

Creation of Specialized Datasets for Satellite Monitoring of the State of Agricultural Lands

Grishkin Valery(✉), Zhivulin Evgeniy, and Khokhriakova Anastasiia

Saint-Petersburg State University, St.Petersburg 199034, Russia
v.grishkin@spbu.ru

Abstract. Currently, there is significant progress in the interpretation of remote images of the earth's surface using various recognition systems. To train these systems, large labeled datasets are required. The creation of such datasets is carried out by experts and is quite a time-consuming task. This paper proposes a method for creating specialized datasets for their use in monitoring the state of agricultural lands undergoing degradation. The method allows you to automate the process of creating such datasets The datasets are generated from freely available data obtained by the Sentinel satellites as part of the European Copernicus space program. The method is based on the processing of the results of the preliminary classification of images of the earth's cover, previously produced by the Sentinel Hub service.

Keywords: Remote sensing · Satellite image segmentation · Multispectral images · Scene classification · Training dataset

1 Introduction

Currently, remote sensing technologies allow for monitoring the state of the earth's surface. Such monitoring makes it possible to assess the state of the environment, explore natural resources, and carry out development planning.

The most successful recognition systems are based on the supervised learning method, which requires a large enough labeled dataset. Currently existing datasets for recognition of remote images are not universal and, as a rule, are focused on use in specialized applications. Therefore, to solve specific monitoring tasks, one has to either try to use existing data sets, or form specialized data sets intended for one or another type of monitoring.

There are two main approaches to solving the problems of identification and segmentation of satellite images. The first one is a method based on a pixel analysis of the intensity of spectral characteristics and corresponding derivative features. The second approach is based on the spatial characteristics of spectral images and is implemented using convolutional neural networks. To implement both approaches, a tagged set of multispectral images is required, taken with a certain periodicity. This is necessary to derive the various characteristics of the

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12956, pp. 406–416, 2021.
https://doi.org/10.1007/978-3-030-87010-2_30

images specific to certain seasons. Currently, there are a number of datasets [1,2] that contain satellite images of selected regions and are designed to meet the challenges of urban and road planning. There are also datasets [3–6] used for assessing and monitoring agricultural, forestry and water resources. They contain satellite images obtained in the visible and near infrared ranges, and some sets also contain multispectral images. Among these sets, the universal dataset SEN12MS [7] deserves special consideration, as it contains a large number of triplets of data types for various regions of the earth's surface. Each triplet of data obtained during all meteorological seasons includes polarimetric images, multispectral images and maps of land cover areas associated with them. These sites are fully georeferenced and are located on all inhabited continents. The information about these areas also contains their markup into land surface classes in accordance with such classification schemes as the International Geosphere-Biosphere Program (IGBP) and the Land cover classification system (LCCS).

Universal data sets are primarily designed to test and compare the performance of different algorithms for the identification of areas of the earth's surface. In the case of monitoring a specific region with its own specific land cover, universal datasets cannot provide sufficient characteristics of the quality of identification of areas in a given region, because this particular region may not be represented in these datasets. Therefore, to solve these problems, it is advisable to use specialized datasets generated for a specific region.

In this paper, we consider the creation of specialized datasets for their use in monitoring the state of agricultural land in certain areas undergoing sand degradation. Within the framework of this monitoring, it is necessary to identify zones of sands (deserts), considering that these zones are completely degraded and cannot be used in agriculture. It is also necessary to identify both transition zones in these areas, subject to varying degrees of degradation, and zones fully suitable for use in agriculture.

2 Data Acquisition and Preprocessing

2.1 Input Data

As input data, on the basis of which a specialized set is built, the data of multispectral imagery from the Sentinel-2 [8] satellites and the results of their preliminary processing are used. These data are freely available on the Copernicus Open Access Hub [9]. Images of the Earth's surface are taken at intervals of 5 days. Surface imagery from satellites is carried out in different spectral ranges with a resolution not better than 10×10 m in one pixel of the image. Table 1 provides a list of spectral channels used to construct the dataset. In addition, it contains channels that reflect the results of preliminary processing of images taken in the Sentinel data center. It should be noted that for the formation of specialized data sets, we use only 11 channels out of 23 channels available.

The first three channels represent reflected light in the optical range. The next three channels are infrared radiation. Then there are two near-infrared channels. The next spectral range is associated with reflection from water vapor.

Table 1. Wave range and maximum resolution of the channels used.

Channel	Description	Wavelength	Resolution
B02	Blue	492.4 nm	10 m
B03	Green	559.8 nm	10 m
B04	Red	664.6 nm	10 m
B05	Vegetation red edge	704.1 nm	20 m
B06	Vegetation red edge	740.5 nm	20 m
B07	Vegetation red edge	782.8 nm	20 m
B08	NIR	832.8 nm	20 m
B08A	Narrow NIR	864.7 nm	20 m
B09	Water vapour	945.1 nm	20 m
B11	SWIR	1613.7 nm	20 m
SNW	Snow probability	–	20 m
CLD	Cloud probability	–	20 m
SCL	Scene classification data	–	20 m

The last spectral channel is shortwave infrared data. Then three channels follow, reflecting the results of data processing by the Sentinel-hub service itself. The first one is the probability that a given pixel in the image is part of a snow-covered area, and the second one is the probability that it belongs to a cloud cover. The last of these channels is an image mask describing the belonging of each pixel to one of the 12 classes. Among these classes are the following: vegetation, bare soil, water, clouds, cloud shadows, snow.

2.2 Loading Input Data

Interaction with the Copernicus Open Access Hub website is carried out using requests that are supported by the corresponding Sentinel-hub service. The request itself is a Javascript code, which indicates: a date range (the service provides the latest snapshot from the specified range), coordinates of the area of interest (coordinates of the lower left and upper right corners of the rectangular area of interest), the height and width of the resulting image in pixels, a list of channels from which data is requested and the format in which they will be loaded.

The coordinates of the requested area are indicated in the UTM system. It is a coordinate system that divides the northern and southern hemispheres into 60 zones. The accuracy of this system is higher than the usual latitude and longitude. The size of the requested area is chosen equal to 10×10 km, which corresponds to the dimensions of the resulting image equaling to 512×512 pixels when using a channel resolution of 20 m.

Before loading data, it is possible to pre-process it directly on the site by performing arithmetic operations on it. These operations allow for normalizing

the values in different channels. Data from channels is requested in the form of standard PNG images. Each image includes 3 channels of data displayed as red, green and blue. Since 13 channels are used, 5 images are loaded for any given area and date. The first image represents the visible part of the spectrum in natural colors. The next three images represent the data from the remaining channels, displayed on these images in pseudo colors. The last image is the grayscale displayed classification mask.

Downloading the data required for exploring large areas also requires a large number of downloads. Therefore, a specially developed script is used to load data. This script generates a request for satellite data for a specified location within a specified date period.

To automate loading, an array with coordinates saved in the configuration file is passed to the script, indicating areas on the map from which images are to be loaded. These coordinates are created by another script that splits the specified area into a set of squares. The division of the area of interest into sections is carried out after translating the boundary coordinates of the entire area in the form of latitude and longitude into the UTM system. This improves the precision of the split. Otherwise, when translating each section into UTM separately, there may be errors in neighboring sections, and the points that should coincide will differ slightly, which will be noticeable in the final images.

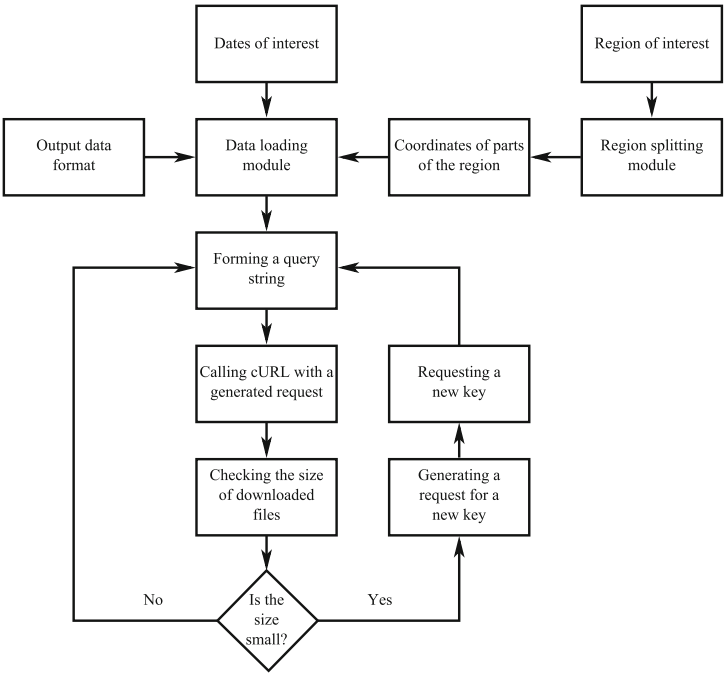

Fig. 1. The algorithm for downloading data from the Sentinel-hub service.

When requesting data from the service, the access key is transmitted which has been obtained earlier. This key is updated hourly and can be obtained as a result of another request to the service. Therefore, when exchanging with the hub, the size of uploaded images is checked, and if it is too small, this indicates that the service did not provide data, therefore, you need to request a new key is to be requested, and the script takes care of that.

For interaction with the site, the standard cURL utility [10] is used, which is part of almost all operating systems. After the request is generated, the script is passed to this utility as a parameter. The algorithm for downloading data from the Sentinel-hub service is shown in Fig. 1

2.3 Preliminary Processing

For various reasons, the resulting images of adjacent parts of the region may be slightly displaced relative to each other. Therefore, they are loaded with a small partial intersection of neighboring images. This is done in order to have a safety margin in case the images are displaced in different directions relative to each other, so as not to lose some of the information. In addition, such an intersection allows for use of the reference point method for the alignment of adjacent images. The essence of this method is to combine a set of points on both images so that all points coincide pairwise. The image displacement is determined based on the calculation of a two-dimensional discrete correlation function of the intersecting parts of adjacent images. In this case, the maximum value of the correlation function corresponds to the true shift of the images relative to each other.

However, the calculation of the correlation function for all intersecting points of the images is too time-consuming, therefore such a calculation was performed only for each n-th pixel, where n was taken equal to 20. This yields reliable results when combining images with edges of uneven color distribution, for example, when green and sandy zones alternate on the earth's surface. In case of a relatively uniform coverage, for example, for a sand only area, this solution yields a wrong result. Therefore, before calculating the correlation function of the image, it is necessary to increase the contrast at the edges of the combined images. We use the following algorithm. First, the median values of all the color channels in the area where the images intersect are determined. Then the pixel values for each color channel are compared to the median values. Pixel values less than the median value are multiplied by a correction factor k less than 1 ($k < 1$). For all other pixel values, the correction factor k is chosen to be larger than 1 ($k > 1$). Such processing made it possible to sharply increase the contrast, which, in turn, made it possible to compare even fairly uniform images at low costs for calculating the correlation function.

The procedure for combining adjacent images is applied only to images obtained in the visible range. The resulting shift is then applied to all images that display data for all other channels. The result of combining adjacent images with increased edge contrast is shown in Fig. 2. In this merged image, one can see that the roads in the contrasting area merge well into roads to the left and to the right from the contrasting area.

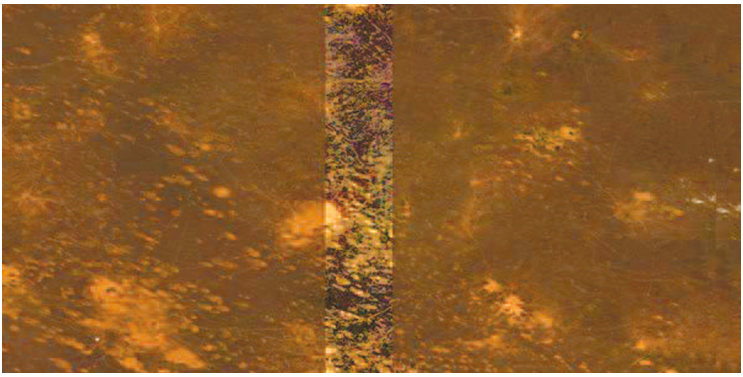

Fig. 2. The result of the algorithm for combining two adjacent images with a demonstration of increasing the contrast of the edges.

3 Creating a Specialized Dataset

To form a dataset suitable for the monitoring purposes, we use satellite data from the spring-summer and late autumn periods. The first period was chosen because it is in this period that the vegetation cover is the strongest and is well determined from satellite images using standard methods. The second period is characterized by the completed development of vegetation, and in these images, the surface that was covered with vegetation looks like soil. Thus, by comparing these images, it is possible to draw conclusions about the location and presence of productive soils in the area under observation, since the soil on which nothing grows is either degraded or unsuitable for agricultural purposes.

The requested data contains, among other things, the results of the preliminary classification of images obtained during data processing by the Sentinel Hub [7] service. This classification result is tied to each requested area of the earth's surface and is represented as a pixel mask M_{scl} of the image of this area. The values of each pixel in this mask correspond to specific types of ground surface or types of clouds. Mapping of mask pixel values to recognized types of objects is shown in Table 2.

This mask is used to segment areas containing agricultural soils. These areas are covered with vegetation in the spring and summer. Therefore, by segmenting such areas on images from the late autumn period, we will obtain data on the distribution of fertile soils. Thus, it is necessary to form two new masks out of the original summer classification mask. The $M_{\text{summer_veg}}$ mask represents the vegetation cover only, while the $M_{\text{summer_soil}}$ mask represents the soil. Two similar masks are also formed out of the autumn classification mask. The first $M_{\text{autumn_soil}}$ mask represents the soil cover, and the second $M_{\text{autumn_veg}}$ mask represents the vegetation cover. The first mask is directly used for the generation of a dataset for recognizing fertile soil cover. The M_{goodsoil} fertile soil mask is calculated by applying a bitwise AND to the summer vegetation mask and the autumn soil distribution mask.

Table 2. Scene classification data.

Value	Recognized type	Mask color
0	No data	
1	Saturated/Defective	
2	Dark Area Pixels	
3	Cloud Shadows	
4	Vegetation	
5	Bare Soils	
6	Water	
7	Clouds low probability/Unclassified	
8	Clouds medium probability	
9	Clouds high probability	
10	Cirrus	
11	Snow/Ice	

$$M_{\text{goodsoil}} = M_{\text{summer_veg}} \,\&\, M_{\text{autumn_soil}}$$

The second mask contains additional information about the location of the fertile soil hidden under the vegetation. This mask can be used to obtain a complete mask of fertile soil $M_{\text{f_goodsoil}}$.

$$M_{\text{f_goodsoil}} = M_{\text{goodsoil}} \Big| \left(M_{\text{summer_soil}} \,\&\, M_{\text{autumn_veg}}\right)$$

Figure 3 shows satellite images of the same region in the visible range in the spring-summer and autumn periods, the corresponding classification masks and the resulting masks of fertile soils. Figure 3c shows a snapshot of an area whose surface is partially covered by clouds. Therefore, the masks shown in Fig. 3e and Fig. 3f do not fully describe the distribution of fertile soils. It is impractical to use such masks and images for monitoring purposes without filtering cloud images. However, if clouds and their shadows from them do not take up too much of the total area of the image, then these images and masks can be used to create a specialized dataset for fertile soil recognition systems. This is possible because the resulting mask shows the distribution of fertile soils in an area not covered by clouds, and thus it can be used as a markup for training the recognition system. The algorithm for creating a specialized dataset for training recognition systems that can be used when monitoring fertile soils is shown in Fig. 4.

Data for a selected area of the earth's surface is requested for both the period of greatest vegetation development and the period when plant development stops and harvest is completed. Typically, data for the first period is for late July and early August, and data for the second period is requested for the end of October. Since the selected area is divided into subareas with a slight overlap, preprocessing of the data is performed, as a result of which the boundaries of neighboring

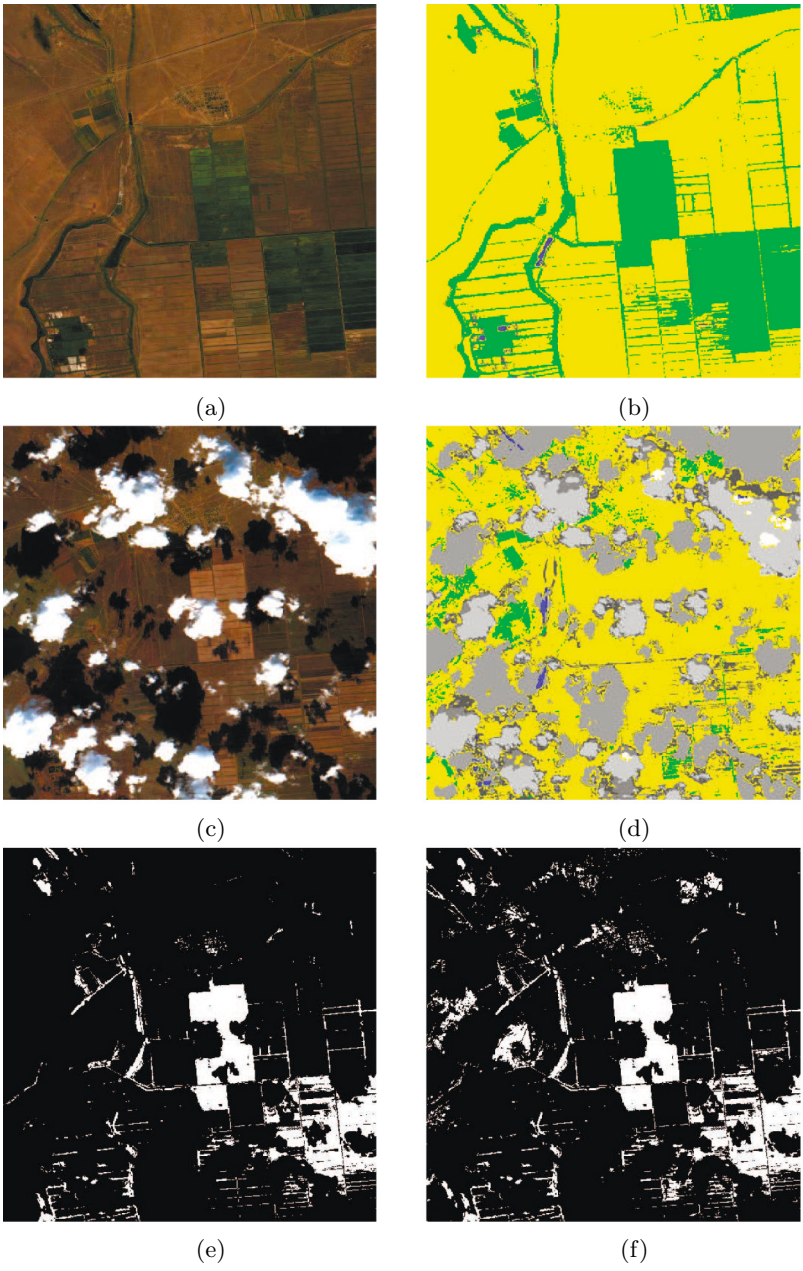

Fig. 3. Images of the region and corresponding masks: (a) image of the region—August; (b) classification mask—August; (c) image of the region—October; (d) classification mask—October; (e) the fertile soil mask without autumn vegetation; (f) the complete fertile soil mask vegetation.

subareas are aligned. Then, using cloud masks, the possibility of including this data in a specialized dataset is determined. If clouds cover more than half of the sub-area, then the data of this sub-area will not be included in the dataset. Further, on the basis of vegetation masks and soil masks, masks of fertile soils are calculated. The computed masks of fertile soils along with the spectral images of the subdomain are added to the dataset.

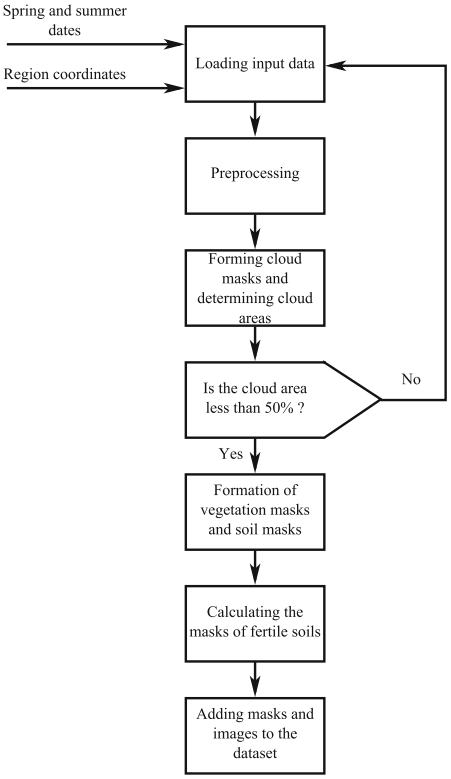

Fig. 4. Algorithm for creating a specialized dataset.

4 Experimental Results

The proposed algorithm was used to generate a specialized dataset. The data were requested for a specific region, broken down into subregions. The dataset was formed from 379 pairs of combined spectral images of subregions. The dataset was formed from 379 pairs of sets of combined spectral images of subregions. Each pair of sets for the same subregion consists of four combined images taken in August and four combined ones taken at the end of October. As a result of filtering out images with sufficiently large clouds, 159 such pairs of sets were

included in the dataset itself. For these pairs of images, masks of fertile soils were calculated, which were also included in the dataset.

The created dataset is intended for a monitoring system for the state of agricultural land. It was used to train land cover type recognition systems. Two recognition systems were investigated. Both recognition systems are based on the support vector machine (SVM) [11]. As a markup of the input images, the corresponding masks of fertile soils were used. The training was carried out on 93 pairs of image sets. The rest of the pairs from the dataset were used for testing.

In the first system, the normalized values of 10 spectral channels B02–B11 were used as the feature vector. In the second system, normalized differential indices V_{ij} were used as features, calculated for combinations of different spectral channels. The calculation of the corresponding index V_{ij} was carried out similarly to the procedure for calculating the known vegetation index NDVI.

$$V_{ij} = \frac{B_i - B_j}{B_i + B_j}.$$

The recognition accuracy of the soil type for the first system was 82%. In an experimental study of the second system, it turned out that the best soil type recognition accuracy was 94%. This accuracy was achieved using a feature vector of 8 indices. In this case, the following indices were used: $V_{47}, V_{46}, V_{89}, V_{38}, V_{92}, V_{01}, V_{48}, V_{57}$. These indices are calculated using the following channel pairs: (B06 B8A), (B06 B08), (B09 B11), (B05 B09), (B8A B04), (B02 B03), (B06 B09), (B07 B8A).

5 Conclusion

The paper proposes a method for the automated creation of specialized datasets of satellite images for their use in monitoring the state of agricultural lands undergoing degradation. The datasets are generated from freely available data obtained by the Sentinel satellites as part of the European Copernicus space program. This data includes multispectral imagery and preliminary classification of land surface types in the region of interest. Based on the preliminary classification masks of the summer and autumn periods, the mask of fertile soils is calculated, which, together with the spectral images of the subregion, is added to the generated dataset.

To test the generated dataset, a soil type recognition system was built and investigated. The system works on the basis of a support vector machine. The results of an experimental study of the system showed a sufficiently high accuracy of recognition of fertile and infertile soils. This allows us to speak about the adequacy of the generated dataset, which can be used to develop and test soil type recognition systems based on other principles. Thus, the proposed method made it possible to automatically create specialized datasets for monitoring agricultural lands.

References

1. Ji, S., Wei, S., Lu, M.: Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set. IEEE Trans. Geosci. Remote Sens. **57**(1), 574–586 (2018)
2. Volpi, M., Ferrari, V.: Semantic segmentation of urban scenes by learning local class interactions. In: Proceedings IEEE Conference Computer Vision and Pattern Recognition Workshop, pp. 1–9, June 2015
3. Xia, G.-S., et al.: AID: a benchmark data set for performance evaluation of aerial scene classification. IEEE Trans. Geosci. Remote Sens. **55**(7), 3965–3981 (2017)
4. Helber, P., Bischke, B., Dengel, A., Borth, D.: EuroSAT: a novel dataset and deep learning benchmark for land use and land cover classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **12**(7), 2217–2226 (2019)
5. Cheng, G., Han, J., Lu, X.: Remote sensing image scene classification: benchmark and state of the art. Proc. IEEE **105**(10), 1865–1883 (2017)
6. Tong, X.-Y., et al.: Land-cover classification with high-resolution remote sensing images using transferable deep models. Remote Sens. Environ. **237**, 111322 (2020)
7. Schmitt, M., Hughes, L.H., Qiu, C., Zhu, X.X.: SEN12MS- a curated dataset of georeferenced multi-spectral sentinel-1/2 imagery for deep learning and data fusion. arXiv preprint arXiv:1906.07789 (2019)
8. ESA Earth Observation Portal. https://directory.eoportal.org/web/eoportal/satellite-missions/c-missions/copernicus-sentinel-2. Accessed 20 Apr 2021
9. Copernicus Open Access Hub. https://scihub.copernicus.eu/dhus/. Accessed 8 Apr 2021
10. Command line tool and library for transferring data with URLs. https://curl.se/. Accessed 15 Apr 2021
11. Cristianini, N., Shawe, T.J.: An Introduction to Support Vector Machines and other Kernel-based Learning Methods. Cambridge University Press, Cambridge (2000)