



Funded by
the European Union

This paper was submitted at ISSCS 2023

AI-Based Visualization of Remotely-Sensed Spectral Images

Ioana Cristina Plajer^{*,1}, Alexandra Băicoianu² and Luciana Majercsik³

Abstract

With the increase in multispectral and hyperspectral satellite data availability, the necessity of interpreting and processing such data is also growing. Satellite imagery can be used in a wide range of fields, from military and defence applications to ecology, agriculture and forest management. As multi- and hyperspectral images cannot be directly interpreted either by the human eye or by usual computer displays, a visually-consistent mapping of these images is necessary. In this paper we propose an approach based on an artificial intelligence (AI) model for spectral image visualisation in the RGB color space. The visualization is performed by a fully-connected neural network trained on the popular CAVE dataset which we consider being suitable for visualization, as it has a significant color diversity in the visible domain. The coloring method was applied on a hyperspectral PRISMA image. The study offers a visual interpretation of the results obtained with the proposed architecture. The results are promising and will be further used for the true mapping of agricultural areas.

Key words: spectral images, satellite data, RGB, neural network, PRISMA satellite, CAVE dataset, linear interpolation

1 Introduction

Each type of material has a unique spectral fingerprint, which means that light is absorbed differently by objects with different properties. Multispectral (MS) and hyperspectral (HS) sensors can capture tens or even hundreds of spectral bands therefore they are much more sensitive to small changes in an object's reflectance or radiance. In the images produced by such sensors, objects are described by additional parameters besides the descriptive geometric data, as each pixel also contains spectral information about the chemical composition of the objects compared to RGB images. For this reason, the former is increasingly used in many remote sensing fields such as agriculture [1, 2, 3], forestry [4, 5], ecology and environmental monitoring [6], and military and industrial applications.

^{1*} *Corresponding author*, Faculty of Mathematics and Informatics, *Transilvania University of Braşov*, Romania, e-mail: ioana.plajer@unitbv.ro

²Faculty of Mathematics and Informatics, *Transilvania University of Braşov*, Romania, e-mail: a.baicoianu@unitbv.ro

³Faculty of Mathematics and Informatics, *Transilvania University of Braşov*, Romania, e-mail: luciana.carabaneanu@unitbv.ro

The visualization of MS and HS data as RGB images is extremely important because an RGB image serves as an interface between the human eye and the multidimensional data space, helping the viewer to correlate pixel areas to the surface features they represent. To generate realistic RGB images using spectral images, a number of visualization methods have been proposed in literature, such as band selection [7][8], PCA based methods [9] or, more recently, machine learning [10][11]. However, research in this field is still quite scarce. For this reason, one of the main objectives of our investigation was to obtain meaningful corresponding RGB images for MS and HS images using a fully-connected neural network (FCNN) trained for this task.

The spectral image sensors provide images with a large number of contiguous spectral channels per pixel and the sources from which these images come are diverse, therefore visualizing these massive datasets is not simple and straightforward. The process of visualization has a particular value for users who need to evaluate the importance of data, which is why we have proposed in this study a suitable architecture for this assignment. As the human visual perception is not necessarily a linear process, the use of a linear color space to represent images might result in an unpleasant effect on the human observer, generating the impression of high contrast between the darker and the lighter areas. Furthermore, pairs of colors which in the RGB color cube are at the same Euclidean distance might be perceived totally different by humans, the RGB cube thus being a perceptually-nonlinear space [12]. Taking this into consideration, the nonlinearity of a neural network (NN) could provide a good way of modeling the transform function from multispectral to tridimensional RGB, in order to offer a perceptually proper interpretation for human users.

The aim of this article is to provide answers to some research questions within the context of spectral image visualization. One of the open questions is if an AI model is capable of learning the correspondence between MS or HS reflectance curves and RGB triplets. The results of our experiments confirm that by training a NN to learn the RGB equivalent for spectral pixels, visually consistent RGB images can be generated. Two important considerations should be mentioned here. The first one is that at the moment, the interpretation of the results has been done from a visual point of view and considering some common metrics. The second is that the network testing was done using a single HS satellite image, and this should be extended to test other satellite images. Another open concern in the field of spectral image visualization is to discover what are the main characteristics of a data set suitable for optimal coloring results. As our tests have shown, in order to have a good coloring, it is important to use for the training of the NN a dataset with a significant color diversity in the visible domain. The amount of data on which training is done must be large enough and correct standardization before training is extremely important. A third issue considered was to feature the advantages of coloring spectral images using a NN over other alternative visualization techniques. We can mention among these advantages the relatively easy adaptability through an interpolation process to inputs having a different spectral resolution. Another advantage of a NN is its non-linearity, the network

being thus able to emulate the nonlinear visual perception of human users and so to enhance the coloring output. Furthermore enhancing the training data set could be an option to achieve improved visual results.

2 Datasets

The high-resolution spectral information provided by multi- and hyperspectral imagery has changed the way we think about environmental and ecosystem phenomena. More and more such data are becoming available for scientific and practical purposes, and analysing them is a first step in better understanding the phenomena mentioned above. Various datasets are available and known including CAVE [13], UGR [14], UEA Colour Group datasets [15]. For this study we chose the CAVE dataset, which is well known in literature, because it offers a good diversity of colors, the images were acquired in controlled environment and for each MS image an RGB correspondent is available. For testing we used a HS image provided by the relatively new PRISMA satellite.

2.1 CAVE Dataset

The CAVE high-resolution MS image dataset [13] contains 32 images of indoor scenes. Each MS image has a resolution of 512×512 pixels and covers a wavelength range in [400 - 700 nm], sampled at 10 nm intervals, providing a total of 31 channels. Each scene has also a unique corresponding color image representation, rendered under a neutral daylight illuminant and displayed using sRGB values. Some samples from CAVE dataset are displayed in Figure 1.



Figure 1: RGB images from the CAVE database.

2.2 PRISMA Image

The PRISMA image used for testing is one of the images captured by the Italian Space Agency (ASI)'s PRISMA hyperspectral satellite on 18th of October 2022 in the north of Brasov county, Romania. The hyperspectral sensors of the satellite are able to recall images in a wavelength range of 239 spectral bands

between [400 - 2500 nm], 66 in the Visible Near Infra Red [400 - 1010 nm], and 173 in the Short Wave Infra Red [920 - 2500 nm], with a spectral sampling interval smaller than 12 nm. The images have a spatial resolution of 1000×1000 pixels, with a ground sample distance of 30 m [16]. The spectral bands used in the experiments are in the visible domain from 406 nm to 713 nm, roughly 8 nm sampled.

3 The Proposed AI Model Architecture

In practice, a number of classical algorithms are commonly used to visualize MS images, for example by using band selection or linear color formation models. These approaches presume the choice of the appropriate illuminant and the modelling of the nonlinearity of human perception by gamma correction. The results of these algorithms often present low visual quality. As mentioned in the introduction, a NN approach could as well simulate the nonlinear human perception as offer a general solution for MS images acquired by different systems.

3.1 Model Description

As the coloring for visualization of the MS images can be formulated as a regression problem, it seems appropriate to use for this purpose a FCNN, inspired by the one proposed in [17]. As the number of wavelength present in the CAVE dataset used for training the model are 31, this is the number of neurons in the input layer of the network. The network is constructed with 3 fully connected hidden layers and an output layer containing three neurons, one for each of the RGB color channels. Thus, the model estimates for each MS pixel input an RGB output, see Figure 2.

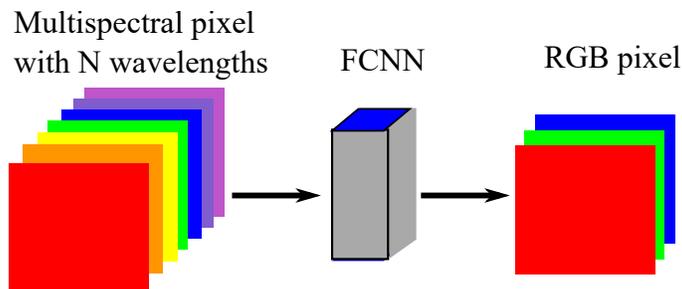


Figure 2: Model pipeline.

3.2 Training of the Model

The FCNN model was trained on all the pixels of all the images in the CAVE dataset, collected into a single randomly shuffled set. This set was partitioned

into train and test set and the model was trained until the loss decay on both train and test set was stabilized. During each training epoch randomly selected pixel batches of the training set were passed through the network, using PyTorch data loader. The training stages can be synthesized by the following methodology.

Training Steps

1. Loading of all the pixels from all the images into one dataframe [18].
2. Random shuffling of all the pixels.
3. Partitioning of the set into train set - 75% of the pixels - and test set - 25% of the pixels.
4. Standardization of the training using the PyTorch standard scaler and using the transformation on test dataset [19].
5. Training the model with random batches of 2048 pixels.

4 Analysis and Interpretation of Results

Following a set of experiments it proved sufficient to train the model on 150 epochs. To validate the correctness of the training method, k-fold cross validation with 10 folds was used. On each fold the average losses on the training and on the test were calculated and plotted. The results on most folds were similar, indicating the correctness of the approach. The loss decay during training on one of the folds is presented in Figure 3. It can be seen, that the loss on the training set has a steep decay, while for similar results on the test set, the model needs more training.

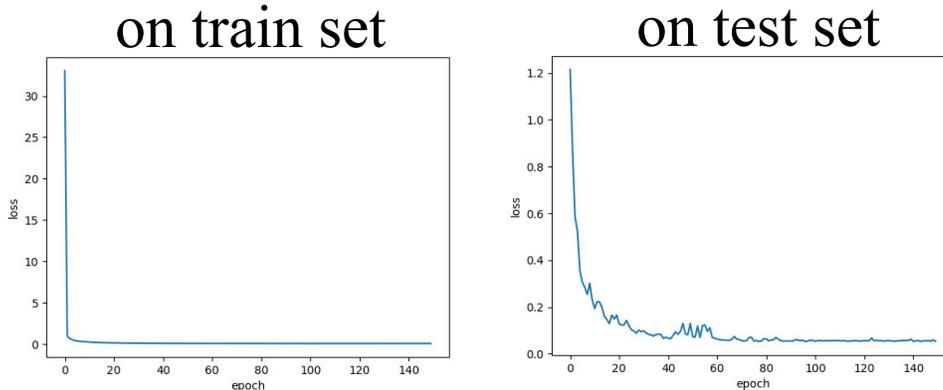


Figure 3: Loss decay on CAVE dataset.

During the k-fold process all the weights were saved and the best ones were used for coloring all the MS images of the CAVE dataset. The visual outputs were accurate, further validating the chosen model. Figure 4 is presenting the coloring result on two samples of the dataset together with the provided RGB label image.

One may notice that the two images are visually identical. The similarity measures for the pairs of input-output images in the train set confirm the visual results. The considered quality metrics were Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR). For MSE the values for the pairs are between [0.09-0.17] and the values for PSNR are between [55.71-58.45].

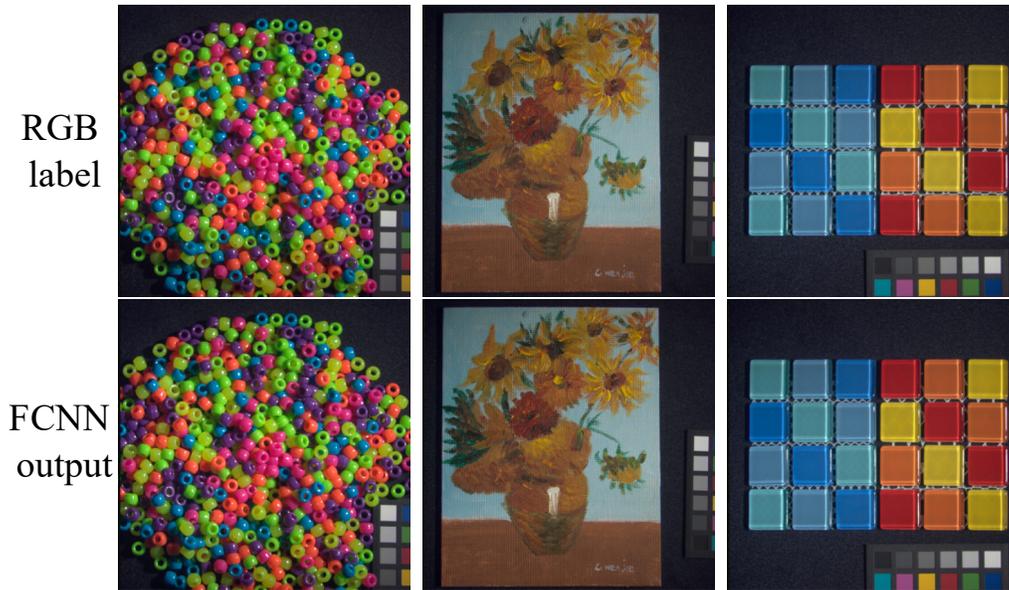


Figure 4: Coloring of two MS images from the CAVE dataset using the trained FCNN model.

The trained model was also used to visualize HS satellite images obtained from the PRISMA satellite. It must be mentioned, that the spectral bands of the PRISMA images differ slightly from those of the images in the CAVE dataset. We considered from the PRISMA images only the wavelength in the visual range and linearly interpolated them to fit those of the CAVE dataset. The coloring of the HS images was performed by the following procedure.

Inference

1. Load the pixels of the PRISMA image into a dataframe.
2. Linearly interpolate the values as to fit the wavelength to those of the CAVE images. As the spectral range of the PRISMA image is very similar to the one of CAVE images, we considered linear interpolation as accurate enough.
3. Standardize the pixels of the dataframe relative to their mean and variance using the standard PyTorch scaler.
4. Pass each pixel through the model to predict the corresponding (R, G, B) triplet.
5. Construct the RGB-image with respect to the original size of the PRISMA image and save as PNG file.

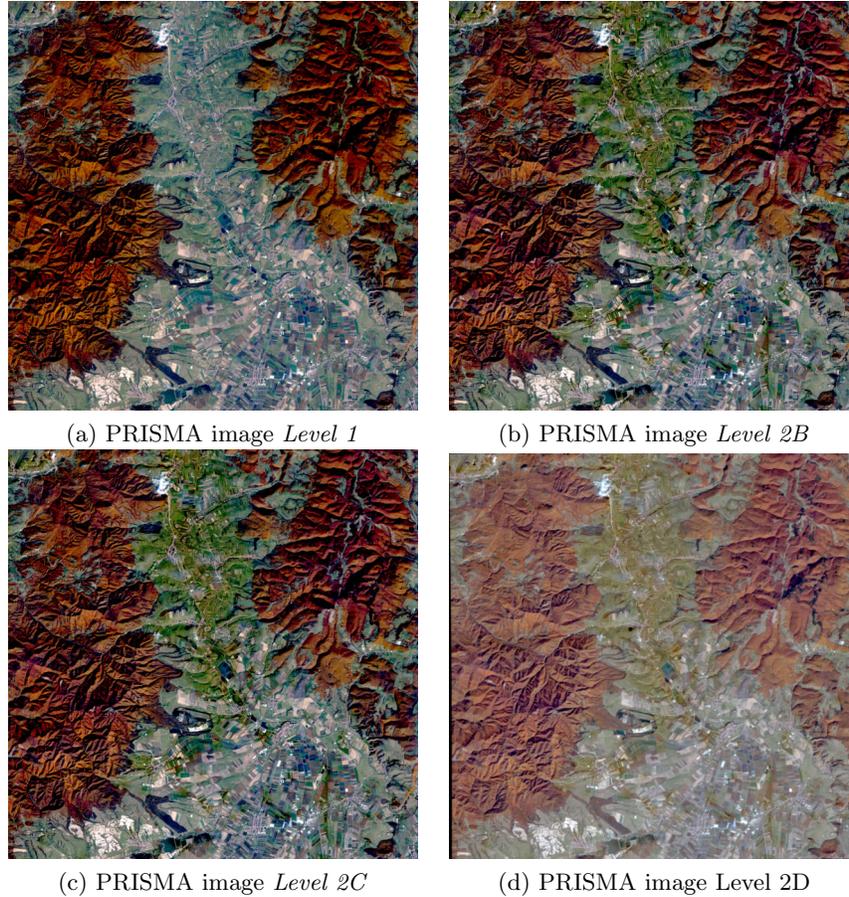


Figure 5: Coloring of four products of PRISMA image using the FCNN model with weights trained on CAVE dataset before and after corrections.

A result of this coloring on a PRISMA image is presented in Figure 5. Four different product levels of the same PRISMA image were visualized by the network, each of them presenting the initial HS image, while the other three present subsequent levels of correction. *Level 1* is a top of the atmosphere radiance imagery. After the atmospheric correction and geolocation of *Level 1*, *Level 2B* image contains the information about the reflected radiance of the Earth’s surface, and *Level 2C* has the information about the boundary reflection coefficient, aerosol optical thickness and water vapor map. The last level, *Level 2D*, represents the image after all the transforms plus orthorectification [16, 5]. According to [16], “the orthorectification process foresees the correction of all image distortions caused by the collection geometry (this includes the optical sensor characteristics) and the variable terrain”.

As can be seen in Figure 5, the results of the coloring are promising. The expected natural coloring of the image is achieved in all four cases, given that the image was acquired late October. The region of the valley is accurately rendered and the agricultural parcels can be clearly distinguished. There are still some

artefacts present, probably due to highly reflective surfaces. This might be due to the fact that the training dataset was acquired in the indoor environment and has a specific spectral signature. Another aspect might concern the interpolation method, opening the research possibility of other interpolation or mapping methods.

It also can be noticed that after corrections in *Level 2*, the slightly blueish coloring of the Level 1 HS image is attenuated, and the green color of the vegetation is enhanced. The most pronounced difference in color to the original sample can be observed after orthorectification. *Level 2C* leads to a higher contrast and more vivid colors. *Level 2B* and *Level 2C* images are the most colorful.

5 Conclusions and Future work

This paper presents a novel fully connected model for the task of HS/MS images coloring. The given architecture has yielded encouraging results. This study was developed on two known datasets in the field, CAVE and PRISMA, but we are interested in running additional tests on several datasets with different characteristics, the indoor category which is a controlled environment, and also the outdoor variety which is a natural environment. Such an approach can bring generality to the solution already offered, but also optimal and diversified results. However, the proposed network can further be optimized, for instance, one minor restriction that we have noticed through testing is that the current model still generates errors for highly reflective surfaces. In addition, at this point of the study, visually it can be seen that the results are meaningful. Still, there is necessary to validate the results with other different classical (Euclidean Distance, Mahalanobis Distance, Cosine Similarity, etc) and/or specific (Structural Similarity Index, Universal Quality Image Index, etc) metrics. We also want to improve the performance and results of the model by enhancing the input datasets.

Acknowledgment

This work was funded from the AI4AGRI project entitled “Romanian Excellence Center on Artificial Intelligence on Earth Observation Data for Agriculture”. The AI4AGRI project received funding from the European Union’s Horizon Europe research and innovation program under the grant agreement no. 101079136. The hyperspectral image from the PRISMA satellite presented in this paper was kindly provided by the Italian Space Agency (ASI).

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

References

- [1] T. Adão et al., “Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry”, *Remote Sensing*, 2017, Volume 9(11). [Online]. <https://doi.org/10.3390/rs9111110>.

- [2] M. Weiss, F. Jacob, G. Duveiller, "Remote sensing for agricultural applications: A meta-review", *Remote Sensing of Environment*, Volume 236, 2020, pp. 111402, ISSN 0034-4257, [Online]. <https://doi.org/10.1016/j.rse.2019.111402>.
- [3] J. Zhao et al., "Deep-Learning-Based Multispectral Image Reconstruction from Single Natural Color RGB Image—Enhancing UAV-Based Phenotyping", *Remote Sensing*, 2022, Volume 14, pp. 1272. [Online]. <https://doi.org/10.3390/rs14051272>.
- [4] P. Rodríguez-Veiga et al., "Forest biomass retrieval approaches from earth observation in different biomes", *Int. J. Appl. Earth Obs. Geoinf*, 2019, Volume 77, pp. 53–68.
- [5] E. Vangi et al., "The New Hyperspectral Satellite PRISMA: Imagery for Forest Types Discrimination," *Sensors*, 2021, Volume 21, pp. 1182, [Online]. <https://doi.org/10.3390/s21041182>.
- [6] L. Du et al., "A comprehensive drought monitoring method integrating MODIS and TRMM data," *Int. J. Appl. Earth Obs. Geoinf*, 2013, Volume 23, pp. 245–253.
- [7] B. Demir, A. Celebi, S. A. Erturk, "Low-complexity approach for the color display of hyperspectral remote-sensing images using one-bit-transform-based band selection", *IEEE Trans. Geosci. Remote Sens*, 2008, Volume 47, pp. 97–105.
- [8] S. Le Moan, A. Mansouri, Y. Voisin, J.Y. Hardeberg, "A constrained band selection method based on information measures for spectral image color visualization", *IEEE Trans. Geosci. Remote Sens.*, 2011, Volume 49, pp. 5104–5115.
- [9] H.A. Khan, M.M. Khan, K. Khurshid, J. Chanussot, "Saliency based visualization of hyper-spectral images", *Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, Italy, 2015, pp. 1096–1099.
- [10] P. Duan, X. Kang, S. Li, "Convolutional neural network for natural color visualization of hyperspectral images", *Proceedings of the IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, Japan, 2019, pp. 3372–3375, doi: 10.1109/IGARSS.2019.8900359.
- [11] R. Tang, H. Liu, J. Wei, W. Tang, W. "Supervised learning with convolutional neural networks for hyperspectral visualization", *Remote Sens. Lett*, 2020, 11, pp. 363–372.
- [12] C. A. Poynton, *A Technical Introduction to Digital Video*, John Wiley & Sons, Inc., 1996, ISBN: 047112253X.

- [13] F. Yasuma, T. Mitsunaga, D. Iso, and S.K. Nayar, "Generalized Assorted Pixel Camera: Post-Capture Control of Resolution, Dynamic Range and Spectrum", Technical Report, Department of Computer Science, Columbia University CUCS-061-08, 2008.
- [14] J. Eckhard, T. Eckhard, E.M. Valero, J.L. Nieves, E. Garrote Contreras, "Outdoor scene reflectance measurements using a Bragg-grating-based hyperspectral imager", *Applied Optics*, Volume 54 (13), pp. D15-D24, 2015.
- [15] UEA Colour Group Datasets, [Online]. Available: <https://colour.cmp.uea.ac.uk/datasets/multispectral.html>
- [16] ASI. Prisma Products Specification Document Issue 2.3 Date 12/03/2020. [Online]. Available: <http://prisma.asi.it/missionselect/docs/PRISMA%20Product%20Specifications.Is2.3.pdf>
- [17] R.M. Coliban, M. Marincas, C. Hatfaludi, M. Ivanovici, "Linear and Non-Linear Models for Remotely-Sensed Hyperspectral Image Visualization", *Remote Sensing*, Volume 12(15), 2020, [Online]. <https://doi.org/10.3390/rs12152479>.
- [18] DataFrame, [Online]. Available: <https://pytorch.org/torcharrow/beta/dataframe.html> (accessed on 17.12.2022)
- [19] StandardScaler, [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html> (accessed on 10.11.2022)