URSI GASS 2023, Sapporo, Japan, 19 – 26 August 2023

ROAD: Radio Observatory Anomaly Detector

Michael Mesarcik*(1), Elena Ranguelova(2), Albert-Jan Boonstra(3),
Marco Iacobelli(3) and Rob V. van Nieuwpoort(1) (2)
(1) Informatics Institute, Science Park 900, 1098 XH, University of Amsterdam, The Netherlands
(2) Netherlands eScience Center, Science Park 140, 1098 XG Amsterdam, The Netherlands
(3) ASTRON, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands

Observational astronomy is as precise as the measurements performed by its instruments. Observatories require data-quality inspection by domain experts to ensure the correctness of observations. However, rapidly growing data-rates from modern instruments make manual data inspection increasingly infeasible. This especially is of concern to distributed, phased array-based radio telescopes, both in terms the amount of generated data and the number of potential points of failure, as these instruments combine thousands of receivers, long-distance networks, large-scale compute hardware, and intricate software [1, 2]. For this reason, we propose a machine learning-based anomaly detection framework using autocorrelation-based spectrographic data from radio telescopes to cope with the increasing data-rates and system complexity.

The Radio Observatory Anomaly Detector or ROAD, is a self-supervised learning framework that models time-frequency representations of data from normal telescope operations. It consists of three components: a ResNet-based backbone, a context prediction classifier and a frequency-band classifier. The model is trained on spectrographic data that has been decomposed into $n \times n$ patches and projected to a latent representation using the backbone network. Then, using the latent representation, the context prediction classifier determines the relative location of the input patches, while the frequency-band classifier predicts which band each patch is sampled from. The gradients are back-propagated through both the classifiers and the backbone, allowing the backbone to learn high-quality representations of the normal telescope behaviour. During inference, we distinguish normal from anomalous behaviour by measuring the distance from the normal representations to a given sample under test [3].

We train and evaluate the performance of our model using an expert-labelled dataset consisting of measurements from the LOFAR telescope (publicly available soon). It comprises of 6500 autocorrelation-based spectrograms, from an observation window between 2019 and 2022, of which 3000 show the presence of anomalies. For purposes of evaluation we identify six different classes of commonly occurring anomalies in the LOFAR telescope as shown in Figure 1. The classes were selected to trace failures in the electronics, data processing systems and environmental effects such as solar and lightning storms.

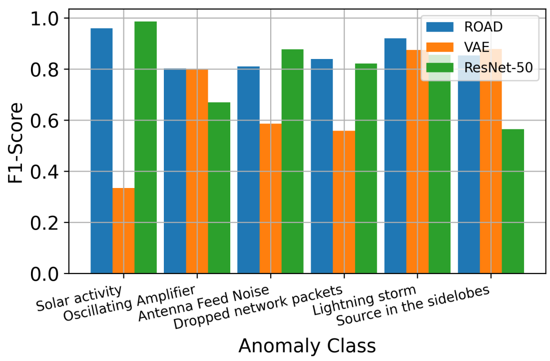

Figure 1. Anomaly detection performance per model

The ResNet-50 is evaluated in a multi-class setting when trained on 50% of the labelled data, whereas the single-class performance of ROAD and the Variational Autoencoder (VAE) are shown. In all cases we evaluate with F1-score using the remaining 50% of the data. The VAE and ROAD, are trained using only normal data whereas the ResNet-50 trained on all classes. We find that our method outperforms its supervised counterpart by $\sim 7\%$ on average, without using any labels. Furthermore we show that ROAD learns better representations than the VAE.

References

[1] M. P. van Haarlem, M. W. Wise, A. W. Gunst, G. Heald, *et al.*, "LOFAR: The LOw-Frequency ARray," *Astronomy & Astrophysics*, vol. 556, p. A2, 8 2013.

[2] P. Dewdney, P. Hall, R. Schilizzi, and T. Lazio, "The Square Kilometre Array," *Proceedings of the IEEE*, vol. 97, pp. 1482–1496, 8 2009.

[3] J. Yi and S. Yoon, "Patch SVDD: Patch-Level SVDD for Anomaly Detection and Segmentation," in *Lecture Notes in Computer Science*, pp. 375–390, 2021.