Contrastive learning for regression on hyperspectral data

Paul Honeine, Mohamad Dhaini, Maxime Berar, Antonin van Exem

CONTRASTIVE LEARNING FOR REGRESSION ON HYPERSPECTRAL DATA

*Mohamad Dhaini[1,2] *, Maxime Berar[1], Paul Honeine[1], Antonin Van Exem[2]*

[1] Univ Rouen Normandie, INSA Rouen Normandie, Université Le Havre Normandie, Normandie Univ, LITIS UR 4108, F-76000 Rouen, France
[2] Tellux, 76650 Petit-Couronne, France

ABSTRACT

Contrastive learning has demonstrated great effectiveness in representation learning especially for image classification tasks. However, there is still a shortage in the studies targeting regression tasks, and more specifically applications on hyperspectral data. In this paper, we propose a contrastive learning framework for the regression tasks for hyperspectral data. To this end, we provide a collection of transformations relevant for augmenting hyperspectral data, and investigate contrastive learning for regression. Experiments on synthetic and real hyperspectral datasets show that the proposed framework and transformations significantly improve the performance of regression models, achieving better scores than other state-of-the-art transformations.

Index Terms— Contrastive Learning, Hyperspectral Data, Regression, Data Augmentation

1. INTRODUCTION

Hyperspectral imagery offers valuable insights into the physical properties of an object or area without the need for physical contact. Hyperspectral sensors capture a broad range of information within the light spectrum, often spanning hundreds of contiguous bands across a wavelength range of approximately 500 nm to 2500 nm. This capability allows each material to possess its own distinct spectral signature. This type of data has been gaining significant attention from the signal processing and machine learning community, and became a direct application for classification [1], regression [2], unmixing [3] and object detection [4] tasks.

Recently, due to the limitation of supervised learning techniques to labeled data, self-supervised learning methods have been gaining popularity to learn general representations from unlabeled data thus having more discriminative features in the used neural networks. In this context, contrastive learning [5] is a self-supervised approach that provides such discriminative features by maximizing the similarity between similar data examples and minimizing it between dissimilar ones. The use of such approach on hyperspectral data is still encountering some challenges especially regarding the augmentation techniques to be used. Data augmentation techniques often used for general images (e.g., image rotation) are not applicable to hyperspectral data. In this article, we investigate the use of contrastive learning to improve regression results on hyperspectral data, with application in hyperspectral unmixing and pollution estimation. The contributions can be seen as following:

1. We revisit some popular augmentation techniques, often used in computer vision, to fit into hyperspectral data. Besides, we make use of well-known radiative transfer models that simulate the atmospheric effect on hyperspectral data to generate augmented spectra.
2. We adapt the cross-entropy based contrastive loss to fit regression tasks by incorporating the use of a ball of given radius to select positive and negative pairs.
3. We show the performance of our method in real-world scenario, including unmixing and prediction of pollution concentration in soil data.

The rest of the paper is organized as follows. Section 2 highlights some of the related work on contrastive learning with hyperspectral data. In Section 3, we present the core ideas behind our proposed method. The experimental studies with the obtained results are presented in Section 4. The contributions and future steps are summarized in Section 5.

2. RELATED WORK

Recently, there have been some studies investigating the use of contrastive learning on hyperspectral data. However, the majority of these studies were targeting classification tasks. [6] introduced a contrastive learning network based on a nearest neighbor augmentation scheme, by extracting similarities from nearest neighbor samples to learn enhanced semantic relationships. In [7], three data augmentation methods were introduced to enhance the representation of features extracted by contrastive learning. These methods were band erasure, gradient mask and random occlusion. In the same context, [8] introduced a spectral-spatial contrastive clustering network, with a set of spectral-spatial augmentation techniques that includes random cropping, resizing, rotation, flipping, and blur-

* Corresponding Author
This work is funded by Tellux Company and ANRT (Association Nationale de la Recherche et de la Technologie).

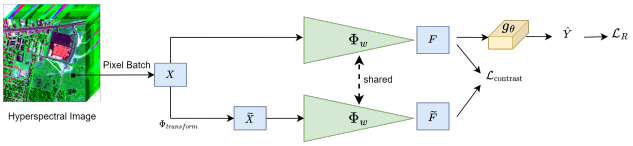


Fig. 1. Architecture of the proposed method.

ring for the spatial domain, as well as band permutation and band erasure for the spectral domain. [9] introduced a framework for detecting surface changes in hyperspectral images using self-supervised learning. Their main contributions involved an augmentation technique based on a Gaussian noise and a contrastive loss based on Pearson coefficient and negative cosine correlation. Similarly, [10] introduced a neighborhood contrastive subspace clustering network for unsupervised classification of large hyperspectral images based on a superpixel pooling autoencoder. Besides, [11] demonstrated that by pre-training an encoder on unlabeled pixels using the proposed Barlow-Twins algorithm, accurate models can be obtained even with a small number of labeled samples. For Unmixing, the authors in [12] introduces in addition to the standard reconstruction error loss often used, a contrastive loss is applied to the endmember matrix to promote separability between endmembers and another regularization loss to encourage the minimum simplex volume constraint in endmembers.

3. METHOD

3.1. Proposed Framework

The task we are addressing is a pixel-level regression on hyperspectral data. The process starts by extracting a batch of N pixels $X = [x^1(\lambda), x^2(\lambda), \ldots, x^N(\lambda)]^\top \in \mathbb{R}^{N\times b}$ from a hyperspectral image, where λ is the wavelength with b total number of wavelengths. X is then transformed using a defined spectral transformation $\Phi_{transform}$ to get $\widetilde{X}$, which will be passed along with the original batch to a shared feature extractor Φ_w to get $\widetilde{F}$ and $F = [f^1, f^2, \ldots, f^N]^\top$ respectively. These features are passed into a regression network g_θ that will generate the regression labels $\hat{Y} = [\hat{y}^1, \hat{y}^2, \ldots, \hat{y}^N]^\top \in \mathbb{R}^{N\times s}$ for s prediction variables. The network will be trained with a joint contrastive and regression losses simultaneously. The architecture of the proposed method can be seen Figure 1.

3.2. Spectral Data Augmentation Methods

Most of data augmentation techniques focus on the spatial domain of images such as geometric transformations, noise injection, and color distortions. These techniques cannot be adapted directly to the spectral domain as the used transformations should not create strong deformations to the original spectrum in order not to loose useful information. In the following, we propose a new list of transformations that can be adequate for the spectral domain:

1. The Spectral Shift involves shifting the spectrum in the wavelength, such as

$$\widetilde{x}(\lambda) = x(\lambda - \Delta), \tag{1}$$

   By applying random shifts, the model can learn to be more robust to spectral variations in the input data.
2. The Spectral Flipping involves reversing the order of spectral bands in a spectrum according to the following:

$$\widetilde{x}(\lambda) = x(\lambda_{min} + \lambda_{max} - \lambda). \tag{2}$$

   where λ_{min} and λ_{max} are the minimum and maximum wavelengths in the spectrum, respectively. This augmentation technique can help the model learn the invariance to the order of spectral bands and handle potential inconsistencies in the spectral ordering across different datasets or sensors.
3. The Scattering Hapke's Model [13] simulates realistic scattering effects caused by light interaction with surfaces according to the following formula:

$$\widetilde{x} = \frac{\omega}{(1 + 2\mu_1\sqrt{1-\omega})(1 + 2\mu_2\sqrt{1-\omega})}, \tag{3}$$

$$w = 1 - \left(\frac{\sqrt{\mu_0^4 x^2 + (1 + 4\mu_0^2 x)(1-x)} - 2\mu_0 x}{1 + 4\mu_0^2 x} \right)^2 \tag{4}$$

   where ω is the single scattering albedo of the material, and μ_1 (resp. μ_2) is the cosine of the angle between the incoming (resp. outgoing) radiation and the normal to the surface and μ_0 is the initial cosine angle of the incoming radiation. By varying μ_1 and μ_2, we can generate spectra with different scattering effect.
4. The Atmospheric Compensation model [14] helps the model learn to handle atmospheric effects. Assuming full visibility and that the adjacency effect is negligible, this model is given by:

$$\widetilde{x} = x \frac{E_{\text{sun-gr}}\,\mu_1 + E_{\text{sky}}}{E_{\text{sun-gr}}\,\mu_2 + E_{\text{sky}}}, \tag{5}$$

   where $E_{\text{sun}-\text{gr}}$ denotes the solar radiance observed at the ground level, and E_{sky} denotes the skylight. The parameters μ_1 and μ_2 are the cosines of the angles between the surface normal and the direction of the sun at each pixel and at the calibration panel, respectively.
5. The Elastic Distortion consists in a displacement grid on the wavelength axis, such as

$$\widetilde{x}(\lambda) = x(\lambda + \epsilon(\lambda)) = x\Big(\lambda + \sum_{i=1}^{N_G} A_i e^{-\frac{(\lambda-\lambda_i)^2}{2\sigma^2}}\Big) \tag{6}$$

where $\epsilon(\lambda)$ is the random displacement function applied to the input signal, N_G is the number of Gaussian kernels used to generate the distortion, A_i is the amplitude of the i-th Gaussian kernel, λ_i its center wavelength, and σ controls its width. This technique can help the model learn to handle spectral variations caused by distortions or misalignments.

In addition to this proposed list, there exist some augmentation techniques that are already presented in literature:

1. The Band Erasure [8] randomly removing certain wavelength from the spectral data.
2. The Band Permutation [8] involves randomly permuting the order of the spectral bands.
3. The Nearest Neighbor [6] involves creating new synthetic samples based on the average of the closest samples.

3.3. Contrastive Learning

After augmenting the data X with different views $\widetilde{X}$, both are forwarded to the feature extractor Φ_w, which extracts the corresponding features, F and $\widetilde{F}$ respectively, and then will be later on passed to a regression head g_θ. To optimize the feature extractor, we train the neural network so that two spectral data transformed via different augmentation techniques and with close regression labels should share similar features in the latent space, while different spectra with different labels should be far away. This type of training can be achieved using a self-supervised contrastive loss [5].

In classification, selecting similar data pairs (referred to as positive pairs) can be done by taking the transformed version of an image as well as other images that belong to the same class, while dealing with the rest as negative ones. In regression, as we do not have class labels, the alternative is to define for the i-th sample a ball $\mathbf{B}^i$ of radius r where positive pair j is selected as following:

$$r \geq \left\| y^i - y^j \right\|_2 . \tag{7}$$

The common contrastive loss used in most recent work is based on the cross entropy, which can be written as

$$\mathcal{L}_{Contrastive} = -\frac{1}{N} \sum_{i=1}^{2N} \sum_{j \in \mathbf{B}^i} \log \frac{\exp\left(\operatorname{sim}\left(f^i, f^j\right)/\tau\right)}{\sum_{k \notin \mathbf{B}^i} \exp\left(\operatorname{sim}\left(f^i, f^k\right)/\tau\right)} \tag{8}$$

where $\operatorname{sim}(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^T \boldsymbol{v}/(\|\boldsymbol{u}\|\|\boldsymbol{v}\|)$ is the cosine similarity between two vectors, and τ is a temperature scalar. By minimizing this loss, the similarity between samples i-th and j-th is maximized while minimizing the similarity between i-th and k-th samples. For training, the contrastive loss is combined with a standard mean squared error regression loss according to the following:

$$\mathcal{L}_{total} = \mathcal{L}_{\text{R}} + \alpha\, \mathcal{L}_{\text{Contrastive}}, \tag{9}$$

$$\mathcal{L}_{\text{R}} = \frac{1}{N} \sum_{i=1}^{N} \left\| y^i - g_\theta\left(f^i\right) \right\|^2 . \tag{10}$$

4. EXPERMIMENTS & RESULTS

4.1. Synthetic Data

For the synthetic data, four random endmembers were selected from the USGS digital spectral library [15]. Each endmember is composed of 224 contiguous bands. A total of 100×100 pixels were generated with abundances following a Dirichlet distribution. Additive zero-mean Gaussian noise was added to the data with a signal-to-noise ratio of 20 dB. We considered a polynomial post-nonlinear mixing of the endmembers where the nonlinearity is represented by the element-wise product of two linear mixtures as following:

$$x = Ma + Ma \odot Ma + n, \tag{11}$$

with M,a and n being endmember matrix, abundances and noise vector respectively and $\odot$ denotes the element-wise product. The model was trained to predict the abundances from the input of mixed spectra. For the architecture of Φ_w, we used three fully connected layers mapping the input shape to 128, 64 and 32 nodes respectively. For the regression network g_θ, we use two fully connected layers that further reduces the dimension to 16 then to $s = 4$ abundances to be estimated in this case. We trained our network with 100 epochs and a batch size of 32. We evaluated the influence of each transformation with two regression metrics, R^2 score and the mean absolute error (MAE). The results are shown in Table 1. We can see that the all the transformations presented in the table improved the results compared to the baseline model where no contrastive loss was applied. Besides, we can see that the shift and elastic transformations provided the top two results. These can be justified as elastic transform can create similar adjustments to spectral data to that created by the shift transform. Besides, in Figure 3(a) we compare the prediction error distribution with and without the use of the contrastive loss. We can see that the mean of the error distribution is shifted closer to zero due to the contrastive learning, which highlights the improvement in the prediction values.

4.2. Real Soil Data

To evaluate our model on a more challenging environment, we considered a dataset provided by Tellux for soil pollution analysis using hyperspectral imaging [16]. The used soil is formed of a mix between different soil matrices (sand, clay, silt, organic material ...). The dataset is composed 10000 spectra with a spectral range [1130-2450 nm] and containing hydrocarbon pollution concentration ranging [0-20000 mg/kg] where each soil mixture is mixed with hydrocarbon pollutants. An example of these spectral data as well as their

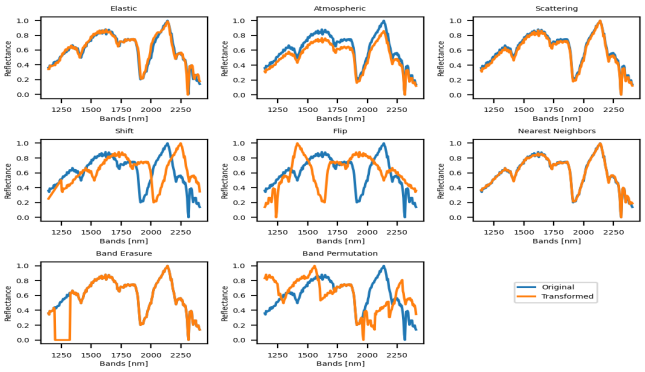
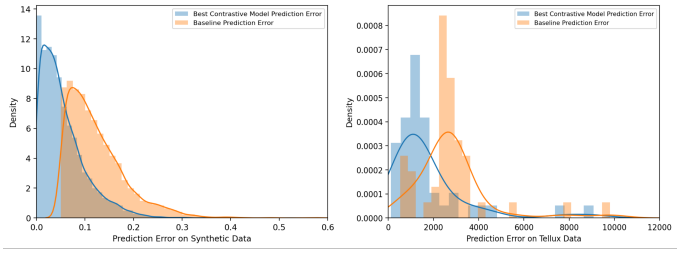


Fig. 2. Original and Transformed Examples from Real Data.



(a) Synthetic Data (b) Tellux Data

Fig. 3. Prediction Error Distribution.

transformed versions obtained by the proposed transformations can be seen in Figure 2. We used the same settings (architecture, batch size, ...) as for the synthetic data and trained it with 1000 epochs. In the same way as with the synthetic data, the obtained results given in Table 2 show a clear improvement in R^2 and MAE scores for all models with contrastive loss compared to the baseline model. Besides, shift, flip and elastic transformations provided the best results, which can be justified as these methods give invariance to the model for spectral bands order while maintaining the spectral profile. In addition, Figure 3(b) shows the shift in the distribution error thanks to the use of the contrastive loss.

Table 1. Regression Results on Synthetic Data.

	R^2	MAE
Baseline (No Contrastive)	0.55 ± 0.004	0.073 ± 0.07
Band erasure [8]	0.62 ± 0.003	0.064 ± 0.05
Band Permutation [8]	0.63 ± 0.003	0.063 ± 0.05
Nearest Neighbor [6]	0.61 ± 0.004	0.065 ± 0.05
Scattering	0.64 ± 0.004	0.061 ± 0.06
Atmospheric	0.65 ± 0.004	0.059 ± 0.06
Flipping	0.62 ± 0.005	0.062 ± 0.07
Elastic	0.66 ± 0.003	0.058 ± 0.05
Shift	0.75 ± 0.003	0.053 ± 0.05

Table 2. Regression Results on Real Data.

	R^2	MAE
Baseline (No Contrastive)	0.45 ± 0.002	2274.04 ± 21.2
Band erasure [8]	0.54 ± 0.003	1620.22 ± 18.1
Band Permutation [8]	0.53 ± 0.003	1700.04 ± 20.5
Nearest Neighbor [6]	0.54 ± 0.003	1850.40 ± 34.5
Scattering	0.56 ± 0.003	1737.26 ± 18.2
Atmospheric	0.55 ± 0.002	1796.63 ± 16.1
Elastic	0.58 ± 0.002	1709.33 ± 18.3
Flip	0.59 ± 0.002	1708.52 ± 18.4
Shift	0.59 ± 0.002	1380.37 ± 16.5

4.3. Combination Study

To provide more robustness of the regression model to various types of variability that might be present in the spectral data, we combine several spectral transformations from the ones presented before. To reduce the number of all possible combinations from the eight presented transformations, we propose an incremental procedure where we start by taking the transformation that provided the best result (namely the shift transformation from Table 1 and Table 2) and then we do all the 2-element combinations. After selecting the best pair, we repeat the process to select the third transformation, and so on. Table 3 and Table 4 provide the R^2 scores of the regression model, as well as the difference ΔR^2 for each incremental update. For simplification, we only show the incremental settings that led to the best combinations. We can see that combining the shift, atmospheric, elastic and scattering transforms can improve the regression metrics when combined together.

Table 3. Combination Study Results on Synthetic Data.

	R^2	ΔR^2
Shift	0.7522	–
Shift + Atmospheric	0.7774	0.0252
Shift + Atmospheric + Scattering	0.7921	0.0147
Shift + Atmospheric + Scattering + Elastic	0.7922	0.0001

Table 4. Combination Study Results on Real Data.

	R^2	ΔR^2
Shift	0.59000	–
Shift + Atmospheric	0.60639	0.01639
Shift + Atmospheric + Elastic	0.61791	0.01152
Shift + Atmospheric + Elastic + Scattering	0.61793	0.00002

5. CONCLUSION

In this paper, we investigated the ability of using contrastive learning for regression tasks on hyperspectral data. We presented a set of spectral transformations adequate for hyperspectral data. Besides, a contrastive loss was added to the training and a clear improvement was seen on the results of both synthetic and real datasets. Future work involves combining the presented framework with another domain adaptation frameworks to generalize knowledge on unseen domains.

6. REFERENCES

[1] Swalpa Kumar Roy, Gopal Krishna, Shiv Ram Dubey, and Bidyut B Chaudhuri, "Hybridsn: Exploring 3-d–2-d cnn feature hierarchy for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 2, pp. 277–281, 2019.

[2] Chao Niu, Kun Tan, Xiuping Jia, and Xue Wang, "Deep learning based regression for optically inactive inland water quality parameter estimation using airborne hyperspectral imagery," *Environmental Pollution*, vol. 286, pp. 117534, 2021.

[3] Mohamad Dhaini, Maxime Berar, Paul Honeine, and Antonin Van Exem, "End-to-end convolutional autoencoder for nonlinear hyperspectral unmixing," *Remote Sensing*, vol. 14, no. 14, pp. 3341, 2022.

[4] Longbin Yan, Min Zhao, Xiuheng Wang, Yuge Zhang, and Jie Chen, "Object detection in hyperspectral images," *IEEE Signal Processing Letters*, vol. 28, pp. 508–512, 2021.

[5] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*. PMLR, 2020, pp. 1597–1607.

[6] Meng Wang, Feng Gao, Junyu Dong, Heng-Chao Li, and Qian Du, "Nearest neighbor-based contrastive learning for hyperspectral and lidar data classification," *IEEE Transactions on Geoscience and Remote Sensing*, 2023.

[7] Jinhui Li, Xiaorun Li, and Yunfeng Yan, "Unlocking the potential of data augmentation in contrastive learning for hyperspectral image classification," *Remote Sensing*, vol. 15, no. 12, pp. 3123, 2023.

[8] Xiang Hu, Teng Li, Tong Zhou, and Yuanxi Peng, "Deep spatial-spectral subspace clustering for hyperspectral images based on contrastive learning," *Remote Sensing*, vol. 13, no. 21, pp. 4418, 2021.

[9] Xianfeng Ou, Liangzhen Liu, Shulun Tan, Guoyun Zhang, Wujing Li, and Bing Tu, "A hyperspectral image change detection framework with self-supervised contrastive learning pretrained model," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 7724–7740, 2022.

[10] Yaoming Cai, Zijia Zhang, Pedram Ghamisi, Yao Ding, Xiaobo Liu, Zhihua Cai, and Richard Gloaguen, "Superpixel contracted neighborhood contrastive subspace clustering network for hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.

[11] Nassim Ait Ali Braham, Lichao Mou, Jocelyn Chanussot, Julien Mairal, and Xiao Xiang Zhu, "Self supervised learning for few shot hyperspectral image classification," in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2022, pp. 267–270.

[12] Abian Hernandez-Guedes, Ryodai Fukushima, Toshihiro Takamatsu, Himar Fabelo, samuel ortega, Juan Ruiz-Alzola, Hiroshi Takemura, and Gustavo M. Callico, "Contrastive Learning approach for blind Hyperspectral Unmixing (CLHU)," 8 2023.

[13] Bruce Hapke, "Bidirectional reflectance spectroscopy: 1. theory," *Journal of Geophysical Research: Solid Earth*, vol. 86, no. B4, pp. 3039–3054, 1981.

[14] Tatsumi Uezato, Richard J Murphy, Arman Melkumyan, and Anna Chlingaryan, "A novel endmember bundle extraction and clustering approach for capturing spectral variability within endmember classes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 11, pp. 6712–6731, 2016.

[15] GA Swayze, RN Clark, TVV King, A Gallagher, and WM Calvin, "The us geological survey, digital spectral library: Version 1: 0.2 to 3.0 mum," in *Bulletin of the American astronomical society*, 1993, vol. 25, p. 1033.

[16] Mohamad Dhaini, François-Joseph Roudaut, Antonin Garret, Ronan Arzur, Audrey Chereau, Fanny Buhler-Varenne, Paul Honeine, Mélanie Mignot, and Antonin van Exem, "Hyperspectral imaging for the evaluation of lithology and the monitoring of hydrocarbons in environmental samples," in *RemTech (International event on Remediation, Coasts, Floods, Climate, Seismic, Regeneration Industry)*, 2021.