Software framework for Pulsar Detection using Machine Learning and Digital Twin

Tanumoy Saha, PhD HTW Berlin, *Email: sahat@htw-berlin.de*
Co-contributors: Marcel Trattner (HTW Berlin), Prof. Dr. Hermann Heßling (DZA Görlitz)

Introduction

Identifying pulsar signals from radio telescope data archives poses a major big data challenge. Although several efficient algorithms have been developed to tackle this problem, our software package introduces an innovative approach: a machine learning–based framework that employs training data generated through Digital Twins derived from theoretical physics models, combined with a U-Net–based neural network for pulsar signal segmentation. This framework aims to fascilitate the verification and discovery of unknown pulsars and other astronomical signals grounded by physical theory, thereby advancing the capabilities of astronomers and physicists in their exploration of the universe. This poster shows the status of the framework.

Digital Twin

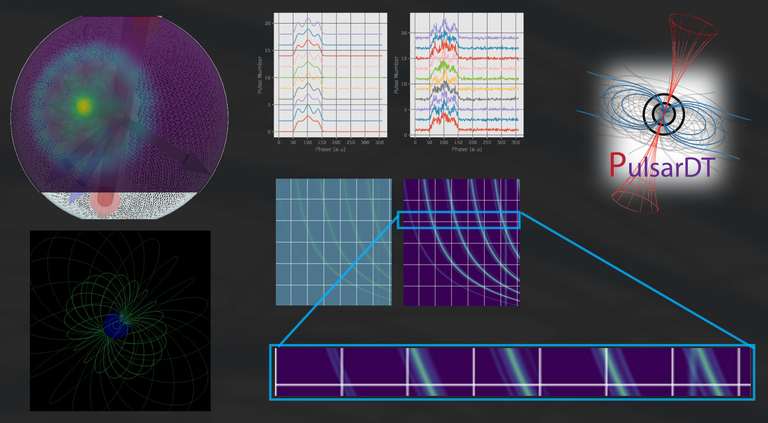


Fig 2.1: Spark pattern based Pulsar Digital Twin

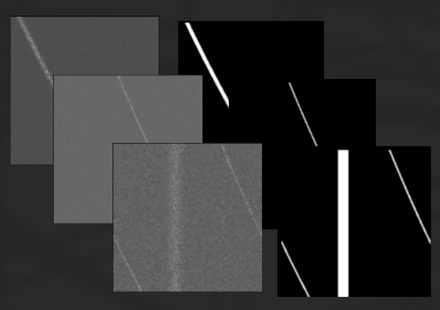


Fig 2.2: Synthetic dispersion image-mask pair for training

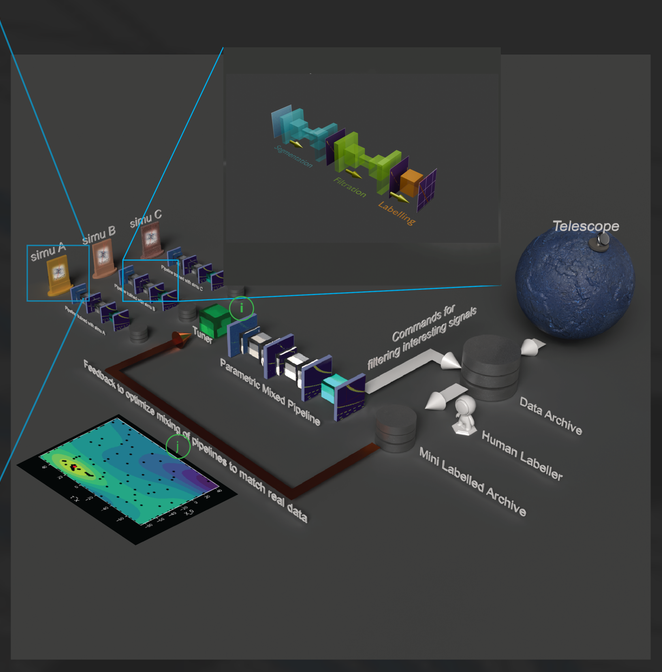


Fig 1.1: Overview of the framework

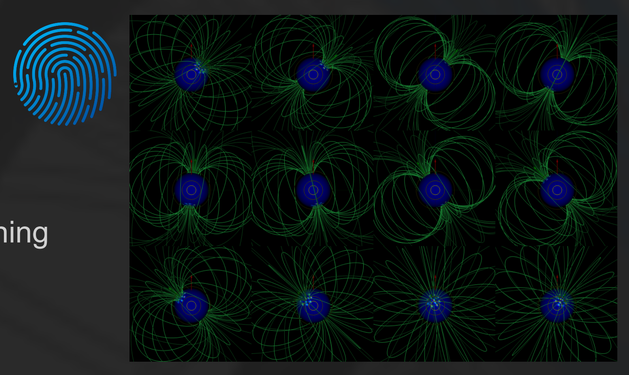


Fig 1.2: Rendering of the PulsarDT virtual oscillator

Pulse Segmentation

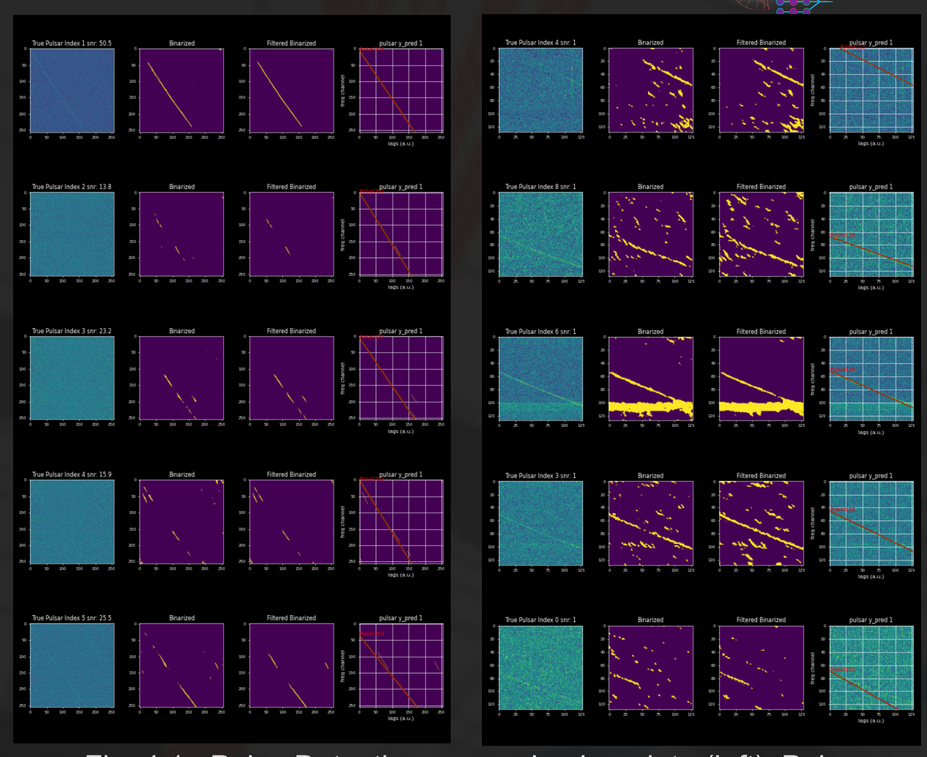
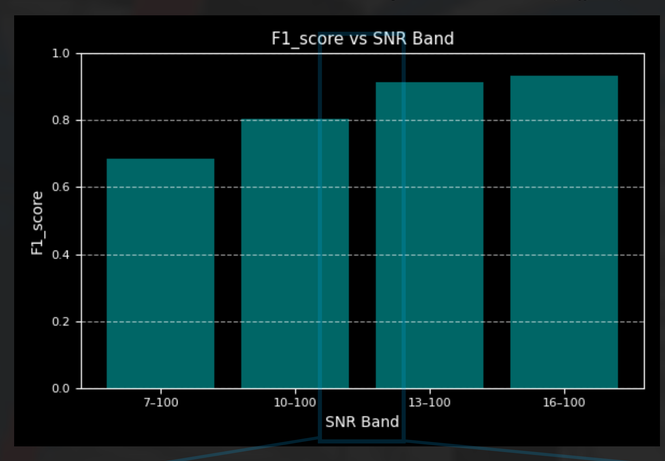
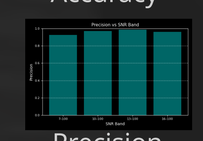


Fig 4.1: Pulse Detections on real pulsar data (left). Pulse Detections on fake real pulsar data (right)





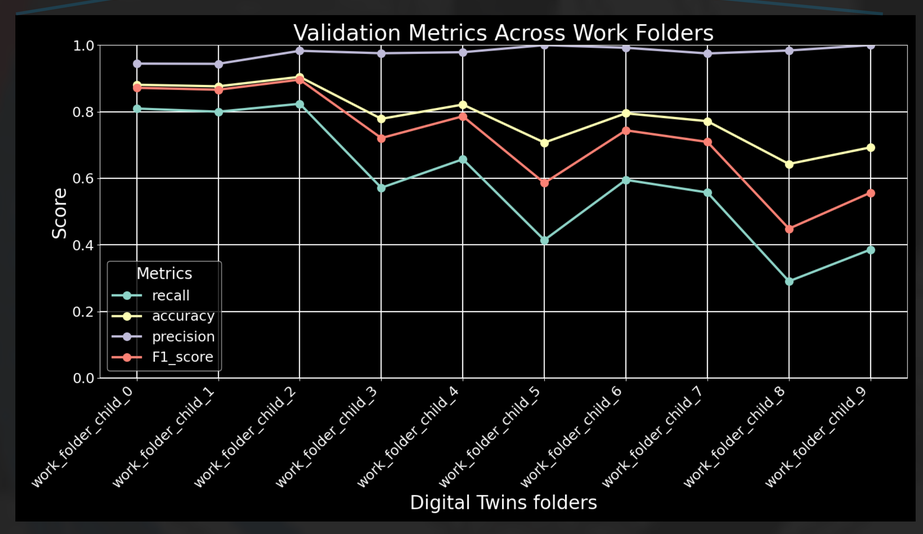


Fig 4.2: Performance scores of the pipeline trained with synthetic data from work_folder_child_0 for different SNR ranges on real pulsar data (top). Performance metrics for SNR band 12-100 for all pipelines trained using corresponding digital twins (bottom) with wide DM range of 30 pc/cm^3 and with 5000 samples each.

Pipeline Steps

Step 1: Segmentation by CNN (Simplified UNet)
Step2: Filtration by CNN (optional)
Step3: Labelling using Connected Component Analysis

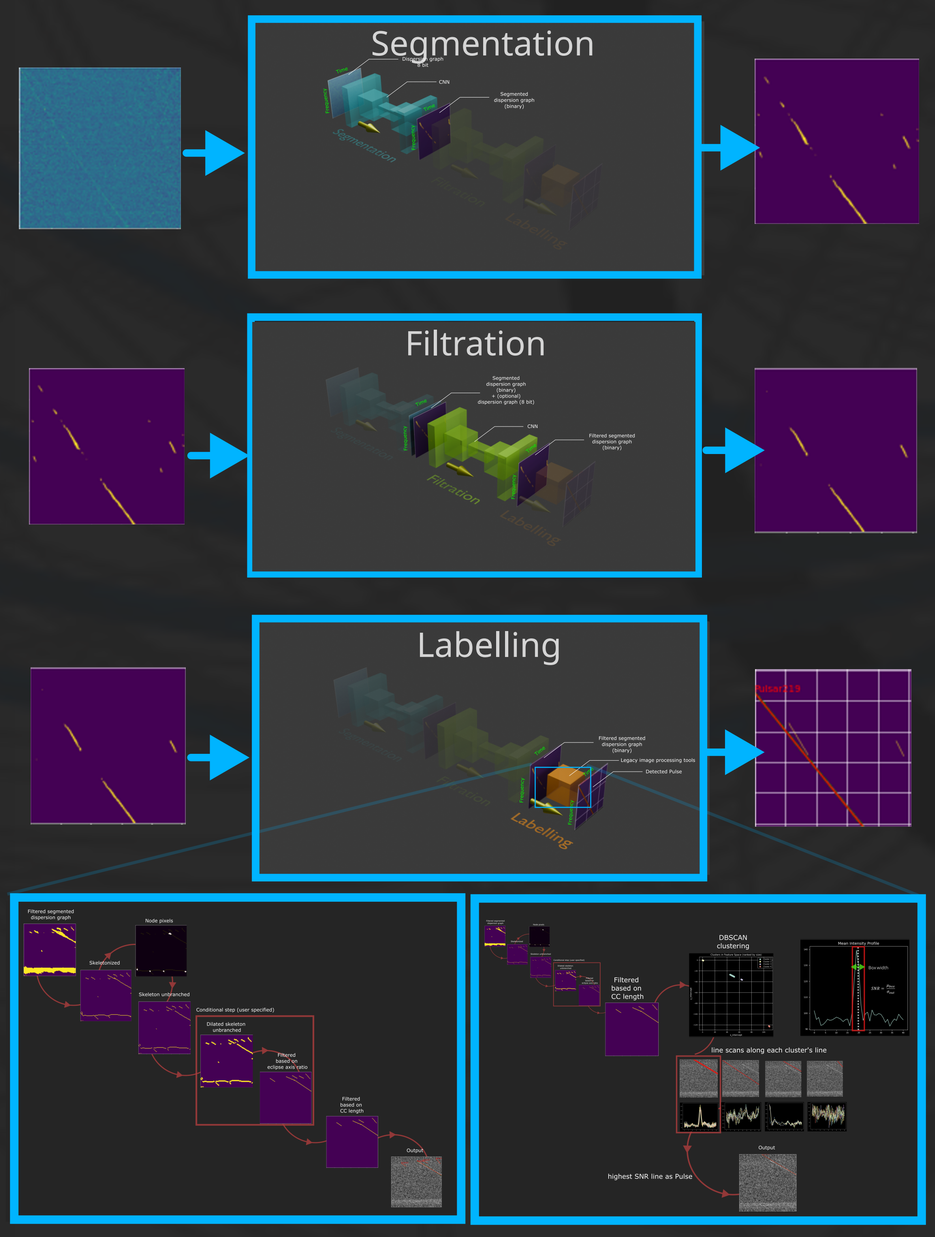
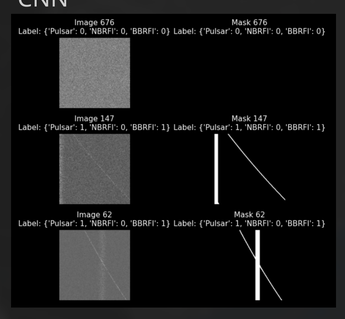


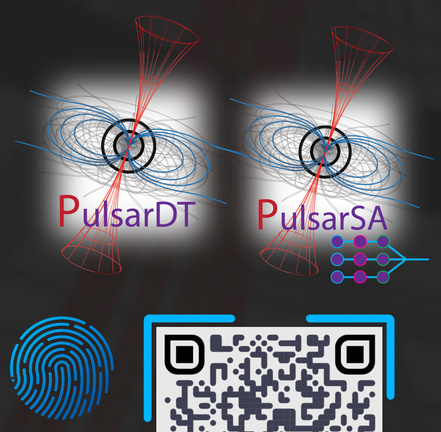
Fig 3: Pipeline steps used to segment and detect pulse

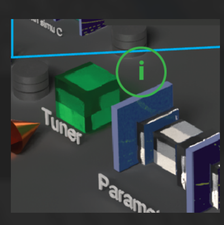
Training set for segmentation CNN



Training set for filter CNN

Software Package





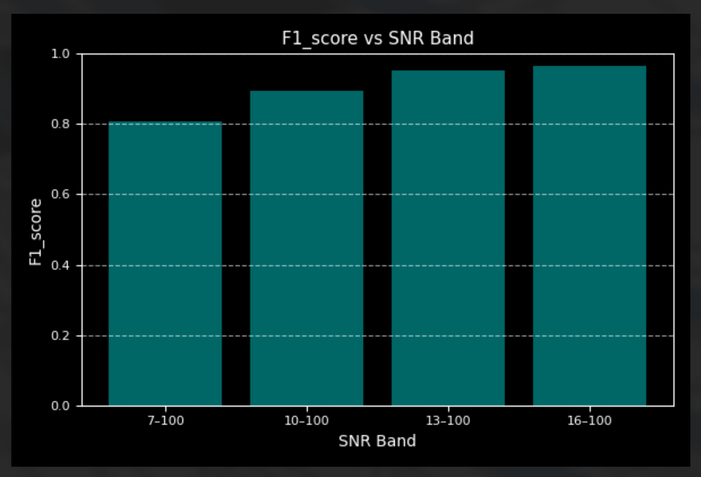


Fig 4.3: Performance after mixing the pipelines using the Tuner module shows improvement in performance (Beta)

Acknowledgement

Andrei Kazantsev (MPIfRBonn), Lars Haupt (DZA Görlitz), Yurii Pidopryhora (MPIfRBonn), Marc Drobek (DZA Görlitz), Sagar Borra (DZA Görlitz), Elsa Buchholz (DZA Görlitz)