

ESTIMATING NDVI FROM SAR IMAGES USING DNN

Iulia Calota¹, Daniela Faur¹, Mihai Datcu^{1,2}

¹Research Center for Spatial Information (CEOSpaceTech), University POLITEHNICA of Bucharest (UPB), ²Earth Observation Center (EOC), German Aerospace Center (DLR)

ABSTRACT

The Normalized Difference Vegetation Index (NDVI) is an important factor to be considered in vegetation tracking and analysis, which can be easily derived from multispectral (MS) images. However, the limitation imposed by the atmospheric conditions makes the calculation of this index difficult. Because of the clouds, only a limited number of multispectral bands can capture the land appropriately. Furthermore, the multispectral sensors are dependent on the sunlight, which makes the acquisition of data more limited. These limitations do not hinder other types of Earth Observation (EO) data, like the scenes captured by the Synthetic Aperture Radar (SAR). However, SAR images cannot be used in NDVI calculation. In this article, we propose a deep learning (DL) based method for NDVI estimation from SAR data. Using a database with corresponding MS and SAR patches, we calculate the NDVI for each sample, then use a convolutional neural network (CNN) for predicting the NDVI of SAR images. This simple method leads to a precision of 70% in NDVI estimation from SAR images.

Index Terms— Normalized Difference Vegetation Index, Synthetic Aperture Radar, Multispectral images, Convolutional Neural Networks

1. INTRODUCTION

Atmospheric conditions cause difficulties in the analysis of multispectral EO data. Lately developed methods for automatic cloud detection perform filtering of the obstructed scenes. The authors of [1], implemented an object-oriented cloud and cloud shadow matching method. It is derived from the automatic cloud cover assessment (ACCA) algorithm. The presented method has three steps. The first step is to modify the ACCA algorithm to match the data provided by the Chinese satellites. The second step is to detect the cloud shadows. The last step is to correlate the clouds and the cloud shadows to refine the cloud map provided by the modified ACCA algorithm.

In [2], the authors propose a pixel-based cloud detection method. This method implies the extraction of features for

each pixel, based on brightness or whiteness. These features and the patches are then classified using several supervised classification algorithms.

Algorithms for cloud removals have also been proposed. In [3], the authors suggest as a starting point the radiative transfer model, initially developed for Landsat, combined with the reflectance of the Cirrus band of the sensor within the Landsat-8 satellite. This ensures that the thinner clouds are also removed. The authors of [4] extend the radiative transfer model to Sentinel-2 images to remove thin clouds. In [5], the authors use a convolutional neural network with multi-scale prediction scheme to remove the clouds from Sentinel-2 images.

The images synthesized by these methods suffer from inherent losses. In [6], the authors study the effects of Cirrus clouds on the NDVI and conclude that the error in NDVI calculation is directly proportionate to the optical thickness and particle distribution in the Cirrus clouds. For thicker clouds, such methods of cloud removal may not deliver appropriate results.

This leads to the development of algorithms for NDVI estimation in cloud-contaminated images. Such algorithms use time-series to estimate the NDVI in cloudy images similarly with the method presented in [7].

A few authors have found working with SAR imagery appealing. In [8], the correlation between SAR features and NDVI on maize fields is described. The authors of [9] use SAR data to compensate for the loss of information in optical images due to cloudy areas.

In this paper we submit a simple, CNN-based method to estimate NDVI from SAR data, at patch level. Our database contains both MS and SAR data of the same scene. The MS patches enable NDVI's calculation while SAR patches serve to calculate SAR features which are added, as additional channels, to the SAR data. We train different architectures of CNNs, using the SAR patches enhanced by the additional features as input, the estimated NDVI representing the output.

Moreover, due to their already proved effectiveness we use the algorithms described in [10] and [11], that reduce the input dimension of a CNN by using Bag-of-Words instead of patches to reduce the training time. We started with SAR patches and additional features as CNN's input. While testing various formats of data inputs, we determined that

Bag-of-Words derived from the data is the optimal tradeoff to obtain both an optimal training time and highest accuracy.

Bag-of-Words is used here as a dimension reduction technique for the convolutional neural network. A downside of CNNs is the large training time, especially when dealing with large datasets with large dimensions. Pruning is one of the developed methods used for faster training. However, the authors in [10] demonstrated that using BoW leads to a reduction of training time without significant impact in performance.

Moreover, the BoW algorithm used in this paper is faster than the original one [19], as it does not rely on K-Means. It is based on the work described in [18] that demonstrates the optimal size of BoW/ Random images are selected and divided into 3x3 patches. From these patches, we randomly select 250 patches. We iterate through each image and divide it in 3x3 patches, that are compared to the patches in the dictionary using nearest neighbor algorithm. Thus, a histogram-like representation of each image is created, resulting in the Bag-of-Words.

The results show that both methods reach high precision.

2. METHODOLOGIES

Supervised learning techniques, such as training with CNNs, require a large amount of labeled data. To validate our method, we should use a database contain both MS and SAR data, with a considerable number of patches revealing vegetation classes

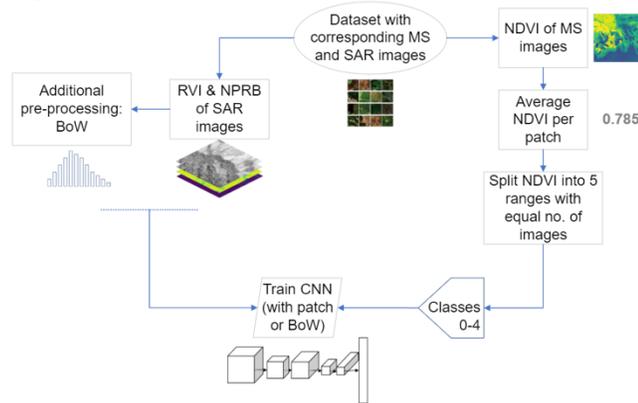


Figure 1: The proposed methodology.

Figure 1 presents the diagram revealing the stages of our method. SAR images enables the pre-process the input of the CNN. From the MS images, we calculate the average NDVI per patch, then we split the range of NDVI into 5 intervals, considering the equal distribution of the images throughout them. Each interval becomes one of the 5 classes. Afterwards, we train the CNN (either the CNN in fig. 2 or VGG-19).

BigEarthNet [12] [13], a dataset with 590326 corresponding Sentinel-1 (SAR) and Sentinel-2 (MS) patches fits this profile. This labeled dataset has 43 classes, most of them being related to vegetation.

We have included in our experiment only the patches relevant for vegetation from the dataset. Consequently a total of 17 labels have been ignored from BigEarthNet, our final dataset included 328586 patches containing vegetation-related classes.

To estimate the density of green over the landcover we have used the Normalized Difference Vegetation Index-NDVI that describes the difference between -infrared (NIR) and the red (RED) reflectance of vegetation cover, using the following formula:

$$NDVI = \frac{NIR-RED}{NIR+RED} \quad (1)$$

In the case of Sentinel-2, the near-infrared band is band 8 and the red band is band 4 [14]. In BigEarthNet, each multispectral image contains all the bands present in the original products, except for band 10.

The Sentinel-1 patches in BigEarthNet are preprocessed from ground range detected (GRD) products. The obtained patches correspond both geographically and temporally with the Sentinel-2 patches.

The NDVI of the 328586 patches selected from BigEarthNet ranges between -0.2 and 0.9. To simplify the operations, we chose to turn the problem into a classification task, rather than a prediction task – we divided the above range into 5 intervals, each interval being an NDVI class. The range is not uniformly divided, as this would lead to unbalanced patch distribution among classes. The Table 1 describes this distribution of the NDVI in the 5 classes and the number of patches corresponding to the classes.

Table 1. The distribution of the NDVI values and of the number of patches in the 5 NDVI classes.

NDVI class	NDVI range	No. of patches
1	-0.192 → 0.368	65708
2	0.368 → 0.5325	70054
3	0.5325 → 0.7047	64978
4	0.7047 → 0.7769	63900
5	0.7769 → 0.9291	63953

We observed that the convolutional neural networks are more sensitive to the number of samples than to the similarity among classes. That is why we chose equal distribution of the patches among the classes, rather than equal NDVI distribution. for example, class 1 covers NDVI values between -0.192 and 0.368, while class 4 covers only values between 0.7047 and 0.7769.

For more robustness, we increased the number of features of SAR images. The original SAR patches contain only the two polarities present in all Sentinel-1 products. Apart from these bands, we added a band for the Radar

Vegetation Index (RVI) adapted for Sentinel-1 [15] and the Normalized Ratio Procedure between Bands (NRPB) described in [16]. The authors have correlated the NDVI with the above indices, however the relationship is not straight-forward on the entire NDVI range.

We chose to train two different CNN architectures, a simple three-layered CNN and VGG-19 [17]. The three-layered architecture is presented in Fig. 2.

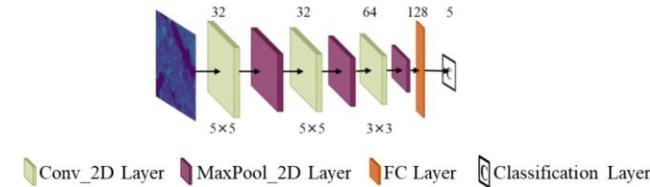


Figure 2: Three-layered architecture used for NDVI estimation in SAR images.

Furthermore, we wanted to verify, whether smaller dimensions deliver better result. Instead of the patches, at the input of the CNN we placed the BoW derived from the SAR patches with the additional bands. This method reduces the training time and benefits from the lower number of parameters, which also influence the performance of a CNN. BoW is a good method to preserve both spatial and spectral information, although the dimension is reduced in a drastic manner. The CNNs must be modified to use this method – instead of 2D layers, 1D layers have to be used.

3. RESULTS

We used the same settings in all runs: we chose stochastic gradient descent as an optimizer, and we trained each network for 100 epochs. The results of training patches are seen in Table 2 and the results of training BoW are seen in Table 3.

In Table 2, we can see that the precision reached a value of 0.7 with this simple method. It is worth mentioning that the training of patches lasted around 800 minutes for the three-layered architecture and around 1400 minutes for VGG-19. However, as expected, the results for VGG-19 are better, as it is a more robust network that the simple three-layered architecture.

Table 2: Results of training SAR images to estimate NDVI.

Architecture	Precision	Recall	F1	F2
3-layered CNN	0.6528	0.6012	0.6259	0.6208
VGG-19	0.7018	0.6615	0.6811	0.6771

In Table 3, we can see that the performance is like the training of patches, with little differences. However, the training time of VGG-19 was reduced to almost 400

minutes, making this approach a faster alternative. As in the previous case, VGG-19 delivered better results.

Table 3: Results of training SAR images to estimate NDVI.

Architecture	Precision	Recall	F1	F2
3-layered CNN	0.644	0.6211	0.6323	0.6301
VGG-19	0.7201	0.6545	0.6857	0.6793

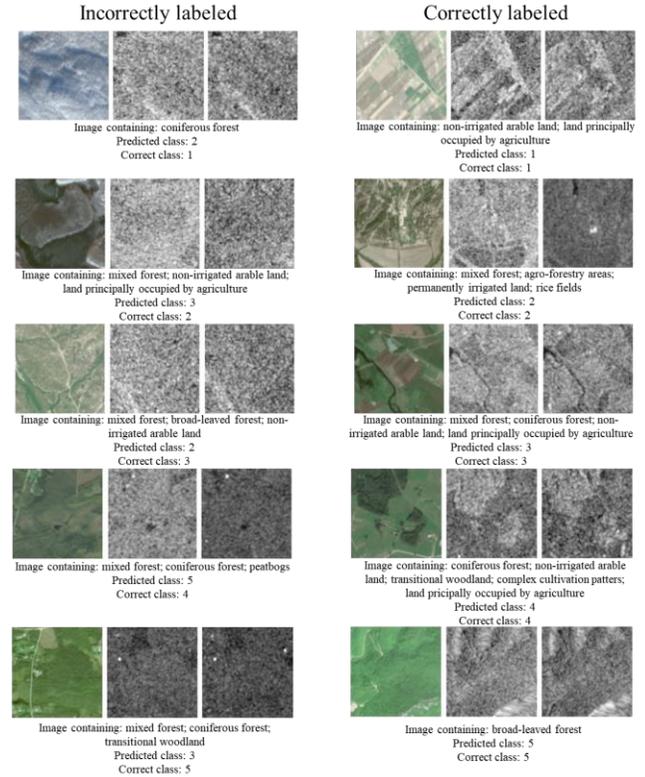


Figure 3: Patches samples that were correctly (right) classified in opposition to those misclassified (left).

Figure 3 shows random images that were misclassified (left column) or accurately (right column) classified. For each example, we specified what it contains (labels taken from the original dataset) and the predicted and correct NDVI classes. Each example includes the RGB image and its correspondent SAR images. Each SAR image represents one of the polarizations (VH and VV). We observe that classes with structures and patterns are correctly classified, as these are distinguishable also in the SAR representation.

4. CONCLUSIONS

This paper addressed the challenge to estimate the NDVI from cloudy patches. We proposed a simple deep learning-based method to estimate the NDVI from SAR images. We benefitted from correlated SAR and MS patches from the BigEarthNet dataset, so that each SAR image has a

corresponding NDVI value calculated from MS images. We succeeded to increase result's accuracy by adding two more bands calculated from SAR features, RVI and NPRB to the SAR patches .

To facilitate this process, we decided to switch to a classification task, rather than a prediction class. However, prediction or estimation of NDVI with a similar method is also possible.

Moreover, by using BoW instead of the patches themselves. we offer a faster alternative for training. This approach reduces the training time with the benefits of achieving similar results.

In future work, the target is to make this method more robust, as it is dependent on a multitude of factors: learning rate, number of samples, architecture and other hyperparameters specific to convolutional neural networks.

5. ACKNOWLEDGEMENT

This work was supported by a grant of the Romanian Ministry of Education and Research, CNCS-UEFISCDI, project number PN-III-P4-ID-PCE-2020-2120, within PNCDI III.

6. REFERENCES

- [1] B. Zhong, W. Chen, S. Wu, L. Hu, X. Luo and Q. Liu, "A Cloud Detection Method Based on Relationship Between Objects of Cloud and Cloud-Shadow for Chinese Moderate to High Resolution Satellite Imagery," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 11, pp. 4898-4908, Nov. 2017.
- [2] L. Gómez-Chova, G. Mateo-García, J. Muñoz-Marí and G. Camps-Valls, "Cloud detection machine learning algorithms for PROBA-V," 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2017, pp. 2251-2254.
- [3] B. Zhou and Y. Wang, "A Thin-Cloud Removal Approach Combining the Cirrus Band and RTM-Based Algorithm for Landsat-8 OLI Data," *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 1434-1437.
- [4] Y. Gao, Y. Wang and H. Lv, "Extendibility of a Thin-Cloud Removal Algorithm to Hi-Resolution Visible Bands of Sentinel-2 Data," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 927-930.
- [5] K. Lee and J. Sim, "Cloud Removal of Satellite Images Using Convolutional Neural Network With Reliable Cloudy Image Synthesis Model," 2019 IEEE International Conference on Image Processing (ICIP), 2019, pp. 3581-3585.
- [6] K. Rajitha, M. M. Prakash Mohan and M. R. R. Varma, "Effect of cirrus cloud on normalized difference Vegetation Index (NDVI) and Aerosol Free Vegetation Index (AFRI): A study based on LANDSAT 8 images," 2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR), 2015, pp. 1-5.
- [7] X. Ling and R. Cao, "A New Spatiotemporal Data Fusion Method to Reconstruct High-Quality Landsat Ndvi Time-Series Data," 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021, pp. 2564-2567.
- [8] J. Alvarez-Mozos, J. Villanueva, M. Arias and M. Gonzalez-Audicana, "Correlation Between NDVI and Sentinel-1 Derived Features for Maize," 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021, pp. 6773-6776.
- [9] A. Mazza, M. Gargiulo, G. Scarpa and R. Gaetano, "Estimating the NDVI from SAR by Convolutional Neural Networks," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 1954-1957.
- [10] I. Calota, D. Faur and M. Datcu, "DNN-Based Semantic Extraction: Fast Learning from Multispectral Signatures," *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, 2020, pp. 3672-3675.
- [11] I. Calota, D. Faur and M. Datcu, "Bag-of-Words for Transfer Learning," 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021, pp. 808-811.
- [12] G. Sumbul, M. Charfuelan, B. Demir and V. Markl, "Bigearthnet: A Large-Scale Benchmark Archive for Remote Sensing Image Understanding," *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 5901-5904.
- [13] G. Sumbul et al., "BigEarthNet-MM: A Large-Scale, Multimodal, Multilabel Benchmark Archive for Remote Sensing Image Classification and Retrieval [Software and Data Sets]," in *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 3, pp. 174-180, Sept. 2021.
- [14] <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm>
- [15] https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-1/radar_vegetation_index/#
- [16] R. Filgueiras, E. C. Mantovani, D. Althoff, E. I. Fernandes Filho, and F. F. da Cunha, "Crop NDVI Monitoring Based on Sentinel 1," *Remote Sensing*, vol. 11, no. 12, p. 1441, Jun. 2019.
- [17] K. Simonyan and A. Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." *CoRR*, (2015).
- [18] S. Cui, G. Schwarz and M. Datcu, "Remote Sensing Image Classification: No Features, No Clustering", in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 11, pp. 5158- 5170, 2015.
- [19] Fei-Fei Li; Perona P., "A Bayesian Hierarchical Model for Learning Natural Scene Categories", 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 2. p. 524, 2005.