, D. Marcos, and D. Tuia, "On the relation between landscape beauty and land cover: a case study in the U.K. at Sentinel-2 resolution with interpretable AI," *ISPRS J. Int. Soc. Photo. Remote Sens.*, vol. 177, pp. 194–203, 2021.
- [276] D. Marcos, S. Lobry, R. Fong, N. Courty, R. Flamary, and D. Tuia, "Contextual semantic interpretability," in *Asian Conference on Computer Vision (ACCV)*, Kyoto, Kapan, 2020.
- [277] D. Tuia, C. Persello, and L. Bruzzone, "Domain adaptation for the classification of remote sensing data: An overview of recent advances," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 41–57, 2016.
- [278] C. Doldirina, "Open data and earth observations the case of opening up access to and use of earth observation data through the global earth observation system of systems," *JIPITEC*, vol. 6, no. 1, pp. 73–85, 2015.
- [279] S. Stall, L. Yarmey, J. Cutcher-Gershenfeld, B. Hanson, K. Lehnert, B. Nosek, M. Parsons, E. Robinson, and L. Wyborn, "Make scientific data fair," *Nature*, vol. 570, no. 7759, pp. 27–29, 2019.

- [280] S. David, "Perspectives on free and open source software. joseph feller, brian fitzgerald, scott a. hissam, karim r. lakhani (eds.)," *Journal of Information Technology & Politics*, vol. 4, no. 3, pp. 97–99, 2008.
- [281] G. Boulton, "The challenges of a big data earth," *Big Earth Data*, vol. 2, no. 1, pp. 1–7, 2018.
- [282] C. E. Woodcock, R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, S. N. Goward, D. Helder, E. Helmer, R. Nemani, L. Oreopoulos, J. Schott, P. S. Thenkabail, E. F. Vermote, J. Vogelmann, M. A. Wulder, and R. Wynne, "Free access to landsat imagery," *Science*, vol. 320, no. 5879, pp. 1011–1011, 2008.
- [283] T. Dhu, B. Dunn, B. Lewis, L. Lymburner, N. Mueller, E. Telfer, A. Lewis, A. McIntyre, S. Minchin, and C. Phillips, "Digital earth australia – unlocking new value from earth observation data," *Big Earth Data*, vol. 1, no. 1-2, pp. 64–74, 2017.
- [284] M. Mahecha, F. Gans, G. Brandt, R. Christiansen, S. Cornell, N. Fomferra, G. Kraemer, J. Peters, G. Camps-Valls, J. Donges, W. Dorigo, L. Estupinan-Suarez, V. Gutierrez-Velez, M. Gutwin, M. Jung, M. Londono, D. Miralles, P. Papastefanou, , and M. Reichstein, "Earth system data cubes unravel global multivariate dynamics," *Earth System Dynamics*, vol. 11, pp. 201–234, Feb 2020.
- [285] G. Giuliani, B. Chatenoux, A. D. Bono, D. Rodila, J.-P. Richard, K. Allenbach, H. Dao, and P. Peduzzi, "Building an earth observations data cube: lessons learned from the swiss data cube (sdc) on generating analysis ready data (ard)," *Big Earth Data*, vol. 1, no. 1-2, pp. 100–117, 2017.
- [286] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google earth engine: Planetary-scale geospatial analysis for everyone," *Remote Sensing of Environment*, vol. 202, pp. 18 – 27, 2017.
- [287] M. W. Scown, "The Sustainable Development Goals need geoscience," *Nature Geoscience*, vol. 13, no. 11, pp. 714–715, 10 2020.