MOUNTAIN GREEN COVER INDEX CALCULATION AT A NATIONAL SCALE USING WEAK AND SPARSE DATA

Natalia Verde[1], Petros Patias[1], Giorgos Mallinis[1]

[1]School of Rural and Surveying Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

ABSTRACT

Mountain ecosystems, essential for global biodiversity and human well-being, undergo rapid changes due to climate change and human influence. The Mountain Green Cover Index (MGCI) measures the percentage of green vegetation across mountain areas, aiming to assess and monitor the health and conservation of mountain ecosystems. MGCI requires ongoing monitoring and distinction by land cover (LC) type, however, challenges arise from the limited temporal and spatial resolution of available LC products. In addition, existing methods face issues due to complex terrain in mountains. This study leverages Deep Learning on Sentinel-2 data to overcome these limitations, and map LC categories for the MGCI at a national scale. Employing weak and sparse labels from the CORINE Land Cover dataset, the approach alleviates the need of large-scale manual sample collection and provides a roadmap for national statistical services to compute the MGCI.

Index Terms— land cover, Sentinel-2, sparse annotations, weakly supervised learning, deep learning

1. INTRODUCTION

Mountain ecosystems embody most of the world's biodiversity and provide many ecosystem services, contributing directly and indirectly to the well-being of at least half of the world's human population [1]. Under the influence of both climate change and human activities, mountain ecosystems are undergoing rapid changes and therefore monitoring their health is of significant importance for their conservation [2], [3].

The United Nations (UN) targets the preservation of mountain ecosystems by 2030 through Sustainable Development Goal (SDG) 15, emphasizing biodiversity conservation and sustainable development. The pivotal monitoring tool is the Mountain Green Cover Index (MGCI), specifically Indicator 15.4.2, which involves determining the percentage of green vegetation across mountain areas (forests, grasslands, croplands and wetlands). It is calculated using the following formula [4]:

$$MGCI = \frac{mountain\ green\ cover\ area}{total\ mountain\ area} \times 100$$

While MGCI data is available on national, regional, and global scales, current calculation systems have limitations. The Indicator's static nature necessitates continuous monitoring for trend analysis. The temporal and spatial resolution of off-the-shelf land cover (LC) products used impedes accurate tracking of mountain changes. In addition, defining "green" poses challenges, as cultivated areas may misrepresent mountain health, emphasizing the importance of distinguishing between LC types for precise monitoring [2], [4]–[6].

Remote sensing satellite data is pivotal for assessing vegetation in challenging mountainous terrains and can allow spatial and temporal separation of the MGCI [2]. However, despite these advantages, mapping LC in mountainous areas remains a challenge due to topographic complexities causing shadows and illumination variations. Challenges include limited samples in rugged, inaccessible areas, data loss from cloud cover, and differences in vegetation phenological cycles at different elevations [7]–[10]. Consequently, traditional Machine Learning (ML) methods face accuracy issues due to the diverse mountainous spectra. To address these challenges, previous studies integrated Digital Elevation Models, Synthetic Aperture Radar, and advanced time-series algorithms [7], [11]–[13], however requiring temporal data and feature engineering. Recently developed Deep Learning (DL) methods can overcome these limitations and alleviate the need of manual sample collection to cover spectral variations within classes. In contrast to traditional ML methods, DL algorithms can handle complex terrain objects with significant variation and are invariant to relatively large distortions, including different shapes, sizes, rotations, and lighting conditions. They can discern objects belonging to different categories but appearing similar in some other aspect [14], [15]. However, the lack of a large, tailored training dataset for mountainous regions remains a challenge for DL methods [8]. Weak supervision in DL, involves building predictive models with either i) a sparse number of reference data, ii) inexact information, or iii) reference data containing errors [16].

Specifically, for MGCI, two recent studies were conducted on a global and national scale using Landsat satellite images and phenology based on the NDVI index [2], [5]. Another study [6] utilized LC CCI products and the mountainous areas level from [17], as recommended by the UN methodology. Few studies have explored the contribution of DL algorithms to LC mapping in mountainous areas in general, either using additional image segmentation techniques [8] or multi-sensor temporal images [15].

This study focuses on mapping LC categories required for Indicator 15.4.2 in mountainous areas using Sentinel-2 data at a national scale. DL is employed for classifying the images, utilizing weak and sparse labels from the CORINE Land Cover (CLC) product.

To facilitate the demanding training of the DL network in this study, freely available computational resources from Google were used, specifically through Google Colaboratory which provides direct access to the Google Earth Engine (GEE) cloud platform. The Python code developed utilizes the Tensorflow and keras libraries.

2. MATERIALS AND METHODS

2.1. Study Area and Data

The study area was the entire terrestrial region of Greece, extending by 2km to cover parts of the sea and land that extend into neighboring regions. Greece covers an approximate area of 131,957 km^2, and geomorphologically, 78% of the area is covered by mountains according to [18].

Manual collection of samples for national-scale mapping, especially for DL algorithms that require a large amount of data, would require tremendous efforts. However, available LC products can be used as data sources to address the problem, with a suitable class correspondence and a careful sampling process. In this direction, for the creation of reference samples, the most recent (2018) CORINE Land Cover (CLC) dataset was utilized, for automated reference data creation, consistent with other national-scale studies. The CLC dataset covers the entire European region, including Greece, it is updated approximately every 4 years and is also freely accessible. With a Minimum Mapping Unit of 25 hectares, it describes land use and land cover with a set of 44 categories organized into three hierarchical levels. The specific choice of this land cover product over others was motivated by the simultaneous coverage of all categories required by Indicator 15.4.2, with high reliability, despite its relatively low spatial accuracy.

Copernicus Sentinel-2 satellite data, providing continuous free data, was used in this study to develop a methodology that could be easily applied in other reference periods. The Sentinel-2 data is available in Analysis Ready Data (ARD) format in GEE. All L2A images with cloud cover below 20% were selected, covering the study area from May 2018 to September 2018, to align with year of the CLC data used to create reference data. The late-spring/summer period was considered sufficient for capturing general land cover categories, such as those of the IPCC classification system, and for Greece offered snow-free and minimal shadow effects, ideal for mapping land cover in mountainous areas [19], [20].

2.2. Data preprocessing

Processing of satellite images was carried out in GEE. A cloud and shadow mask algorithm, following the methodology of [21], was applied. To achieve spatially homogeneous images for the entire country and allow a unified processing framework, 2 seasonal temporal composites based on median reflectance values per pixel and per band were created. For the spring composite, May and June images were used while for the summer composite, the images used were from July to September [22].

The CLC categories were matched with those of the CCI LC product, used in the UN's Indicator methodology. They were then grouped into broader categories of the IPCC classification scheme, based on the class correspondence in [23]. Classes with mixed coverage were excluded from the analysis (Classes 2.4, "Heterogeneous Agricultural Areas," and 3.2.4, "Transitional Wooded Shrubland Areas"), leaving gaps and thus making the labeled data "sparse". The omission of information from reference samples, combined with the use of reference samples with reduced spatial accuracy compared to the satellite data, characterizes this as a learning problem under the paradigm of "weak supervision". Specifically, the CLC samples can be described as "incomplete" due to their "sparse" nature but also "inexact" due to the reduced spatial accuracy they provide [16].

2.3. Mountain Green Cover Index

The primary variable for mapping MGCI was "green cover," essentially referring to LC mapping into 6 classes (forests, grasslands, croplands, wetlands, settlements, and other land). Mapping LC at the national level using satellite imagery poses a complex problem, especially in mountainous regions where rugged terrain can affect reflectance values. In addition, collecting reference samples on such large scales and model feature selection are also significant challenges.

Regarding total mountain area, the UN document for Indicator 15.4.2 utilizes the Kapos product [17] as a definition and was therefore used in this study. For Greece, the total mountainous area according to the Kapos product is 74,454.673km^2, or 56.42% of the country.

2.3.1. Sample generation

To minimize the difficult and time-consuming process of manual sampling, information from open geospatial data was utilized in an automated manner. Additionally, a DL algorithm was chosen over traditional classification algorithms, providing not only relief from the difficult parameter selection process but also achieving high accuracy given the low quality of reference samples.

The creation of the reference samples was based on spatial partitioning into fully independent regions for training, validation, and control sample acquisition [24], [25]. At first, a 50km grid was used to partition the country into independent training/validation and test regions. Subsequently, for each set of regions, within a 5120m grid, 28% of the total polygons (14% each) were randomly selected to create the training/validation and 18% to create the test patches [26] (Figure 1). Sample images of 256 pixels were randomly generated in each patch. After performing data augmentation, 10,000 training and 10,000 validation images were used in the model. Only for the test samples, 10m-LC information generated by the LIFE-IP 4 NATURA[1] project was utilized. This data was created through a semi-automated process and validated via photointerpretation.

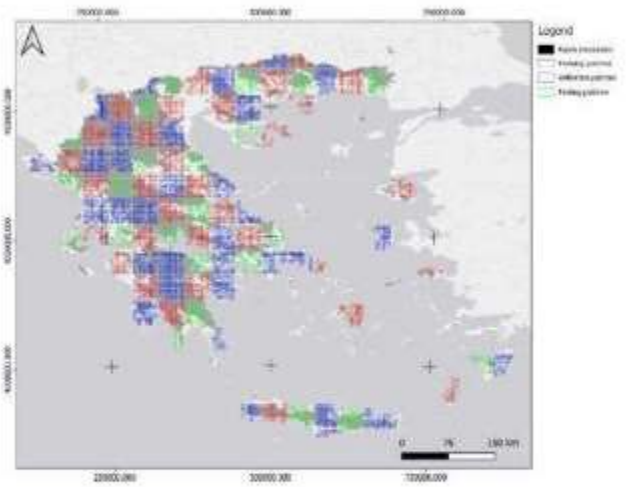

Figure 1. Training, validation and test regions and patches. Basemap: Esri.

2.3.2. Features and model

The 10 features used in the model were based on a similar study [8] for LC mapping in mountain regions using DL and can be seen in Table 1.

We chose to use a model whose code is available [27] and was developed to work for S2 data and sparse reference data, as in our case. We used the Adam optimizer [28], which achieves faster convergence in cases of "sparse" data and in complex models [29], and categorical cross-entropy loss function [30]. More specifically, we used weighted categorical cross-entropy loss to manage unevenly distributed per-class reference data [27]. The weights for each class were calculated by dividing the log of the total area percentage of all classes in the reference data by the area percentage of each class. A further normalization was performed on the resulting weight values, by dividing each value by the total weights, to avoid large variations in the weight values, due to the large differences in LC areas. The network was trained for 10 epochs in total, with a batch size of 6. A learning rate of 0.001 was set for the first 5 epochs, and 0.0001 for the other 5. Finally, the model with the lowest validation loss was used to predict LC at 10m spatial resolution.

Table 1 Features used in the classification model.

Image	Features
Spring composite	B2, B3, B4, B8, NDVI
Summer composite	B2, B3, B4, B8, NDVI
Sum	10

2.3.3. Accuracy Assessment

The model's learning performance was measured using the loss function value at each epoch. Additionally, categorical accuracy was calculated [27] and the error matrix, using 2500 randomly generated control points within and outside mountainous regions. From the error matrix, standardized accuracy metrics were derived, including Overall Accuracy (OA), Producer's Accuracy (PA), and User's Accuracy (UA) [31]. Finally, a cross-tabulation matrix was created to measure the agreement between the classified pixels and those in the ESA CCI LC product.

3. RESULTS AND DISCUSSION

The final value of Indicator 15.4.2 calculated for 2018, was 93.54%. For the same year, the United Nations Economic Commission for Europe (UNECE), for Greece, reports a value of 97.42%[2], while Our World in Data, reports a value of 97.53%[3].

The total green cover, by aggregating all sub-categories (forests, grassland, cropland, and wetlands), sums up to 69,643,776km^2, while the non-green categories occupy an area of 4,810,897km^2. The total area of each green cover sub-category, for the mountainous areas, was:

- Forests: 29,225,738km^2
- Shrublands & Grasslands: 22,520,327km^2
- Cultivated land: 17,680,949km^2
- Wetlands: 216,762km^2

[1] https://edozoume.gr/

[2] https://w3.unece.org/SDG/en/Indicator?id=68

[3] https://ourworldindata.org/grapher/mountain-green-cover-index?region=Europe

The OA was 68.65% for all classes, while considering only the "green" classes, it reached 73.04%. When reducing the problem to a binary classification of "green" and "non-green" areas, the OA was 88.21%. These results are in agreement with other studies using weak labels [32]. Settlements, wetlands and "other land", proved to be quite difficult to map, due to high inter-class variations, weak class representation or low spatial resolution of the reference data. Overall, it can be noted that the very generalised classification scheme created difficulties in mapping such an extensive area. The PA and UA metrics (Table 2) further revealed that large omission errors were present for wetlands and 'other land', while commission errors were present, but to a lesser extent, for the categories 'croplands', 'grasslands' and 'wetlands'.

Table 2. Producer's and User's accuracies derived from the error matrix.

	Class	PA	UA
Green cover	Forest	0.75	0.83
	Shrubland & Grassland	0.75	0.59
	Cropland	0.75	0.58
	Wetland	0.11	0.63
Non-green	Settlement	0.74	0.76
	Other	0.55	0.86

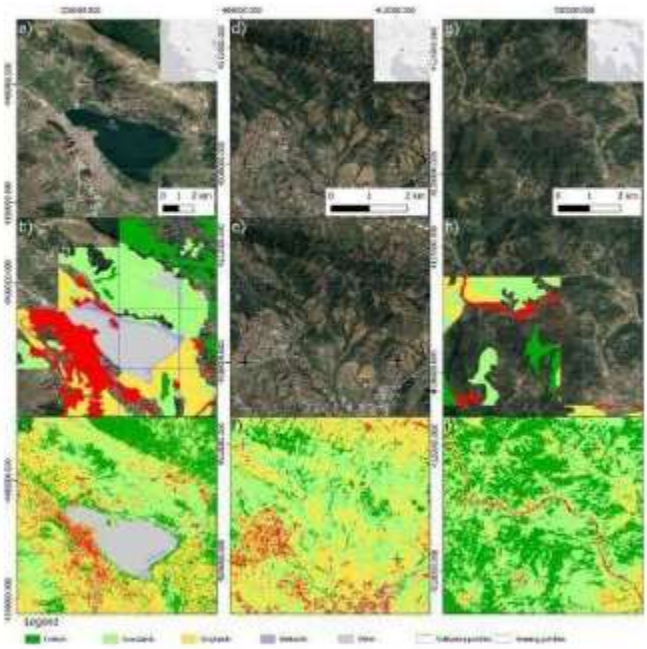

Figure 2 Classification results (c,f,i), compared to the CLC reference data (b,e,h), and Google Earth imagery of each case (a,d,g).

The cross-tabulations between the LC categories derived from the ESA CCI LC product and the categories derived from the model revealed an overall agreement of 62.91% for the country-wide classification and 57.23% in the mountainous areas for the 6 categories, with discrepancies especially in wetlands and settlements, which may be due to the significant difference in their spatial resolution.

Visual evaluation of the results (Figure 2) showed improved spatial resolution, compared to the reference data from CLC. In particular, the result is similar to that of a traditional pixel-based classifier (e.g. Random Forest). At the locations from which the sample images were generated, the classification contributed by improving the spatial resolution, compared to the reference "inexact" data, and by predicting the coverage, in pixels that lacked information due to "incomplete supervision" of the CLC data (Figure 2 - a-c, g-i). In areas with a complete lack of information, it also mapped the LC satisfactorily (Figure 2 - d-f).

Regarding the DL model's accuracy and loss (Figure 3), for the epochs trained, the validation loss was lower than the training loss, indicating that the model was underfitted and thus, required more training epochs [33]. Other reasons for this may be the dropout layers of the model, or the fact that the validation loss is calculated at the end of each epoch, while the training loss is calculated in each batch [34]. The loss function curves, decreased only slightly per epoch, suggesting that the number of data is too small for the complexity of the problem [35], however, the model was learning in a correct way.

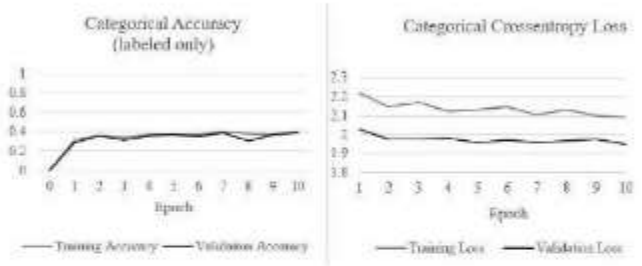

Figure 3 Model accuracy and loss curves.

DL algorithms, combined with available open data for sample creation, have been used in other studies. In two recent studies [25], [36], LC mapping at the national level was performed using transfer learning from pre-trained models with data from the CLC product. The differentiation of this study from the mentioned ones lies not only in its application to mountainous areas with challenging terrain but also in the use of weak and sparse reference data, without using transfer learning from pre-trained models.

The developed methodology can serve as a roadmap by national statistical services for calculating Indicator 15.4.2, bypassing the time-consuming and costly process of collecting reference samples.

The code used in this study is available in the GITHUB repository, at the following link: https://github.com/n-verde/Indicator_15.4.2.

4. ACKNOWLEDGEMENTS

This work has been supported by the European Commission LIFE Programme and Green Fund, LIFE EL-BIOS Project "hELlenic BIOodiversity Information System: An innovative tool for biodiversity conservation", under grant number LIFE20 GIE/GR/001317. This research work was also co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the Act "Enhancing Human Resources Research Potential by undertaking a Doctoral Research" Sub-action 2: IKY Scholarship Programme for PhD candidates in the Greek Universities.

5. REFERENCES

[1] J. M. Foggin, "Conservation Issues: Mountain Ecosystems," *Ref. Modul. Earth Syst. Environ. Sci.*, Jan. 2016.

[2] J. Bian, A. Li, X. Nan, G. Lei, and Z. Zhang, "Dataset of the mountain green cover index (SDG15.4.2) over the economic corridors of the Belt and Road Initiative for 2010-2019," *Big Earth Data*, vol. 00, no. 00, pp. 1–13, 2021.

[3] A. Grêt-Regamey, S. Brunner, J. Altwegg, M. Christen, and P. Bebi, "Integrating Expert Knowledge into Mapping Ecosystem Services Trade- offs for Sustainable Forest Management," *Soc. Ecol.*, vol. 18, no. 3, pp. 34-, 2013.

[4] United Nations, "SDG Indicators Metadata - Indicator 15.4.2: Mountain Green Cover Index," 2021.

[5] J. Bian, A. Li, G. Lei, Z. Zhang, and X. Nan, "Global high-resolution mountain green cover index mapping based on Landsat images and Google Earth Engine," *ISPRS J. Photogramm. Remote Sens.*, vol. 162, no. September 2019, pp. 63–76, 2020.

[6] L. De Simone, D. Navarro, P. Gennari, A. Pekkarinen, and J. de Lamo, "Using standardized time series land cover maps to monitor the sdg indicator 'mountain green cover index' and assess its sensitivity to vegetation dynamics," *ISPRS Int. J. Geo-Information*, vol. 10, no. 7, 2021.

[7] K. Cheng and J. Wang, "Forest-type classification using time-weighted dynamic timewarping analysis in mountain areas: A case study in southern China," *Forests*, vol. 10, no. 11, pp. 1–18, 2019.

[8] L. Gao, J. Luo, L. Xia, T. Wu, Y. Sun, and H. Liu, "Topographic constrained land cover classification in mountain areas using fully convolutional network," *Int. J. Remote Sens.*, vol. 40, no. 18, pp. 7127–7152, 2019.

[9] D. P. Shrestha and J. A. Zinck, "Land use classification in mountainous areas: integration of image processing, digital elevation data and field knowledge (application to Nepal)," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 3, no. 1, pp. 78–85, Jan. 2001.

[10] G. Singh and A. Pandey, "Evaluation of classification algorithms for land use land cover mapping in the snow-fed Alaknanda River Basin of the Northwest Himalayan Region," *Appl. Geomatics*, vol. 13, no. 4, pp. 863–875, 2021.

[11] M. A. Mulders, "Advances in the application of remote sensing and GIS for surveying mountainous land," *ITC J.*, vol. 3, no. 1, pp. 3–10, 2001.

[12] S. Mondal and C. Jeganathan, "Mountain agriculture extraction from time-series MODIS NDVI using dynamic time warping technique," *Int. J. Remote Sens.*, vol. 39, no. 11, pp. 3679–3704, 2018.

[13] S. Vanonckelen, S. Lhermitte, and A. Van Rompaey, "The effect of atmospheric and topographic correction on pixel-basedimage composites: Improved forest cover detection in mountainenvironments," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 35, no. PB, pp. 320–328, 2015.

[14] L. Zhang, L. Zhang, and V. Kumar, "Deep learning for Remote Sensing Data," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, 2016.

[15] F. Song *et al.*, "Multi-scale feature based land cover change detection in mountainous terrain using multi-temporal and multi-sensor remote sensing images," *IEEE Access*, vol. 6, pp. 77494–77508, 2018.

[16] M. Schmitt, J. Prexl, P. Ebel, L. Liebel, and X. X. Zhu, "Weakly supervised semantic segmentation of satellite images for land cover mapping-challenges and opportunities," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 5, no. 3, pp. 795–802, 2020.

[17] V. Kapos, J. Rhind, M. Edwards, M. F. Price, and C. Ravilious, "Developing a map of the world's mountain forests.," in *Forests in sustainable mountain development: a state of knowledge report for 2000. Task Force on Forests in Sustainable Mountain Development.*, vol. 5, no. October 2017, Wallingford: CABI, 2000, pp. 4–19.

[18] M. Schuler, E. Stucki, O. Roque, and M. Perlik, "Mountain Areas in Europe : Analysis of mountain areas in EU member states, acceding and other European countries," European Commission, Brussels, 2004.

[19] J. Hepinstall-Cymerman, S. Coe, and M. Alberti, "Using urban landscape trajectories to develop a multi-temporal Land cover database to support ecological modeling," *Remote Sens.*, vol. 1, no. 4, pp. 1353–1379, 2009.

[20] H. Saadat, J. Adamowski, R. Bonnell, F. Sharifi, M. Namdar, and S. Ale-Ebrahim, "Land use and land cover classification over a large area in Iran based on single date analysis of satellite imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 5, pp. 608–619, 2011.

[21] A. Hollstein, K. Segl, L. Guanter, M. Brell, and M. Enesco, "Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images," *Remote Sens.*, vol. 8, no. 8, p. 666, Aug. 2016.

[22] G. Kotsias, C. J. Lolis, N. Hatzianastassiou, P. Lionello, and A. Bartzokas, "A comparison of different approaches for the definition of seasons in the Mediterranean region," *Int. J. Climatol.*, vol. 42, no. 3, pp. 1954–1974, Mar. 2022.

[23] V. Reinhart, C. C. Fonte, P. Hoffmann, B. Bechtel, D. Rechid, and J. Boehner, "Comparison of ESA climate change initiative land cover to CORINE land cover over Eastern Europe and the Baltic States from a regional climate modeling perspective," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 94, no. April 2020, p. 102221, Feb. 2021.

[24] C. Karakizi, K. Karantzalos, M. Vakalopoulou, and G. Antoniou, "Detailed Land Cover Mapping from Multitemporal Landsat-8 Data of Different Cloud Cover," *Remote Sens.*, vol. 10, no. 8, p. 1214, Aug. 2018.

[25] S. Scepanovic, O. Antropov, P. Laurila, Y. Rauste, V. Ignatenko, and J. Praks, "Wide-Area Land Cover Mapping with Sentinel-1 Imagery Using Deep Learning Semantic Segmentation Models," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 10357–10374, 2021.

[26] H. Ning, Z. Li, C. Wang, and L. Yang, "Choosing an appropriate training set size when using existing data to train neural networks for land cover segmentation," *Ann. GIS*, vol. 26, no. 4, pp. 329–342, 2020.

[27] A. Stoian, V. Poulain, J. Inglada, V. Poughon, and D. Derksen, "Land Cover Maps Production with High Resolution Satellite Image Time Series and Convolutional Neural Networks: Adaptations and Limits for Operational Systems," *Remote Sens.*, vol. 11, no. 17, p. 1986, Aug. 2019.

[28] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.

[29] S. Ruder, "An overview of gradient descent optimization algorithms," pp. 1–14, Sep. 2016.

[30] Y. Hua, Di. Marcos, L. Mou, X. X. Zhu, and D. Tuia, "Semantic Segmentation of Remote Sensing Images with Sparse Annotations," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, no. d, 2022.

[31] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.*, vol. 37, no. September, pp. 34–46, Jul. 1991.

[32] M. Schmitt, L. H. Hughes, C. Qiu, and X. X. Zhu, "SEN12MS – A CURATED DATASET OF GEOREFERENCED MULTI-SPECTRAL SENTINEL-1/2 IMAGERY FOR DEEP LEARNING AND DATA FUSION," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. IV-2/W7, no. 2/W7, pp. 153–160, Sep. 2019.

[33] M. Talo, U. B. Baloglu, Ö. Yıldırım, and U. Rajendra Acharya, "Application of deep transfer learning for automated brain abnormality classification using MR images," *Cogn. Syst. Res.*, vol. 54, pp. 176–188, 2019.

[34] A. Soleymani, "Your validation loss is lower than your training loss? This is why!," *Towards Data Science*, 2022. [Online]. Available: https://towardsdatascience.com/what-your-validation-loss-is-lower-than-your-training-loss-this-is-why-5e92e0b1747e. [Accessed: 02-Jun-2023].

[35] J. Brownlee, "How to use Learning Curves to Diagnose Machine Learning Model Performance," *Machine Learning Mastery*, 2019. [Online]. Available: https://machinelearningmastery.com/learning-curves-for-diagnosing-machine-learning-model-performance/.

[36] P. Ulmas and I. Liiv, "Segmentation of Satellite Imagery using U-Net Models for Land Cover Classification," pp. 1–11, Mar. 2020.