Real-time detection of Solar and Jovian radio bursts with NenuFAR: advancing astrophysical data mining with the EXTRACT project.

A distributed data-mining software platform for extreme data across the compute continuum

An overview of TASKA – A use cases

Emilie Mauduit, Baptiste Cecconi, Cedric Viou, Louis Bondonneau, Julien Girard, Fadi Nammour

The EXTRACT Project has received funding from the European Union's Horizon Europe programme under grant agreement number 101093110

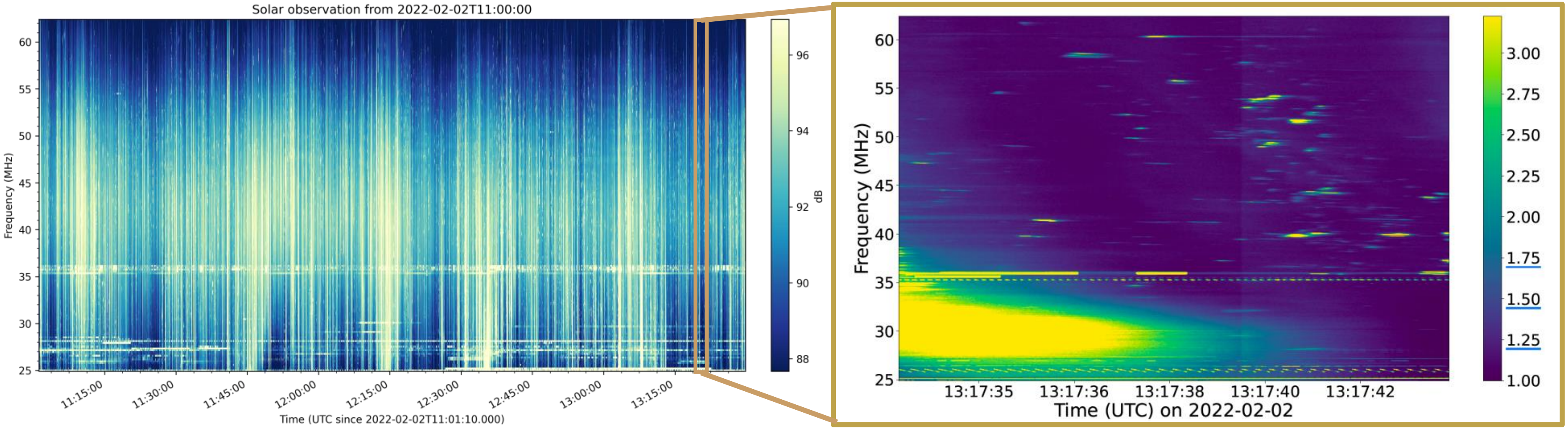
Transient Astrophysics with a Square Kilometer array (TASKA)

- Increasing need for ultra-high-resolution radio observations with high sensitivity → surge in data volumes from next-generation radiotelescopes such as NenuFAR or SKA.
- ⟹ Need for efficient tools for data management, processing, and storage optimization
- **TASKA-A** use case, takes advantage of the technologies developed within **EXTRACT** to handle massive data streams (~100GB/hr in beamformed mode) produced by NenuFAR
- Two projects focusing on real-time detection of radio emissions :
  - **A1** : dedicated to Solar radio spikes (C.Viou),
  - **A2** : dedicated to Jovian millisecond bursts (E.Mauduit)

TASKA – A1 : Solar spikes

Solar radio spikes

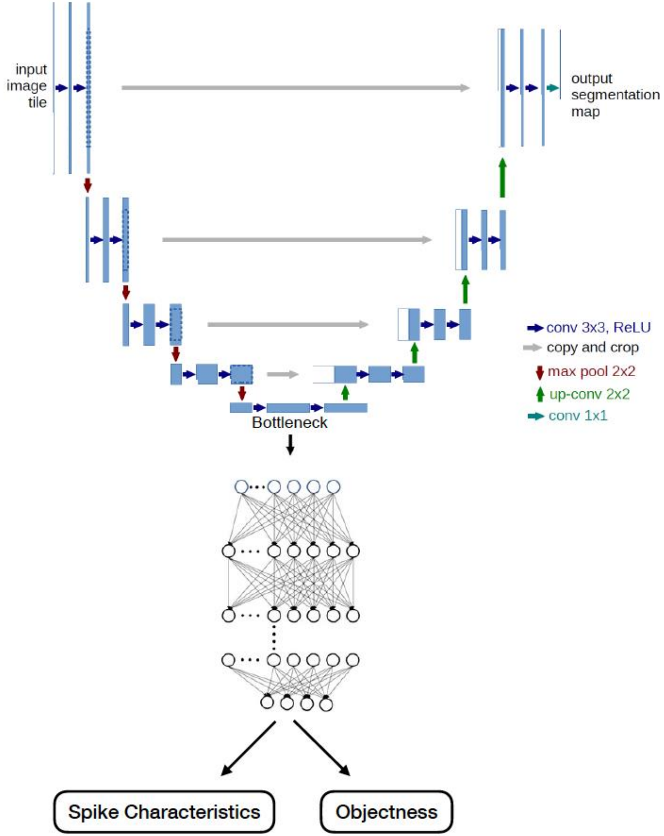
- Dedicated to automatically detect solar radio spikes using SpikeNet

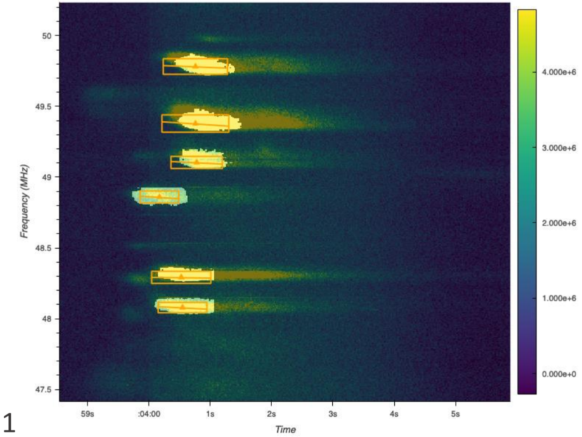


[Murphy et al, OJA, 2024]

Detection method : SpikeNet

- ML model trained on ~100 000 samples of solar spikes (64 by 64 pixels)
- Produced segmentation masks for radio spikes in the validation set and predicted the spike characteristics, i.e. location in time and frequency, duration, spectral width and drift rate
- Bursts that are not fully in one tile can not be detected



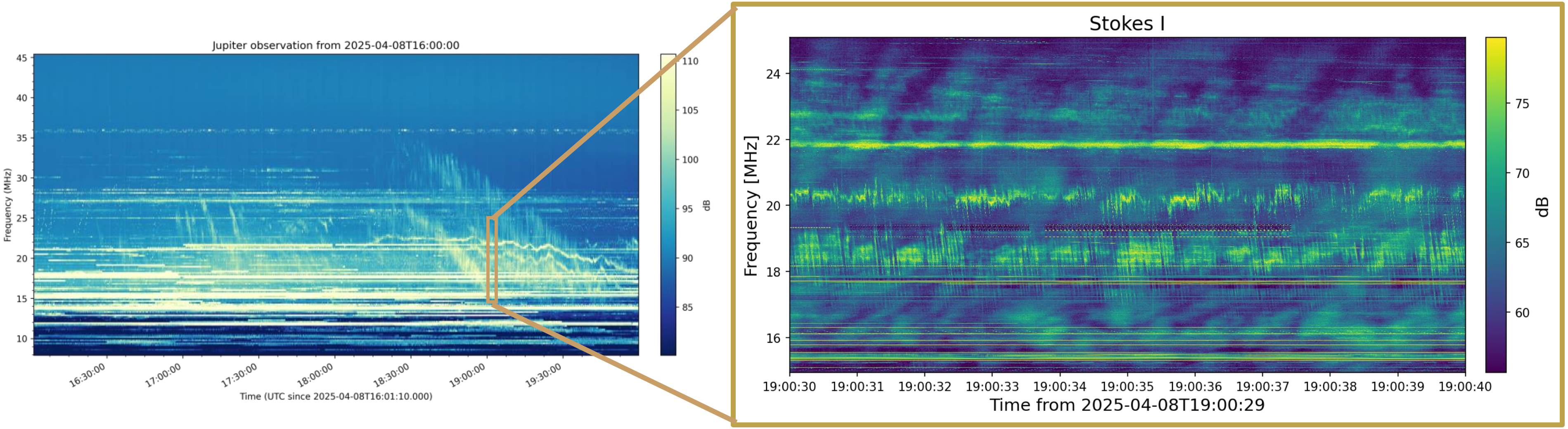


[Murphy et al, OJA, 2024]

TASKA – A2 : Jovian S-bursts

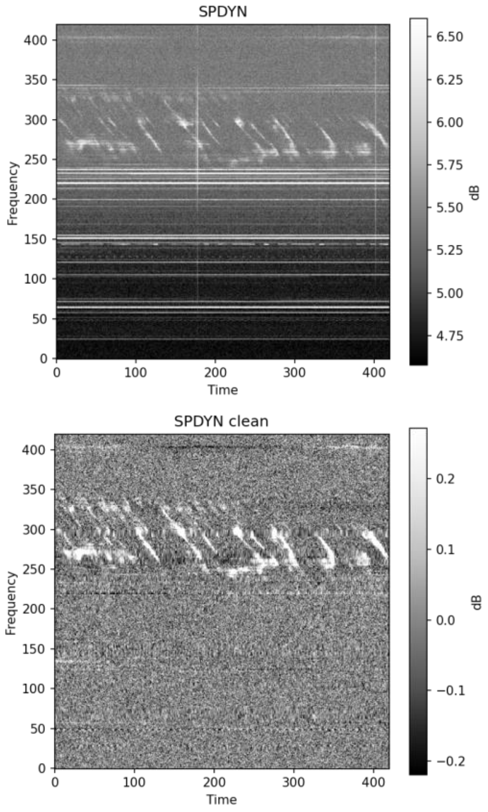
The TASKA – A2 Use case

- Developing an algorithm for real-time detection of fine drifting structures during radio observations with NenuFAR
- Use a method originally developed for Jupiter observations with NDA-JunoN data *[Mauduit et al, 2023, Nat Coms]*

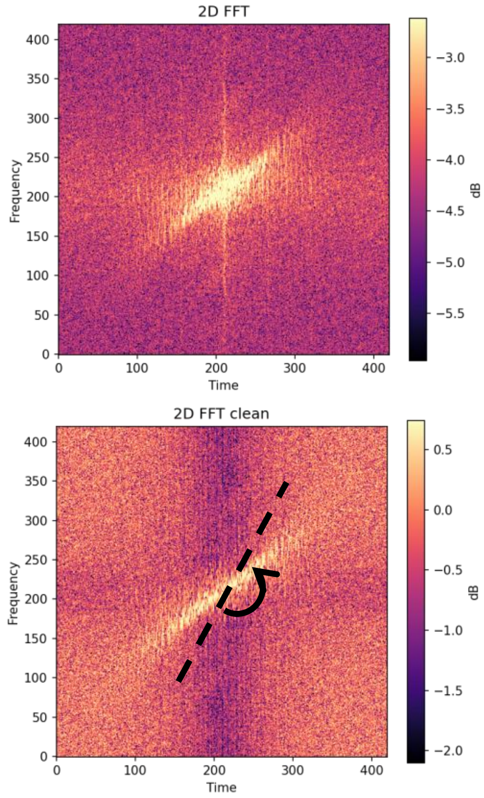


Detection algorithm

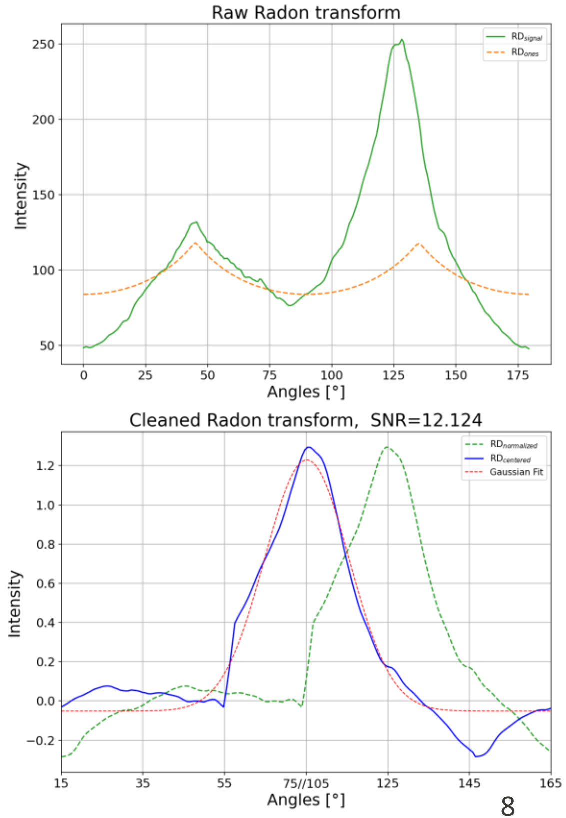
1 - RFI mitigation



2 – FFT2D

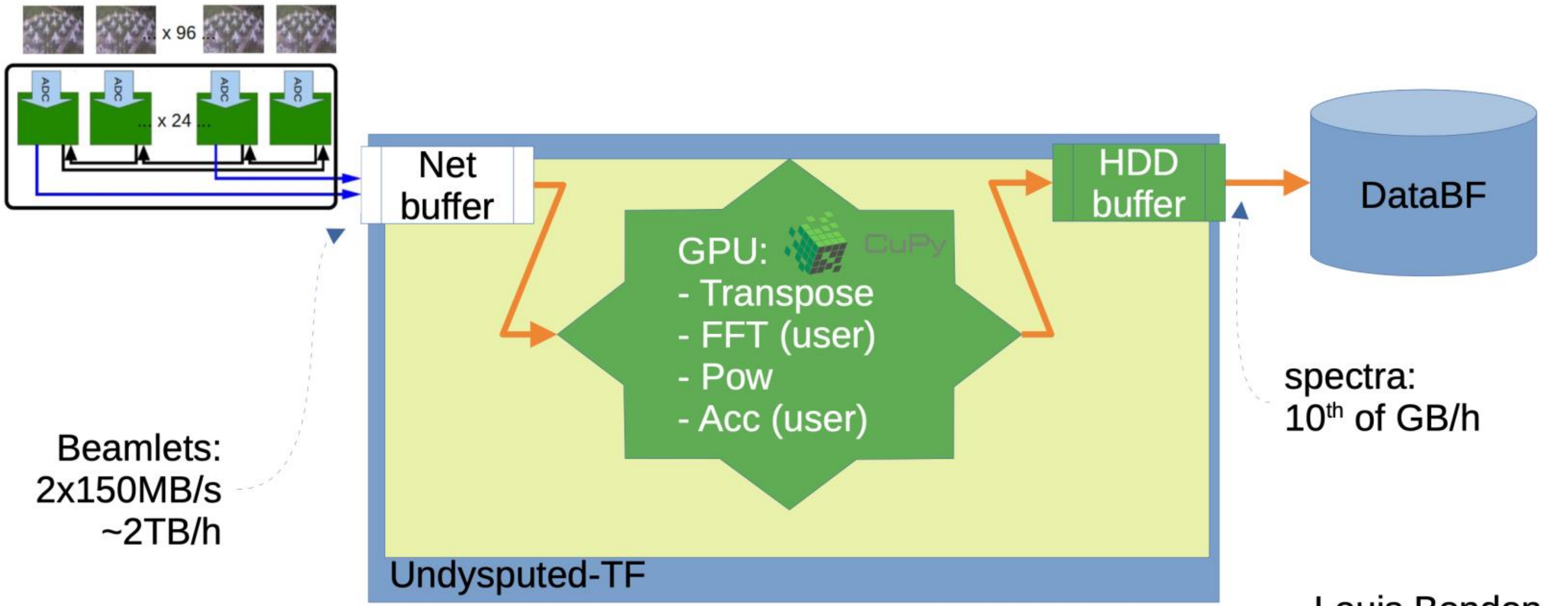


3 - Radon transform



MurMuRe pipeline and EXTRACT interfaces

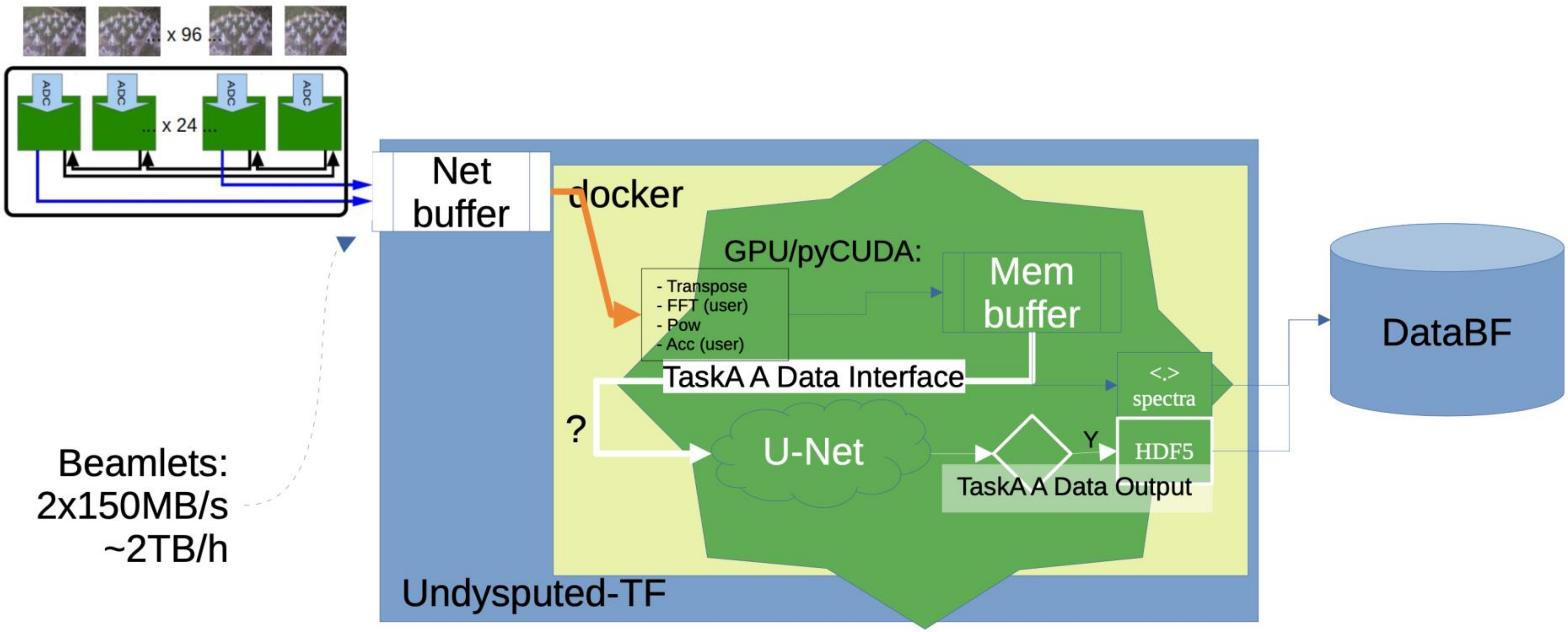
Undysputed-tf with MurMuRe (Modular Multicast Reciever)



Louis Bondonneau

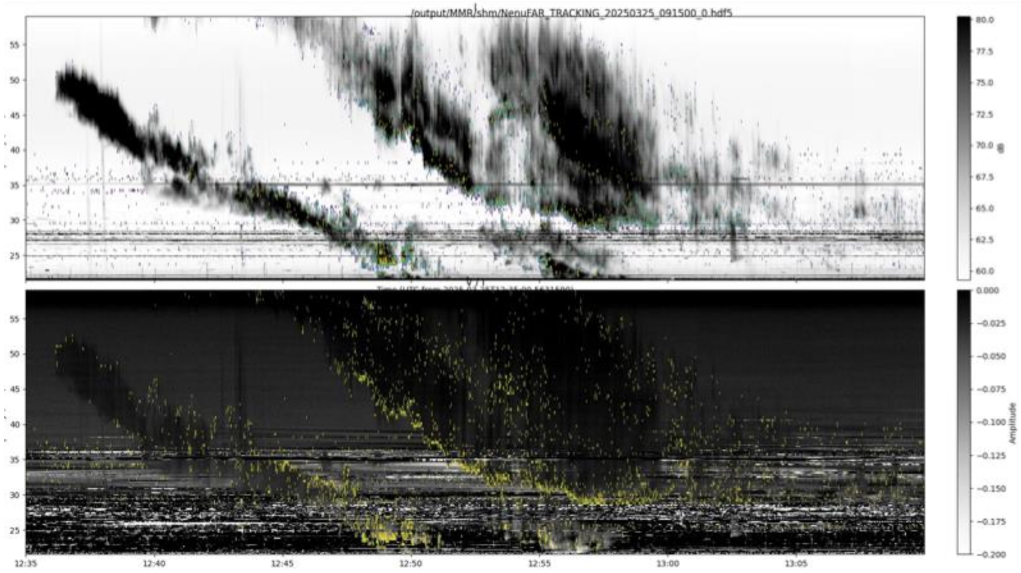
https://gitlab.obspm.fr/receiver/modular_multicast_receiver

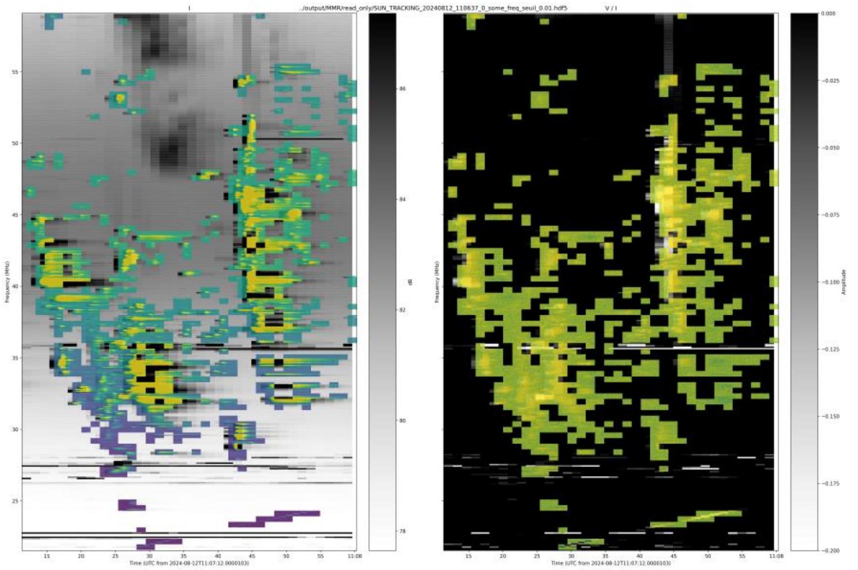
Undysputed-tf with MurMuRe and EXTRACT



Real-time detection with NenuFAR

SpikeNet successfully applied to real-time observations





Observation campaign : from 22 to 29/03/2025, 09h15-14h50, 21-84 MHz

→10x data volume reduction !

	Original pipeline	TASKA-A1 pipeline
(df,dt)	(6.1 kHz, 21 ms)	(98 kHz, 1.34 s)
Spectra	27 GB/hr	0.037 GB/hr
HDF5	--	2.5 GB/hr

S-bursts detection applied to NenuFAR observations

	Original pipeline	TASKA-A2 pipeline
(df,dt)	(3.05 kHz, 2.6 ms)	(24 kHz, 1.1 s)
Spectra	250 GB/hr	0.014 GB/hr
HDF5	--	2.5 GB/hr

→ expect 10x data volume reduction !

Conclusions

Conclusions

- Important step toward smart data filtering for next generation radiotelescopes such as NenuFAR or SKA.
- It also paves the way for “analog to information” processing, which would drastically reduce the storage needs.
- The type of emissions studied in this work require a high time-frequency resolution, but they are embedded within larger slowly-varying emissions that can be studied at a lower resolution.
- This approach helps optimizing data storage while maintaining its value for scientific analysis, thus preparing for scalable solutions in the SKA era.
- Code available on GitLab, easy to clone and use with notebook tutorials provided.

[https://gitlab.obspm.fr/extract]

Next steps

- Develop a CNN for S-bursts based on anomaly detection.
- Optimize existing algorithms to have better detections
- Find an even more efficient way to store the outputs.

Thank you !

emilie.mauduit@obspm.fr

Follow us on social media:

www.extract-project.eu

EXTRACT

A distributed data-mining software platform for extreme data across the compute continuum

The EXTRACT Project has received funding from the European Union's Horizon Europe programmeunder grant agreement number 101093110