Transfer Learning in Spatial–Temporal Forecasting of the Solar Magnetic Field

Eurico Covas

CITEUC - Center for Earth and Space Science Research of the University of Coimbra, Geophysical and Astronomical Observatory of the University of Coimbra, 3040-004 Coimbra, Portugal

Abstract

We attempt to improve on the forecast of the solar surface's longitudinal average of the absolute value of the magnetic field, by using a form of spatial-temporal neural networks. Given that the recording of this dataset only started in 1975 and is therefore quite short, we employ another machine learning technique called transfer learning which has received considerable attention in the literature. This approach consists in first training the spatial-temporal neural network on the much longer dataset (sunspot area data), which starts in 1874, then transferring the trained set of neural network layers and continue training the network on the magnetic flux dataset.

Introduction

While sunspots have been observed since ancient times and have been recorded systematically since the introduction of the telescope in the early 1600s, the actual physical phenomena that presumably originates the sunspots, the solar surface magnetic field, has only been observed consistently since the early 1970s. Both datasets now encompass spatial-temporal dimensions, and the sunspot set is usually depicted in time versus latitude, the so called sunspot butterfly diagram. The solar surface magnetic field "butterfly diagram" equivalent dataset is available since 1974. While most solar cycle forecasting focus on the temporal dimension only, there are some examples of attempts to forecast the sunspot butterfly diagram in both latitude and time, i.e., spatial-temporal forecast [Covas, 2017, Jiang et al., 2018, Covas et al., 2019]. Here we apply the technique of transfer learning to forecasting the solar surface's longitudinal average of the absolute value of the magnetic field by first training a deep neural network on the larger source dataset (sunspot areas), and then transferring wholly or partially the weights of the trained network to the target one and applying it to the target dataset (magnetic field).

Model

Our neural network follows on the approach introduced in [Covas et al., 2019], which draws on an technique based on spatial-temporal embeddings [Parlitz and Merkwirth, 2000, Covas, 2017]. We start with a spatial-temporal series s_m^n. The embedding vectors $\mathbf{x}(s_m^n)$ are constructed using:

$$\mathbf{x}(s_m^n) = \{s_{m-IK}^n, \dots, s_m^n, \dots, s_{m+IK}^n, s_{m-IK}^{n-L}, \dots, s_m^{n-L}, \dots, s_{m+IK}^{n-L}, \dots, s_{m-IK}^{n-JL}, \dots, s_m^{n-JL}, \dots, s_{m+IK}^{n-JL}\}, \qquad (1)$$

where K and L represent the spatial and temporal delays and $2I$ is the number of neighbours in space and J is the number of neighbours in time. The neural network takes as input these embedding vectors and as target the value of s_m^{n+1}. Transfer learning [Caruana, 1995] is a technique whereby one trains a neural network on a larger dataset, and then transfer part or the whole set of layers to another network which is then re-trained on a smaller dataset. The intuition is that once trained the higher layers have acquired the ability to detect generic features and this will be useful on the subsequent task.

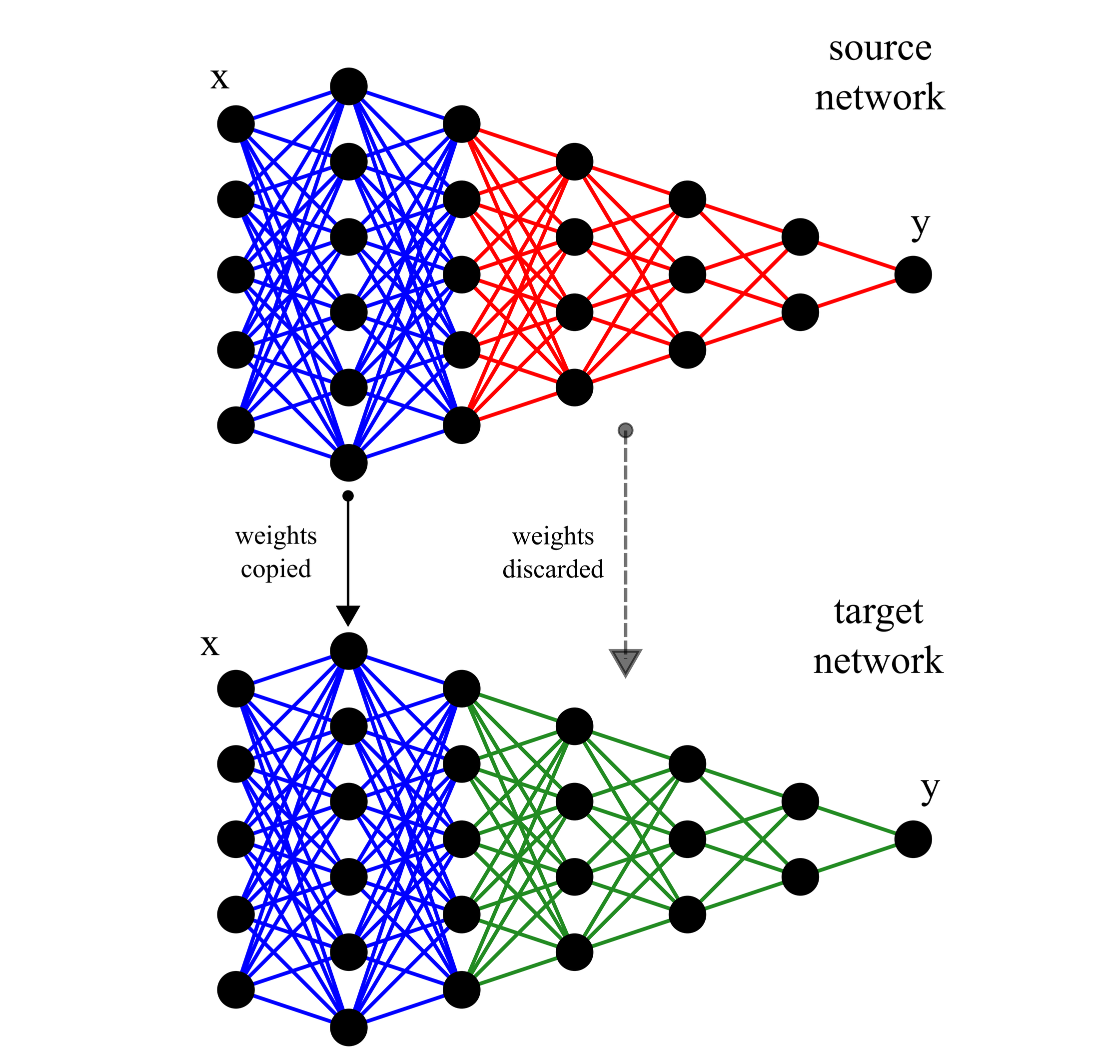

Figure 1:Schematic transfer learning architecture. The source neural network is trained first (on the sunspot area data). Then the first few layers are copied to the target network (in blue), while the other weights (in red) are discarded and new randomly initialized weights are created (in green). The target network is then re-trained on the magnetic field data.

Results

Figure 2 shows our first result, showing the global error during training, as measured by the global cost function $\mathcal{L}_g = \frac{1}{2}\sum\left\|y^{\text{pred}} - y\right\|^2$, where y^{pred} is the predicted value of the neural network and $y = s_m^{n+1}$ is the target value. It shows how the global error convergence improves with transfer learning.

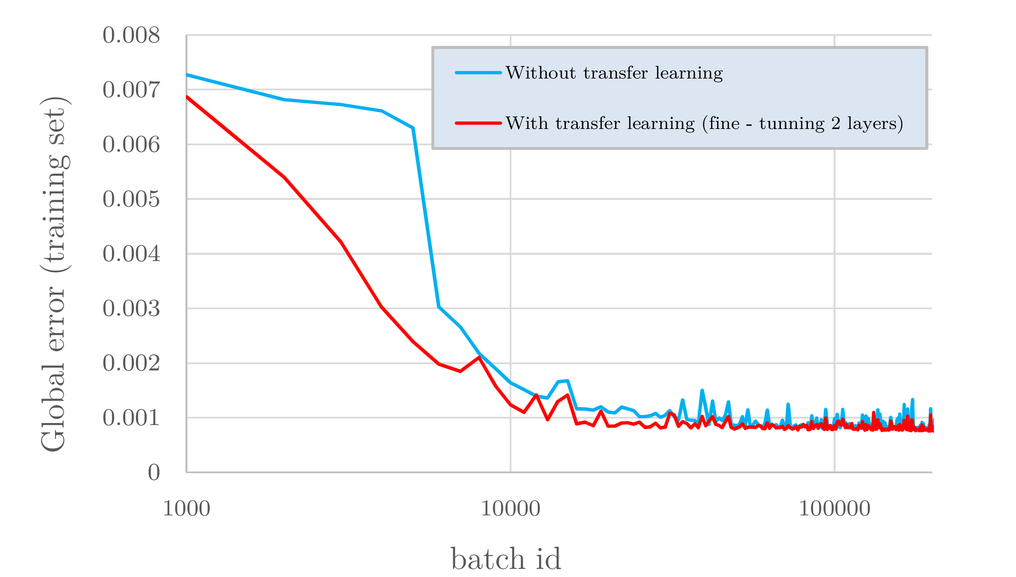

Figure 2:Error function (for the whole training set) on the target network with and without transfer learning.

In order to quantify the accuracy of our forecasts, we use the Structural Similarity Index (SSIM), which is widely used in computer vision [Wang et al., 2004]. It has values SSIM $\in [0, 1]$ and a value of one occurs when one calculates it between two identical images or datasets. The results in Figure 3, obtained using different transfer learning techniques, show that one can improve the forecast by using prior knowledge obtained using the sunspot dataset.

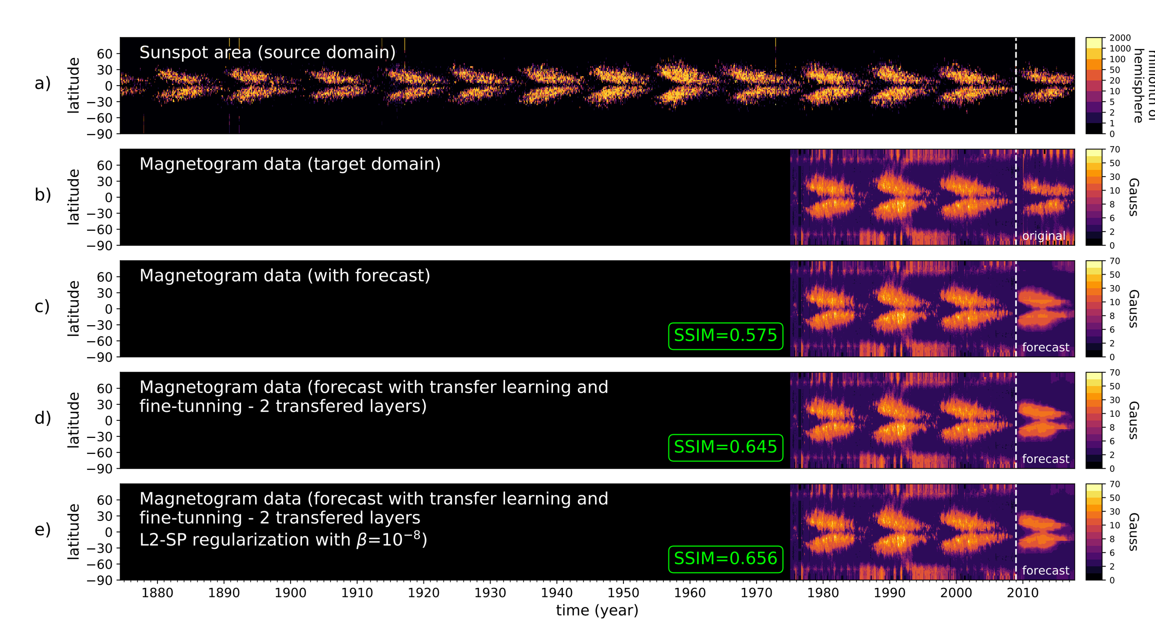

Figure 3:Results of transfer learning against the pure forecast without it. Results improve as one starts using transfer learning, and they improve even further with fine-tuning allowed and with a regularization that forces the fine-tuning not to deviate too much from the source network knowledge [Li et al., 2018].

Conclusion

We use the technique of transfer learning to enhance the performance of spatial-temporal forecasts of the solar surface's longitudinal average of the absolute value of the magnetic field. As the length of this data is quite short, it is quite difficult to forecast with a reasonable precision. However, if we use the sunspot area dataset, available for much longer than the magnetic field data, to first train a source network and then transfer the whole or part of the layers to the target network, we can enhance the forecast and obtain higher values of the SSIM accuracy index.

References

[Caruana, 1995] Caruana, R. (1995).
Learning many related tasks at the same time with backpropagation.
In Tesauro, G., Touretzky, D. S., and Leen, T. K., editors, *Advances in Neural Information Processing Systems 7*, pages 657–664. MIT Press.

[Covas, 2017] Covas, E. (2017).
Spatial-temporal forecasting the sunspot diagram.
Astron. Astrophys., 605:A44.

[Covas et al., 2019] Covas, E., Peixinho, N., and Fernandes, J. (2019).
Neural network forecast of the sunspot butterfly diagram.
Solar Physics, 294(3):24.

[Jiang et al., 2018] Jiang, J., Wang, J.-X., Jiao, Q.-R., and Cao, J.-B. (2018).
Predictability of the solar cycle over one cycle.
The Astrophysical Journal, 863(2):159.

[Li et al., 2018] Li, X., Grandvalet, Y., and Davoine, F. (2018).
Explicit inductive bias for transfer learning with convolutional networks.
CoRR, abs/1802.01483.

[Parlitz and Merkwirth, 2000] Parlitz, U. and Merkwirth, C. (2000).
Prediction of Spatiotemporal Time Series Based on Reconstructed Local States.
Physical Review Letters, 84:1890–1893.

[Wang et al., 2004] Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E. (2004).
Image quality assessment: From error visibility to structural similarity.
IEEE Transactions on Image Processing, 13(4):600–612.

Acknowledgements

We would like to thank David Hathaway from NASA's Ames Research Centre for providing the solar surface's longitudinal average of the absolute value of the magnetic field dataset and the sunspot area data on which all the results of this article are based upon. CITEUC is funded by National Funds through FCT - Foundation for Science and Technology (project: UID/Multi/00611/2013) and FEDER - European Regional Development Fund through COMPETE 2020 – Operational Programme Competitiveness and Internationalization (project: POCI-01-0145-FEDER-006922).

Contact Information

- Web: http://www.euricocovas.com
- Email: eurico.covas@mail.com

CITEUC
CENTRO DE INVESTIGAÇÃO DA TERRA E DO ESPAÇO
UNIVERSIDADE DE COIMBRA