A Meta-Analysis of Convolutional Neural Networks for Remote Sensing Applications

Hamid Ghanbari [ID], Masoud Mahdianpari [ID], *Member, IEEE*, Saeid Homayouni [ID], *Senior Member, IEEE*, and Fariba Mohammadimanesh

Abstract—Since the rise of deep learning in the past few years, convolutional neural networks (CNNs) have quickly found their place within the remote sensing (RS) community. As a result, they have transitioned away from other machine learning techniques, achieving unprecedented improvements in many specific RS applications. This article presents a meta-analysis of 416 peer-reviewed journal articles, summarizes CNN advancements, and its current status under RS applications. The review process includes a statistical and descriptive analysis of a database comprised of 23 fields, including: 1) general characteristics, such as various applications, study objectives, sensors, and data types, and 2) algorithm specifications, such as different types of CNN models, parameter settings, and reported accuracies. This review begins with a comprehensive survey of the relevant articles without considering any specific criteria to give readers an idea of general trends, and then investigates CNNs within different RS applications to provide specific directions for the researchers. Finally, a conclusion summarizes potentialities, critical issues, and challenges related to the observed trends.

Index Terms—Convolutional neural network (CNN), deep learning (DL), meta-analysis, remote sensing (RS).

I. INTRODUCTION

MACHINE learning (ML) is a subset of artificial intelligence that has been used in various applications to train a system using provided data [1]. The ML paradigm covers numerous areas of study, through which a variety of algorithms have been introduced [2]. In the past decade, various ML methods and computational algorithms have been developed to analyze remote sensing (RS) data [3]. In a survey of the ISI Web-of-Science (WoS) database, Scheunders *et al.* [2] report that between 2004–2015 about 60 000 papers applying ML in an RS context were published, of which 10 000 applied classification algorithms and 3000 applied regression models. Given the recent developments in RS data acquisition technologies and

Hamid Ghanbari is with the Department of Geography, Université Laval, Québec, QC G1V 0A6, Canada (e-mail: hamid.ghanbari.1@ulaval.ca).

Masoud Mahdianpari is with the C-CORE, St. John's, NL A1B 3X5, Canada, and also with the Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada (e-mail: m.mahdianpari@mun.ca).

Saeid Homayouni is with the Institut National de la Recherche Scientifique, Centre Eau Terre Environnement, Quebec City, QC G1K 9A9, Canada (e-mail: saeid.homayouni@ete.inrs.ca).

Fariba Mohammadimanesh is with the C-CORE, St. John's, NL A1B 3X5, Canada (e-mail: fm7550@mun.ca).

the rising diversity of objectives capable of being resolved, the use of ML in RS applications is expected to increase [4].

ML methods in the context of RS cover a vast range of applications, including land use and land cover (LULC) classification, change detection, object detection, feature selection, and extraction, etc. [2]. LULC, in particular, benefits from advancements in ML methods [5]. The growing number of satellite platforms with various revisit times has increased the ability to capture nature accurately, and human-made changes to the Earth's surface [6], and ML methods have increasingly been used to address related change detection problems [7]. Similarly, the development of high spatial resolution instruments installed on airborne and spaceborne platforms has resulted in an increase in applications of ML for special object detection [8]. ML also plays a vital role in dimensionality reduction [9] of hyperspectral images composed of many essential features for several scientific applications [10]. Other roles ML plays in RS applications include spectral unmixing, regression, image fusion, etc. [1]. Within each category, several ML algorithms have been introduced based on sensor type, study objective, ancillary data, and limitations such as spatial resolution and training sample size [11]. The applicability and effectiveness of these algorithms have been demonstrated in many geosciences and RS tasks [12], [13].

The most commonly used ML algorithms in RS are artificial neural networks (ANN), support vector machines (SVM), decision trees (DT), and ensemble methods, such as random forest (RF) [14]. Each of these methods carries specific advantages. For example, SVM can best tackle high-dimensionality problems and limited training data [15], while RF does not require the fine-tuning of a large number of hyperparameters and can easily be used for both simple and complex computations [16], [17]. These two methods share the advantage of lower computational complexity and higher interpretability capabilities [14].

Over the past few years, however, there has been an ongoing shift toward using deep learning (DL) methods in ML applications [18]. DL, which is characterized by neural networks (NN), is the fastest-growing trend in big RS data analysis and is regarded as a breakthrough technology [19]–[22]. DL has been used in many areas of research, such as speech recognition [23], stereo vision [24], biomedicine [25], time-series analysis [26], agriculture [27], and medical image recognition [28]. Although DL has the disadvantages of: 1) being a "black box" naturally, which mitigates its interpretability, and 2) requiring greater amounts of training samples compared to other ML methods, it

has become a hotspot in the realm of ML and has been approved by many researchers in the geoscience and RS community [29].

To date, several DL architectures have been introduced, of which the stacked autoencoder (SAE), convolutional NN (CNN), generative adversarial network (GAN), deep belief network (DBN), and recurrent NN (RNN) have become mainstream [30]. Of these DL networks, CNN is the most popular and the most published [31]. Along with the development of DL methods, CNNs have emerged as an incredibly powerful tool by providing both remarkable performances in image processing and the ability to work in a wide variety of applications in the vision community [32]. In the past few years, biologically inspired CNNs have emerged and proven to be effective in a diverse range of fields to which image processing is fundamental, from social media [33] to precision medicine [34] and robotics [35].

A particularly beneficial characteristic of CNNs is data processing in multiple arrays and automatic feature extraction ability, which has received acknowledgment in the RS community [18], [36]. Moreover, the inherent characteristics of CNNs, such as local connectivity and weight sharing, allow this DL method to tackle the drawbacks of artificial feature extraction by considering the 2-D structures and reducing network parameters using convolutional filters [32]. CNN-based approaches have benefited from the recent exponential increase in RS technologies that includes various image types (optical, RADAR, temperature and microwave radiometer, altimeter, etc.) with complex characteristics (high dimensionality, multiple scales, and nonstationary) [37].

CNNs are composed of a set of blocks that make them particularly suitable for image analysis. The multiple layers of operations, such as convolution, pooling, and nonlinear activation functions, allow for the hierarchical extraction of high-level abstract features [31], [38], [39]. Therefore, CNNs have been successfully used in image preprocessing, scene classification, pixel-based classification and image segmentation, and object detection [40]–[44]. For example, CNNs have been used in numerous studies to improve image classification results [36], [46]–[48], to extract buildings and nonbuilding regions automatically [49], and to detect areas of build-up [50]. Scarpa *et al.* [51] proposed and analyzed CNN-based methods to estimate spectral features when optical data are missing. In another example, a CNN regression was proposed to develop a model applicable to hyperspectral imagery for estimation of concentrations of phycocyanin and chlorophyll-a [52]. CNNs also have been used in OpenStreetMap data quality assessment [53], oil spill segmentation [54], ship position detection and direction prediction [55], multimodal RS image registration [56], road extraction [57], and many other areas of study [58]–[60].

Several related review papers have been published due to DL methods' significant performance over other state-of-the-art methods in RS. For example, a review by Ball *et al.* [31] focused on the theories, tools, and challenges of using DL algorithms in the RS community. In another review, Zhang *et al.* [45] and Zhu *et al.* [19] summarized recent advances in DL methods and discussed related challenges in RS applications. Following the explosive growth of new DL methods in different

RS applications and their striking achievements, some review papers have focused on studying the task-based reports of DL methods [61]–[65]. Liu *et al.* [66] presented a systematic review of the application of DL techniques in the field of pixel-based image fusion. A recent comprehensive review by Tsagkatakis *et al.* [30] was conducted on RS image enhancement, including super-resolution, denoising, restoration, pan-sharpening, and fusion.

Most of the existing review manuscripts covering major DL concepts related to RS applications consider all DL architectures. Early efforts reviewing CNNs generally were performed by Rawat and Wang [67] in which the authors focused on the application of CNNs in image classification tasks and debated their rapid advancement in recent years and the contribution CNNs had made to DL developments. Many other articles provide readers with a summary of the CNNs' basic concepts in different applications such as radiology [68], biology [69], and action recognition [70].

The majority of the CNN review papers are descriptive, often with no quantitative assessment, and tend to focus on applications other than RS. Accordingly, this study's main objective is to describe and discuss the RS-based applications of CNN through a meta-analysis of published papers and to provide RS experts with a "big picture" summary of current research in this field. As a whole, the contributions of this article are: 1) whereas almost all the review papers in RS applications cover all related DL network structures, this article reviews the publications dedicated to the use of CNNs for RS applications alone and summarizes hotspots based on the paper frequency and accuracy. Moreover, it discusses trends and specific setups for different subtasks, and 2) this review defines a complete framework for professionals and even nonexperts, outlining ongoing research and the architectures that receive the most attention in each application.

To fulfill the proposed meta-analysis task and to construct a database of case studies, more than 400 peer-reviewed articles are reviewed, and many other papers are cited. To create a context for what follows, we first summarize the performed systematic literature search query using the preferred reporting items for systematic reviews and meta-analyses (PRISMA) in Section II. After presenting the general characteristics of CNNs in Section III, the application of CNNs in different study objective is discussed in Section IV. Finally, in Section V, concluding remarks are presented.

II. METHODS

A systematic literature search query was performed using the WoS to identify relevant articles for this comprehensive review. The WoS is one of the biggest bibliographic databases covering scholarly literature from approximately any discipline. Notably, the PRISMA methodology was followed for study selection [71]. After some trials, a title/abstract/keyword search was performed in WoS using a search query of: "convolutional neural network∗" OR "CNN" OR "FCNN" OR "fully convolutional neural network∗" OR "deep learning" for the title and "Remote sensing" for abstract/keyword, checking to include papers that

TABLE I
RS JOURNALS USED TO COLLECT RESEARCH STUDIES FOR THIS LITERATURE REVIEW

Journal	Publication	Impact factor (2019)	# papers
Remote Sensing	MDPI	4.118	107
Transaction on Geoscience and Remote Sensing	IEEE	5.63	36
Geoscience and Remote Sensing Letters	IEEE	3.534	37
ISPRS Journal of Photogrammetry and Remote Sensing	Elsevier	6.942	28
Journal of Selected Topics in Applied Earth Observations and Remote Sensing	IEEE	3.392	24
International Journal of Remote Sensing	Taylor & Francis	2.493	21
Journal of Applied Remote Sensing	SPIE	1.344	20
IEEE ACCESS	IEEE	4.098	16
SENSORS	MDPI	3.510	16
Remote Sensing Letters	Taylor & Francis	2.024	14
ISPRS International Journal of Geo-Information	MDPI	2.350	9
Remote Sensing of Environment	Elsevier	8.218	7

used data from the most common RS platforms (search date: June 12, 2020). This research resulted in 1038 papers, which served as the basis for further paper surveys.

Of the 1038 initial number of studies, 664 papers were related to the peer-reviewed journals, and the remaining were majority proceedings papers. When investigating the 664 Journal papers in detail and after eligibility assessment, 248 papers were determined unrelated to this meta-analysis and removed. The journals with more than five papers are listed in Table I.

The 416 eligible papers were included in the meta-analysis and described using the following—title, year, journal, citation, first author's institution's country, RS data, study type, application, CNN model, processing unit, training sample, area of the data, spatial resolution, geographical coverage, framework, learning strategy, number of layers, dataset, CPU/GPU run, processing time, convolutional kernel dimension, accuracy, and accuracy metric. A summary of the literature search is demonstrated in Fig. 1.

III. RESULTS AND DISCUSSION

A. General Characteristics of Studies

There has been a steep upward trend in the use of CNNs from 2014, the point at which the first RS-related application of CNNs was introduced [72] (see Fig. 2). The exponential trend of annual publication frequency peaks in 2019 and includes more than one-third of the database articles. The expansion in the use of CNNs continued in the current year (2020). In the first half of 2020, the number of papers published exceeds the number of published papers for the equivalent period in 2019.

Studies were conducted in 34 different countries and regions from six continents, the majority of which are based in Asia (72%), Europe (15%), and North America (9%). Among the different countries, most of the contributions were carried out
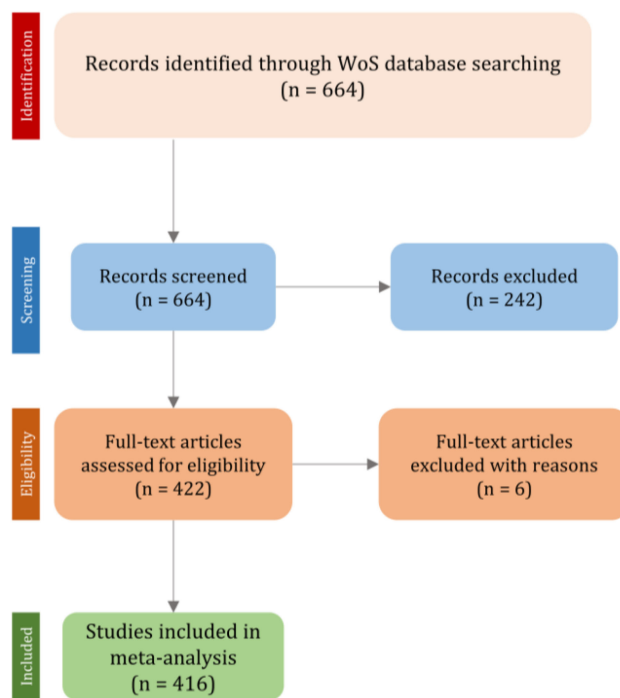

Fig. 1. PRISMA flowchart demonstrating the selection of studies.
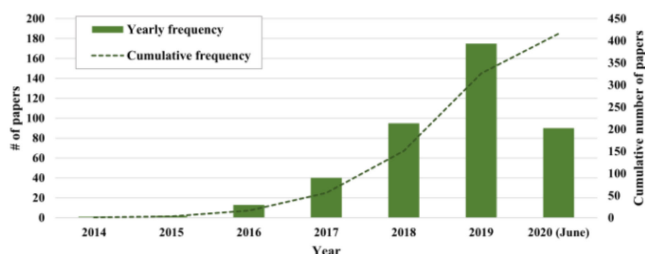

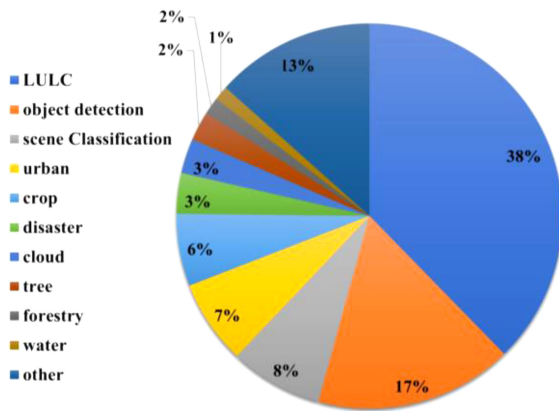Fig. 2. Annual and cumulative frequency of CNN studies in RS applications.

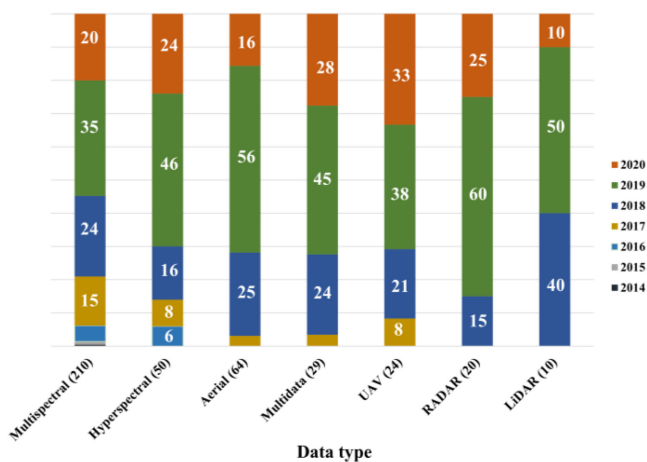Fig. 3.   Number of studies per RS application.

Fig. 4.   Usage rate of different RS data (the numbers in parenthesis show the number of studies).

in China, with about 63% of total studies, followed by the USA (7%) and Germany (4%). Further analysis revealed that only six countries had published more than ten papers and 13 countries contributed only one paper.

Fig. 4 demonstrates the usage frequency of CNNs in a vast range of applications. Of the reviewed 416 articles, LULC was the most frequented application (with about 155 studies), followed by object detection, scene classification, and urban studies with 68, 32, and 25 studies, respectively. The remaining articles include specific applications in crop- (25), disaster- (14), cloud- (12), tree- (10), forestry- (6), and water- (5) related research. Some other applications also benefited from CNNs, clustered in Fig. 3 as "other" as the number of associated articles was fewer than five, including sea ice, agriculture, and wetland mapping (see Fig. 3).

B. Sensors and Data Types

The first published RS-related work using CNNs occurred in 2014 for vehicle detection using multispectral satellite images [72]. Since then, CNNs have been applied in research using numerous RS data types (see Fig. 4). The largest share of this research has used multispectral satellite images, making up 51% of the database, mainly Landsat archives, Worldview-4, and Quickbird-2 imageries. Since early 2016, CNNs have been used for hyperspectral data analysis in about 12% of the studies. Through the development of CNN architectures and their achievements in research using different types of data, by 2017, CNNs were increasingly used to analyze other data types, including aerial (16%), unmanned aerial vehicle (UAV) (6%), RADAR (5%), and light detection and ranging (LiDAR) (2%). Almost 7% of the studies have also used a combination of different data types.

C. CNN Frameworks and Models

The DL community's framework and library development is highly dynamic and offers different possibilities to speed up the training process with interactive interfaces [73]. A graphical representation of the most used frameworks and libraries and their annual publication frequency in CNN studies is shown in Fig. 5. The presented libraries provide built-in classes of NNs, fast numerical computation, and automated gradients for both CPU and GPU. These libraries include TensorFlow (Google Brain team), MATLAB (MathWorks), Caffe (Berkeley Artificial Intelligence Research), Torch (Berkeley Software Distribution), MXNET (University of Washington), and Theano (Berkeley Software Distribution). A detailed search through the literature revealed that TensorFlow is the most significant CNN implementation source, with about 43% of total usages. TensorFlow is a free and open-source end-to-end DL framework for numerical computations using data-flow graphs [74]. It is designed to be highly portable, running on various platform scales, from a single CPU to a GPU or GPUs cluster [75].

Built on top of TensorFlow, Keras library has been employed for CNN implementation as well. Keras, which supports almost all models of CNNs, was executed on both CPU and GPU. The second most used framework is Caffe (22% of studies). Caffe is well suited for machine vision and forecasting applications, which permits a network with sophisticated configurations [76]. Caffe's specific properties include suitability for image processing tasks with CNNs, accessibility to pretrained networks, and easy coding on Python and MATLAB [73], [77]. MATLAB, the third most used platform for CNN implementation, was utilized in almost 19% of the studies. MATLAB's beneficial characteristics are its simplicity, especially for practitioners, various visualizing tools, and the capability of deploying models on a variety of servers and devices [78]. The remaining libraries include Torch (10%), Theano (5%), and MXNET (1%). A comparison of the frameworks' annual usage frequency shows that TensorFlow and Torch have the steepest growth rates since they were introduced for CNN-based RS applications, followed by MATLAB with a lower rate of increase. On the other hand, Theano reached its peak in 2018, falling in numbers with no cases thus far in 2020. Caffe, which was used first in 2015, experienced a steep increasing number of usages in 2017 and afterward held its position in CNN implementations.

To date, different CNN models have been developed starting from 1989 with the introduction of the first CNN-based feature
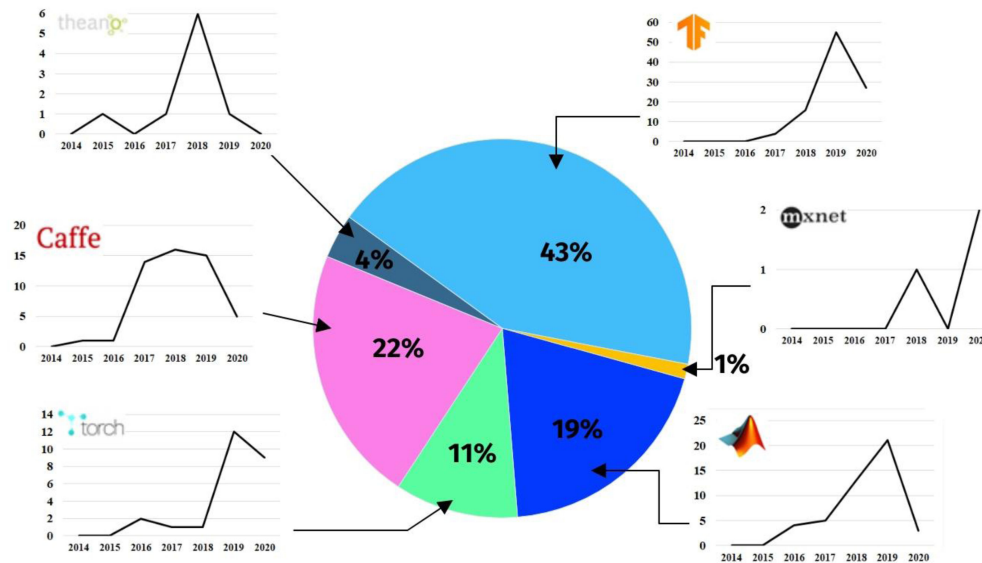
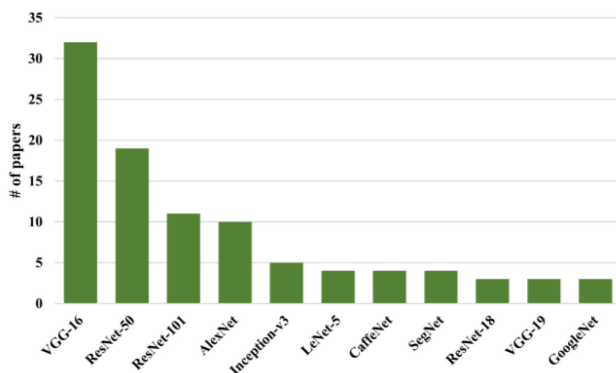Fig. 5.   Frequency of the used CNN frameworks and libraries.

Fig. 6.   Frequency of different CNN models in the article database.
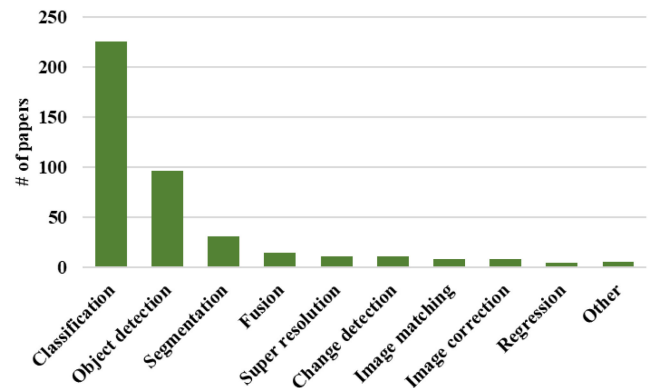

Fig. 7.   Distribution of study objectives.

encoder, LeNet [79]. Several factors have contributed to the development of new CNN models, including efficient training on modern processing units, introduction of the rectified linear unit (ReLU) activation function, availability of new datasets, and innovation related to depth and spatial inclusion [80]–[82]. Fig. 6 shows the CNN models with three or more usage frequency in different problems as a backbone architecture or a starting point with the pretrained parameters (i.e., transfer learning). The most frequently used model, employed in 33% of studies, was VGG-16 [83], followed by ResNet-50 (19%), ResNet-101 (11%) [84], AlexNet (10%) [81], and Inception-V3 (5%) [85]. The remaining models, comprising about 22% of the case studies, mainly consist of LeNet-5, CaffeNet [77], SegNet [86], ResNet-18, VGG-19, and GoogleNet [87].

IV. CNNs for Different Study Objectives

A survey among RS-related studies revealed that CNNs could be applied to almost any significant RS task, making them a promising option for handling various problems. Fig. 7 shows that 54% of the reviewed papers (225 papers) are devoted to classification problems, including LULC and scene classification

studies. CNNs have been used in several other study objectives, including object detection (23%), image segmentation (7%), data fusion (4%), image super-resolution (3%), image matching (2%), image correction (2%), and regression (1%). Recently, CNNs have been applied to other study objectives, such as image retrieval, prediction, quality assessment, and unmixing, making up about 4% of the studies.

Following the review of the publications in different research domains, an in-depth review of different study types and their respective findings are provided in the following sections. Based on the proximity of the research domains and their paper frequency, papers were categorized into three distinctive groups (i.e., image classification, object detection and segmentation, and others).

A. Classification

The number of classification-based studies focusing on different applications (the number of papers is shown in the parenthesis) alongside the statistical analyses reported by overall accuracy (OA) is shown in Fig. 8. Most of the classification tasks focused on LULC, scene classification, and crop-, urban-,
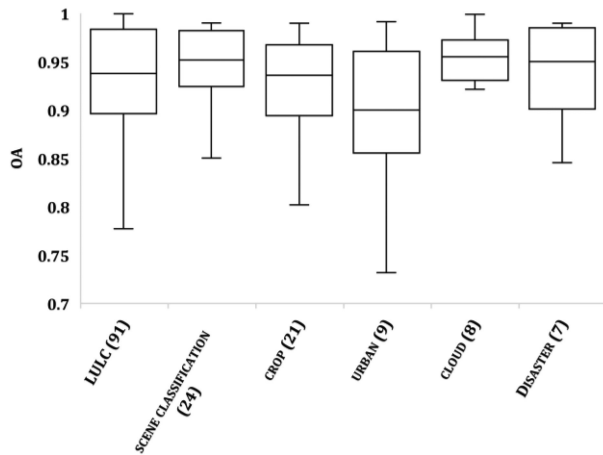
Fig. 8. OA distribution of CNN classification studies based on different applications.

Fig. 9. Distribution of OA for different data types.

Fig. 10. Frequency and average OA of CNN classification by spatial resolution.

cloud-, and disaster-classification, with about 52%, 14%, 12%, 5%, 5%, and 4%, respectively. Other classification tasks, including agriculture, forestry, wetland, and sea ice, comprise about 8% of the total cases but are not shown in Fig. 8. The classification accuracy assessment shows the maximum average accuracy and the lowest variability for cloud studies, with almost 95.5%, followed by scene classification with about 95.2%. Urban studies have the lowest average accuracy among the applications, with an average OA of 90%.

As mentioned earlier, many CNN classification studies are devoted to LULC applications, potentially because of the extensive scope of relevant datasets available for training the networks. The papers in this group mostly focused on hyperspectral image classification by using benchmark datasets, including Salinas [airborne visible/infrared imaging spectrometer (AVIRIS) sensor], Pavia University [reflective optics system imaging spectrometer (ROSIS-03) sensor], and Indian Pines (AVIRIS sensor) [88]. The second most studied classification task is scene classification, which is generally defined as a procedure to categorize a specific scene theme, e.g., a part of a forest, an agricultural landscape, a river, etc. [89]. A majority of these studies apply high-resolution RS images because of the availability of many large-scale high-resolution datasets in recent years [90], [91]. In scene classification, the most commonly used datasets are the UC-Merced dataset [92], the aerial image data set (AID) [93], and NWPU-RESISC 45 (RS image scene classification) datasets [94].

A detailed survey of the article database shows that about 46% of classification studies used spaceborne multispectral RS images for classification tasks (see Fig. 9). The most frequent multispectral satellite data sources are Landsat 8, Gaofen 1-2, and Sentinel 2. The least used data type is related to LiDAR datasets, making up about 2% of the whole database. The remaining types of RS data are aerial images (18%), hyperspectral (13%), UAV (7%), multidata (7%), and RADAR (6%). For the aerial images, the Vaihingen Semantic Labelling dataset [95] and Potsdam Semantic Labelling dataset [96] were the most used datasets, and for hyperspectral images, the most used
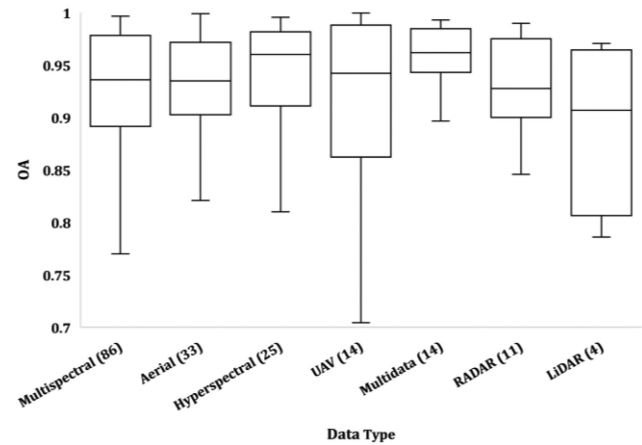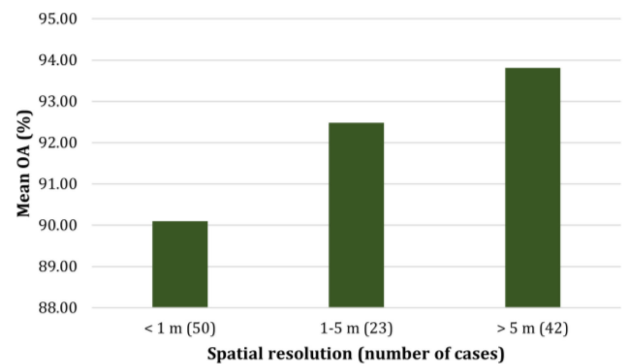
datasets were the Salinas, Indian Pines, and Pavia University datasets. In recent years, UAV data have been deployed in some CNN classification tasks, and it is expected that their usage will increase because the cost of UAV has lowered in recent years [97]. The capability of automatic data acquisition using UAV has made them a convenient tool for some classification tasks such as geological mapping [97], crop yield prediction [98], and wetland mapping [99]. Integration of different RS data types (i.e., multidata) is beneficial for different tasks such as tree species diversity mapping using LiDAR and high-resolution multispectral images [100], soybean yield prediction by fusion of weather data and MODIS products [101], and coastal land cover classification by integration of optical and RADAR satellite images [102].

A statistical analysis of different data types showed that using multidata resulted in maximum average OA (96.2%) and lower variability. In the case of single-data research, the mean classification accuracy of hyperspectral datasets is the highest at 96%, followed by UAV (94.20%), multispectral (93.58%), aerial (93.48%), RADAR (92.75%), and LiDAR (90.67%).

Fig. 10 shows the average obtained accuracy for CNN classification based on the spatial resolution of the remotely sensed image dataset and their respective number of published papers (the number of papers is written in the parenthesis). The papers were categorized based on the spatial resolution into very high

(<1 m), high (between 1 and 5 m), and medium and low resolution (>5 m).

Data with a spatial resolution of <5 m (very high and high) were used in 63% of the publications, mainly composed of studies using 1 m resolution. Based on this analysis, it can be concluded that there is an agreement between the mean accuracy of the classification and the spatial resolution. The mean accuracy for very high- and high-resolution datasets is 90.10% and 92.48%, respectively. The maximum mean accuracy is related to datasets with spatial resolutions > 5 m, with about 93.82% mean OA.

A survey among the usage frequency of different CNN models for classification tasks revealed that VGG variants were the most frequented backbones, followed by ResNet-50 and AlexNet. In about 52% of the studies, stochastic gradient descent (SGD) was used for parameter optimization of the CNN models [89], [103], while the remaining studies used adaptive moment estimation (Adam) optimizer [104]. Moreover, the training process was conducted over 100 epochs in about 68% of the cases.

B. Object Detection and Image Segmentation

As shown in Fig. 7, about 30% of CNN's RS studies involve image segmentation and object detection tasks. A close inspection of the usage frequency of different sensor types in image segmentation-based studies shows that the largest share use multispectral satellite images (60%), aerial (18%), and multidata (8%). The other articles were devoted to UAV, RADAR, LiDAR, and panchromatic images with almost 7%, 4%, 2%, and 1%, respectively. Open datasets and their respective applications significantly influenced these topics by providing information related to various land covers. A survey among the article database showed that two publicly available datasets, including Potsdam Semantic Labelling [96] and NWPU VHR-10 [105], established themselves as baseline datasets for CNN-based image segmentation and object detection. They are followed by DOTA (dataset for object detection in aerial images) [105], UC-Merced [92], and the Massachusetts buildings and roads datasets [106]. However, to investigate method development in large-scale research areas, some studies used custom spaceborne datasets. Google Earth was the most employed spaceborne data source, followed by Gaofen 1-2, Quickbird-2, and Worldview-4. Of all the reviewed papers for object detection and segmentation tasks, 68% of the CNN models are applied to analyze datasets with 1 m or finer spatial resolutions. SGD was used as an optimizer in about 75% of the studies, while Adam optimizer was used in the remaining 25%. In 72% of the cases, the number of epochs was set less than 100 iterations, and in the remaining cases (i.e., 28%), the number of epochs was more than 100.

An overview of the popular designs in the publications showed a focus on region-based CNN (R-CNN)-inspired architectures [107] and its series of improvements, including fast R-CNN [108], faster R-CNN [109], and mask R-CNN [110]. The most frequently used model, representing 39% of studies, was related to land cover mapping, with other categorial applications, including agriculture (15%), urban (11%), forest (10%), wetland (12%), disaster (3%), and soil (2%). The remaining applications, comprising about 8% of the case studies, mainly consist of mining area classification, water mapping, benthic habitat, rock types, and geology mapping.

Concerning the different architectures designed for image segmentation and object detection, the VGG variants have been the most used backbone models (34%), followed by the ResNet family (30%). The Inception, SegNet, LeNet, and GoogleNet backbones were much less commonly used.

C. Other Applications

Along with the above-mentioned conventional applications in RS, CNNs have also been applied in other research areas, such as data fusion, super-resolution, change detection, image registration, etc. Because of their specific capability of feature extraction and learning, CNNs demonstrate an outstanding possibility to delineate the relationship between different data, which has been used in panchromatic/multispectral data fusion applications [111]–[113]. The first papers that were introduced in this category were motivated by the impressive performance of CNNs in a large number of closely related super-resolution problems [113], [114]. A comprehensive overview of the article database's fusion studies shows a trend of using pixel-based processing units with residual learning strategy with SGD optimizers mostly implemented in MATLAB [115].

Super-resolution, which aims to enhance spatial resolution, is an ongoing research topic in computer vision and RS [114]. The latest super-resolution trend focused on example (learning)-based techniques, including a training phase between low-resolution and high-resolution pairs of images [116]. Example-based techniques have seen enhanced accuracies by the introduction of CNNs to generic super-resolution problems [117]. However, RS imageries exhibit a different level of complexity than images in other fields such as computer vision, which delayed the use of CNNs in RS image super-resolution until 2018 by introducing a specific super-resolution CNN architecture to adapt with multispectral satellite imagery [116]. An overview of the related papers shows that all the CNN models were 2-D structured, in which Adam and SGD equally were used for parameter optimizations with the epoch numbers ranging from 80 to 600.

In recent years, DL methods have been successfully applied in natural image change detection-based applications [118]. Previously, different DL-based methods have been applied to various change detection tasks, such as urban dynamics [119], LULC applications [120], or landslides [121]. First, CNN models were employed for high-resolution remotely sensed image change detection in 2018 using faster R-CNN [119] and have gained attention since then. With a review of the database, it is observed that most of the studies devoted to change detection applications applied a range of data types, from spaceborne and airborne optical to RADAR images.

Image registration is a fundamental part of many RS tasks, such as image fusion and change detection [122], [123]. Like the first study, in 2018, Ye et al. [122] fine-tuned the VGG-16 model

using custom RS data to obtain deep CNN features to build an automatic registration algorithm. CNNs showed powerful performance in the registration of RGB and infrared [124], SAR [125], multimodal [56], and aerial RS images [126].

Another active area that uses CNNs is image correction, which includes categories such as image denoising [127], image reconstruction [128], and image compensation [129]. CNNs were introduced to this field in 2017 to build a nonparametric color-correcting scheme for multispectral images [130] and the year after that for removing haze from RS images [131]. Overall, eight studies in the article database used CNN models for correction purposes, of which most of them were applied to multispectral satellite images.

Besides the above applications of CNNs in RS image analysis, CNNs have also been applied in other areas, including regression [52], image retrieval [132], prediction [133], quality assessment [134], and hyperspectral unmixing [135]. CNN models have achieved outstanding performances in each of these applications by presenting a novel way to solve them. Considering the high-accuracy of CNN, it is expected that they will continue to find their use in other research areas.

V. CONCLUSION AND PROSPECTS

This review presents a comprehensive review of CNNs in RS data analysis. It summarizes their progression and advancement since their emergence in the RS field in terms of general characteristics and technical specifications. Based on the detailed analysis of the article database, we come to the following conclusion.

1) There is an increasing trend of using CNNs in RS applications beginning in 2014, and since then, their use has expanded into many new research areas.

2) The use of CNNs in the context of RS started through its employment to analyze multispectral images. As CNNs were increasingly used for efficient problem solving based on different data types and platforms, researchers began to incorporate CNNs in their projects using other data types such as LiDAR and RADAR.

3) Advances in new and freely open-source frameworks and libraries with highly dynamic interfaces allowed the RS community to research new study objectives. The survey among different frameworks identifies TensorFlow as the most used framework. Based on the yearly usage frequency, it is expected to hold its position in the coming years.

4) A survey among the CNNs' parameters shows that SGD and Adam are the most frequented optimizers, and, in most cases, the number of epochs was set more than 100 iterations.

5) Classification tasks using various data types and sensors focus on most of the studies (54%) using CNNs. Classification results have shown to be better when using multisource data and using images with a spatial resolution of more than 5 m. As classification tasks require large quantities of training samples, researchers tackled this problem by improving their training dataset's efficiency using a transfer learning strategy. For this aim, most of the case studies used VGG variants.

6) We could not further analyze and summarize the processing time because it was neither available nor specified if the entire time is for optimizing metaparameters or not. However, in contrast to the general belief, numerous cases report a training time of less than 1 h. It is generally true that deep networks need considerably more processing time for training (though the testing/simulation process is generally quick). However, with continuous increases in processing power, deep networks are readily usable, particularly by incorporating both CPUs and GPUs. It would be interesting to evaluate the time saved by using pretrained networks and fine-tuning them, but currently, there were no statistics reported to extract conclusive information.

7) CNNs have been typically utilized for data with dimensions up to three which is common for RS data. However, 1-D CNNs have shown promising results in discovering intricate patterns in high-dimensional data, especially for sequential data. This capability of extracting meaningful information from the high-dimensional vector data has indicated 1-D CNNs as promising alternatives to conventional methods in some regression problems.

8) The majority of studies focused more on the architecture designs and few studies elaborated on the time efficiency while training networks. By the growth of big RS data and their applications in practical productions, which require much more time to train rather than typical research, much more attention is required to develop time-efficient networks that meet the practical projects' requirements. However, this problem has been partly solved by using public online free or commercial cloud computing platforms, such as Google, Amazon, Microsoft, and IBM. Cloud computing can be used for the development of computationally intensive CNNs by providing high speed and flexible facilities that can handle huge amount of data.

Based on the 416 reviewed articles in this survey, it is evident that CNNs have pervaded every aspect of RS image analysis. This has happened very fast as over 96% of the contributions, a total of 400 papers, were published starting from 2017. For example, the first application of CNNs in object detection studies happened in 2014, whereas they have been used for image registration tasks in 2018. However, the growth rate of applying CNNs to different tasks and data types is challenged by the lack of large training datasets. Although CNNs can be considered newly introduced algorithms in RS, they are now clearly among the top performers in most RS applications. Despite this progress, the study of CNN-based approaches is currently at its beginning stages, and there is still much potential for new developments, particularly in applications such as hyperspectral unmixing, image retrieval, and image quality assessment. Another striking conclusion is that a few studies are conducted in new application areas. As a result, there is a gap in examining different aspects of CNNs. Therefore, in order to get the best

results, researchers should consider investigating new CNN architectures. In this perspective, the design of new network architectures for specific tasks, the generation of large-scale datasets for network training, the integration of conventional techniques according to the RS data, the advancement and analysis of existing networks concerning their architectures, optimization techniques, and the regularization strategies are still open topics which are in close relation with each other and should be jointly considered.

REFERENCES

[1] D. Dhall, R. Kaur, and M. Juneja, "Machine learning: A review of the algorithms and its applications," in *Proc. 2nd Int. Conf. Recent Innov. Comput.*, 2020, pp. 47–63.

[2] P. Scheunders, D. Tuia, and G. Moser, "Contributions of machine learning to remote sensing data analysis," in *Data Processing and Analysis Methodology*. Amsterdam, The Netherlands: Elsevier, 2018, pp. 199–243.

[3] J. Holloway and K. Mengersen, "Statistical machine learning methods and remote sensing for sustainable development goals: A review," *Remote Sens.*, vol. 10, no. 9, p. 1365, 2018.

[4] G. Camps-Valls, "Machine learning in remote sensing data processing," in *Proc. IEEE Int. Workshop Mach. Learn. Signal Process.*, 2009, pp. 1–6.

[5] G. M. Foody and A. Mathur, "Toward intelligent training of supervised image classifications: Directing training data acquisition for SVM classification," *Remote Sens. Environ.*, vol. 93, no. 1/2, pp. 107–117, 2004.

[6] Y. Ban and O. Yousif, "Change detection techniques: A review," in *Multitemporal Remote Sensing*. New York, NY, USA: Springer, 2016, pp. 19–43.

[7] L. Bruzzone and F. Bovolo, "A novel framework for the design of change-detection systems for very-high-resolution remote sensing images," *Proc. IEEE*, vol. 101, no. 3, pp. 609–630, Mar. 2012.

[8] N. M. Nasrabadi, "Hyperspectral target detection: An overview of current and future challenges," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 34–44, Dec. 2013.

[9] L. Zhang and F. Luo, "Review on graph learning for dimensionality reduction of hyperspectral image," *Geo-spatial Inf. Sci.*, vol. 23, no. 1, pp. 98–106, 2020.

[10] J. A. Benediktsson, M. Pesaresi, and K. Amason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1940–1949, Sep. 2003.

[11] F. E. Fassnacht et al., "Review of studies on tree species classification from remotely sensed data," *Remote Sens. Environ.*, vol. 186, pp. 64–87, 2016.

[12] D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, "Machine learning in geosciences and remote sensing," *Geosci. Frontiers*, vol. 7, no. 1, pp. 3–10, 2016.

[13] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support vector machine vs. random forest for remote sensing image classification: A meta-analysis and systematic review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 6308–6325, Sep. 2020.

[14] Y. Shao and R. S. Lunetta, "Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points," *ISPRS J. Photogramm. Remote Sens.*, vol. 70, pp. 78–87, 2012.

[15] M. Sheykhmousa, N. Kerle, M. Kuffer, and S. Ghaffarian, "Post-disaster recovery assessment with machine learning-derived land cover and land use information," *Remote Sens.*, vol. 11, no. 10, 2019, Art. no. 1174.

[16] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 3, pp. 247–259, 2011.

[17] M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogramm. Remote Sens.*, vol. 114, pp. 24–31, 2016.

[18] 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 J. Photogramm. Remote Sens.*, vol. 152, pp. 166–177, 2019.

[19] X. X. Zhu et al., "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.

[20] G. Cheng, X. Xie, J. Han, L. Guo, and G.-S. Xia, "Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3735–3756, 2020.

[21] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.

[22] P. Ghamisi, B. Höfle, and X. X. Zhu, "Hyperspectral and LiDAR data fusion using extinction profiles and deep convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 6, pp. 3011–3024, Dec. 2016.

[23] G. Hinton et al., "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.

[24] A. Tonioni, M. Poggi, S. Mattoccia, and L. Di Stefano, "Unsupervised adaptation for deep stereo," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 1605–1613.

[25] G. Lu and B. Fei, "Medical hyperspectral imaging: A review," *J. Biomed. Opt.*, vol. 19, no. 1, 2014, Art. no. 010901.

[26] J. C. B. Gamboa, "Deep learning for time-series analysis," 2017, *arXiv:1701.01887*.

[27] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agriculture*, vol. 147, pp. 70–90, 2018.

[28] G. Litjens et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.

[29] P. Liu, K.-K. R. Choo, L. Wang, and F. Huang, "SVM or deep learning? A comparative study on remote sensing image classification," *Soft Comput.*, vol. 21, no. 23, pp. 7053–7065, 2017.

[30] G. Tsagkatakis, A. Aidini, K. Fotiadou, M. Giannopoulos, A. Pentari, and P. Tsakalides, "Survey of deep-learning approaches for remote sensing observation enhancement," *Sensors*, vol. 19, no. 18, 2019, Art. no. 3929.

[31] J. E. Ball, D. T. Anderson, and C. S. Chan, "Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community," *J. Appl. Remote Sens.*, vol. 11, no. 4, 2017, Art. no. 042609.

[32] S. Srinivas, R. K. Sarvadevabhatla, K. R. Mopuri, N. Prabhu, S. S. Kruthiventi, and R. V. Babu, "An introduction to deep convolutional neural nets for computer vision," in *Deep Learning for Medical Image Analysis*. Amsterdam, The Netherlands: Elsevier, 2017, pp. 25–52.

[33] M. Ebrahimi, C. Y. Suen, and O. Ormandjieva, "Detecting predatory conversations in social media by deep convolutional neural networks," *Digit. Investigation*, vol. 18, pp. 33–49, 2016.

[34] V. S. Parekh and M. A. Jacobs, "Deep learning and radiomics in precision medicine," *Expert Rev. Precis. Med. Drug Develop.*, vol. 4, no. 2, pp. 59–72, 2019.

[35] U. C. Allard et al., "A convolutional neural network for robotic arm guidance using sEMG based frequency-features," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 2464–2470.

[36] Z. Xu, K. Guan, N. Casler, B. Peng, and S. Wang, "A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 144, pp. 423–434, 2018.

[37] J. Song, S. Gao, Y. Zhu, and C. Ma, "A survey of remote sensing image classification based on CNNs," *Big Earth Data*, vol. 3, no. 3, pp. 232–254, 2019.

[38] F. Xie, M. Shi, Z. Shi, J. Yin, and D. Zhao, "Multilevel cloud detection in remote sensing images based on deep learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3631–3640, Apr. 2017.

[39] Y. Wang, Z. Li, C. Zeng, G.-S. Xia, and H. Shen, "An urban water extraction method combining deep learning and Google Earth Engine," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 768–781, 2020.

[40] S. S. Heydari and G. Mountrakis, "Meta-analysis of deep neural networks in remote sensing: A comparative study of mono-temporal classification to support vector machines," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, pp. 192–210, 2019.

[41] W. Zhao, S. Du, and W. J. Emery, "Object-based convolutional neural network for high-resolution imagery classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 7, pp. 3386–3396, Mar. 2017.

[42] L. Dong *et al.*, "Very high resolution remote sensing imagery classification using a fusion of random forest and deep learning technique—Subtropical area for example," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 113–128, 2020.

[43] Y. Chen, L. Tang, X. Yang, R. Fan, M. Bilal, and Q. Li, "Thick clouds removal from multitemporal ZY-3 satellite images using deep learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 143–153, 2020.

[44] W. Fang *et al.*, "Recognizing global reservoirs from Landsat 8 images: A deep learning approach," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 9, pp. 3168–3177, Aug. 2019.

[45] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state-of-the-art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[46] X. Lu, H. Sun, and X. Zheng, "A feature aggregation convolutional neural network for remote sensing scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7894–7906, Jun. 2019.

[47] C. Zhang *et al.*, "A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 140, pp. 133–144, 2018.

[48] X. Xu, W. Li, Q. Ran, Q. Du, L. Gao, and B. Zhang, "Multisource remote sensing data classification based on convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 937–949, Oct. 2017.

[49] F. Alidoost and H. Arefi, "A CNN-based approach for automatic building detection and recognition of roof types using a single aerial image," *PFG–J. Photogramm., Remote Sens. Geoinf. Sci.*, vol. 86, no. 5/6, pp. 235–248, 2018.

[50] Y. Tan, S. Xiong, and Y. Li, "Automatic extraction of built-up areas from panchromatic and multispectral remote sensing images using double-stream deep convolutional neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 11, pp. 3988–4004, Oct. 2018.

[51] G. Scarpa, M. Gargiulo, A. Mazza, and R. Gaetano, "A CNN-based fusion method for feature extraction from sentinel data," *Remote Sens.*, vol. 10, no. 2, 2018, Art. no. 236.

[52] J. Pyo *et al.*, "A convolutional neural network regression for quantifying cyanobacteria using hyperspectral imagery," *Remote Sens. Environ.*, vol. 233, 2019, Art. no. 111350.

[53] X. Xie, Y. Zhou, Y. Xu, Y. Hu, and C. Wu, "OpenStreetMap data quality assessment via deep learning and remote sensing imagery," *IEEE Access*, vol. 7, pp. 176884–176895, 2019.

[54] D. Cantorna, C. Dafonte, A. Iglesias, and B. Arcay, "Oil spill segmentation in SAR images using convolutional neural networks. A comparative analysis with clustering and logistic regression algorithms," *Appl. Soft Comput.*, vol. 84, 2019, Art. no. 105716.

[55] X. Yang, H. Sun, X. Sun, M. Yan, Z. Guo, and K. Fu, "Position detection and direction prediction for arbitrary-oriented ships via multitask rotation region convolutional neural network," *IEEE Access*, vol. 6, pp. 50839–50849, 2018.

[56] H. Zhang *et al.*, "Registration of multimodal remote sensing image based on deep fully convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 8, pp. 3028–3042, Jun. 2019.

[57] L. Gao, W. Song, J. Dai, and Y. Chen, "Road extraction from high-resolution remote sensing imagery using refined deep residual convolutional neural network," *Remote Sens.*, vol. 11, no. 5, 2019, Art. no. 552.

[58] Y. Sun, Q. Xin, J. Huang, B. Huang, and H. Zhang, "Characterizing tree species of a tropical wetland in Southern China at the individual tree level based on convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 11, pp. 4415–4425, Nov. 2019.

[59] Y. Guo, H. Cao, J. Bai, and Y. Bai, "High efficient deep feature extraction and classification of spectral-spatial hyperspectral image using cross domain convolutional neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 1, pp. 345–356, Jan. 2019.

[60] M. Qin, F. Xie, W. Li, Z. Shi, and H. Zhang, "Dehazing for multispectral remote sensing images based on a convolutional neural network with the residual architecture," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1645–1655, Mar. 2018.

[61] M. Paoletti, J. Haut, J. Plaza, and A. Plaza, "Deep learning classifiers for hyperspectral imaging: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 158, pp. 279–317, 2019.

[62] H. Parikh, S. Patel, and V. Patel, "Classification of SAR and PolSAR images using deep learning: A review," *Int. J. Image Data Fusion*, vol. 11, no. 1, pp. 1–32, 2020.

[63] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6690–6709, Apr. 2019.

[64] M. Pashaei, H. Kamangir, M. J. Starek, and P. Tissot, "Review and evaluation of deep learning architectures for efficient land cover mapping with UAS hyper-spatial imagery: A case study over a wetland," *Remote Sens.*, vol. 12, no. 6, 2020, Art. no. 959.

[65] Y. Li, H. Zhang, X. Xue, Y. Jiang, and Q. Shen, "Deep learning for remote sensing image classification: A survey," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discov.*, vol. 8, no. 6, 2018, Art. no. e1264.

[66] Y. Liu, X. Chen, Z. Wang, Z. J. Wang, R. K. Ward, and X. Wang, "Deep learning for pixel-level image fusion: Recent advances and future prospects," *Inf. Fusion*, vol. 42, pp. 158–173, 2018.

[67] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural Comput.*, vol. 29, no. 9, pp. 2352–2449, 2017.

[68] M. Mahdianpari, B. Salehi, M. Rezaee, F. Mohammadimanesh, and Y. Zhang, "Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery," *Remote Sens.*, vol. 10, no. 7, Jul. 2018, Art. no. 1119, doi: 10.3390/rs10071119.

[69] C. J. Spoerer, P. McClure, and N. Kriegeskorte, "Recurrent convolutional neural networks: A better model of biological object recognition," *Frontiers Psychol.*, vol. 8, pp. 1551, 2017.

[70] G. Yao, T. Lei, and J. Zhong, "A review of convolutional-neural-network-based action recognition," *Pattern Recognit. Lett.*, vol. 118, pp. 14–22, 2019.

[71] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and PRISMA Group, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *PLoS Med.*, vol. 6, no. 7, 2009, Art. no. e1000097.

[72] X. Chen, S. Xiang, C.-L. Liu, and C.-H. Pan, "Vehicle detection in satellite images by hybrid deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 10, pp. 1797–1801, Mar. 2014.

[73] G. Nguyen *et al.*, "Machine learning and deep learning frameworks and libraries for large-scale data mining: A survey," *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 77–124, 2019.

[74] T. Hope, Y. S. Resheff, and I. Lieder, *Learning Tensorflow: A Guide to Building Deep Learning Systems*. Sebastopol, CA, USA: O'Reilly Media, 2017.

[75] A. Sergeev and M. Del Balso, "Horovod: Fast and easy distributed deep learning in tensorflow," 2018, *arXiv:1802.05799*.

[76] S. Chetlur *et al.*, "cuDNN: Efficient primitives for deep learning," 2014, *arXiv: 1410.0759*.

[77] Y. Jia *et al.*, "Caffe: Convolutional architecture for fast feature embedding," in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 675–678.

[78] P. Kim, *MATLAB deep learning: With Machine Learning, Neural Networks and Artificial Intelligence*, vol. 130. Berkeley, CA, USA: Apress, 2017, pp. 21.

[79] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[80] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artif. Intell. Rev.*, vol. 53, pp. 5455–5516, 2020.

[81] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Neural Inf. Process. Syst. Conf.*, 2012, pp. 1097–1105.

[82] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 248–255.

[83] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learning Representations*, 2015.

[84] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1492–1500.

[85] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2818–2826.

[86] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Jan. 2017.

[87] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9.

[88] P. Ghamisi, J. Plaza, Y. Chen, J. Li, and A. J. Plaza, "Advanced spectral classifiers for hyperspectral images: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 1, pp. 8–32, Mar. 2017.

[89] Q. Zou, L. Ni, T. Zhang, and Q. Wang, "Deep learning based feature selection for remote sensing scene classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2321–2325, Sep. 2015.

[90] Z. Chen, Y. Wang, W. Han, R. Feng, and J. Chen, "An improved pretraining strategy-based scene classification with deep learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 5, pp. 844–848, Aug. 2019.

[91] P. Du, E. Li, J. Xia, A. Samat, and X. Bai, "Feature and model level fusion of pretrained CNN for remote sensing scene classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 8, pp. 2600–2611, Nov. 2018.

[92] M. Rezaee, M. Mahdianpari, Y. Zhang, and B. Salehi, "Deep convolutional neural network for complex wetland classification using optical remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 9, pp. 3030–3039, Sep. 2018.

[93] G.-S. Xia *et al.*, "AID: A benchmark data set for performance evaluation of aerial scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3965–3981, Apr. 2017.

[94] G. Cheng, J. Han, and X. Lu, "Remote sensing image scene classification: Benchmark and state-of-the-art," *Proc. IEEE*, vol. 105, no. 10, pp. 1865–1883, Apr. 2017.

[95] *ISPRS 2D Semantic Labeling–Vaihingen* Data. [Online]. Available: http://www2.isprs.org/commissions/comm3/wg4/2d-sem-label-vaihingen.html

[96] *ISPRS 2D Semantic Labeling–Potsdam* Data. [Online]. Available: http://www2.isprs.org/commissions/comm3/wg4/2d-sem-label-potsdam.html

[97] X. Sang, L. Xue, X. Ran, X. Li, J. Liu, and Z. Liu, "Intelligent high-resolution geological mapping based on SLIC-CNN," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 2, 2020, Art. no. 99.

[98] P. Nevavuori, N. Narra, and T. Lipping, "Crop yield prediction with deep convolutional neural networks," *Comput. Electron. Agriculture*, vol. 163, 2019, Art. no. 104859.

[99] T. Liu, A. Abd-Elrahman, J. Morton, and V. L. Wilhelm, "Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system," *GISci. Remote Sens.*, vol. 55, no. 2, pp. 243–264, 2018.

[100] Y. Sun, J. Huang, Z. Ao, D. Lao, and Q. Xin, "Deep learning approaches for the mapping of tree species diversity in a tropical wetland using airborne LiDAR and high-spatial-resolution remote sensing images," *Forests*, vol. 10, no. 11, 2019, Art. no. 1047.

[101] J. Sun, L. Di, Z. Sun, Y. Shen, and Z. Lai, "County-level soybean yield prediction using deep CNN-LSTM model," *Sensors*, vol. 19, no. 20, 2019, Art. no. 4363.

[102] R. Taghizadeh-Mehrjardi *et al.*, "Multi-task convolutional neural networks outperformed random forest for mapping soil particle size fractions in central Iran," *Geoderma*, vol. 376, 2020, Art. no. 114552, doi: 10.1016/j.geoderma.2020.114552.

[103] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, 2015.

[104] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learning Representations*, 2015.

[105] G.-S. Xia *et al.*, "DOTA: A large-scale dataset for object detection in aerial images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3974–3983.

[106] V. Mnih, *Machine Learning For Aerial Image Labeling*. Ph.D. dissertation, Graduate Dept. Comput. Sci., Univ. Toronto, Toronto, ON, Canada, 2013.

[107] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 580–587.

[108] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1440–1448.

[109] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Neural Inf. Process. Syst. Conf.*, 2015, pp. 91–99.

[110] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2961–2969.

[111] H. Shen, M. Jiang, J. Li, Q. Yuan, Y. Wei, and L. Zhang, "Spatial–spectral fusion by combining deep learning and variational model," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6169–6181, Aug. 2019.

[112] M. Mahdianpari, B. Salehi, F. Mohammadimanesh, and M. Motagh, "Random forest wetland classification using ALOS-2 L-band, RADARSAT-2 C-band, and TerraSAR-X imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 130, pp. 13–31, Aug. 2017, doi: 10.1016/j.isprsjprs.2017.05.010.

[113] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, 2016, Art. no. 594.

[114] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Jul. 2015.

[115] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[116] C. Tuna, G. Unal, and E. Sertel, "Single-frame super resolution of remote-sensing images by convolutional neural networks," *Int. J. Remote Sens.*, vol. 39, no. 8, pp. 2463–2479, 2018.

[117] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 184–199.

[118] Q. Wang, X. Zhang, G. Chen, F. Dai, Y. Gong, and K. Zhu, "Change detection based on Faster R-CNN for high-resolution remote sensing images," *Remote Sens. Lett.*, vol. 9, no. 10, pp. 923–932, 2018.

[119] H. Lyu *et al.*, "Long-term annual mapping of four cities on different continents by applying a deep information learning method to Landsat data," *Remote Sens.*, vol. 10, no. 3, 2018, Art. no. 471.

[120] H. Lyu, H. Lu, and L. Mou, "Learning a transferable change rule from a recurrent neural network for land cover change detection," *Remote Sens.*, vol. 8, no. 6, 2016, Art. no. 506.

[121] Y. Liu and L. Wu, "Geological disaster recognition on optical remote sensing images using deep learning," *Procedia Comput. Sci.*, vol. 91, pp. 566–575, 2016.

[122] F. Ye, Y. Su, H. Xiao, X. Zhao, and W. Min, "Remote sensing image registration using convolutional neural network features," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 232–236, Jan. 2018.

[123] H. He, M. Chen, T. Chen, and D. Li, "Matching of remote sensing images with complex background variations via Siamese convolutional neural network," *Remote Sens.*, vol. 10, no. 2, 2018, Art. no. 355.

[124] E. R. DeLancey, J. F. Simms, M. Mahdianpari, B. Brisco, C. Mahoney, and J. Kariyeva, "Comparing deep learning and shallow learning for large-scale wetland classification in Alberta, Canada," *Remote Sens.*, vol. 12, no. 1, Dec. 2019, Art. no. 2, doi: 10.3390/rs12010002.

[125] T. Bürgmann, W. Koppe, and M. Schmitt, "Matching of TerraSAR-X derived ground control points to optical image patches using deep learning," *ISPRS J. Photogramm. Remote Sens.*, vol. 158, pp. 241–248, 2019.

[126] F. Mohammadimanesh, B. Salehi, M. Mahdianpari, E. Gill, and M. Molinier, "A new fully convolutional neural network for semantic segmentation of polarimetric SAR imagery in complex land cover ecosystem," *ISPRS J. Photogramm. Remote Sens.*, vol. 151, pp. 223–236, May 2019, doi: 10.1016/j.isprsjprs.2019.03.015.

[127] Q. Yuan, Q. Zhang, J. Li, H. Shen, and L. Zhang, "Hyperspectral image denoising employing a spatial–spectral deep residual convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1205–1218, Sep. 2018.

[128] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4274–4288, Mar. 2018.

[129] Z. Zhaoxiang, A. Iwasaki, and G. Xu, "Attitude jitter compensation for remote sensing images using convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 9, pp. 1358–1362, Mar. 2019.

[130] J. Guo, Z. Pan, B. Lei, and C. Ding, "Automatic color correction for multisource remote sensing images with Wasserstein CNN," *Remote Sens.*, vol. 9, no. 5, 2017, Art. no. 483.

[131] H. Jiang and N. Lu, "Multi-scale residual convolutional neural network for haze removal of remote sensing images," *Remote Sens.*, vol. 10, no. 6, 2018, Art. no. 945.

[132] X. Yang, N. Wang, B. Song, and X. Gao, "BoSR: A CNN-based aurora image retrieval method," *Neural Netw.*, vol. 116, pp. 188–197, 2019.

[133] M. Han, Y. Feng, X. Zhao, C. Sun, F. Hong, and C. Liu, "A convolutional neural network using surface data to predict subsurface temperatures in the Pacific Ocean," *IEEE Access*, vol. 7, pp. 172816–172829, 2019.

[134] G. Chen, Q. Pei, and M. Kamruzzaman, "Remote sensing image quality evaluation based on deep support value learning networks," *Signal Process., Image Commun.*, vol. 83, 2020, Art. no. 115783.

[135] X. Zhang, Y. Sun, J. Zhang, P. Wu, and L. Jiao, "Hyperspectral unmixing via deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 11, pp. 1755–1759, Aug. 2018.

Hamid Ghanbari received the B.Sc. degree in geomatics engineering and the master's degree in remote sensing from the School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran, in 2014 and 2016, respectively. He is currently working toward the Ph.D. degree with the Department of Geography, Université Laval, Québec, QC, Canada.

His research interests include hyperspectral image classification, multitemporal and multisensor image analysis, and machine learning.

Masoud Mahdianpari (Member, IEEE) received the B.S. degree in surveying and geomatics engineering and the M.Sc. degree in remote sensing engineering from the School of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran, in 2010 and 2013, respectively, and the Ph.D. degree in electrical engineering from the Department of Engineering and Applied Science, Memorial University, St. John's, NL, Canada, in 2019.

In 2019, he was an Ocean Frontier Institute (OFI) Postdoctoral Fellow with Memorial University and C-CORE. He is currently a Remote Sensing Technical Lead with C-CORE and a Cross Appointed Professor with the Faculty of Engineering and Applied Science, Memorial University. His research interests include remote sensing and image analysis, with a special focus on PolSAR image processing, multisensor data classification, machine learning, geo big data, and deep learning.

Dr. Mahdianpari was the recipient of the Top Student Award from the University of Tehran for four consecutive academic years from 2010 to 2013, the Research and Development Corporation Ocean Industries Student Research Award, organized by the Newfoundland Industry and Innovation Center, amongst more than 400 submissions, in 2016, the T. David Collett Best Industry Paper Award organized by IEEE, in 2016, The Com-Adv Devices, Inc., Scholarship for Innovation, Creativity, and Entrepreneurship from the Memorial University of Newfoundland, in 2017, the Artificial Intelligence for Earth Grant organized by Microsoft, in 2018, and the Graduate Academic Excellence Award organized by Memorial University, in 2019. He is currently an Editorial Team Member for the journal *Remote Sensing* and the *Canadian Journal of Remote Sensing*.

Saeid Homayouni (Senior Member, IEEE) received the B.Sc. degree in surveying and geomatics engineering from the University of Isfahan, Isfahan, Iran, in 1996, the M.Sc. degree in remote sensing and geographic information systems from Tarbiat Modares University, Tehran, Iran, in 1999, and the Ph.D. degree in signal and image from Télécom Paris Tech, Paris, France, in 2005.

From 2006 to 2007, he was a Postdoctoral Fellow with the Signal and Image Laboratory, University of Bordeaux Agro-Science, Bordeaux, France. From 2008 to 2011, he was an Assistant Professor with the Department of Surveying and Geomatics, College of Engineering, University of Tehran, Tehran, Iran. From 2011 to 2013, through the Natural Sciences and Engineering Research Council of Canada (NSERC) Visitor Fellowship Program, he worked with the Earth Observation Group of the Agriculture and Agri-Food Canada (AAFC), Ottawa Center of Research and Development, Ottawa, ON, Canada. In 2013, he joined the Department of Geography, Environment, and Geomatics, University of Ottawa, Ottawa, ON, Canada, as a Replacing Assistant Professor of Remote Sensing and Geographic Information Systems. Since April 2019, he has been an Associate Professor of environmental remote sensing and geomatics with the Centre Eau Terre Environnement, Institut National de la Recherche Scientifique, Quebec, QC, Canada. His research interests include optical and radar earth observations analytics for urban and agro-environmental applications.

Fariba Mohammadimanesh received the B.Sc. degree in surveying and geomatics engineering and the M.Sc. degree in remote sensing engineering from the School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran, in 2010 and 2014, respectively, and the Ph.D. degree in electrical engineering from the Department of Engineering and Applied Science, Memorial University, St. John's, NL, Canada, in 2019.

She is currently a Remote Sensing Scientist with C-CORE's Earth Observation Team, St. John's, NL, Canada. Her research interests include image processing, machine learning, segmentation, and classification of satellite remote sensing data, including SAR, PolSAR, and optical, in environmental studies, with a special interest in wetland mapping and monitoring, as well as geo-hazard mapping using the interferometric SAR technique (InSAR).

Dr. Mohammadimanesh was the recipient of several awards, including the Emera Graduate Scholarship (2016–2019), CIG-NL Award (2018), and the IEEE NECEC Best Industry Paper Award (2016).