We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Deep Learning Training and Benchmarks for Earth Observation Images: Data Sets, Features, and Procedures

Mihai Datcu, Gottfried Schwarz and Corneliu Octavian Dumitru

Abstract

Deep learning methods are often used for image classification or local object segmentation. The corresponding test and validation data sets are an integral part of the learning process and also of the algorithm performance evaluation. High and particularly very high-resolution Earth observation (EO) applications based on satellite images primarily aim at the semantic labeling of land cover structures or objects as well as of temporal evolution classes. However, one of the main EO objectives is physical parameter retrievals such as temperatures, precipitation, and crop yield predictions. Therefore, we need reliably labeled data sets and tools to train the developed algorithms and to assess the performance of our deep learning paradigms. Generally, imaging sensors generate a visually understandable representation of the observed scene. However, this does not hold for many EO images, where the recorded images only depict a spectral subset of the scattered light field, thus generating an indirect signature of the imaged object. This spots the load of EO image understanding, as a new and particular challenge of Machine Learning (ML) and Artificial Intelligence (AI). This chapter reviews and analyses the new approaches of EO imaging leveraging the recent advances in physical process-based ML and AI methods and signal processing.

Keywords: Earth observation, synthetic aperture radar, multispectral, machine learning, deep learning

1. Introduction

This chapter introduces the basic properties, features, and models for very specific Earth observation (EO) cases recorded by very high-resolution (VHR) multispectral, Synthetic Aperture Radar (SAR), and multi-temporal observations. Further, we describe and discuss procedures and machine learning-based tools to generate large semantic training and benchmarking data sets. The particularities of relative data set biases and cross-data set generalization are reviewed, and an algorithmic analysis frame is introduced. Finally, we review and analyze several examples of EO benchmarking data sets.

In the following, we describe what has to be taken into account when we want to benchmark the classification results of satellite images, in particular the classification capabilities, throughputs, and accuracies offered by modern machine learning and artificial intelligence approaches.

Our underlying goal is the identification and understanding of the semantic content of satellite images and their application-oriented interpretation from a user perspective. In order to determine the actual performance of automated image classification routines, we need to find and select test data and to analyze the performance of our classification and interpretation routines in an automated environment.

A particular point to be understood is what type of data exists for remote sensing images that we want to classify. We are faced with long processing chains for the scientific analysis of image data, starting with uncalibrated "raw" sensor data, followed by dedicated calibration steps, subsequent feature extraction, object identification and annotation, and ending with quantitative scientific research and findings about the processes and effects being monitored in the geophysical environment of our planet with respect to climate change, disaster risks, crop yield predictions, etc.

In addition, we have to mention that free and open-access satellite products have revolutionized the role of remote sensing in Earth system studies. In our case, the data being used are based on multispectral (i.e., multi-color) sensors such as Landsat with 7 bands, Sentinel-2 [4] with 13 bands, Sentinel-3 with 21 bands, and MODIS with 36 bands but also SAR sensors such as Sentinel-1 [6], TerraSAR-X [26] or RADARSAT. For a better understanding of their imaging potential, we will describe the most important parameters of these images. For multispectral sensors, there exists several well-known and publicly available land cover benchmarking data sets comprising typical remote sensing image patches, while comparable SAR benchmarking data sets are very scarce and dedicated.

The main aspects being treated are:

- ML paradigms to support the semantic annotation of very large data sets, that is, using hybrid methods integrating Support Vector Machines (SVMs), Bayesian, and Deep Neural Networks (DNNs) algorithms in active learning paradigms by using initially small and controllable training data sets, and progressively growing the volume of labeled data by transfer learning.
- Proposing solutions to the semantic aspects of the spatial annotations for different sensor resolutions and spatial scales.
- Discussing the implications of the sensory and semantic gaps.

In this chapter, we assume that we can rely on already processed data with sufficient calibration accuracy and accurate annotation allowing us to understand all imaging parameters and their accuracy. We also assume that we can profit from reliably documented image data and that we can continue with data analytics for image understanding and high-level interpretation without any further precautions.

The latter steps have to be organized systematically in order to guarantee reliable results. A common strategy is to split these tasks into three phases, namely initial basic software functionality testing; second, training and optimizing of the software parameters by means of selected reference data, and finally, benchmarking of the overall software functionality such as processing speed and attainable results. This systematic approach leads to quantifiable and comparable results as described in the following sections.

During the last years, the field of deep learning had an explosive expansion in many domains with predominance in computer vision, speech recognition, and text analysis. For example, during 2019, more than 500 articles per month have been published in the field of deep learning. Thus, any reports on the state of the art hardly can follow this development. In Ref. [1], published in January 2019, more than 330 references were analyzed reviewing the theoretical and architectural aspects for Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), including Long Short-Term Memories (LSTMs) and Gated Recurrent Units (GRUs), Auto-Encoders (AEs), Deep Belief Networks (DBNs), Generative Adversarial Networks (GANs), and Deep Reinforcement Learning (DRL). The review paper [1] also summarizes 20 deep learning frameworks, two standard development kits, 49 benchmark data sets in all domains, from which three are dedicated to hyperspectral remote sensing. In addition, Ball et al. [2] describe the landscape of deep learning from all perspectives, theory, tools, applications, and challenges as of 2017. This article analyzes 419 references. A more recent overview from April 2019 [3] summarizes more than 170 references reporting on applications of deep learning in remote sensing.

2. Remote sensing images

Typical remote sensing images acquired by aircraft or satellite platforms can be characterized based on the operational capabilities of these platforms (such as their flight path, their capabilities for instrument pointing, and the on-board data storage and data downlink capacities), the type of instruments and their sensors (such as optical images with distinctive spectral bands [4, 5] or radar images such as synthetic aperture radars [6]), and opportunities for the repetitive acquisition of geographically overlapping image time series (for instance, for vegetation monitoring to predict optimal crop harvesting dates).

Current images can provide raw data with more than eight bits per sample, can perform initial data processing and annotation already on board, and can downlink compressed data with error correcting codes. After downlinking the image data to ground stations, the received data will be stored and processed by dedicated computing facilities. A common remote sensing strategy is to perform a systematic level-by-level processing (generating so-called products that comprise image data together with metadata documenting relevant image acquisition and processing parameters).

A common conventional approach is to follow a unified concept, where Level-0 products contain unprocessed but re-ordered detector data; Level-1 data represent radiometrically calibrated intensity images, while Level-2 data are geometrically corrected and map-projected data. Level-3 data are higher level products such as semantic maps or overlapping time-series data. In general, users have access to different product levels and can access and download selected products from databases via image catalogs and so-called quick-look (also called thumb-nail) images.

Some additional products have to be generated interactively by the users. Typical examples are image content classifications and trend analyses following mathematical approaches. Today, these interactive steps migrate from purely interactive and simple tools to commonly accepted machine learning tools. At the moment, the majority of machine learning tools use “deep” learning approaches; here, the problem is decomposed into several layers to find a good representation of image content categories [7]. These aspects will be dealt with in more detail in Section 4.

What we have to outline first are some important parameters of remote sensing images. One critical point of typical remote sensing images is their enormous size

calling for big data environments with powerful processors and large data stores. A second important point is the geometrical and radiometrical resolution of the image pixels, resulting in different target types that can be identified and discriminated during classification. While the typical pixel-to-pixel spacing of air-borne cameras corresponds to centimeters on the ground, space-borne instruments with high resolution flown on low polar orbits mostly lie in the range of half a meter to a few meters. In contrast, imaging from more distant geostationary or geosynchronous orbits results in low-resolution images. As for the number of brightness levels of each pixel, modern cameras often provide more than eight bits of resolution. **Table 1** shows some typical parameters of current satellites with imaging instruments.

Further, the pixels of an image can be complemented by additional information obtained by feature extraction and automated object identification (used as image content descriptors) as well as publicly available information from auxiliary external databases (e.g., geographical references or geophysical parameters). These data allow the provision of accurate quantitative results in physical units; however, one has to be aware of the fact that while many phenomena become visible, some internal relationships may remain invisible without dedicated additional investigations. **Table 1** shows some typical parameters of current satellite images.

In addition to the standard image products as described above, any additional automated or interactive analysis and interpretation of remote sensing images calls for intelligent strategies how to quickly select distinct and representative images, how to generate image time series, to extract features, to identify objects, to recognize hitherto hidden relationships and correlations, to exploit statistical descriptive models describing additional relationships, and to apply techniques for the annotation and visualization of global/local image properties (that have to be stored and administered in databases).

While typical traditional image content analysis tools either use full images, sequences of small image patches, collections of mid-size image segments or countless individual pixels together with routines from already established toolboxes (e.g., Orfeo [9]), or advanced machine learning approaches exploiting innovative machine learning strategies, as for instance, transfer learning [8] or the use of adversarial networks [10]. However, any use of advanced approaches requires the

High-resolution imaging instruments	**Optical cameras and spectrometers**	**SAR instruments**
Image size (typ. lines × columns)	$10^4 \times 10^4$ pixels	$10^4 \times 10^4$ pixels
Bands (typ.)	300 to 1000 nm and infrared bands	C-band, X-band, L-band, etc.
Spatial resolution (typ.)	Sub-meter to tens of m	Meters to several meters
Target areas (typ.)	Land, ocean, ice, atmosphere	Land, ocean, ice
Special modes (typ.)	Dynamical targeting stereo views fusion of bands	V and H polarization scan modes interferometry
Pixel types (typ.)	Detector counts reflectances	Complex-valued "detected data" (amplitudes or intensities)
Important parameters (typ.)	Number of overlapping bands	Viewing/incidence angle polar or geo. orbit

Table 1.
Typical imaging parameters of current satellites.

Clipping of outliers and de-noising
Color coding of brightness levels
Histogram manipulation (e.g., stretching)
Normalization and contrast enhancement
Box-car filtering (e.g., high-pass filtering, smoothing)
Transformations and filtering of coefficients
Analysis of pixel statistics and use of computer vision algorithms (e.g., histograms of gradients, local binary patterns, speeded-up robust features)
Feature extraction and classification (edges, corners, ridges, texture, color, interest points, shapes)
Extraction of content-oriented regions and objects

Table 2.
Typical capabilities of traditional image content analysis tools.

preparation and conduction of tests that allow a benchmarking of the new software routines, notably methods and tools to generate and analyze data for testing, training, verification, and final benchmarking. These testing activities have to be supported by efficient visualization tools.

As can be seen from **Table 2**, there exist already quite a number of traditional image content analysis tools. Some of them generate pre-processed images for subsequent analysis by human image interpreters, while others allow the identification and extraction of objects. However, these tools do not yet exploit the most recent automated machine learning techniques.

3. Machine learning, artificial intelligence, and data science for remote sensing

Currently, we see a lot of public interest in machine learning (ML), artificial intelligence (AI), and data science (DS). We have to make sure what we mean by these buzzwords:

- ML is often used if we describe technical developments where a computer system is trained and used to find and classify objects in data sets. A prominent example is the identification and interpretation of traffic signs for automated driving, typically use cases where a computer system is coupled with a camera and other sensors, and the traffic signs have to be recognized independent of different illumination and weather conditions, a vast range of potential driving speeds, varying distances and perspectives, other cars moving within the field of view of the camera, supplementary information provided by text panels or adjacent traffic signs, and constraints to be observed such as the maximum reasonable processing time. In essence, we can consider these applications as a reduction of many image pixels into single features (from a given list of cases and options) or a combination of features (e.g., max speed of 30 mph except on weekends). In most cases, the ML software is tested and trained by many typical examples as well as counterexamples.

- AI combines the full functionality of ML with additional decision-making and reaction capabilities. This additional decision-making can be implemented by continuous understanding of the current overall situation, the extraction of reactions from given rule sets (supported by continuously updated

parameters), and the handling of unexpected emergency cases. In the case of autonomous driving, one can think of a lane change on a motorway after a reason for a lane change has been found, and from a number of alternative reactions, when a lane change appears as the best reaction. Then the current situation has to be checked when a lane change becomes possible, and a sequence of subsequent actions is executed.

- DS as a scientific and technical discipline of its own shall provide all guiding principles that are needed from end-to-end system design up to data analytics and image understanding—including the system layout and verification, the selection of components and tools, the implementation and installation of the components and their verification, and the benchmarking of the full functionality. In the case of remote sensing applications, we also have to include all aspects of sensor calibration, comparisons with the findings of other researchers via Internet, and traceable scientific data interpretation.

As our applications mostly use cases dealing with remote sensing images, we can limit ourselves to the main ML paradigms that support the semantic annotation of very large data sets. Based on the current state-of-the art developments, we consider that there are three currently important fundamental and internationally accepted image classification approaches for remote sensing applications and two additional learning principles useful for satellite images:

- *Bayesian networks*: a Bayesian network consists of a probabilistic graphical model representing a set of variables together with their conditional dependencies. It can be used for parameter learning and is based on traditional formulas derived by Bayes [11].

- *Support Vector Machines* (SVMs): SVMs support classification and regression tasks by identifying basic support points that are used to define a robust separation plane between all sample points. In general, the resulting separation plane is a hyperplane with nonlinear characteristics. In order to obtain a separation plane with linear characteristics, the sample points are mapped into a higher-dimensional system with linear characteristics. This mapping exploits so-called kernel functions [12]. A well-known SVM software package is [13], which also explains how to train and verify a new SVM.

- *Neural Networks*: neural networks follow the concept of biological neurons that trigger a positive response if the input signal corresponds to a known object. Thus, technical implementations mostly consist of three levels, namely a visible input layer followed by an internal processing layer that is not visible to the user (in principle, an artificial neural network), and a visible output layer. An extension of general neural networks are deep neural networks; here, the processing layer is split into several linked internal sublayers that allow a more detailed analysis of the input data (e.g., on selected scales). The internal network parameters (i.e., the filter coefficients) are derived ("learned") by means of typical (and atypical) image samples and manual labeling by users [11].

- *Active learning*: this learning strategy combines automated learning with interactive steps involving the user during important decisions. This can be

accomplished by a visualization interface where a user can select or deselect image patches that do belong to or do not belong to a specific target class. For further details, see [14].

- *Transfer learning*: the idea of transfer learning is to train a network for a given task and then to exploit or "translate" the resulting network parameters to another use case. A typical example cited in [8] is the use of knowledge gained, while learning to recognize cars in images is applied when trying to recognize trucks.

One of the most critical points for satellite image classification is the dependence of the classification results on the resolution (pixel spacing) of the images. Experiences gained by many authors demonstrate that the identified classes and their local assignment within image patches are strongly resolution-dependent as higher resolution will often lead to a higher number of visible and identified semantic categories. Thus, the performance of any semantic interpretation of images must be considered as a data-dependent metric: this potential difficulty should prevent us from blind-folded direct performance comparisons.

Another similar point to be mentioned is the risk of sensory and semantic gaps encountered during image classification. Sensory gaps result from cases where a sensing instrument cannot measure the full range of potential cases with all their physical effects and details that could exist in a real-world scene and that we cannot record and identify with uniform confidence. A similar potential pitfall for image understanding can result from semantic gaps. For instance, during interactive labeling by test persons, different people could assign different categories to image patches due to their educational background, professional experiences, etc. For further details, see [15].

The number of available approaches, algorithms, and tools is growing continuously. Some examples have become very widespread in academia such as Caffe [16], TensorFlow [17], and PyTorch [18]. In contrast to these established solutions, a large number of fresh publications are submitted every day. As an example, the ArXiv preprint repository [19] collects in its "computer science" and "statistics" directories hundreds of new machine learning papers per day.

4. Networks for deep learning

Many experiments with image classification systems have shown that traditional single-level ("shallow") algorithms are less performant than multi-level ("deep") concepts where distinct filtering operations are applied on each level, and the results of the previous levels can be used on each deeper level; the final result will be obtained by combining the specific results of each separate level. The reason for the better performance of multi-level algorithms is that one can apply distinct filters specifically tailored to each level. Typical examples are multi-resolution filters that detect image characteristics on several scales: when we look at satellite images of urban settlements, then a business district normally has larger high-rise buildings and broader streets than a residential suburb with interspersed low-rise buildings and individual gardens.

From a high-level perspective, we can say that learning works best with deep learning approaches exploiting dedicated "network" structures. Here, we understand networks as design structures of the data flows and the arrangement of pixel handling steps governing the processing of our images. This concept also supports

more intricate label assignment concepts such as primary labels defining the main category of an image patch supplemented by secondary labels that provide additional information about "mixed classes" or supplementary spatial details of a given image patch.

In the meantime, some types of networks have emerged that have proven their robustness in the case of satellite images to be annotated semantically. In the following, we list four types of networks that have proven their usefulness for satellite image interpretation:

- *Deep Neural Networks* (DNNs): as described in [20], these networks consist of several layers and comprise an input layer, an output layer, and at least one hidden layer in between. Each layer performs dedicated pixel processing. The corresponding training phase can be understood as deep learning.

- *Recursive Neural Networks (not to be confused with recurrent neural networks; both network types appear as* RNNs*)*: when we have structured input data, these data can be efficiently handled by recursive neural networks that are often being used for speech processing and understanding. Recursive neural networks can also be used for natural scenes such as images containing recursive structures [23]. RNN algorithms identify the units that an image contains and how the units interact. Thus, one can use RNNs for semantic scene segmentation and annotation.

- *Convolutional Neural Networks* (CNNs): these networks have been conceived for low-error classification of big images with a very large number of classes. As described by [21], one can classify more than a million images and assign more than 1000 different classes. This is accomplished internally by five convolutional layers, three fully connected layers, and a million internal parameters. To reduce overfitting, the method applies regularization by disregarding offending elements ("dropout method").

- *Generative Adversarial Networks* (GANs): an adversarial network allows the mutual training of two competing multilayer perceptron models *G* and *D* following an adversarial process: *G* determines the data distribution, while *D* estimates the probability that a sample comes from training data rather than from *D*. In addition, D maps the high-dimensional input data to semantic category labels. For further details, see [24].

Besides the network types listed above, we also need an overall algorithmic architecture embedding the networks. For our applications, a "U" approach has proven to be a useful concept for satellite image content analysis. A "U" approach contains a descending branch followed by an ascending branch and is conceived for handling a progressively shrinking number of elements until a final core element (a main category) is found, followed by stepwise complementary semantic information. Further details can be found in [21].

In our experience, most general remote sensing applications can be solved efficiently by CNNs or similar approaches. However, quite a number of innovative alternatives have been proposed during the last years, for example, common auto-encoders, recursive approaches for time series, and adversarial networks for fast learning with only a few examples. In our case, we suggest to use CNNs for non-critical satellite image applications, while highly complicated or time-critical applications could call for innovative approaches as already described above.

5. Training and benchmarking

When we train a classification network and verify its performance, the main goal is to train the system for correct category assignments resp. semantic annotations (labels), that is, to add supplementary information to each satellite image patch that we analyze.

The semantic annotations can either be learned in a preparatory phase or be taken from catalogs of already existing categories. If we aim at long-term analyses of satellite images, a good approach is to use the same catalogs during the entire lifetime of the analysis or to re-run the entire system with updated catalogs.

The easiest approach is to select typical examples for each category and to assign the given labels to all new image data. However, if we follow this straightforward approach, we will probably encounter some difficulties when image patches with unexpected content arrive. A first remedy is to add an additional "unknown" category and to assign this label to all image patches that do not fit well to one of the given categories. Further, experience with machine learning systems has shown that good classification results can also be reached when we systematically select positive as well as negative examples (i.e., counterexamples) for each category leading to a comprehensive coverage and understanding of each category. This process can be accomplished manually by knowledgeable operators (i.e., image interpretation experts) [22]. Another approach is data augmentation: If we do not have sufficient examples of a necessary category, one can create additional realistic data by simply flipping or rotating already available images.

This simple example leads us to systematic methods for a database creation. One has to find a comprehensive and fairly balanced set of examples that covers the expected total variety of cases. Thus, we avoid so-called database biases [23]. In addition, one has to make sure that the inclusion of additional examples does not lead to overfitting or excessive runtimes. This can be accomplished by setting up a validation testbed where these potential pitfalls can be tested, trained, and where the final performance of the created database structure can be verified. One has to be aware of the fact that database access times may strongly depend on the available computer systems, their interconnections, and the selected type of database.

These approaches led to a number of publicly available databases with label annotations for civilian remote sensing data. There are several semantically annotated databases based on optical (most often multispectral) data, while there are only a few databases based on SAR data. Some advanced remote sensing database examples are [25–27]. Of course, their general applicability and transferability depend on the actual image resolution, the imaging geometry, and the noise content of the images. Current state-of-the-art systems are being assessed based on end-to-end tests covering also inter alia practical aspects such as the runtime depending on the database design and the selected test images, the amount and organization of available labels, the correctness of the obtained annotations, and the overall implementation and validation effort.

6. Perspectives

As for remote sensing images, there exist already several semantically annotated collections of typical high-resolution satellite images—a number of collections of optical images and a few collections of SAR images. However, these collections often seem to be potpourris of interesting snapshots rather than systematically selected samples based on regionally typical target classes and their visibility as a

function of different instrument types. The situation is aggravated by the current lack of systematically selected benchmarking data that could be used as well-known reference data for quality and performance assessments such as classification tasks or throughput testing.

These deficiencies have to be solved in the near future as more and more high-resolution images become publicly available, while the end-users already expect reliable automated image classification and content understanding results for more and more high-level applications. We can expect that the progress in deep learning will also lead to much progress in many other fields of image processing, even beyond the field of remote sensing; thus, remote sensing should be aware of what is published by the image processing and environmental protection communities at large.

7. Conclusions

While high-resolution imaging has made much progress for many remote sensing applications, standardized image classification benchmarking still deserves more progress. On the one hand, several benchmarking concepts and tools could still be gleaned from other disciplines; on the other hand, an optimal solution of test cases for SAR image interpretation still needs more progress in basic approaches of how to verify actual image classification results and the identification of dubious cases.

Acknowledgements

We appreciate the cooperation with Politehnica University of Bucharest (UPB) in Romania and our project partners from the European H2020 projects CANDELA (under grant agreement No. 776193) and ExtremeEarth (under grant agreement No. 825258).

Author details

Mihai Datcu[1,2*], Gottfried Schwarz[1] and Corneliu Octavian Dumitru[1]

1 German Aerospace Center (DLR), Remote Sensing Technology Institute, Wessling, Germany

2 Politehnica University of Bucharest, Bucharest, Romania

*Address all correspondence to: mihai.datcu@dlr.de

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] Alom Z, Taha T, Yakopcic C, Westberg S, Sidike P, Nasrin S, et al. State-of-the-art survey on deep learning theory and architectures. Electronics. 2019;**8**:292. Available at: https://www.mdpi.com/2079-9292/8/3/292

[2] Ball J, Anderson D, Chan CS. A comprehensive survey of deep learning in remote sensing: Theories, tools and challenges for the community. Journal of Applied Remote Sensing. 2017;**11**(4):042609. Available at: https://arxiv.org/abs/1709.00308

[3] Ma L, Liu Y, Zhang X, Ye Y, Yin G, Johnson B. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing. 2019;**152**:166-177. Available at: https://www.sciencedirect.com/science/article/pii/S0924271619301108

[4] ESA Sentinel-2. Available at: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2 [Accessed: April 2019]

[5] ESA Sentinel-3. Available at: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-3 [Accessed: April 2019]

[6] ESA Sentinel-1. Available at: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1 [Accessed: April 2019]

[7] Deep learning. Available at: https://www.deeplearningbook.org [Accessed: April 2019]

[8] Transfer learning. Available at: https://en.wikipedia.org/wiki/Transfer_learning [Accessed May 2019]

[9] Orfeo Toolbox. Available at: https://www.orfeo-toolbox.org/ [Accessed May 2019]

[10] Adversarial machine learning. Available at: https://en.wikipedia.org/wiki/Adversarial_machine_learning [Accessed May 2019]

[11] Bayesian network. Available at: https://en.wikipedia.org/wiki/Bayesian_network [Accessed May 2019]

[12] Support vector machine. Available at: https://en.wikipedia.org/wiki/Support-vector_machine [Accessed May 2019]

[13] LIBSVM -- A Library for Support Vector Machines. Available at: https://www.csie.ntu.edu.tw/~cjlin/libsvm/ [Accessed May 2019]

[14] Active learning. Available at: https://en.wikipedia.org/wiki/Active_learning [Accessed May 2019]

[15] Bahmanyar R, Murillo A. Evaluating the sensory gap for earth observation images using human perception and an LDA-based computational model. In Image Processing (ICIP), 2015 IEEE International Conference on pp. 566-570

[16] Caffe software. Available at: https://caffe.berkeleyvision.org/ [Accessed May 2019]

[17] TensorFlow. Available at: https://www.tensorflow.org [Accessed: March 2019]

[18] PyTorch. Available at: https://pytorch.org/ [Accessed: March 2019]

[19] arXiv e-Print archive. Available at: https://arxiv.org/ [Accessed: March 2019]

[20] Krizhevsky A, Sutskever I, Hinton GC. ImageNet Classification with Deep Convolutional Neural Networks. Available at: https://papers.nips.cc/4824- imagenet-with-deep-convolutional-neural-networks.pdf

[21] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation, Medical Image Computing and Computer-Assisted Intervention (MICCAI). Vol. 9351. Basel, Switzerland: Springer International Publishing, LNCS; 2015. pp. 234-241, Available at: https://arXiv:150504597

[22] Murillo Montes de Oca A, Bahmanyar R, Nistor N, Datcu M. Earth observation image semantic bias: A collaborative user annotation approach. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS). 2017;**10**(6):2462-2477. DOI: 10.1109/JSTARS.2017.2697003

[23] Socher R, Lin C, Ng AY, Manning, CD. Parsing Natural Scenes and Natural Language with Recursive Neural Networks, 28th International Conference on Machine Learning (IVML 2011). Available at: https://nlp.stanford.edu/pubs/SocherLinNgManning_ICML2011.pdf

[24] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Nets. Available at: http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

[25] Xia G-S, Bai X, Ding J, Zhu Z, Belongie S, Luo J, et al. DOTA: A large-scale dataset for object detection in aerial images. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, Utah, USA. 2018

[26] Dumitru CO, Schwarz G, Datcu M. SAR image land cover datasets for classification benchmarking of temporal changes. IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing (JSTARS). 2018;**11**(5):1571-1592

[27] Sumbul G, Charfuelan M, Demir B, Markl V. BIGEARTHNET: A large-scale benchmark archive for remote sensing image understanding. In: IEEE International Conference on Geoscience and Remote Sensing Symposium (IGARRS), Yokohama, Japan. 2019