Doppler imaging and deep neural networks


[1]*Tomasz Różański, [2]*Maja Jabłońska

[1]Astronomical Institute, University of Wrocław, Wrocław, Poland
[2]Astronomical Observatory of the University of Warsaw, Warsaw, Poland
* Equal contribution

e-mail: tomasz.rozanski@uwr.edu.pl, mj.jablonska3@student.uw.edu.pl

Uniwersytet Wrocławski

Abstract

Doppler imaging (DI) is an effective inversion method for reconstruction of stellar surface maps. The time series of high resolution spectra used in this approach allow to infer about temperature and overabundance spots, magnetic field and non-radial pulsations. Typically, the DI problem is solved with regularized chi-square minimization that produces solutions that fit observed variations in line profiles. Here we propose a different approach to this task based on deep neural networks (DNN). We test the method on a synthetic dataset and show that DNN models can give results comparable to traditional techniques in a much shorter time, which is important due to the rapidly increasing number of available data.

Data description

Data used in this work are composed of abundance maps and related synthetic time series of spectra. To compute synthetic spectra we used stellar atmosphere modeling codes ATLAS/SYNTHE (Kurucz, 1970). Following the work by Kochukhov (2017), atmosphere model with effective temperature, $T_{eff} = 7000\,\mathrm{K}$, surface gravity, $\log g = 4.0$, microturbulence, $v_{mic} = 2\,\mathrm{km/s}$ and solar metallicity was used. We focused our interest on Fe I 5242.49 Å line that we fit using appropriate analytical formula in [Fe/H] abundance range from -7 to -2.

We started from generating abundance maps. For each abundance map background abundance and number of overabundance spots were drawn, and then their diameter, contrast and position.

Then, using spectrum model and abundance maps, the time series of spectra sampled equidistantly in phase were integrated. For each abundance map rotational velocity from 0 km/s to 200 km/s and inclination from 0° to 180° was drawn.

In total 50000 pairs, abundance map – time series, were prepared. Abundance maps have resolution 128×256. Each time series is composed of 16 spectra of 1024 samples each. During models training 10% of data was used as validation data while the rest was used for training.

Results

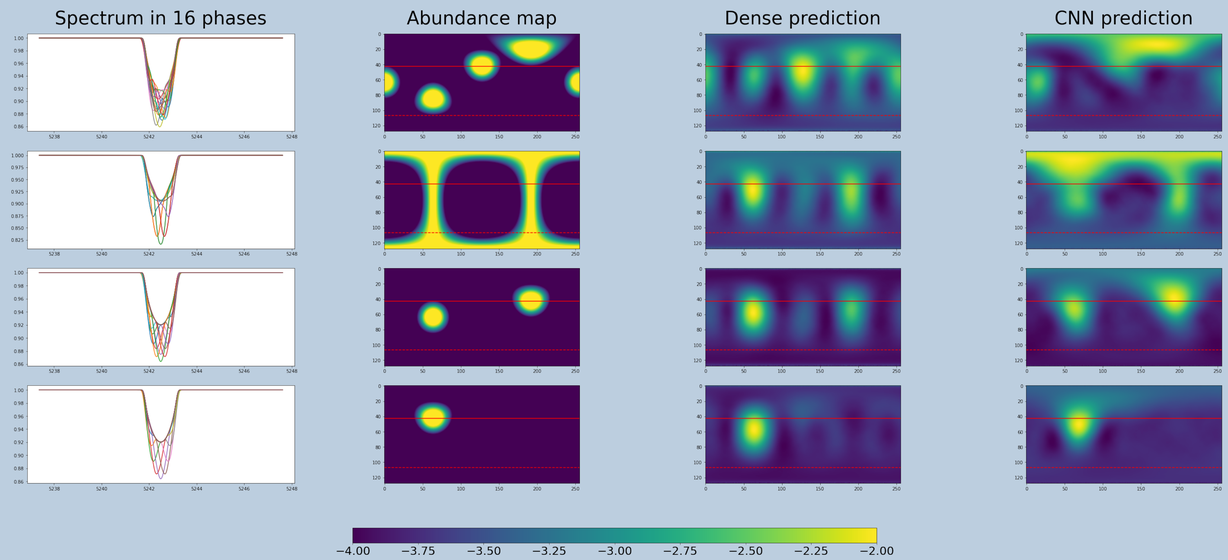

Figure 1: Models' predictions. Each row contains results associated with different abundance map. In columns from left there are: spectra fed into the neural network input, the expected result (abundance map), the prediction from fully connected neural network (dense prediction) and in the last column results from convolutional neural network (CNN prediction). Red lines are lines of equal latitude: solid one for 30° N, and dashed for 60° S. Higher south latitudes are never visible to the observer. For all these experiments the inclination $i = 60°$ and $v \sin i = 40\,\mathrm{km/s}$.

Both architectures performed well for simple geometries with one or two overabundance spots. In more complicated cases, the dense baseline was able to resolve spots with greater accuracy (first row, third column in the Fig. 1), but convolutional neural network generalize better in case of ring-like overabundance pattern (second row, fourth column in the Fig. 1). We believe that convolutional neural networks' predictions can be further improved after introduction of techniques preventing the effects of overtraining, such as dropout or layers normalization.

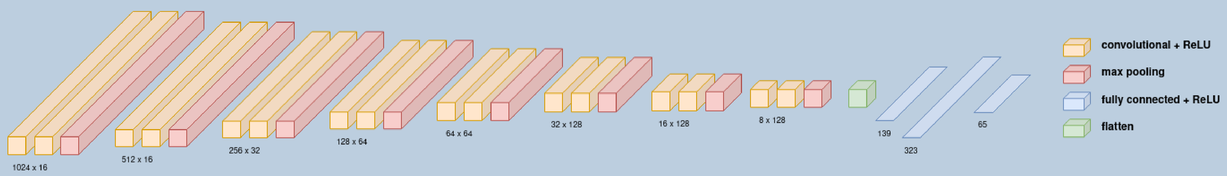

Figure 2: CNN architecture

Conclusions

Deep neural networks can be trained in the task of doppler imaging inversion and give promising results comparable to traditionally used methods. The strengths of DNN's based approach are its speed, simplicity and generalizability. Important directions of further development include taking into account many spectral lines, further adaptation of the neural network architecture, experiments with the regularization of predictions and testing the network with observational data.

Architecture

An extremely important step in the model selection process is to establish a **strong baseline architecture**. A common baseline architecture is a fully connected deep neural network, which is simple in principle, yet can yield satisfactory results.

Even without information about the spatial structure of data and a large number of trainable parameters, which makes this type of network prone to overfitting, it was possible to reconstruct most of the spots in the test data.

The next step was to try a popular architecture for dealing with image data – **a convolutional neural network** (Lecun et al., 1998). We have decided on a hybrid approach – a fully convolutional network with the best found fully connected deep neural network as the prediction head. In such a structure CNN layers are responsible for creation of informative feature maps which are later used by fully connected layers to deliver even better predictions on the validation data.

All experiments presented here were implemented using `TensorFlow` library.

References

O. Kochukhov. Doppler imaging of chemical spots on magnetic Ap/Bp stars. Numerical tests and assessment of systematic errors. *Astronomy & Astrophysics*, 597:A58, Jan. 2017. doi: 10.1051/0004-6361/201629768.

R. L. Kurucz. Atlas: a Computer Program for Calculating Model Stellar Atmospheres. *SAO Special Report*, 309, Feb. 1970.

Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. doi: 10.1109/5.726791.