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

A comparison framework for RFI detection algorithms

Charl D. Du Toit*[(1)], Trienko L. Grobler[(1)], and Danie J. Ludick[(1)]
(1) Stellenbosch University, Stellenbosch, South Africa; e-mail: 21708886@sun.ac.za; tlgrobler@sun.ac.za, dludick@sun.ac.za

1 Introduction

In radio astronomy, radio frequency interference (RFI) signals are unwanted electromagnetic emissions that degrade the performance of a receiver. Before celestial signals can be imaged, the RFI must be flagged in the time-frequency plane and discarded. Many observatories make use of post-correlation statistical based cumulative thresholding techniques to flag data, such as AOFlagger, which is used by the Low Frequency Array (LOFAR). Recently, machine learning (ML) approaches, specifically fully convolutional networks (FCN) has shown to be a viable alternative to traditional flaggers. This study compares such FCN models in terms of complexity and performance.

2 Datasets

The models are fully supervised, which means it requires ground truth flags for training. Reliable ground truth flags are difficult to obtain because the nature, origin and intensity of RFI is often unknown. The models are trained on two datasets. The first set is Hydrogen Epoch of Reionization Array (HERA) data which was generated using a simulator called hera-sim [1]. Simulated data allows us to obtain proper ground truth data. The second dataset is LOFAR data, for which the training masks are generated with AOFlagger. However, the FCNs are tested with expert hand-labelled flags, since AOFlagger tends to overflag.

3 Models and metrics

Seven models are compared. Five models from literature, UNET, AC-UNET, RFI-NET, RNET, DSC-DUAL-RESUNET and two adaptations by the author ASPP-UNET and DSC-MONO-RESUNET. The models take time-frequency images as input and produce images of the same size as output, where the value of a pixel indicates the confidence of it being RFI. The binary classification metrics are: area under receiver operating characteristic (AUROC), area under precision recall curve (AUPRC) and F1-score. The floating point operations (FLOPs) and number of trainable parameters indicate the complexity of the model. The hyper-parameter space is also explored, such as the loss function, number of filters, dropout rate, and the learning rate.

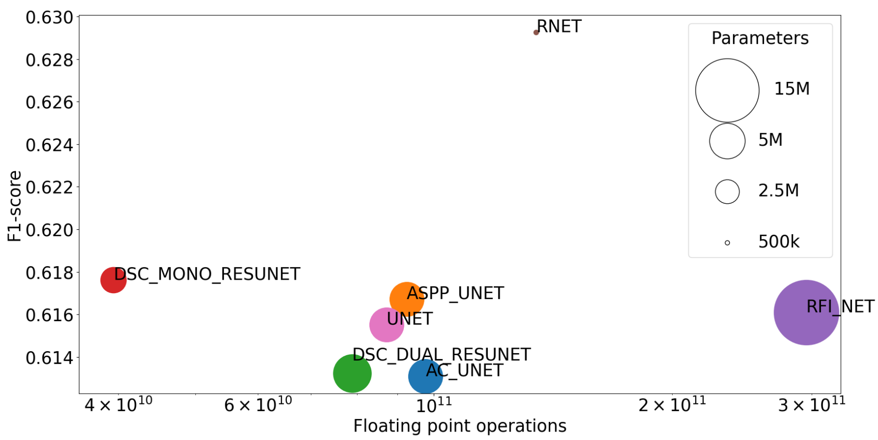

Figure 1. LOFAR test results for binary cross entropy loss, 64 filters and learning rate of 1e-5.

[1]https://hera-sim.readthedocs.io/en/latest/