

# A meta-analysis of convolutional neural networks for remote sensing applications

Hamid Ghanbari, Masoud Mahdianpari, Saeid Homayouni, 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 paper 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, 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**— Deep Learning, Convolutional Neural Network, Remote Sensing, Meta-analysis.

## 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 3,000 applied regression models. Given recent developments in RS data acquisition technologies and the rising diversity of objectives capable of being resolved, the use of ML in RS applications is expected to increase [4].

Machine learning 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 natural 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].

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) [13]. Each of these methods carries specific advantages. For example, SVM can best tackle high-dimensionality problems and limited training data [14], while RF does not require the fine-tuning of a large number of hyper-parameters and can easily be used for both simple and complex computations [15, 16]. These two methods share the advantage of lower computational complexity and higher interpretability capabilities [13].

Over the past few years, however, there has been an ongoing shift toward using deep learning methods in ML applications [17]. Deep learning (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 [18]. DL has been used in many areas of research such as speech recognition [19], stereo vision [20], biomedicine [21], time-series analysis [22], agriculture [23], and medical image recognition [24]. 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 [25].

To date, several DL architectures have been introduced, of which the stacked autoencoder (SAE), convolutional neural network (CNN), Generative Adversarial Network (GAN), deep belief network (DBN), and recurrent NN (RNN) have become mainstream [26]. Of these DL networks, CNN is the most popular and the most published [27]. 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 [28]. In the last 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 [29] to precision medicine [30] and robotics [31].

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 [17, 32]. 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 2D structures and reducing network parameters using convolutional filters [28]. 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 non-stationary) [33].

CNNs are composed of a set of blocks that makes 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 [27]. Therefore, CNNs have been successfully used in image preprocessing, scene classification, pixel-based classification and image segmentation, and object detection [34, 35]. For example, CNNs have been used in numerous studies to improve image classification results [32, 36-38], to extract buildings and non-building regions automatically [39], and to detect areas of build-up [40]. Scarpa, et al. [41] 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 [42]. CNNs also have been used in OpenStreetMap Data Quality Assessment [43], oil spill segmentation [44], ship position detection and direction prediction [45], multimodal RS image registration [46], road extraction [47], and many other areas of study.

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. [27] focused on the theories, tools, and challenges of using DL algorithms in the RS community. In another review, Zhang, et al. [35] and Zhu, et al. [18] 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 [48-52]. Liu, et al. [53] presented a systematic review of the application of DL techniques in the field of pixel-based image fusion. A recent comprehensive review by Tsagakatakis, et al. [26] 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 was performed by Rawat and Wang [54] 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 [55], biology [56], and action recognition [57].

The majority of the CNN review papers are descriptive, often with no quantitative assessment, and tend to focus on applications other than remote sensing. 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 provide remote sensing experts with a "big picture" summary of current research in this field. As a whole, the contributions of this paper are: (1) Whereas almost all the review papers in RS applications cover all related DL network structures, this paper 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 sub-tasks. (2) This review defines a complete framework for professionals and even non-experts, 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 in Section 2. After presenting the general characteristics of the studies in Section 3, the accuracy assessment of different CNN setups is given in Section 4. Finally, in the last section, concluding remarks are presented.

## II. METHODS

A systematic literature search query was performed using the Web of Science (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 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology was followed for study selection [58]. 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 used data from the most common remote sensing 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 journals, and the remaining were majority proceedings papers. Conference papers were excluded from our meta-analysis due to their (typically) lower academic rigor

level relative to peer-reviewed articles. 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.

TABLE I. REMOTE SENSING 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

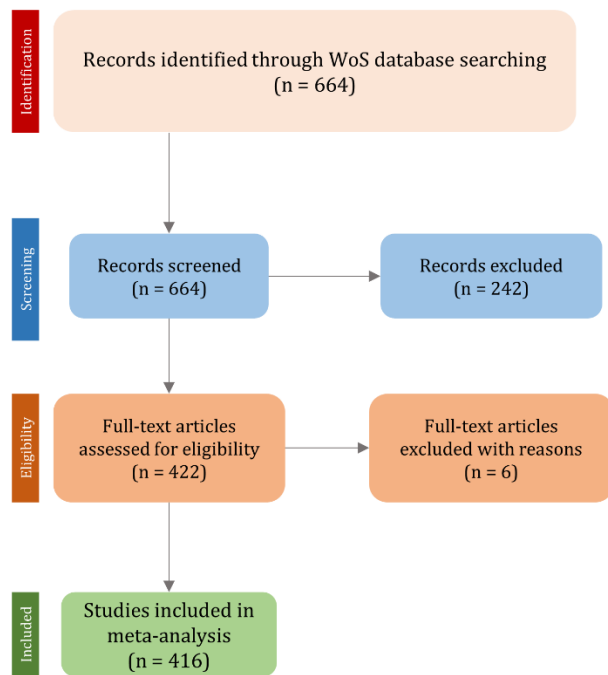


Fig. 8. PRISMA flowchart demonstrating the selection of studies.

The 416 eligible papers were included in the meta-analysis and described using the following: title, year, journal, citation, first author institution's country, remote sensing 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 [59] (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.

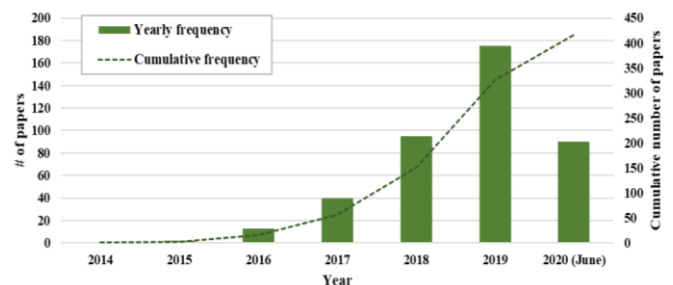


Fig. 2. Annually and cumulative frequency of CNN studies in remote sensing applications.

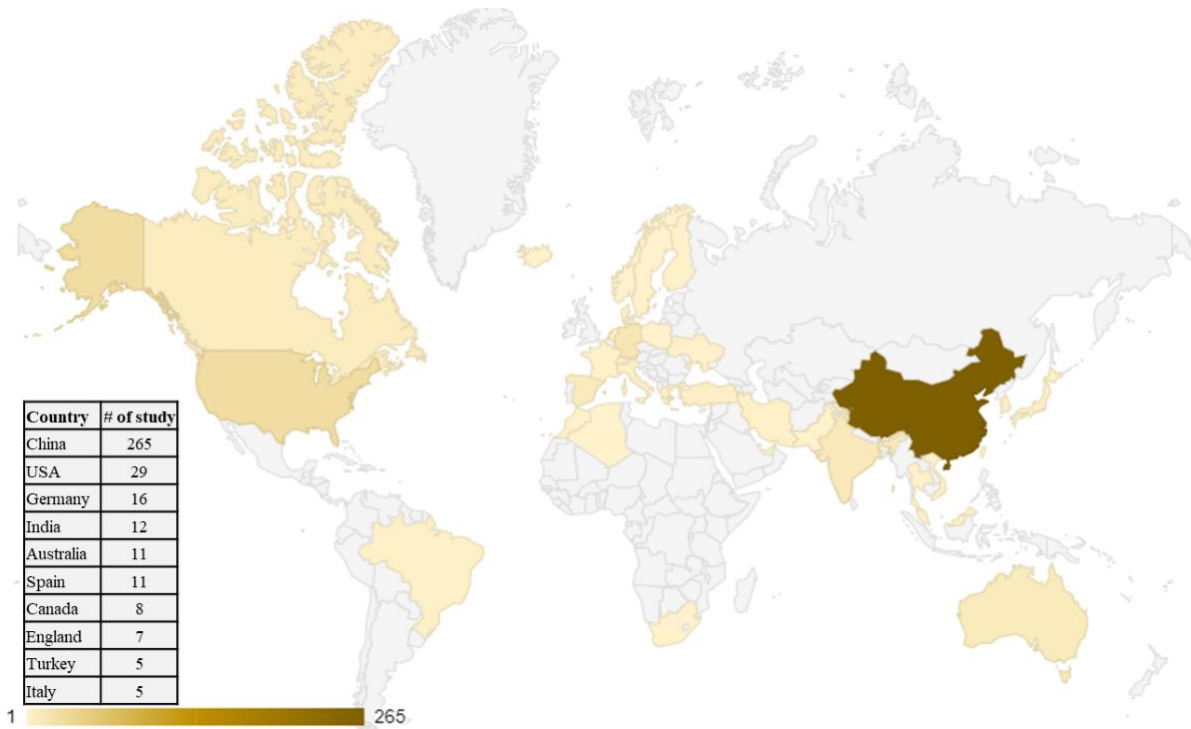


Fig. 3. Distribution of research institutions, according to the country reported in the article. The countries with more than five studies are presented in the figure.

The first authors' associated research institutions' geographical distribution is shown in Fig. 3. Studies were conducted in 34 different countries 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 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 thirteen 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.4 as "other" as the number of associated articles was fewer than five, including sea ice, agriculture, and wetland mapping (Fig. 4).

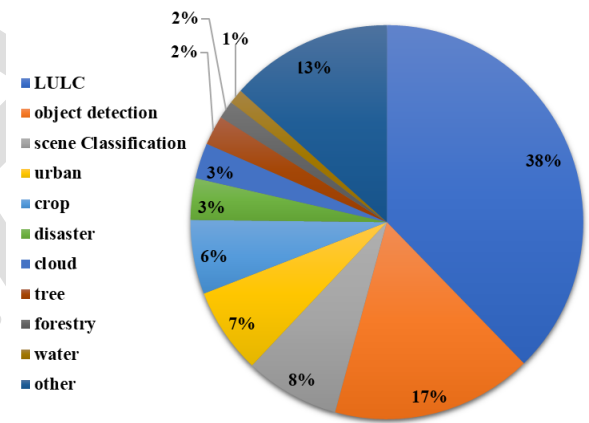


Fig. 4. Number of Studies per each remote sensing application.

*B. Sensors and data types*

The first published RS-related work using CNNs occurred in 2014 for vehicle detection using multispectral satellite images [59]. Since then, CNNs have been applied in research using numerous RS data types (see Fig. 5). 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.

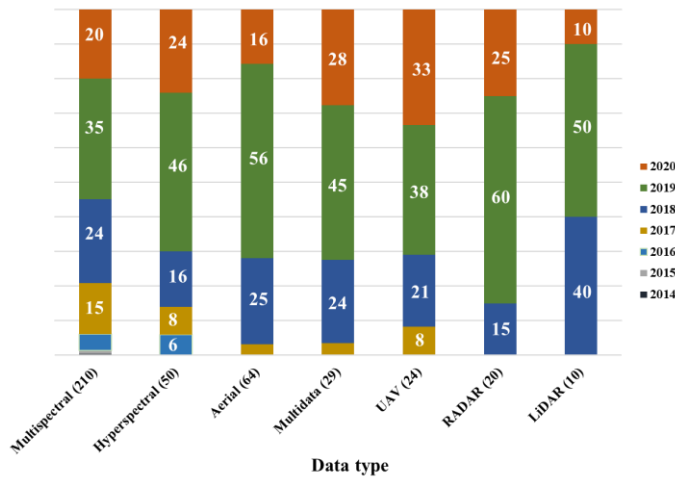


Fig. 5. The usage rate of different remote sensing data (the numbers in parathesis shows the number of studies).

### 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 [60]. A graphical representation of the most used frameworks and libraries and their annual publication frequency in CNN studies is shown in Fig. 6. 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 deep learning framework for numerical computations using data-flow graphs [61]. It is designed to be highly portable, running on various platform scales, from a single CPU to a GPU or GPUs cluster [62].

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 [63]. Caffe's specific properties include suitability for image processing tasks with CNNs, accessibility to pre-trained networks, and easy coding on Python and MATLAB [60, 64]. MATLAB, the third most used platform for CNN implementation, was utilized in almost 19% of the studies. MATLAB's beneficial characteristics are simplicity, especially for practitioners, various visualizing tools, and the capability of deploying models on a variety of servers and devices [65]. 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.

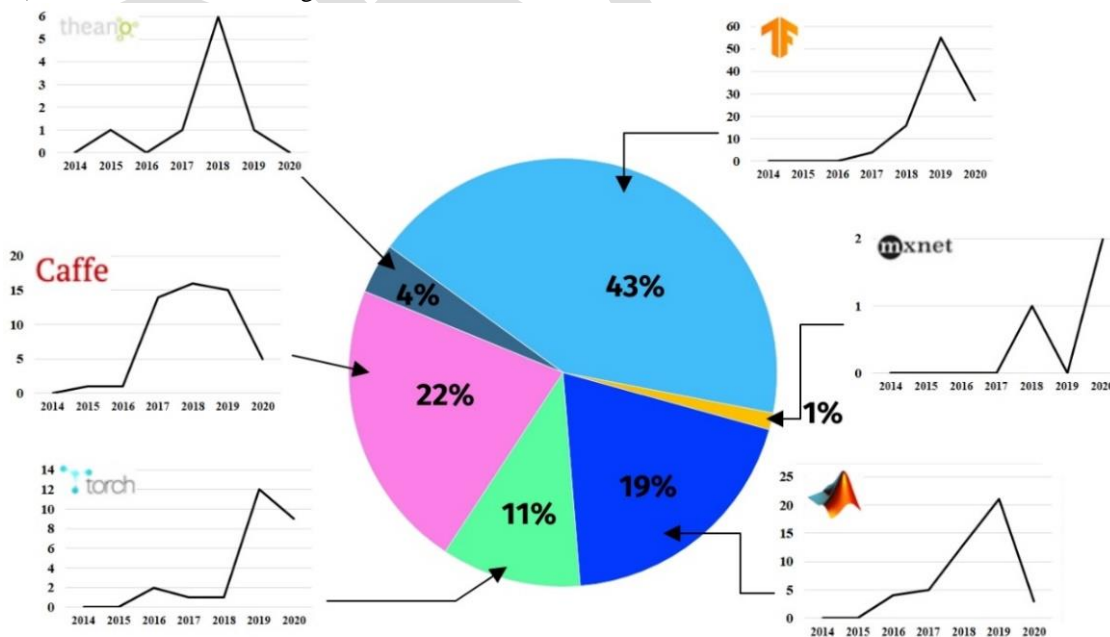


Fig. 6. The frequency of the used CNN frameworks and libraries.

To date, different CNN models have been developed starting from 1989 with the introduction of the first CNN-

based feature encoder, LeNet [66]. 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 [67-69]. Fig. 7 shows the frequency usage of CNN models used in different problems as a backbone architecture or a starting point with the pre-trained parameters (i.e., transfer learning). The most frequently used model, employed in 33% of studies, was VGG-16 [70], followed by ResNet-50 (19%), ResNet-101 (11%) [71], AlexNet (10%) [68], and Inception-V3 (5%) [72]. The remaining models comprising about 22% of the case studies mainly consist of LeNet-5, CaffeNet [64], SegNet [73], ResNet-18, VGG-19, and GoogleNet [74].

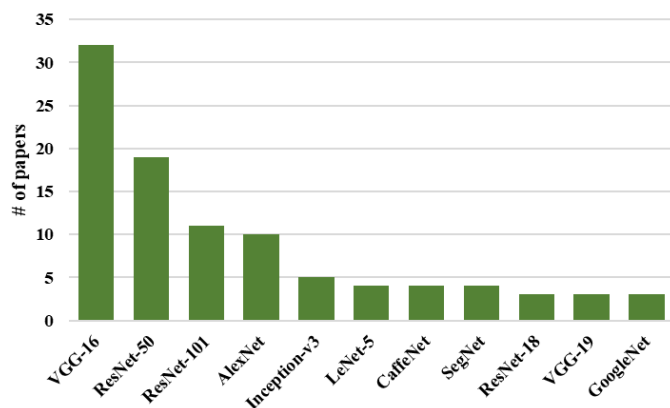


Fig. 7. Frequency of different CNN models in the article database.

#### 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. 8 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).

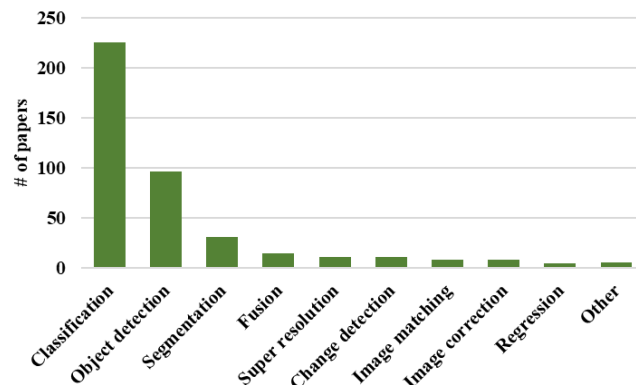


Fig. 8. Distribution of study objectives.

##### 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. 9. Most of the classification tasks focused on LULC, scene classification, and crop-, urban-, 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. 9. 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) [75]. 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. [76]. A majority of these studies apply high-resolution remote sensing images because of the availability of many large-scale high-resolution datasets in recent years [77]. In scene classification, datasets most commonly used are the UC-Merced dataset [78], the aerial image data set (AID) [79], and NWPU-RESISC 45 (Remote Sensing Image Scene Classification) datasets [80].

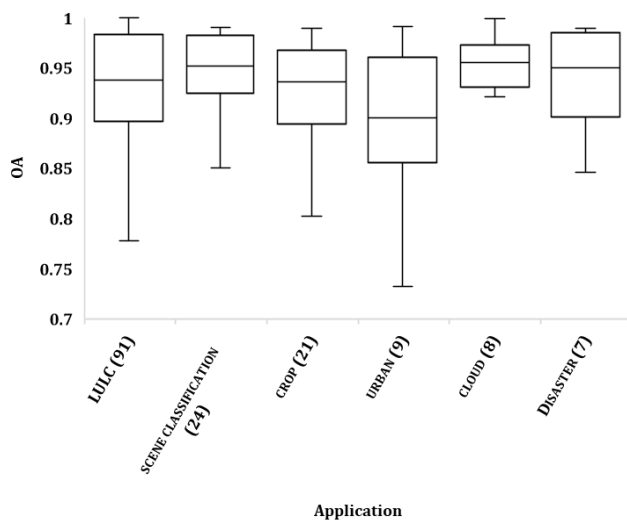


Fig. 9. OA distribution of CNN classification studies based on different applications.

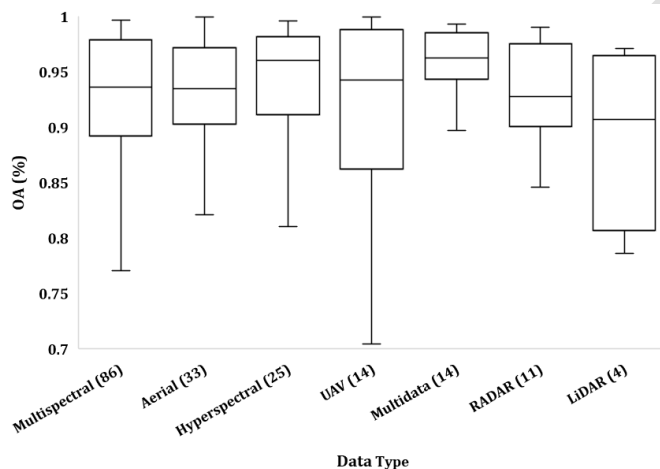


Fig. 10. Distribution of overall accuracies for different data types.

A detailed survey of the article database shows that about 46% of classification studies used spaceborne multispectral remote sensing images for classification tasks (Fig. 10). 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%), multi-data (7%), and RADAR (6%). For the aerial images, the Vaihingen Semantic Labelling dataset [81] and Potsdam Semantic Labelling dataset [82] 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 has 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 [83]. The capability of automatic data acquisition using UAV has made them a convenient tool for some classification tasks such as geological mapping [83], crop yield prediction [84], and wetland mapping [85]. Integration of different remote sensing data types (i.e., multi-data) is beneficial for different tasks

such as tree species diversity mapping using LiDAR and high-resolution multispectral images [86], soybean yield prediction by fusion of weather data and MODIS products [87], and coastal land cover classification by integration of optical and RADAR satellite images [88].

A statistical analysis of different data types showed that using multi-data 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%).

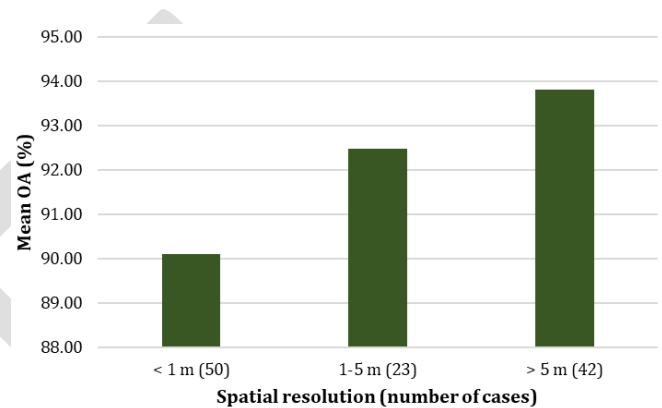


Fig. 11. Frequency and average overall accuracy of CNN classification by the spatial resolution.

Fig. 11 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 m 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 [66, 89], while the remaining studies used Adaptive Moment Estimation (Adam) optimizer [90]. Moreover, the training process was conducted over 100 epochs in about 68% of the cases.

### B. Object detection and image segmentation

As shown in Figure 8, about 30% of CNN's remote sensing 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 the largest share use multispectral satellite images (60%), aerial (18%), and multi-data (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 [82] and NWPU VHR-10 [91], established themselves as baseline datasets CNN-based image segmentation and object detection. They are followed by DOTA (Dataset for Object Detection in Aerial images) [91], UC-Merced [78], and the Massachusetts buildings and roads datasets [92]. 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 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 [93] and its series of improvements, including Fast R-CNN [94], Faster R-CNN [95], Mask R-CNN [96]. 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 remote sensing, 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 [97-99]. 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 [99, 100]. 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 [101].

Super-resolution, which aims to enhance spatial resolution, is an ongoing research topic in computer vision and remote sensing [100]. The latest super-resolution trend focused on example (learning)-based techniques, including a training phase between low-resolution and high-resolution pairs of images [102]. Example-based techniques have seen enhanced accuracies by the introduction of CNNs to generic super-resolution problems [103]. 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 [102]. An overview of the related papers shows that all the CNN models were 2D structured in which Adam and SGD equally were used for parameter optimizations with the epoch numbers ranging from 80 to 600.

In recent years, deep learning methods have been successfully applied in natural image change detection-based applications [104]. Previously, different DL-based methods have been applied to various change detection tasks such as urban dynamics [105], LULC applications [106], or landslides [107]. First, CNN models were employed for high-resolution remotely sensed image change detection in 2018 using faster R-CNN [105] 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 [108, 109]. Like the first study, in 2018, Ye, et al. [108] 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 [110], SAR [111], multimodal [46], and aerial remote sensing images [112].

Another active area that uses CNNs is image correction, which includes categories such as image denoising [113], image reconstruction [114], and image compensation [115]. CNNs were introduced to this field in 2017 to build a nonparametric color-correcting scheme for multispectral images [116] and the year after that for removing haze from remote sensing images [117]. 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 remote sensing image analysis, CNNs have also been applied in other areas, including regression [42], image retrieval [118], prediction [119], quality assessment [120], and hyperspectral unmixing [121]. 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 they will continue to find their use in other research areas.



## V. CONCLUSIONS 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 conclude that:

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 meters. 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 meta-parameters or not. However, in contrast to the general belief, numerous cases report a training time of less than one hour. 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 pre-trained networks and fine-tuning them, but currently, there were no statistics reported to extract conclusive information.

Based on the 416 reviewed articles in this survey, it is evident that CNNs have pervaded every aspect of remote sensing 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 *Proceedings of ICRIC 2019*: Springer, 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*: 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 Sensing*, vol. 10, no. 9, p. 1365, 2018.
- [4] G. Camps-Valls, "Machine learning in remote sensing data processing," in *2009 IEEE international workshop on machine learning for signal processing*, 2009: IEEE, 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 Sensing of Environment*, vol. 93, no. 1-2, pp. 107-117, 2004.
- [6] Y. Ban and O. Yousif, "Change detection techniques: a review," in *Multitemporal Remote Sensing*: 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," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 609-630, 2012.
- [8] N. M. Nasrabadi, "Hyperspectral target detection: An overview of current and future challenges," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 34-44, 2013.
- [9] L. Zhang and F. Luo, "Review on graph learning for dimensionality reduction of hyperspectral image," *Geo-spatial Information Science*, 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 Transactions on Geoscience and Remote Sensing*, vol. 41, no. 9, pp. 1940-1949, 2003.
- [11] F. E. Fassnacht *et al.*, "Review of studies on tree species classification from remotely sensed data," *Remote Sensing of Environment*, 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," *Geoscience Frontiers*, vol. 7, no. 1, pp. 3-10, 2016.
- [13] 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 Journal of Photogrammetry and Remote Sensing*, vol. 70, pp. 78-87, 2012.
- [14] M. Sheykhmousa, N. Kerle, M. Kuffer, and S. Ghaffarian, "Post-disaster recovery assessment with machine learning-derived land cover and land use information," *Remote sensing*, vol. 11, no. 10, p. 1174, 2019.
- [15] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 247-259, 2011.
- [16] M. Belgiu and L. Drăguț, "Random forest in remote sensing: A review of applications and future directions," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24-31, 2016.
- [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, 2019.
- [18] X. X. Zhu *et al.*, "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.
- [19] G. Hinton *et al.*, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal processing magazine*, vol. 29, no. 6, pp. 82-97, 2012.
- [20] A. Tonioni, M. Poggi, S. Mattoccia, and L. Di Stefano, "Unsupervised adaptation for deep stereo," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 1605-1613.
- [21] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *Journal of biomedical optics*, vol. 19, no. 1, p. 010901, 2014.
- [22] J. C. B. Gamboa, "Deep learning for time-series analysis," *arXiv preprint arXiv:1701.01887*, 2017.
- [23] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and electronics in agriculture*, vol. 147, pp. 70-90, 2018.
- [24] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60-88, 2017.
- [25] P. Liu, K.-K. R. Choo, L. Wang, and F. Huang, "SVM or deep learning? A comparative study on remote sensing image classification," *Soft Computing*, vol. 21, no. 23, pp. 7053-7065, 2017.
- [26] 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, p. 3929, 2019.
- [27] 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," *Journal of Applied Remote Sensing*, vol. 11, no. 4, p. 042609, 2017.
- [28] 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*: Elsevier, 2017, pp. 25-52.
- [29] M. Ebrahimi, C. Y. Suen, and O. Ormandjieva, "Detecting predatory conversations in social media by deep convolutional neural networks," *Digital Investigation*, vol. 18, pp. 33-49, 2016.
- [30] V. S. Parekh and M. A. Jacobs, "Deep learning and radiomics in precision medicine," *Expert review of precision medicine and drug development*, vol. 4, no. 2, pp. 59-72, 2019.
- [31] U. C. Allard *et al.*, "A convolutional neural network for robotic arm guidance using sEMG based frequency-features," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016: IEEE, pp. 2464-2470.
- [32] 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 Journal of Photogrammetry and Remote Sensing*, vol. 144, pp. 423-434, 2018.
- [33] 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.
- [34] 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 Journal of Photogrammetry and Remote Sensing*, vol. 152, pp. 192-210, 2019.
- [35] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 22-40, 2016.
- [36] X. Lu, H. Sun, and X. Zheng, "A feature aggregation convolutional neural network for remote sensing scene classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 7894-7906, 2019.
- [37] C. Zhang *et al.*, "A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 140, pp. 133-144, 2018.
- [38] X. Xu, W. Li, Q. Ran, Q. Du, L. Gao, and B. Zhang, "Multisource remote sensing data classification based on convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 937-949, 2017.
- [39] F. Alidoost and H. Arefi, "A CNN-based approach for automatic building detection and recognition of roof types using a single aerial image," *PFG—Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, vol. 86, no. 5-6, pp. 235-248, 2018.
- [40] 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 Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 11, pp. 3988-4004, 2018.
- [41] G. Scarpa, M. Gargiulo, A. Mazza, and R. Gaetano, "A cnn-based fusion method for feature extraction from sentinel data," *Remote Sensing*, vol. 10, no. 2, p. 236, 2018.
- [42] J. Pyo *et al.*, "A convolutional neural network regression for quantifying cyanobacteria using hyperspectral imagery," *Remote Sensing of Environment*, vol. 233, p. 111350, 2019.
- [43] 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.
- [44] 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," *Applied Soft Computing*, vol. 84, p. 105716, 2019.
- [45] 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.
- [46] H. Zhang *et al.*, "Registration of multimodal remote sensing image based on deep fully convolutional neural network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 8, pp. 3028-3042, 2019.

- [47] 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 Sensing*, vol. 11, no. 5, p. 552, 2019.
- [48] M. Paoletti, J. Haut, J. Plaza, and A. Plaza, "Deep learning classifiers for hyperspectral imaging: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 158, pp. 279-317, 2019.
- [49] H. Parikh, S. Patel, and V. Patel, "Classification of SAR and PolSAR images using deep learning: a review," *International Journal of Image and Data Fusion*, vol. 11, no. 1, pp. 1-32, 2020.
- [50] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690-6709, 2019.
- [51] M. Pashaie, 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 Sensing*, vol. 12, no. 6, p. 959, 2020.
- [52] Y. Li, H. Zhang, X. Xue, Y. Jiang, and Q. Shen, "Deep learning for remote sensing image classification: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 6, p. e1264, 2018.
- [53] 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," *Information Fusion*, vol. 42, pp. 158-173, 2018.
- [54] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural computation*, vol. 29, no. 9, pp. 2352-2449, 2017.
- [55] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into imaging*, vol. 9, no. 4, pp. 611-629, 2018.
- [56] C. J. Sporer, P. McClure, and N. Kriegeskorte, "Recurrent convolutional neural networks: a better model of biological object recognition," *Frontiers in psychology*, vol. 8, p. 1551, 2017.
- [57] G. Yao, T. Lei, and J. Zhong, "A review of Convolutional-Neural-Network-based action recognition," *Pattern Recognition Letters*, vol. 118, pp. 14-22, 2019.
- [58] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and P. Group, "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement," *PLoS med*, vol. 6, no. 7, p. e1000097, 2009.
- [59] X. Chen, S. Xiang, C.-L. Liu, and C.-H. Pan, "Vehicle detection in satellite images by hybrid deep convolutional neural networks," *IEEE Geoscience and remote sensing letters*, vol. 11, no. 10, pp. 1797-1801, 2014.
- [60] G. Nguyen *et al.*, "Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 77-124, 2019.
- [61] T. Hope, Y. S. Resheff, and I. Lieder, *Learning tensorflow: A guide to building deep learning systems*. " O'Reilly Media, Inc.", 2017.
- [62] A. Sergeev and M. Del Balso, "Horovod: fast and easy distributed deep learning in TensorFlow," *arXiv preprint arXiv:1802.05799*, 2018.
- [63] S. Chetlur *et al.*, "cudnn: Efficient primitives for deep learning," *arXiv preprint arXiv:1410.0759*, 2014.
- [64] Y. Jia *et al.*, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 675-678.
- [65] P. Kim, "Matlab deep learning," *With Machine Learning, Neural Networks and Artificial Intelligence*, vol. 130, p. 21, 2017.
- [66] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.
- [67] A. Khan, A. Sohail, U. Zahoor, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial Intelligence Review*, pp. 1-62, 2020.
- [68] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.
- [69] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009: Ieee, pp. 248-255.
- [70] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [71] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated transformations for deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1492-1500.
- [72] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818-2826.
- [73] 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*, vol. 39, no. 12, pp. 2481-2495, 2017.
- [74] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1-9.
- [75] 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.
- [76] Q. Zou, L. Ni, T. Zhang, and Q. Wang, "Deep learning based feature selection for remote sensing scene classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 11, pp. 2321-2325, 2015.
- [77] Z. Chen, Y. Wang, W. Han, R. Feng, and J. Chen, "An Improved Pretraining Strategy-Based Scene Classification With Deep Learning," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 5, pp. 844-848, 2019.
- [78] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems*, 2010, pp. 270-279.
- [79] G.-S. Xia *et al.*, "AID: A benchmark data set for performance evaluation of aerial scene classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3965-3981, 2017.
- [80] G. Cheng, J. Han, and X. Lu, "Remote sensing image scene classification: Benchmark and state of the art," *Proceedings of the IEEE*, vol. 105, no. 10, pp. 1865-1883, 2017.
- [81] ISPRS. *ISPRS 2D Semantic Labeling-Vaihingen data*. <http://www2.isprs.org/commissions/comm3/wg4/2d-sem-label-vaihingen.html>.
- [82] ISPRS. *ISPRS 2D Semantic Labeling-Potsdam data*. <http://www2.isprs.org/commissions/comm3/wg4/2d-sem-label-potsdam.html>.
- [83] X. Sang, L. Xue, X. Ran, X. Li, J. Liu, and Z. Liu, "Intelligent High-Resolution Geological Mapping Based on SLIC-CNN,"

- ISPRS International Journal of Geo-Information*, vol. 9, no. 2, p. 99, 2020.
- [84] P. Nevavuori, N. Narra, and T. Lipping, "Crop yield prediction with deep convolutional neural networks," *Computers and electronics in agriculture*, vol. 163, p. 104859, 2019.
- [85] 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," *GIScience & remote sensing*, vol. 55, no. 2, pp. 243-264, 2018.
- [86] 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, p. 1047, 2019.
- [87] 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, p. 4363, 2019.
- [88] Q. Feng *et al.*, "Integrating multitemporal sentinel-1/2 data for coastal land cover classification using a multibranch convolutional neural network: A case of the Yellow River Delta," *Remote Sensing*, vol. 11, no. 9, p. 1006, 2019.
- [89] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85-117, 2015.
- [90] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [91] G.-S. Xia *et al.*, "DOTA: A large-scale dataset for object detection in aerial images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3974-3983.
- [92] V. Mnih, *Machine learning for aerial image labeling*. Citeseer, 2013.
- [93] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580-587.
- [94] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440-1448.
- [95] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91-99.
- [96] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961-2969.
- [97] H. Shen, M. Jiang, J. Li, Q. Yuan, Y. Wei, and L. Zhang, "Spatial-Spectral Fusion by Combining Deep Learning and Variational Model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 6169-6181, 2019.
- [98] J. Zhong, B. Yang, G. Huang, F. Zhong, and Z. Chen, "Remote sensing image fusion with convolutional neural network," *Sensing and Imaging*, vol. 17, no. 1, p. 10, 2016.
- [99] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sensing*, vol. 8, no. 7, p. 594, 2016.
- [100] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295-307, 2015.
- [101] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
- [102] C. Tuna, G. Unal, and E. Sertel, "Single-frame super resolution of remote-sensing images by convolutional neural networks," *International journal of remote sensing*, vol. 39, no. 8, pp. 2463-2479, 2018.
- [103] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *European conference on computer vision*, 2014: Springer, pp. 184-199.
- [104] 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 sensing letters*, vol. 9, no. 10, pp. 923-932, 2018.
- [105] 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 Sensing*, vol. 10, no. 3, p. 471, 2018.
- [106] H. Lyu, H. Lu, and L. Mou, "Learning a transferable change rule from a recurrent neural network for land cover change detection," *Remote Sensing*, vol. 8, no. 6, p. 506, 2016.
- [107] Y. Liu and L. Wu, "Geological disaster recognition on optical remote sensing images using deep learning," *Procedia Computer Science*, vol. 91, pp. 566-575, 2016.
- [108] F. Ye, Y. Su, H. Xiao, X. Zhao, and W. Min, "Remote sensing image registration using convolutional neural network features," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 232-236, 2018.
- [109] H. He, M. Chen, T. Chen, and D. Li, "Matching of remote sensing images with complex background variations via Siamese convolutional neural network," *Remote Sensing*, vol. 10, no. 2, p. 355, 2018.
- [110] R. Zhu, D. Yu, S. Ji, and M. Lu, "Matching RGB and Infrared Remote Sensing Images with Densely-Connected Convolutional Neural Networks," *Remote Sensing*, vol. 11, no. 23, p. 2836, 2019.
- [111] T. Bürgmann, W. Koppe, and M. Schmitt, "Matching of TerraSAR-X derived ground control points to optical image patches using deep learning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 158, pp. 241-248, 2019.
- [112] S. Ji, J. Liu, and M. Lu, "CNN-Based Dense Image Matching for Aerial Remote Sensing Images," *Photogrammetric Engineering & Remote Sensing*, vol. 85, no. 6, pp. 415-424, 2019.
- [113] Q. Yuan, Q. Zhang, J. Li, H. Shen, and L. Zhang, "Hyperspectral image denoising employing a spatial-spectral deep residual convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 2, pp. 1205-1218, 2018.
- [114] 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 Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4274-4288, 2018.
- [115] Z. Zhaoxiang, A. Iwasaki, and G. Xu, "Attitude Jitter Compensation for Remote Sensing Images Using Convolutional Neural Network," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 9, pp. 1358-1362, 2019.
- [116] J. Guo, Z. Pan, B. Lei, and C. Ding, "Automatic color correction for multisource remote sensing images with Wasserstein CNN," *Remote Sensing*, vol. 9, no. 5, p. 483, 2017.
- [117] H. Jiang and N. Lu, "Multi-scale residual convolutional neural network for haze removal of remote sensing images," *Remote Sensing*, vol. 10, no. 6, p. 945, 2018.
- [118] X. Yang, N. Wang, B. Song, and X. Gao, "BoSR: A CNN-based aurora image retrieval method," *Neural Networks*, vol. 116, pp. 188-197, 2019.
- [119] 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.

- [120] G. Chen, Q. Pei, and M. Kamruzzaman, "Remote sensing image quality evaluation based on deep support value learning networks," *Signal Processing: Image Communication*, vol. 83, p. 115783, 2020.
- [121] X. Zhang, Y. Sun, J. Zhang, P. Wu, and L. Jiao, "Hyperspectral unmixing via deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 11, pp. 1755-1759, 2018.

DRAFT