A DEEP LEARNING APPROACH TO AIRGLOW SPECTRA ANALYSIS

Yuguang Zhang

Abstract: In this project, we use a convolution neural network to denoise the airglow spectra obtained by the British Antarctic Survey.

1. Introduction

Spectroscopy of the hydroxyl (OH) airglow has been used to calculate temperatures in the mesopause region (~87 km) by the British Antarctic Survey (cf. [3]). Clouds cause the measured spectra to be noisy which reduces the robustness of the temperature calculation. This project builds a convolution neural network (CNN) to denoise the spectra.

The data consists wintertime observations with one spectrum every minute. For example, Rothera (BAS research station in the Antarctic) has a clear sky at 22-04-2016, and one raw spectrum is the following:

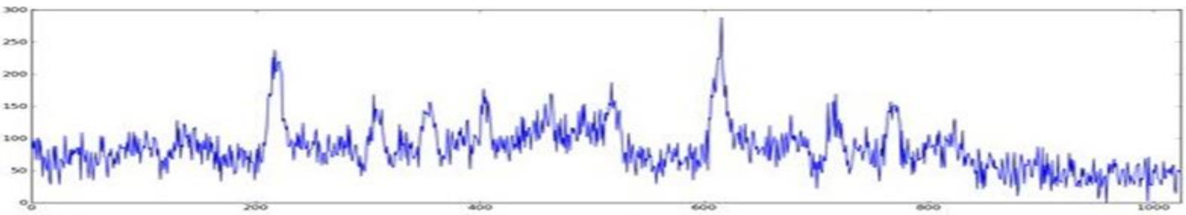

Figure 1 ([3])

In the spectrum smoothed with a 5-point running window, the peaks marked by red arrows are used to calculate the temperature (cf. [3]). Notice that the interested peaks are in the wavelength bands 200-600 and 600-800.

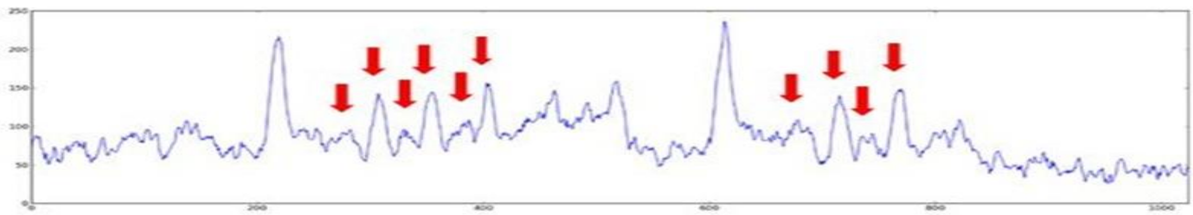

Figure 2 ([3])

In some cloudy nights, running window smoothed spectra still do not show clear peaks ([3]), and a further denoising is required. One way to achieve it is to use the median stacking for several spectra in a row. The signal tends to add up, and the noise tends to cancel out. However, the observed spectra are from the emission of molecules in the atmosphere. Stacking may erase some fine details of short time behaviours, which we would like to preserve as empirical data. Therefore, we use the machine learning techniques, especially the neutral networks, to approach this question.

One difficulty of applying the neutral networks is the lack of the ground-truth. The usual supervised neutral networks need a large amount dataset, including the ground-truth as labels, to train. Then trained networks give the prediction when applying to unseen data. In the current case, clouds form between the molecule emission and the spectrometer in the ground research station. The true spectra without the noise caused by clouds are not accessible.

Fortunately, the deep learning without ground-truth has been actively studied (cf. [1, 2, 4, 5] and references in these papers) due to applications in many disciplines, e.g. astronomy, medical imaging, and engineering etc. The recent work [5], called Deep Image Prior, demonstrates that a CNN with a random input can denoise effectively a noisy image by training only on the given noisy image without any extra training data. Inspired by [5], we build a standard 'decoder-encoder' CNN architecture to denoise the spectra.

2. Result

We use the spectra at 27-07-2019 to demonstrate the performance of our denoising CNN. The following is the raw spectrogram of 27-07-2019.

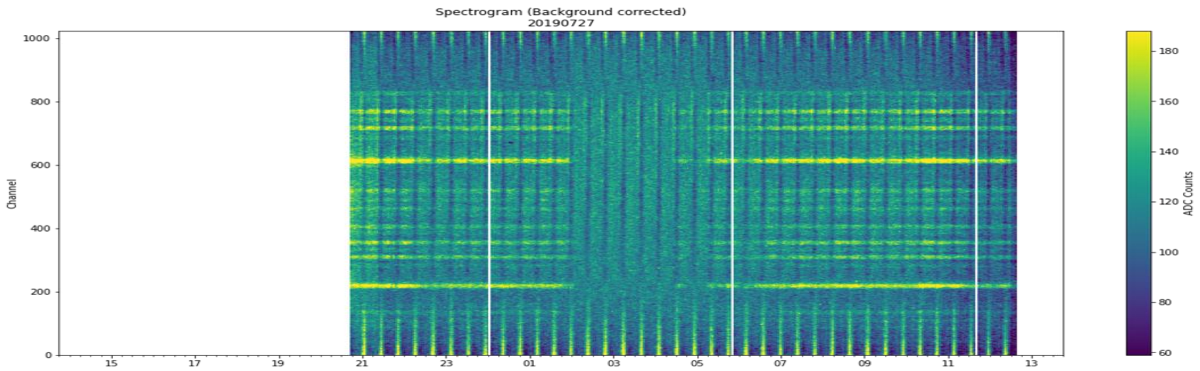

Figure 3

The 30-minutes stack of the raw spectrogram:

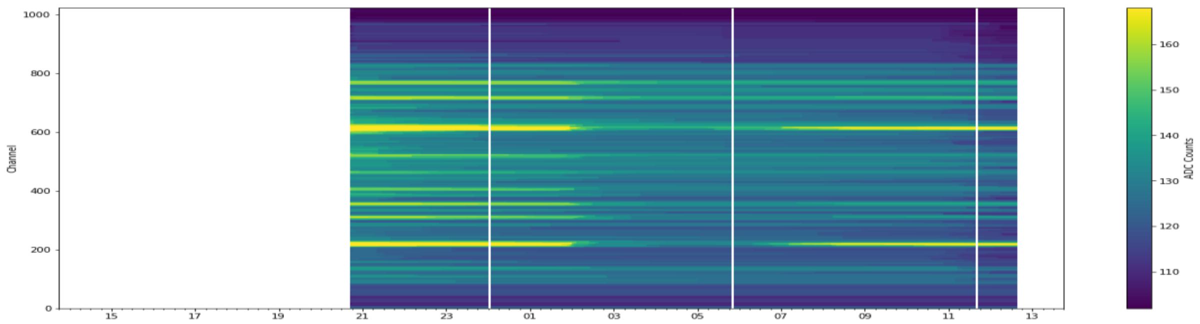

Figure 4

Notice that although the interested peaks appear, many short time features are erased. We use the 30-minutes stack as the input of the CNN instead of a random input as in [5], and wishfully, this deliberately added bias might reduce some running time. The smoothed spectrogram is:

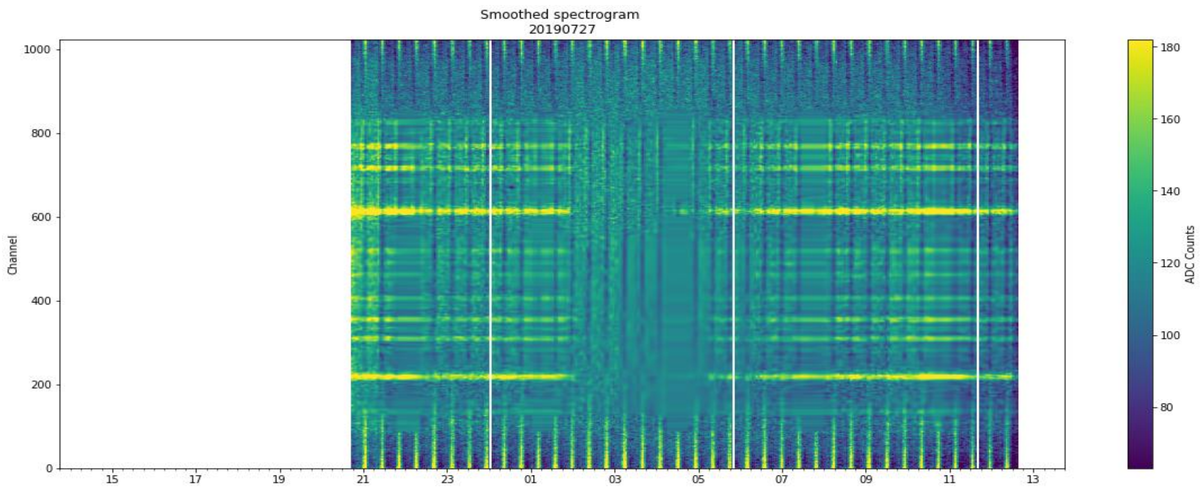

Figure 5

Around 22:00, the sky was clear, and the CNN behaves well. The following is a plot of spectra.

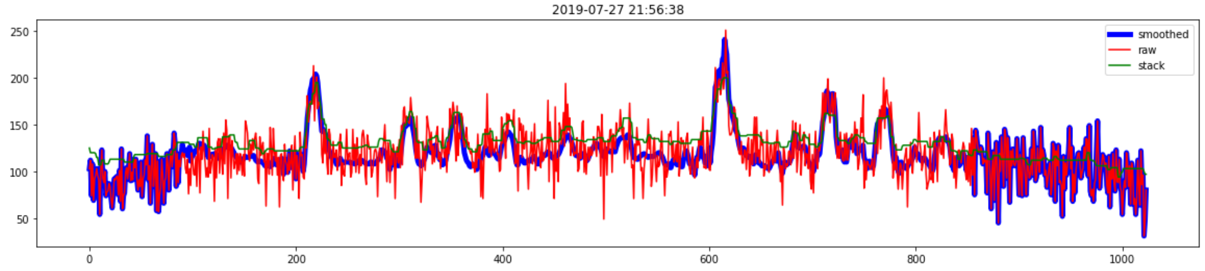

Figure 6

The plot of 20 smoothed spectra from 21:56 to 22:16 is:

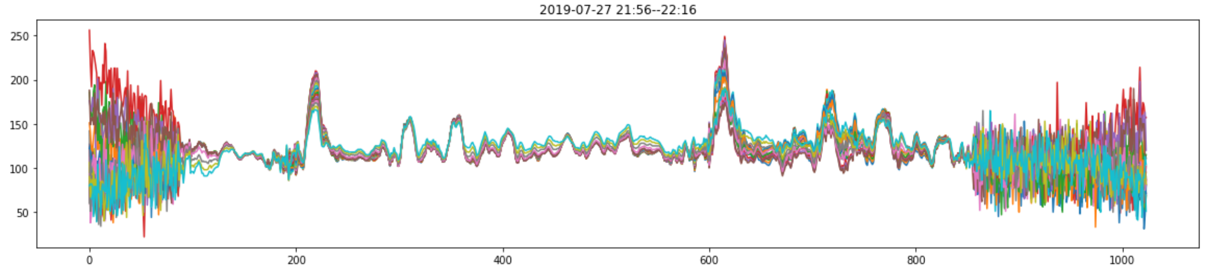

Figure 7

The CNN does not perform ideally when clouds present. The following are two spectra in this case.

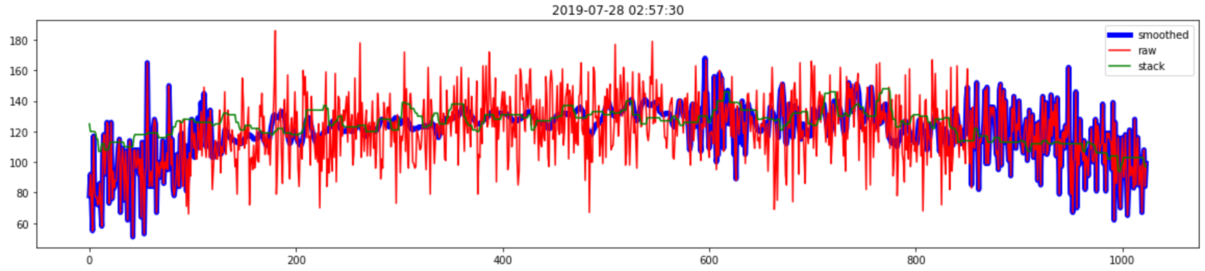

Figure 8

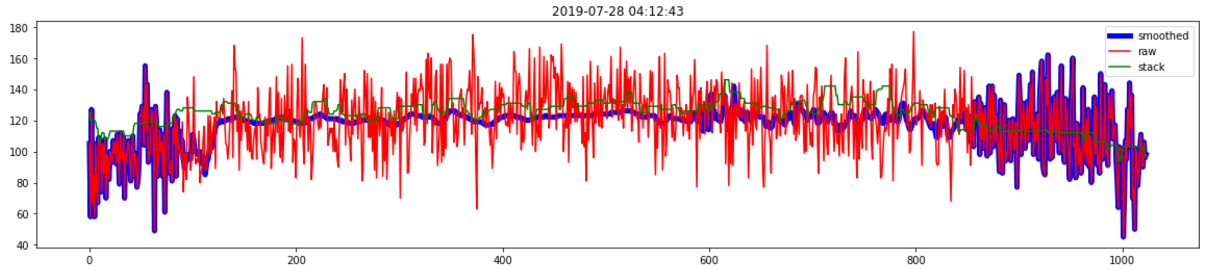

Figure 9

3. CNN architecture

input → Conv2D(128, (3, 3), 'relu', padding='same') → MaxPooling2D((2, 2), padding='same')

→ Conv2D(128, (3, 3), 'relu', padding='same') → MaxPooling2D((2, 2), padding='same')

→ Conv2D(128, (3, 3), 'relu', padding='same') → MaxPooling2D((2, 2), padding='same')

→ Conv2D(128, (3, 3), 'relu', padding='same') → MaxPooling2D((2, 2), padding='same')

→ Conv2D(128, (3, 3), 'relu', padding='same') → UpSampling2D((2, 2), 'bilinear')

→ Conv2D(128, (3, 3), 'relu', padding='same') → UpSampling2D((2, 2), 'bilinear')

→ Conv2D(128, (3, 3), 'relu', padding='same') → UpSampling2D((2, 2), 'bilinear')

→ Conv2D(128, (3, 3), 'relu', padding='same') → UpSampling2D((2, 2), 'bilinear')

→ Conv2D(1, (3, 3), 'relu', padding='same') → output

optimizer='adam', loss='mean_absolute_error', learning_rate=1e-100, epochs=2000, label = original spectra, input = 30-miniutes stack spectra, prediction = output.

4. Conclusions and Discussions

Firstly, the CNN behaves reasonably to denoise random noises, for example in the case of clear sky, and performs less ideally when dealing with the biased noises caused by clouds. There are both underfitted and overfitted parts. Notice that our CNN is the basic sequential 'decoder-encoder' architecture. For a better performance, one may need a more sophisticated CNN as in [5], for example, a U-network CNN architecture etc.

Secondly, this approach is very computational demanding and time consuming. Notice that the CNN does not require any prior training and ground-truth. Every time, it needs to be trained on the unseen noisy data from the beginning, and the algorithm has no memory either. It requires more than 12 hours for denoising one night in a laptop, which is a limitation of this approach.

Thirdly, the CNN fails sometime. A possible reason is that the algorithm optimises a non-convex function, and it sticks at some local minimums. One might overcome it by tuning the learning_rate or simply rerunning the programme. The randomness of choice of initial parameters might help in this case.

Finally, the algorithm is only a mathematical manipulation. The results are neither new empirical data, nor provide a theorical understanding of what are the real spectra hidden in the noisy data. We regard the results as technical cleaned data for the improvement of temperature calculation, but not as the ground-truth, which may require more theoretical and observational studies.

Acknowledgement: The author would like to thank all teachers of Cambridge Spark for an enjoyable bootcamp, which include Elena Chatzimichali, David Illes, Ricardo Monti, Sahan Bulathwela, Ekaterina Kochmar, Yauheniya Shynkevich, Marc Torrellas, Ash D'Cruz, Chris Tegho, Ioannis Gkioulekas, Helen Hu, and Miha Zgubic. He thanks Christina Sevigny for many valuable suggestions for improving his CV. He thanks Dr. Tracy Moffat-Griffin from the British Antarctic Survey for providing an interesting and challenging project, and many suggestions. He thanks his mentor Yauheniya Shynkevich for many discussions.

References:

[1] N. Borodinov, S. Neumayer, S. Kalinin, O. Ovchinnikova, R. Vasudevan, S. Jesse, *Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy*, nature.com/articles/s41524-019-0148-5.

[2] D. Karimi, H. Dou, S. Warfield, A. Gholipour, *Deep learning with noisy labels: exploring techniques and remedies in medical image analysis,* arXiv:1912.02911.

[3] T. Moffat-Griffin, Instruction of the project, and communications.

[4] S. Soltanayev, S. Y. Chun, *Training Deep Learning Based Denoisers Without Ground Truth Data*, arXiv:1803.01314.

[5] D. Ulyanov, A. Vedaldi, V. Lempitsky, *Deep Image Prior*, arXiv:1711.10925.