

Status of results in TA5

PUNCH TA5 Meeting

17/07/2025

M. Kramer A. Redelbach

Deliverables (so far)

Starting point: Finalization of deliverables and recent reporting requests

- D-TA5-WP1-1: Report on the impact of on-line filtering on the discovery potential
- D-TA5-WP2-1: Curation and metadata schemes for dynamic filtering
- D-TA5-WP2-2: Strategy concept for identifying highly complex (multi-parametric) signals in huge data streams
- D-TA5-WP2-3: Test environment for identifying highly complex (multi-parametric) signals in huge data streams
- D-TA5-WP2-4 Generic tools to both convert trained neural networks into efficient HLS/VHDL FPGA firmware optimised for a real-time, low-latency environment and to establish comparable software solutions
- D-TA5-WP3-1: Specifying the concept of a dynamic archive

- Collection in Gitlab:

https://gitlab-p4n.aip.de/punch/intra-docs-content/-/tree/master/files/TA5/Documents_deliverables

Documents (Reports, Zenodo)

- Deliverable reports:
 - D-TA5-WP1-1
 - D-TA5-WP2-1
 - D-TA5-WP2-2
 - D-TA5-WP3-1
 - Available: <https://results.punch4nfdi.de/?md=/docs/Documents/deliverable-reports.md>
- Zenodo:
 - D-TA5-WP1-1 <https://zenodo.org/records/11658437>
 - D-TA5-WP2-1 <https://zenodo.org/records/10692169>
 - D-TA5-WP2-3: Planned, not yet uploaded
 - Assignment to PUNCH4NFDI community in Zenodo

ML on FPGAs – document with results and recommendations

TA5 WP2-4 report

Generic Tools for Artificial Neural Network Implementation on Field Programmable Gate Arrays

Christian Kahra⁴, Ramesh Karuppusamy¹, Andrei Kazantzev¹, Yunpeng Men¹, Andreas Redelbach³, Christian Schmitt⁴, Arno Straessner², and Johann C. Voigt²

¹Max-Planck-Institut für Radioastronomie Bonn

²Technische Universität Dresden

³Frankfurt Institute for Advanced Studies, Universität Frankfurt

⁴Johannes Gutenberg-Universität Mainz

June 23, 2025

Contents

| | | |
|----------|-----------------------------------------------------------------------------------------------------------------------------|-----------|
| 1 | Introduction | 2 |
| 1.1 | Relation to other work in TA5 | 3 |
| 2 | Evaluation of hls4ml for real-time classification of astronomical radio signals | 3 |
| 2.1 | Background | 3 |
| 2.2 | FPGA Solutions | 4 |
| 2.2.1 | ML model | 4 |
| 2.2.2 | Hls4ml framework | 4 |
| 2.2.3 | FPGA Implementation | 5 |
| 3 | VHDL implementation of convolutional neural networks for real-time processing of ATLAS Liquid-Argon Calorimeter data | 6 |
| 3.1 | Overview of LAr calorimeter off-detector upgrade | 6 |
| 3.2 | Requirements and their influence on network training and firmware design | 6 |
| 3.3 | Firmware implementation | 7 |
| 4 | Evaluation of AI hardware engines with AMD Versal AI | 9 |
| 5 | Recommendations for users and developers | 11 |
| 5.1 | General recommendations | 11 |
| 5.2 | Further related tools and libraries | 12 |

Evaluated different tools to implement neural network inference on FPGA for different applications: hls4ml, direct VHDL implementation, specialized hardware (Versal AI engines)

- No feedback/extra comments received
- To be sent to EB/MB tomorrow
- Presentation in FairMAT meeting

Open datasets

Robust realtime identification of dispersed radio astronomical signals that last much less than a second is challenging. Here we explore the utility of machine learning techniques to identify such signals and use data taken on the **Crab pulsar using the Effelsberg 100m Radio Telescope**.

Data corresponds to the frequency range of 1240-1510 MHz, and contains **20 minutes of the pulsar signal**. In addition, the DM-time data generated by the realtime pipeline, the associated Tensorflow CNN model is included and the training dataset are included. The data are intended for machine learning tasks focused on single-pulse detection and classification.

Nice example for a fast ML use case

The screenshot shows the Edmond dataset interface. At the top, the Edmond logo and navigation links (Search, Help, Support, Log) are visible. The dataset title is "Realtime identification of Dispersed Radio signals using ML - A Case Study on the Crab Pulsar". Below the title, the version is "Version 1.0". A citation is provided: "Kazantsev, Andrei; Karuppusamy, Ramesh, 2025, 'Realtime identification of Dispersed Radio signals using ML - A Case Study on the Crab Pulsar', <https://doi.org/10.17617/3.HQYC80>, Edmond, V1". There are buttons for "Cite Dataset" and "Learn about Data Citation Standards". On the right, there is a button for "Access Dataset" and a "Contact Owner" button. Below this, "Dataset Metrics" shows "2 Downloads". The "Description" section states: "Robust realtime identification of dispersed radio astronomical signals that last much less than a second is challenging. Here we explore the utility of machine learning techniques to identify such signals and use data taken on the Crab pulsar using the Effelsberg 100m Radio Telescope. The data corresponds to the frequency range of 1240-1510 MHz, and contains 20 minutes of the pulsar signal. In addition, the DM-time data generated by the realtime pipeline, the associated Tensorflow CNN model is included and the training dataset are included. The data are intended for machine learning tasks focused on single-pulse detection and classification." The "Subject" is "Physics". The "License/Data Use Agreement" is "CC0 1.0". There are tabs for "Files", "Metadata", "Terms", and "Versions". A search bar is present. Below, there are filters for "File Type", "Access", and "File Tag". A "Download" button is visible. The file list shows two files: "AAA_README.txt" (Plain Text - 2.0 KB, Published May 5, 2025, 0 Downloads, MD5: c9e...26c) and "B0531+21_59000_48386.fil" (Unknown - 3.1 GB, Published May 5, 2025, 1 Download, MD5: 3cf...bf5). The second file has a detailed description: "This file contains radio observations of the Crab pulsar (PSR B0531+21) in SIGPROC's filterbank format, recorded with the Effelsberg radio telescope. The data represents rapidly sampled radio spectra spanning the 1210-1530 MHz frequency range. The file contains a minimal metadata encoded in the first 349 bytes, following which the spectra (256 bins) is recorded as 8-bit unsigned integers. For more information, contact the authors." There are buttons for "Data" and "filterbank" below the file description.

ML-PPA

Many developments:

New modules SA, Tutorial
Project

Demonstration of a working
setup for ML-PPA on C4P

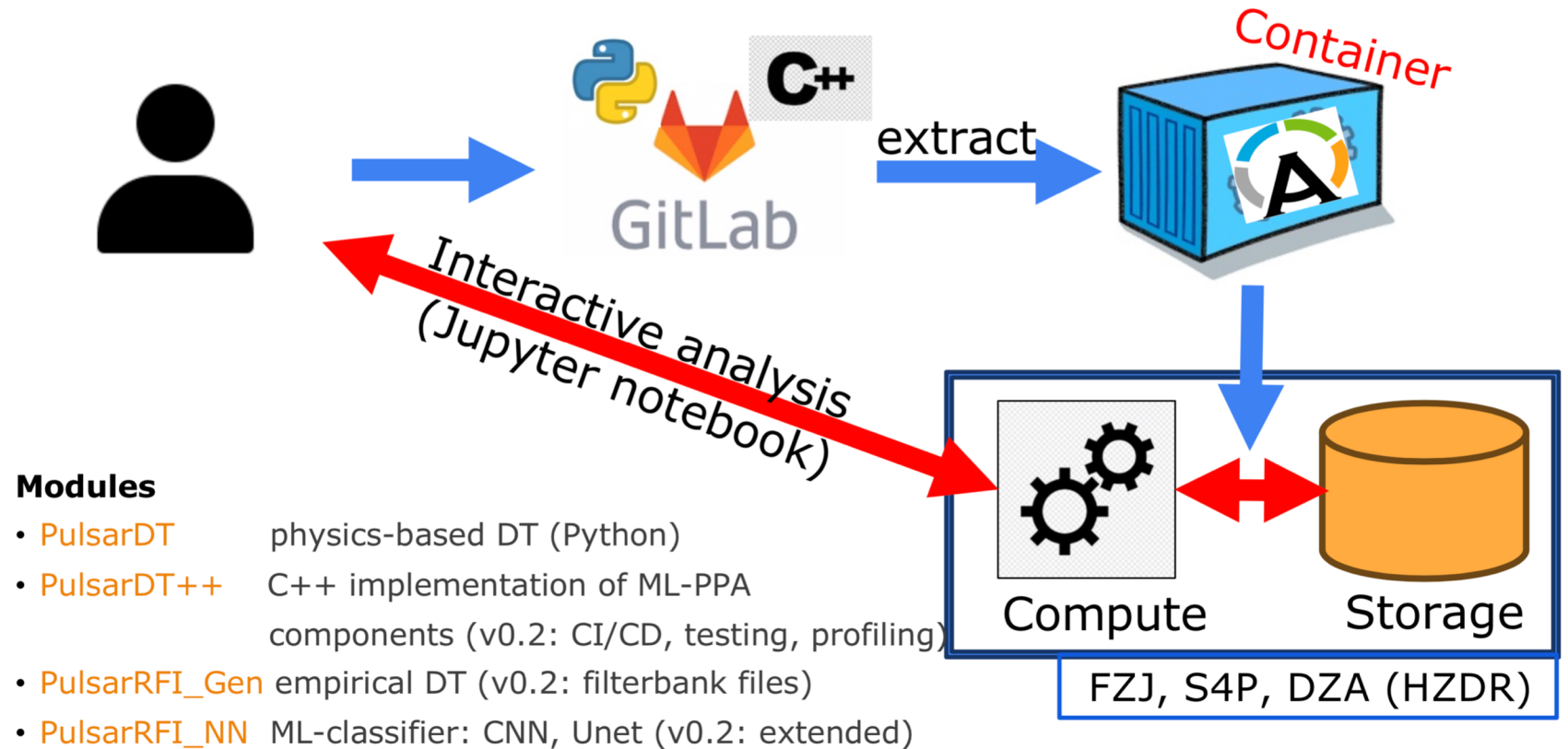
Availability of simulated
datasets?

Connection to TA5-WP4
deliverables

Porting common off-line
packages (e.g. CASA) to a
memory-based computing
prototype to prepare
analysis of “data monster”

ML-PPA v0.2: workflow

(<https://gitlab.dzastro.de/punch/ml-ppa>)



Visibility and availability of results

- Useful for further developments, collaborations, PUNCH 2.0
- PUNCHLunch presentations:
 - Effelsberg prototype
 - Dynamic archiving
 - ML-PPA

For discussion