Section A: Overview of the Research Project Proposal

1. **Title of the research project:** Enhancing Radio Frequency Interference Detection for the MeerKAT Telescope Using Deep Learning and Attention Mechanisms

2. **Broad area of research:** Science

3. **Academic level of research project:** Masters
4. **Abstract:** This project aims to investigate an attention-based deep learning framework for the detection of Radio Frequency Interference (RFI) in MeerKAT data. By integrating attention mechanisms, we aim to localise and highlight precise data regions containing RFI, thereby improving the accuracy of the RFI detection.
5. **Primary supervisors:** Prof. Oleg Smirnov & Dr Marcellin Atemkeng, o.smirnov@ru.ac.za m.atemkeng@ru.ac.za, Rhodes University

Section B: Details of Research Project

1. Scientific merit:

In recent years, we have witnessed the construction of several new radio telescopes designed with specific scientific capabilities. Notable projects include the Square Kilometre Array (SKA) precursors at high and low radio frequencies such as the MeerKAT telescope, the Low-Frequency Array (LoFAR), and the Murchison Widefield Array (MWA). One of the challenges that significantly affects the performance of these instruments is the corruption of their measurements, primarily caused by Radio Frequency Interference (RFI). RFI poses a significant challenge in radio astronomy, primarily due to its diverse sources, including terrestrial transmitters like radar signals and telecommunication systems. This interference typically leads to narrowband flagging during the initial stages of data pre-processing. Additionally, there are wideband RFI sources, such as satellite systems, that require the identification and flagging of substantial portions of the raw data. Removing RFI during data analysis is crucial for upholding the precision of astronomical observations. Machine learning and deep learning methods have shown potential for RFI detection in astronomical data. For example, Mesarcik et al. [1] proposed a supervised segmentation model based on an autoencoder and modified Unet architecture for RFI detection. Mosiane et al. [2] used shallow models such as random forests to detect RFI. The disadvantage of this work is that the importance of RFI features must be studied to determine the most important features before training the random forest model. Generative Adversarial networks and autoencoders are trained for RFI detection in [3], while [4] used a generative machine learning approach to RFI mitigation. All these RFI detection approaches involve training deep neural networks without associating any memory or attention to retain the smeared features caused by several downsampling and pooling layers. Integrating attention-based modules is crucial to mitigate learning bias and preserve useful information for learning intrinsic normal patterns [5]. Using attention-based techniques to learn RFI anomalies, we plan to develop an RFI framework for the MeerKAT telescope, and a multiple-instance learning framework to improve the RFI localization and leverage the attention-based module in both frameworks for comparison and decision-making.

[1] Mesarcik, M., Boonstra, A. J., Ranguelova, E., & van Nieuwpoort, R. V. (2022). Learning to detect radio frequency interference in radio astronomy without seeing it. Monthly Notices of the Royal Astronomical Society, 516(4), 5367-5378.

[2] Mosiane, O., Oozeer, N., & Bassett, B. A. (2016, October). Radio frequency interference detection using machine learning. In 2016 IEEE Radio and Antenna Days of the Indian Ocean (RADIO) (pp. 1-2). IEEE.

[3] Vinsen, K., Foster, S., & Dodson, R. (2019, March). Using machine learning for the detection of radio frequency interference. In 2019 URSI Asia-Pacific Radio Science Conference (AP-RASC) (pp.1-4). IEEE.

[4] Vos, E. E., Luus, P. F., Finlay, C. J., & Bassett, B. A. (2019, October). A generative machine
learning approach to RFI mitigation for radio astronomy. In 2019 IEEE 29th International Workshop on
Machine Learning for Signal Processing (MLSP) (pp. 1-6). IEEE.

[5] Li, Q., Yang, R., Xiao, F., Bhanu, B., & Zhang, F. (2022). Attention-based anomaly detection in multi-view surveillance videos. Knowledge-Based Systems, 252, 109348. Deep learning assisted data inspection for radio astronomy

2. Feasibility: This project will use RFI simulators for the MeerKAT telescope to generate datasets with RFI for both training and testing. Transferability will be expanded by employing real MeerKAT data to detect RFI. This approach ensures a comprehensive evaluation of the RFI detection system across simulated and real data, enhancing its robustness and reliability. Copious amounts of MeerKAT data are available for testing and are already hosted at Rhodes.

Timelines:

Year 1:

Months 1-3:

- Review existing literature on RFI detection methods in radio astronomy.
- Familiarize with MeerKAT data and Measurement sets
- Familiarize with RFI simulation tools

Months 4-6:

- Acquire simulated RFI data for initial training and testing.
- Develop preprocessing pipelines to clean and format the data for training.

Months 7-9:

- Design and implement deep learning models for RFI detection, including attention-based modules.
- Experiment with various architectures, such as autoencoders, generative adversarial networks, and multiple-instance learning frameworks.
- Train and validate the models using simulated RFI data.
- Fine-tune hyperparameters and optimize model performance.

Months 10-12:

- Test the trained models on unseen simulated data to evaluate their generalization capability.
- Analyze the performance metrics, such as accuracy, precision, recall, and F1 score.
- Identify any limitations or shortcomings of the models and propose potential improvements.

Year 2:

Months 13-15:

- Adapt preprocessing pipelines developed in Year 1 to handle real MeerKAT data.
- Incorporate attention-based modules to enhance model performance and interpretability.

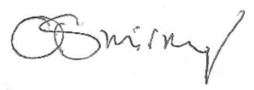
Months 16-21:
- Refine deep learning models based on insights gained from real data experiments.
- Fine-tune parameters and optimize
- Conduct extensive testing and validation

Months 22-24:
- Prepare and submit MSc thesis

3. Link to SARAO research priority areas for 2023: This project will focus on tools that potentially improve the quality of any and all MeerKAT imaging observations.

4. Qualifications, academic abilities, skills and/or experience that a student should have in order to successfully deliver on the objectives of the research proposed: Familiarity with radio interferometry and/or machine learning would be advantageous but not required. A solid understanding of applied mathematics and good software skills are required.

Supervisors



Oleg Smirnov 23 February 2024

Marcellin Atemkeng 23 February 2024