ONLINE ICRC 2021
THE ASTROPARTICLE PHYSICS CONFERENCE
Berlin | Germany
37th International Cosmic Ray Conference
12–23 July 2021

Classification and Denoising of Cosmic-Ray Radio Signals using Deep Learning

Abdul Rehman, Alan Coleman, Frank Schroeder, Dmitriy Kostunin

1- arehman@udel.edu 2- acoleman@udel.edu 3- fgs@udel.edu 4- dmitriy.kostunin@desy.de

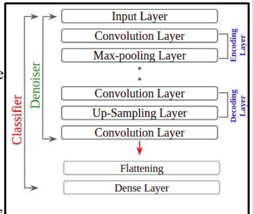
UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

Introduction:

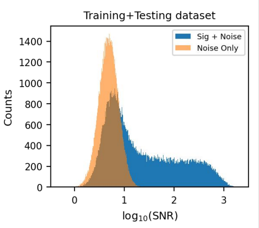
- Radio detection technique is a relatively new and effective technique for the detection of radio signals produced from the Cosmic-Ray (CR) air-showers.
- The continuous, irreducible background of extra-terrestrial origin sets the detection threshold also hinders the ability to reconstruct air-shower properties.
- Convolutional neural networks (CNNs) are used here to create two networks in order to mitigate the effects of background.
- **Classifier:**
  - Used to distinguish the traces with radio signals.
- **Denoiser:**
  - Trained to recover the underlying signals from the noisy traces.

Model Architecture and Data Set:

- The networks are based on an auto-encoder technique.
- Convolutional layer (CL) is paired with max-pooling layer and up-sampling layer to create encoding and decoding layers, respectively.
- An additional CL serves as the final layer for the Denoiser. The Classifier has additional Flattening and dense-layers at the end.
- CoREAS simulations and Cane model are used to produce radio signals and background waveforms, respectively.



Model Architecture



SNR distribution of data set.

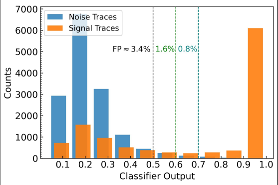
- The signal-to-noise ratio (SNR) used to quantify the signals in the traces is given by:

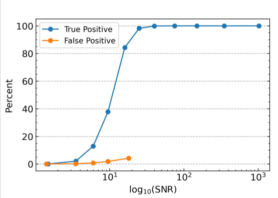
$$\mathrm{SNR} = \left(\frac{\mathrm{Signal}_{\mathrm{Peak}}}{\mathrm{Noise}_{\mathrm{RMS}}}\right)^2$$

- The SNR distribution of data set is shown consisting of 103k and 135k signal and background traces, respectively.
- For validation an additional data set consisting of 11k signal and 15k background traces is produced.

Classifier:

- For training the Classifier, ReLu activation function is used in all the layers except the last layer which uses Sigmoid activation function.
- The output of the trained Classifier is shown in the right plot for the validation data set.
- False positive (FP) and True positive (TP) can be tuned by choosing the threshold on the output.

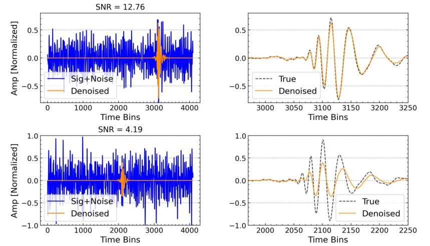




- TP and FP rates (in percent) for threshold of 0.6 are plotted against the SNR values.
- A TP rate of > 80% is achieved for SNR value > 15.
- The FP rate at highest SNR values is below 5%.

Denoiser:

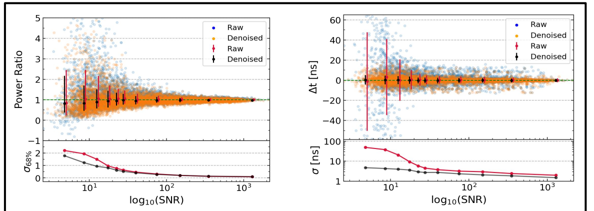
- The classified signal traces from the Classifier are fed to the trained Denoiser to clean them.
- Two result example are shown here.



Example traces from the denoising network.

Accuracy Metrics:

- Power ratio = $\frac{P_{\mathrm{Measured}}}{P_{\mathrm{True}}}$, where P = P(Signal) - P(Noise).
- Time difference, $\Delta t = T_{\mathrm{measured}} - T_{\mathrm{true}}$. where T is the time at which the peak of the Hilbert envelope (= time when the instantaneous amplitude is maximum).
- Both accuracy metrics are computed for denoised traces and compared with raw (not-denoised) traces.



Power ratio and Peak time difference vs SNR

Summary/Outlook:

- CNNs are used in this work in order to distinguish the signal traces from the background and also to remove the background from the noisy traces.
- Time series information is used in this analysis only. Frequency domain information can also be used to improve the results.
- These methods can be applied to real data to better reconstruct the air-shower properties.

Acknowledgements and References:

- This work was supported by **NASA EPSCoR** project, Grant **80NSSC20M01**.
- We also thank the **Tunka-Rex** collaboration for sharing their code for the networks architecture.

References:

[1] F. G. Schröder *Progress in Particle and Nuclear Physics* **93** (2017) 1–68.
[2] D. Shipilov *et al. EPJ Web of Conferences* **216** (2019) 02003.
[3] F. Chollet *et al.*, “Keras.” https://keras.io, 2015.