IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-6032-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IGARSS53475.2024.10640541

HIGH RESOLUTION MAPPING OF VEGETATION BIODIVERSITY BY HYPERSPECTRAL IMAGES AND CONVOLUTIONAL AUTOENCODERS

Giovanni Nico[1], Manilo Monaco[2], Olimpia Masci[3], Mario G.C.A. Cimino[4]

[1] National Research Council of Italy, Institute for Applied Mathematics, Bari, Italy
[2] Italian Space Agency, Matera, Italy
[3] DIAN S.r.l., Matera, Italy
[4] University of Pisa, Department of Information Engineering, Pisa, Italy

ABSTRACT

A methodology is presented to map the vegetation biodiversity based on the hypothesis of the spectral variation (SV) which has been proposed to assess the forest biodiversity by means of Earth Observation (EO) data. Hyperspectral data acquired by the PRecursore Iperspettrale della Missione Applicativa (PRISMA) mission of the Italian Space Agency to spectral signature with a high spectral resolution. The NDVI is computed from PRISMA data and used to identify pixels corresponding to vegetation cover. The spectral signatures at those pixels are then clusterized using the convolutional autoenconders technique and the final map with the location of pixels belonging to the different classes is produce. The methodology is applied to assess the vegetation biodiversity in National Parks of Gargano, Alta Murgia, Cilento-Vallo di Diano-Alburni, Appennino Lucano Val D'Agri Lagonegrese and Pollino, all located in Southern Italy.

Index Terms— Vegetation biodiversity, Spectral diversity, hyperspectral images, PRISMA, convolutional auto-encoder

1. INTRODUCTION

Forest biodiversity is one of the seven thematic programmes established by the Conference of the Parties within the Convention on Biological Diversity. The topic of identification, monitoring, definition of indicators and assessment of biodiversity is one of the cross-cutting issues of the Convention to collect, maintain and organize biodiversity information. In addition, controlling and planning the removal of invasive species are topics of outmost importance in management of natural resources because of the severe ecological damages and economic losses caused by non-native alien species. The huge amount, spectral diversity, regular and dense acquisition plan of current Earth Observation spaceborne missions provides a means to monitor and evaluate the vegetation biodiversity. In the literature, many applications of multispectral images (e.g. Sentinel-2) have been used to compute different vegetation indexes to analyze the phenological properties of plants. Mathematical tools useful for the quantification of biodiversity, including the problem of invasive species, and the optimal resource allocation for conservation planning of critical habitats and control of invasive species have been proposed.

The maps of vegetation indexes are further analyzed in terms of abundance-based metrics or agent-based model to quantify vegetation biodiversity. For instance, the Shannon entropy and Rao's Q metrics haven been implemented and applied to the matrices of vegetation indexes and used as computational tools to describe the diversity of the agro-forestry landspace. The main disadvantage of such an approach is the loss of spatial resolution due the kernel used to compute the above metrics and the presence of artifacts related to the kernel itself (e.g. blurring effects and the kernel size). In the case of the modelling of the landscape heterogeneity by intelligent agents, the selected areas are sampled by agents which collect information on vegetation indices in order to measure their diversity. The main disadvantage of such an approach is the computational load and difficulty to provide a physical-based interpretation to the output.

As a further tool to assess the forest biodiversity, the scientific community is debating on the validity of the spectral variation (SV) hypothesis which has been proposed to assess the forest biodiversity by means of Earth Observation data. The current availability of hyperspectral images acquired by the PRecursore Iperspettrale della Missione Applicativa (PRISMA) mission of the Italian Space Agency and the recently launched Environmental Mapping and Analysis Program (EnMAP) mission of the German Space Agency, as well as the planned missions, e.g., the

Copernicus Hyperspectral Imaging Mission for the Environment (CHIME) of the European Space Agency open unique perspective for the multi-temporal mapping of forest biodiversity and the assessment of SV hypothesis validity. Open issues remain such as the how to related the results obtained by applying the SV hypothesis to the analysis of hyperspectral data to the biodiversity as assessed in the field and till what extent the SV hypothesis can be applied as a proxy to detect vegetation biodiversity and its spatial gradients.

In this work we propose a methodology to map the forest biodiversity using the SV hypothesis and the autoenconders technique. The methodology is applied to assess the vegetation biodiversity in National Parks of Gargano, Alta Murgia, Cilento-Vallo di Diano-Alburni, Appennino Lucano Val D'Agri Lagonegrese and Pollino, all located in Southern Italy.

2. METHODOLOGY

The methodology proposed in this paper combines the potential of hyperspectral data to capture subtle variations in the spectral signatures of plant species, the ability of vegetation indices (such as NDVI) to discriminate vegetation cover from other land covers, the effectiveness of convolutional neural networks in extracting relevant features within highly multidimensional and nonlinear data, and the efficiency of the k-means algorithm to clustering objects in high-dimensional vectorial spaces.

The relationship between the NDVI and vegetation productivity is well established, and the link between this index and the fraction of absorbed photosynthetic active radiation intercepted has been well documented, both theoretically and empirically [7]. In this paper, we have used the NDVI as a filter in order to differentiate pixels from the original image containing dense forest from those containing water, urban areas, non-forest, and agricultural fields.

Convolutional neural networks (CNNs) are hierarchical models whose convolutional layers alternate with sub-sampling layers, reminiscent of simple and complex cells in the primary visual cortex [9]. Convolutional Auto-Encoders (CAEs) differs from conventional Auto-Encoders as their weights are shared among all locations in the input, preserving in general spatial locality and in our specific case spectral locality. Indeed, in our methodology we have used a per-pixel approach, performing the convolutions along only one dimensional (1D) axis, that is, along the spectral bands of the PRISMA signal. The reconstruction of the signal is hence due to a linear combination of basic spectral patches based on the latent vector. The capability of convolutional auto-encoders to provide better feature extraction and dimensionality reduction than other techniques is well known, and scientific research highlights the potential of these models to achieve fine clustering and visualization of vegetation types [10]. The K-means algorithm is used as a basis for clustering of vegetation species and spatial visualization. Fig. 1 shows the data processing workflow used in this work.

The pixels of the original PRISMA hyperspectral image (HSI) are selected on the basis of specific NDVI values in order to segment the region of interest.

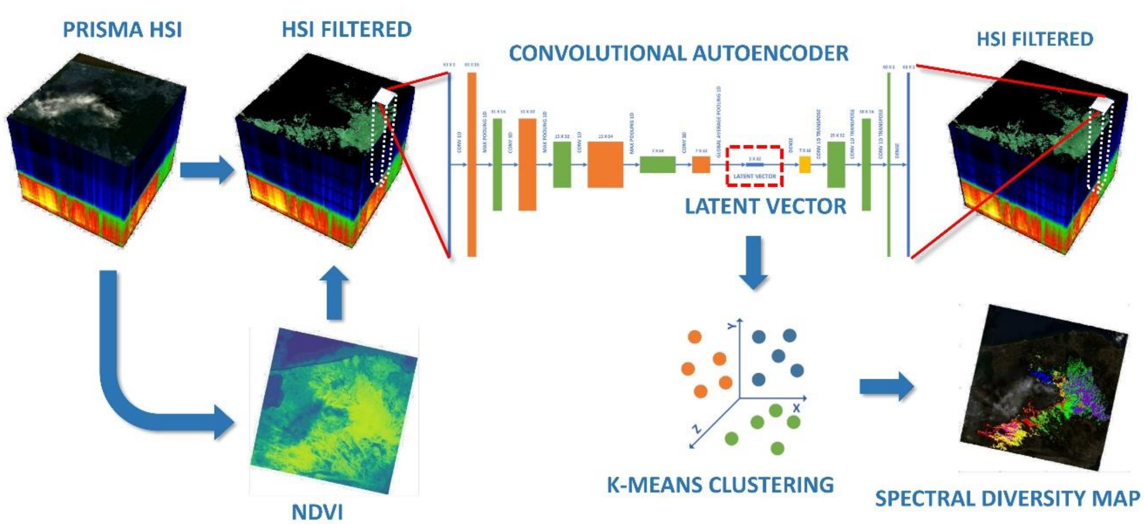

Fig. 1: Overall data processing workflow

A convolutional auto-encoder is employed for the decorrelation of the high-dimensional spectral signal. Finally, an unsupervised learning method is used to identify similar clusters of pixels, which are subsequently visualized on a spatial map.

3. DATA SET

A set of 51 PRISMA images have been used for this study. Table 1 summarizes the number of PRISMA images used in this work, for each study area. Images have been acquired between 2019 and 2022.

Table 1: Number of PRISMA images for each of the national parks

Study area	N. images
Gargano	4
Alta Murgia	21
Cilento-Vallo di Diano-Alburbi	4
Appennino Lucano Val D'Agri Lagonegrese	18
Pollino	4

4. RESULTS

In order to show the potential of the approach of Fig. 1, in this section a pilot case is discussed. Fig. 2 shows the result of the spatial distribution of a cluster analysis. In this analysis, a convolutional window of 5 spectral bands and 9 clusters has been considered.

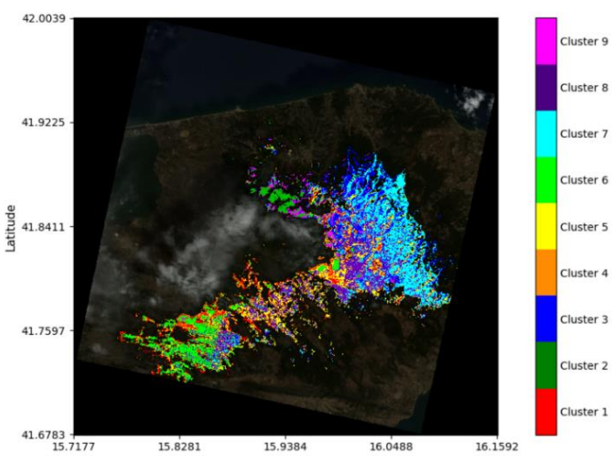

Fig. 2: Map with the spatial distribution of clusters.

Fig. 3 highlights the mean spectral signatures of the original image pixels, grouped based on the cluster composition provided by the methodology.

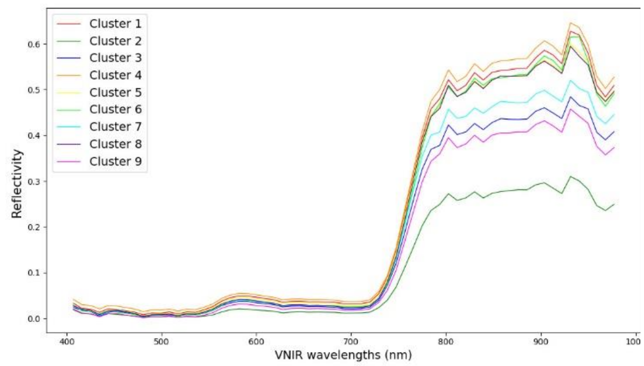

Fig. 3: Mean spectral signature of the clusters.

As can be seen from the graph, the technique has been able to separate the variability of spectral reflectivity within specific portions of the electromagnetic spectrum, probably related to the diversity of plant species.

CONCLUSIONS

This paper proposes a robust methodology based on convolutional autoencoders for high-resolution mapping of vegetation biodiversity via hyperspectral images. The proposed approach leverages the spectral variation (SV) hypothesis to assess forest biodiversity through Earth Observation (EO) data. The methodology is successfully applied to assess vegetation biodiversity in five National Parks in Southern Italy, the approach demonstrates the potential of fine-scale analysis of spectral diversity within specific regions of interest.

The use of convolutional autoencoders, with their ability to extract relevant features from highly multidimensional and nonlinear data, proves to be a valuable contribution to vegetation biodiversity mapping. The effectiveness in differentiating spectral signatures associated with distinct plant species showcases its potential for applications in monitoring and managing biodiversity in natural ecosystems.

The cluster analysis is based on the high spectral resolution of PRISMA images and does not depend on the number of HS images used for the same study area. Nevertheless, the availability of a large number of PRISMA images would facilitate the study of seasonal variations of the cluster characteristics. The cluster analysis performance shall be assessed through field campaigns and related to the alpha and beta diversities, i.e., the local species diversity at a specific site and the ratio between regional and local species diversities.

However, open challenges and discussions remain, especially regarding the validation of results obtained through the SV hypothesis and the extent to which this hypothesis serves as a reliable proxy for detecting vegetation biodiversity and its spatial gradients. Ongoing and future missions are expected to contribute further insights into these debates and enhance the understanding of forest biodiversity dynamics through hyperspectral imaging.

5. ACKNOWLEDGMENTS

Project carried out using ORIGINAL PRISMA Products - © Italian Space Agency (ASI); the Products have been delivered under an ASI License to Use

11. REFERENCES

[1] Somayeh Nezami, Ehsan Khoramshahi, Olli Nevalainen, Ilkka Pölönen, and Eija Honkavaara, "Tree Species Classification of Drone Hyperspectral and RGB Imagery with Deep Learning Convolutional Neural Networks", *Remote Sensing*, 12, 1070, 2020.

[2] Duccio Rocchini, Sandra Luque, Nathalie Pettorelli, Lucy Bastin, Daniel Doktor, Nicolò Faedi, Hannes Feilhauer, Jean-Baptiste Féret, Giles M. Foody, Yoni Gavish, Sergio Godinho, William E. Kunin, Angela Lausch7, Pedro J. Leitão, Matteo Marcantonio. Markus Neteler, Carlo Ricotta, Sebastian Schmidtlein, Petteri Vihervaara, Martin Wegmann, Harini Nagendra, "Measuring β-diversity by remote sensing: A challenge for biodiversity monitoring", *Methods in Ecology and Evolution*, 9:1787–1798, 2018.

[3] Jeannine Cavender-Bares, John A. Gamon, and Philip A. Townsend Editors, Remote Sensing of Plant Biodiversity, Springer Nature Switzerland AG, Cham, Switzerland, 2020.

[4] Duccio Rocchini, Matteo Marcantonio, Carlo Ricotta, "Measuring Rao's Q diversity index from remote sensing: An open source solution", *Ecological Indicators*, 72, 234–238, 2017.

[5] Angela Lausch, Michael E. Schaepman, Andrew K. Skidmore, Eusebiu Catana, Lutz Bannehr, Olaf Bastian, Erik Borg, Jan Bumberger, Peter Dietrich, Cornelia Glässer, Jorg M. Hacker, Rene Höfer, Thomas Jagdhuber, Sven Jany, András Jung, Arnon Karnieli, Reinhard Klenke, Toralf Kirsten, Uta Ködel, Wolfgang Kresse, Ulf Mallast, Carsten Montzka, Markus Möller, Hannes Mollenhauer, Marion Pause, Minhaz Rahman, Franziska Schrodt, Christiane Schmullius, Claudia Schütze, Peter Selsam, Ralf-Uwe Syrbe, Sina Truckenbrodt, Michael Vohland, Martin Volk, Thilo Wellmann, Steffen Zacharias, and Roland Baatz, "Remote Sensing of Geomorphodiversity Linked to Biodiversity—Part III: Traits, Processes and Remote Sensing Characteristics", *Remote Sensing*, 14, 2279, 2022.

[6] Patrick Kacic, and Claudia Kuenzer, "Forest Biodiversity Monitoring Based on Remotely Sensed Spectral Diversity—A Review", *Remote Sensing*, 14, 5363, 2022.

[7] Nathalie Pettorelli, Jon Olav Vik, Atle Mysterud, Jean-Michel Gaillard, Compton J. Tucker, and Nils Chr. Stenseth, "Using the satellite-derived NDVI to assess ecological responses to environmental change", *TRENDS in Ecology and Evolution*, 20, 9, 503-510, 2005.

[8] Ran Wang, and John A. Gamon, "Remote sensing of terrestrial plant biodiversity", *Remote Sensing of Environment*, 231, 111218, 2019.

[9] Antonio Gulli, Amita Kappor, an Sujit Pal, *Deep learning with TensorFlow2 and Keras,* Packt Publishing Ltd., Birmingham, UK, 2019.

[10] Shutao Li, Weiwei Song, Leyuan Fang, Yushi Chen, Pedram Ghamisi, and Jón Atli Benediktsson, "Deep Learning for Hyperspectral Image Classification: An Overview", IEEE Transactions on Geoscience and Remote Sensing, 57, 9, 6690-6709, 2019.