DELAWARE Data Science DARWIN COMPUTING SYMPOSIUM MAR 24, 2022

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

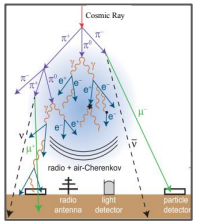
Abdul Rehman, Alan Coleman, Frank Schroeder

1- arehman@udel.edu 2- acoleman@udel.edu 3- fgs@udel.edu

UNIVERSITY OF DELAWARE
BARTOL RESEARCH INSTITUTE

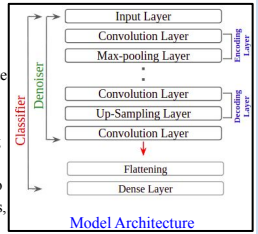
Introduction:

- Cosmic-Ray (CR) produce air showers when they interact with the atmosphere.
- Radio signals emitted from these air showers can be detected using antennas.
- The continuous, irreducible background pose a significant problem for the radio detection.
- We are using Convolutional neural networks (CNNs) to mitigate the effects of background.
- **Classifier:** Used to distinguish the radio signals.
- **Denoiser:** Trained to remove noise from traces.

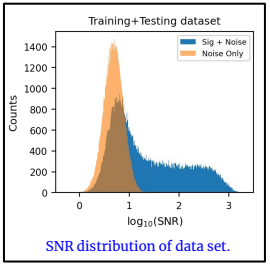


Model Architecture and Data Set:

- 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.

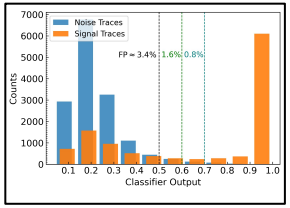
- The signal-to-noise ratio (SNR) used to quantify the signals in the traces is given by:

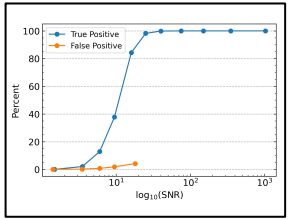
$$\mathrm{SNR} = \left(\frac{\mathrm{Signal}_{\mathrm{Peak}}}{\mathrm{Noise}_{\mathrm{RMS}}}\right)^2$$

- 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.

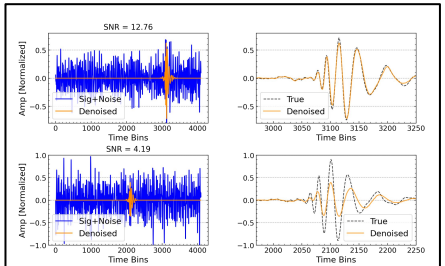




- TP and FP rates (in percent) for threshold of 0.6 are plotted against the SNR values.
- A TP rate of > 80% is achieved for SNR value > 15.
- The FP rate at highest SNR values is below 5%.

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.

DARWIN Allocation:

- We were fortunate to get access of DARWIN during its early testing phase and we utilized it to produce and process the dataset required for this project.
- The software to produce simulations of the air showers is only compatible with the CPUs and can take from few hours up to to a few days to complete one simulation consisting of a few hundred sampling (data) points.
- The training of the CNN network on the other hand are found to run much faster (some time up to 100 times faster) on GPUs as compared to CPUs.
- The results shown here were compiled mostly during the early access period of DARWIN and we continued improving the results during our current research allocations of computing resources on DARWIN.

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.
- The dataset required for this project was produced using the CPUs, whereas the training was done using the GPUs available to us on DARWIN.
- Access to the state-of-the-art computing system, DARWIN, made it possible for us to accomplish our goals for the project and we intended to continue using the available resources to extend our analysis to incorporate real data as well.

Acknowledgements and References:

- This work is supported by **NASA EPSCoR** project, Grant **80NSSC20M01** and U.S. **National Science Foundation-EPSCoR** (RII Track-2 FEC, award #**2019597**)

References:

[1] A. Rehman et al. 37th International Cosmic Ray Conference 2021, **417**.
[2] A. Rehman et al. Workshop on ML for Cosmic-Ray Air Showers, DOI.
[3] D. Kullgren et al. Workshop on ML for Cosmic-Ray Air Showers, DOI.
[4] F. G. Schröder *Progress in Particle and Nuclear Physics* **93** (2017) 1–68.
[5] D. Shipilov *et al. EPJ Web of Conferences* **216** (2019) 02003.
[6] http://docs.hpc.udel.edu/abstract/darwin/darwin.
[7] F. Chollet *et al.*, "Keras." https://keras.io, 2015.