Internship proposal: Domain Adaptation by using Riemannian Geometry for Remote Sensing Data

1 Context and application

Domain Adaptation is a field of machine learning that consists of developing learning techniques with a different set of training and test data [Moreno-Torres et al., 2012]. This type of methods is interesting because it allows either to strongly limit the training phase and thus achieve frugality, or to do supervised learning in domains with very little labeled data. A significant number of techniques are used to address this problem such as methods based on optimal transport [Courty et al., 2017]. On the other hand, Riemannian geometry has shown its interest in learning when the features used in classification are subject to constraints such as covariance matrices in EEG [Barachant et al., 2012]. Similarly, in a recent work, it has been shown that these mathematical tools are robust to transformations of the training data [Collas et al., 2022]. The performance loss is then very small if we consider several features and their associated geometry.

2 Objectives

We propose to apply more specifically the tools of Riemannian geometry to the problem of domain adaption. More particularly, we propose to study the interest of deep networks specific to covariance matrices and their associated layers. These networks are based on different Riemannian geometry tools and have shown good performances in computer vision. In particular, we will rely on the following papers [Li et al., 2017, Huang and Gool, 2017] proposing specific layers for covariance matrices. With the help of these papers, the work of the trainee will first consist in assessing the interest of these networks for domain adaptation.

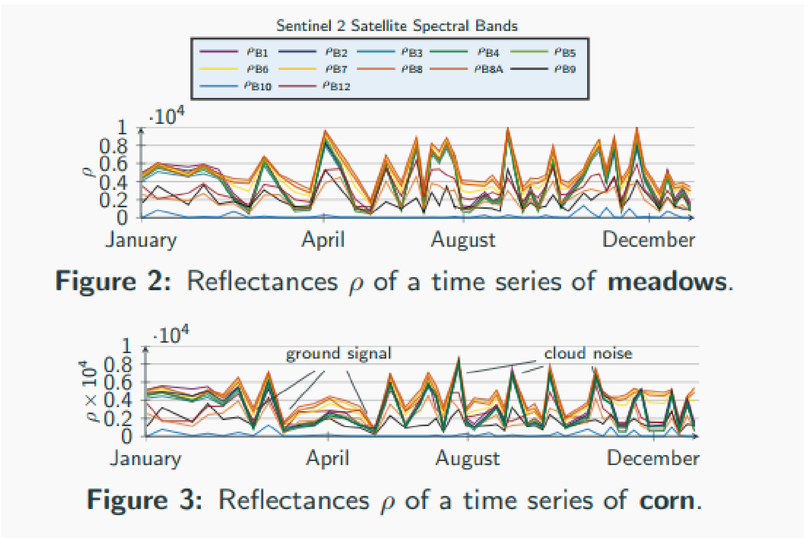

Figure 1: Time series of reflectances of a corn and meadows field. It is easy to see that the spectral information is identical, especially at the beginning of the year. It is the temporal behavior that allows to distinguish them.

Then we will develop new learning methods to study multivariate image time series in remote sensing. Specifically, we will adapt the networks used in [Rußwurm et al., 2020] by adding covariance matrix specific layers inspired by [Li et al., 2017, Huang and Gool, 2017]. The application goal is then to classify agricultural fields which is practically impossible without temporal information as shown in Figure 1.

Finally, we will test the previously developed approaches on the dataset [pas,] which contains time series of optical and SAR images over the same period. We will then be able to perform the learning phase on the optical images and measure the performance loss by applying our algorithms on the SAR data. Indeed, many optical data are labeled which is much rarer for SAR images because they are often a little noisier and a little more difficult to analyze for non-specialists. On the other hand, they are very interesting because they allow a better periodicity and give relevant information even at night or in the presence of clouds.

3 Requirements

Master/Engineering student with knowledge in statistics and machine learning (having followed courses on these topics is strongly advised). Good coding skills in Python.

4 Internship details:

Location: LISTIC laboratory in Univ. Savoie Mont-Blanc, Annecy.

Duration: 4-6 months

Context: This work will constitute the beginning of a collaboration between Univ. Savoie Mont-Blanc and CentraleSupelec (Rennes and Paris)).

Supervision team contacts:

- Ammar Mian, LISTIC, Univ. Savoie-Mont Blanc webpage.
  email: ammar.mian@univ-smb.fr
- Paul Chauchat, IETR, CentraleSupelec Rennes email: paul.chauchat@centralesupelec.fr
- Florent Bouchard, L2S, CentraleSupelec webpage.
  email: florent.bouchard@centralesupelec.fr
- Guillaume Ginolhac, LISTIC, Univ. Savoie-Mont Blanc webpage.
  email: guillaume.ginolhac@univ-smb.fr

References

[pas,] `https://github.com/VSainteuf/pastis-benchmark`.

[Barachant et al., 2012] Barachant, A., Bonnet, S., Congedo, M., and Jutten, C. (2012). Multiclass brain–computer interface classification by riemannian geometry. *IEEE Transactions on Biomedical Engineering*, 59(4):920–928.

[Collas et al., 2022] Collas, A., Breloy, A., Ren, C., Ginolhac, G., and Ovarlez, J.-P. (2022). Riemannian optimization for non-centered mixture of scaled gaussian distributions. *IEEE Transactions on Signal Processing*.

[Courty et al., 2017] Courty, N., Flamary, R., Tuia, D., and Rakotomamonjy, A. (2017). Optimal transport for domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(9):1853–1865.

[Huang and Gool, 2017] Huang, Z. and Gool, L. V. (2017). A riemannian network for spd matrix learning. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, AAAI'17, page 2036–2042. AAAI Press.

[Li et al., 2017] Li, P., Xie, J., Wang, Q., and Zuo, W. (2017). Is second-order information helpful for large-scale visual recognition? In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2089–2097.

[Moreno-Torres et al., 2012] Moreno-Torres, J. G., Raeder, T., Alaiz-Rodríguez, R., Chawla, N. V., and Herrera, F. (2012). A unifying view on dataset shift in classification. *Pattern Recognition*, 45(1):521–530.

[Rußwurm et al., 2020] Rußwurm, M., Pelletier, C., Zollner, M., Lefèvre, S., and Körner, M. (2020). Breizhcrops: A time series dataset for crop type mapping. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences ISPRS (2020).*