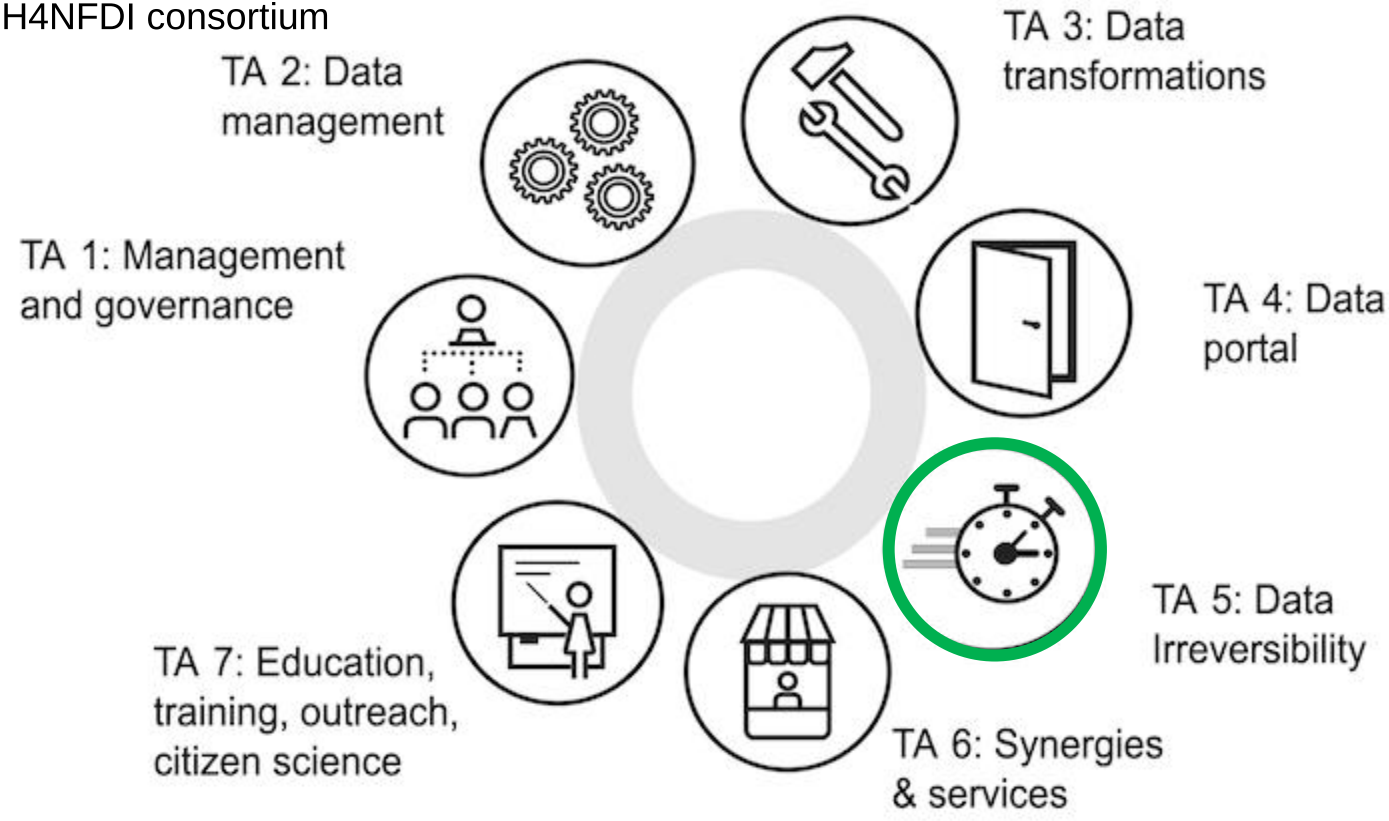
TA3-WP3 Monthly Meeting

Deep Learning for Real-time Classification of Astronomical Radio Signals: Current Status

Andrei Kazantsev, Ramesh Karuppusamy, Yunpeng Men, Rishi Kumar, Michael Kramer
Max Planck Institute for Radio Astronomy
Bonn, Germany

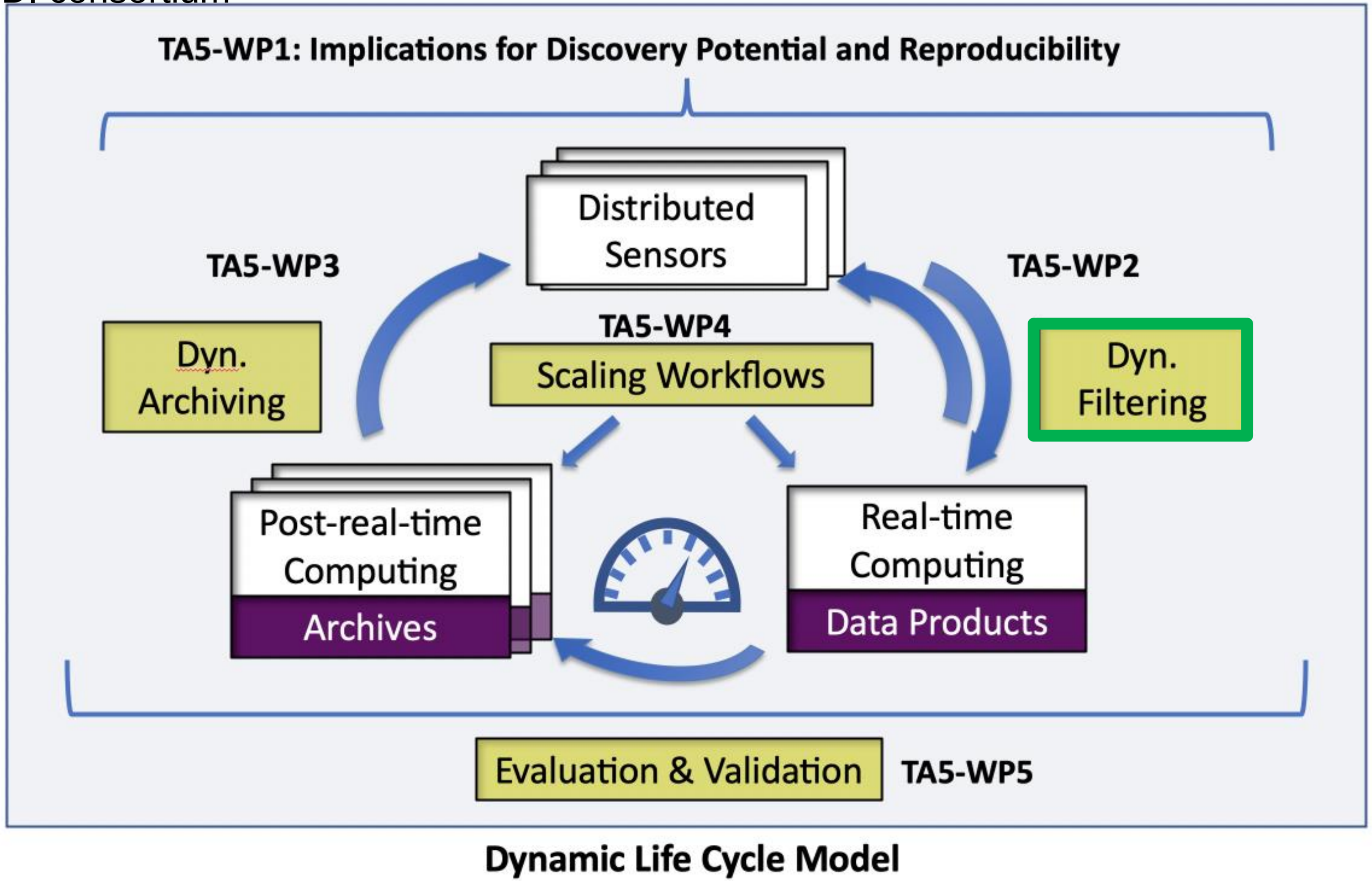
May 19, 2025

PUNCH4NFDI consortium



The task areas of the PUNCH4NFDI consortium

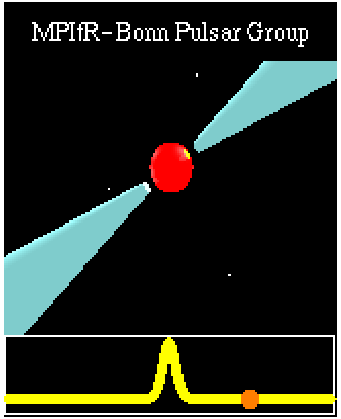
Work Packages in Task Area 5

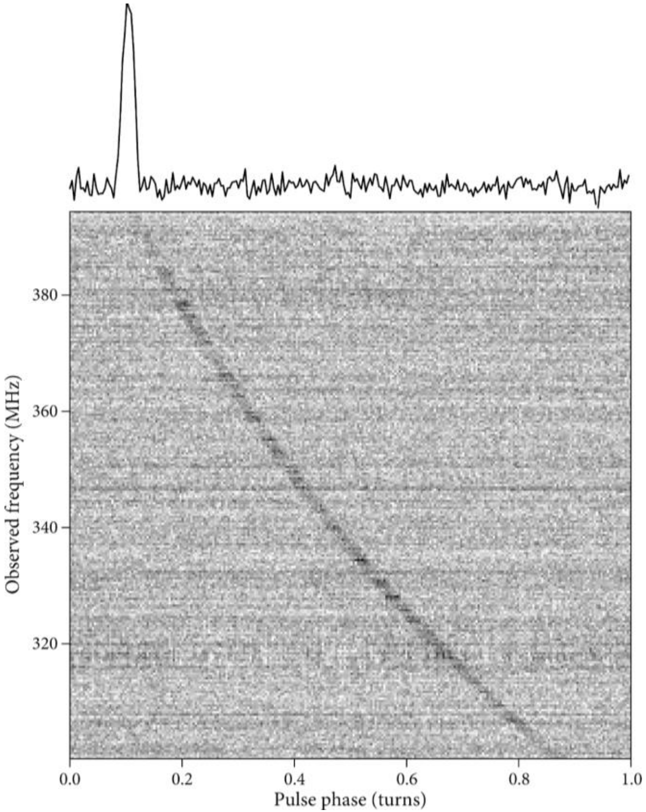


The work packages of TA5 of the PUNCH4NFDI consortium

Astronomical Radio Signals

Pulsars



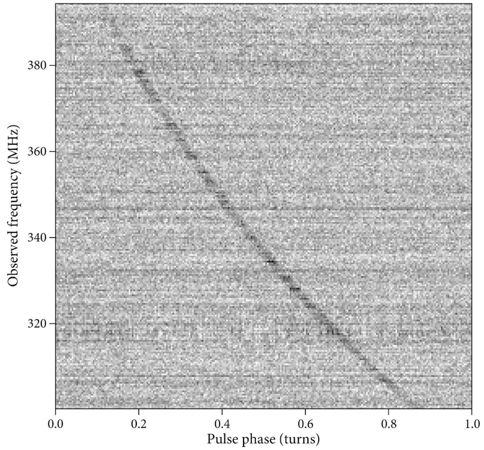


Example of dispersion delay for an individual pulse from J1800+5034

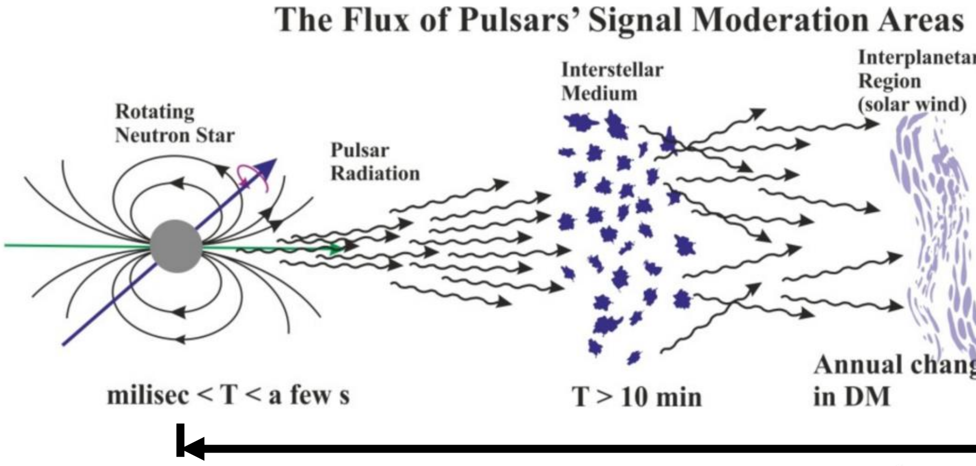
Fast **R**adio **B**ursts

?

Astronomical Radio Signals



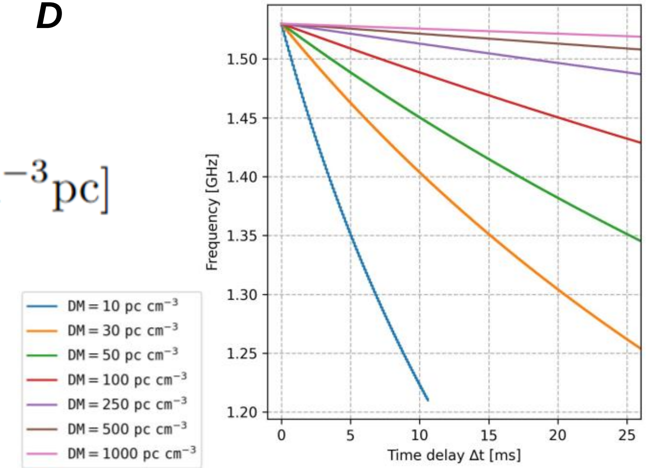
Example of dispersion delay for an individual pulse from J1800+5034



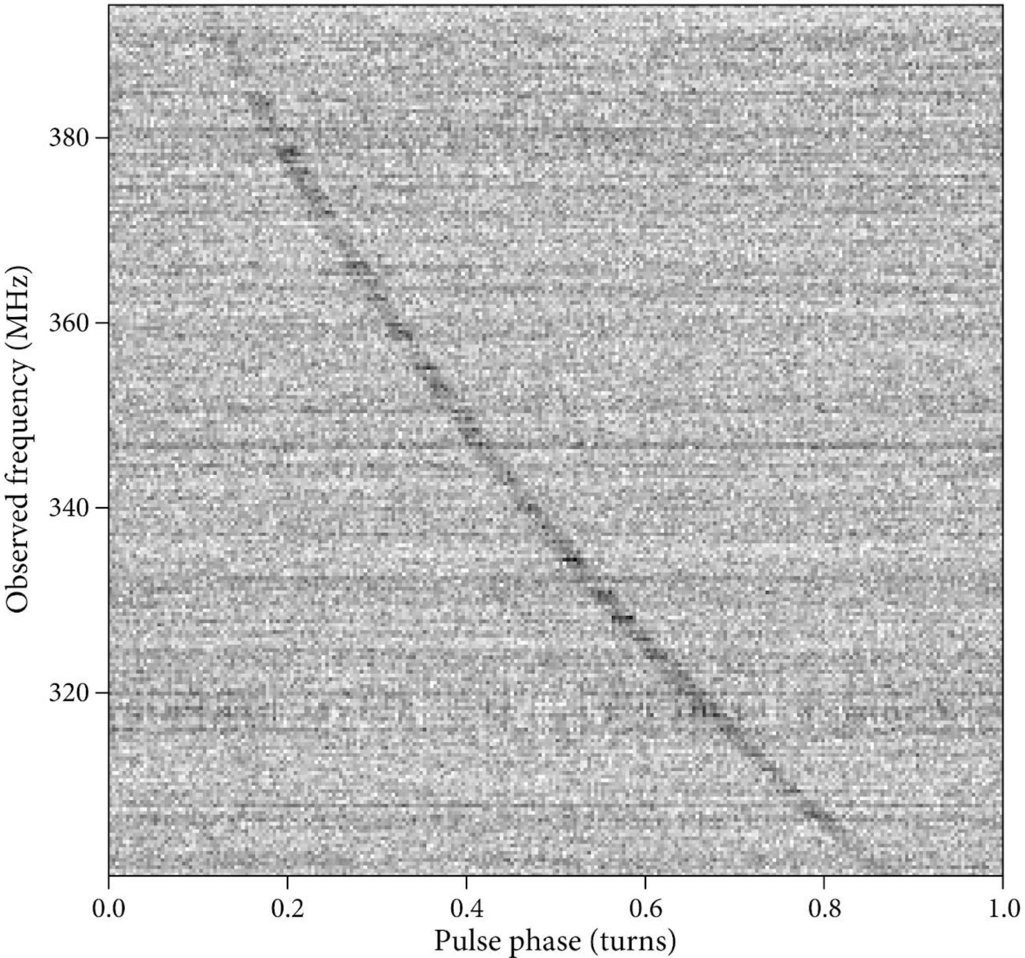
D

$$\Delta t = 4.12\ \text{ms}\ \left((f_{\text{LO}}[\text{GHz}])^{-2} - (f_{\text{HI}}[\text{GHz}])^{-2} \right)\ \text{DM}\ [\text{cm}^{-3}\text{pc}]$$

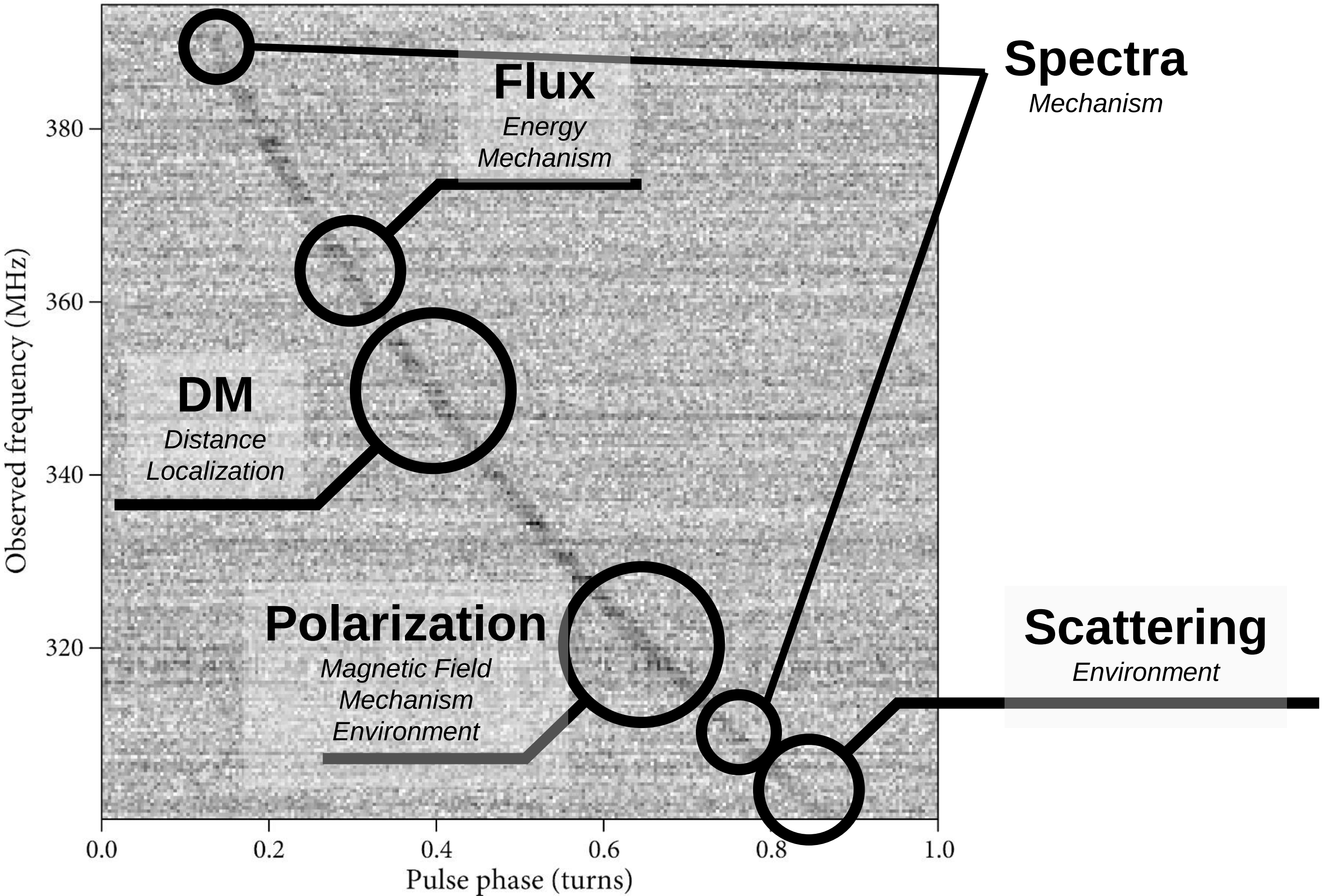
$$\text{DM} = n_e\ [\text{cm}^{-3}]\ D\ [\text{pc}]$$



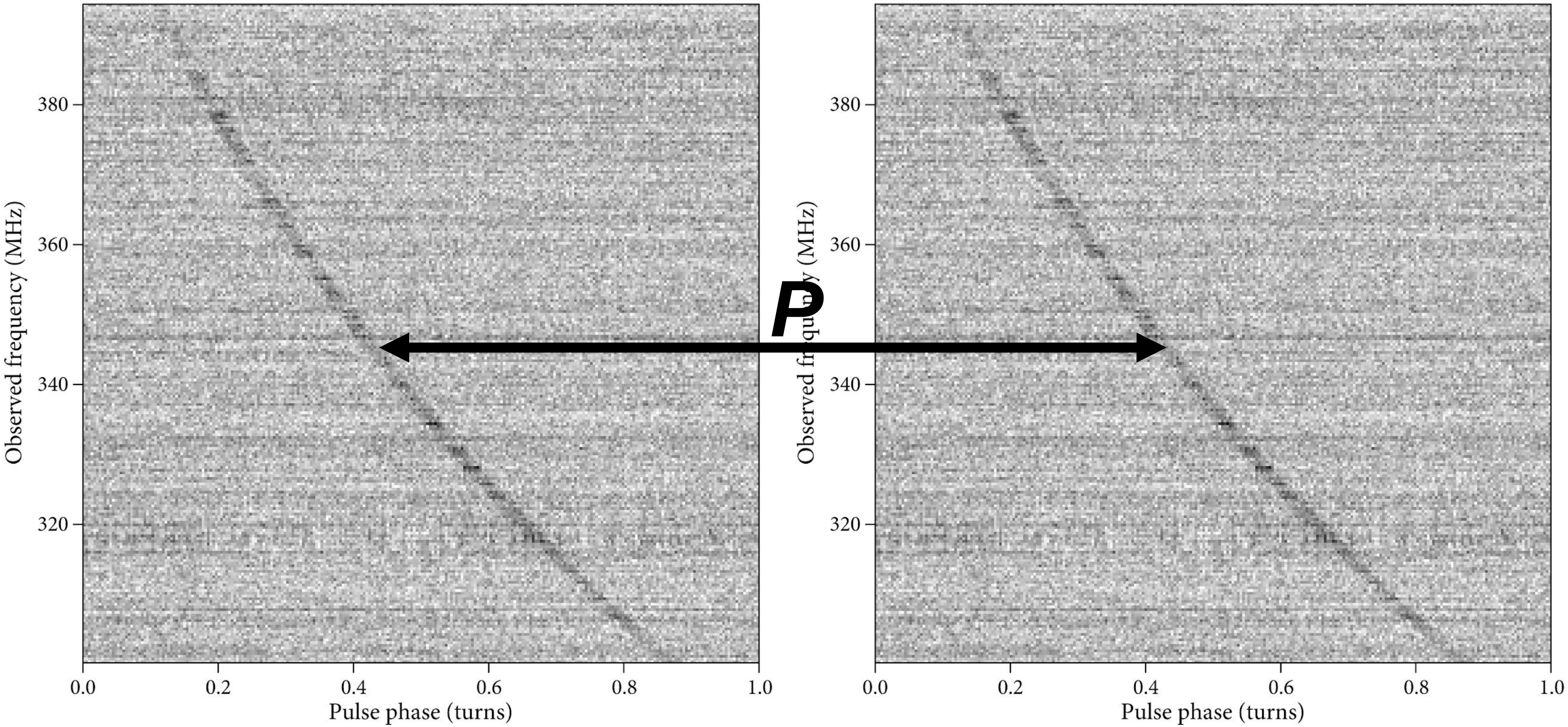
Information from the Signal



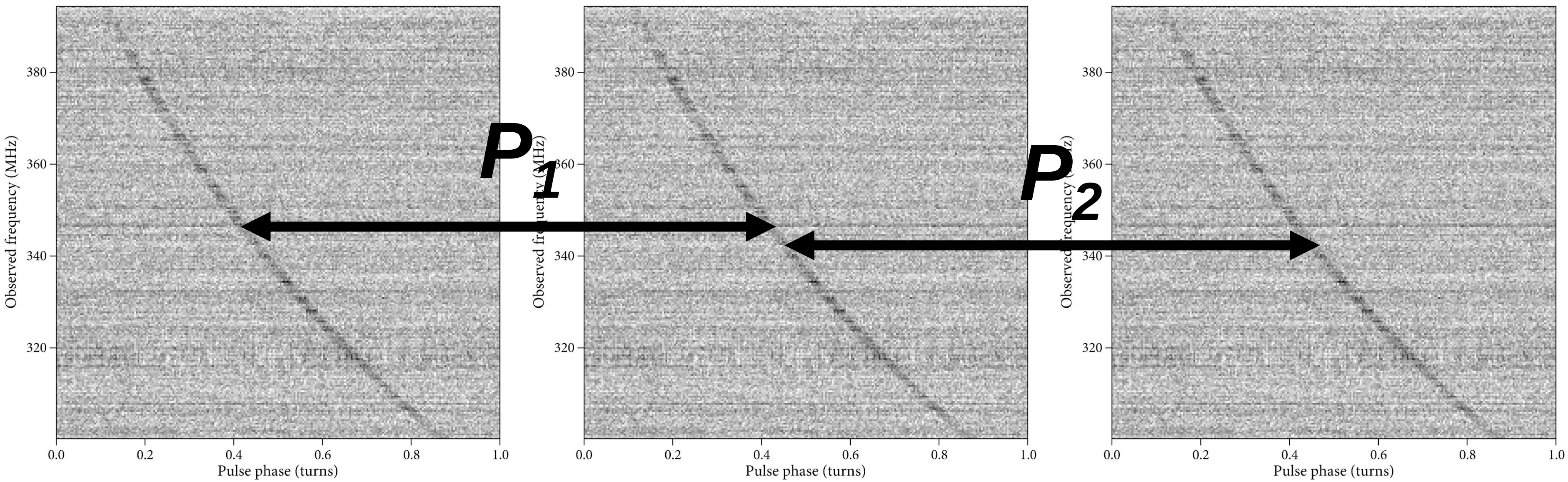
Information from the Signal



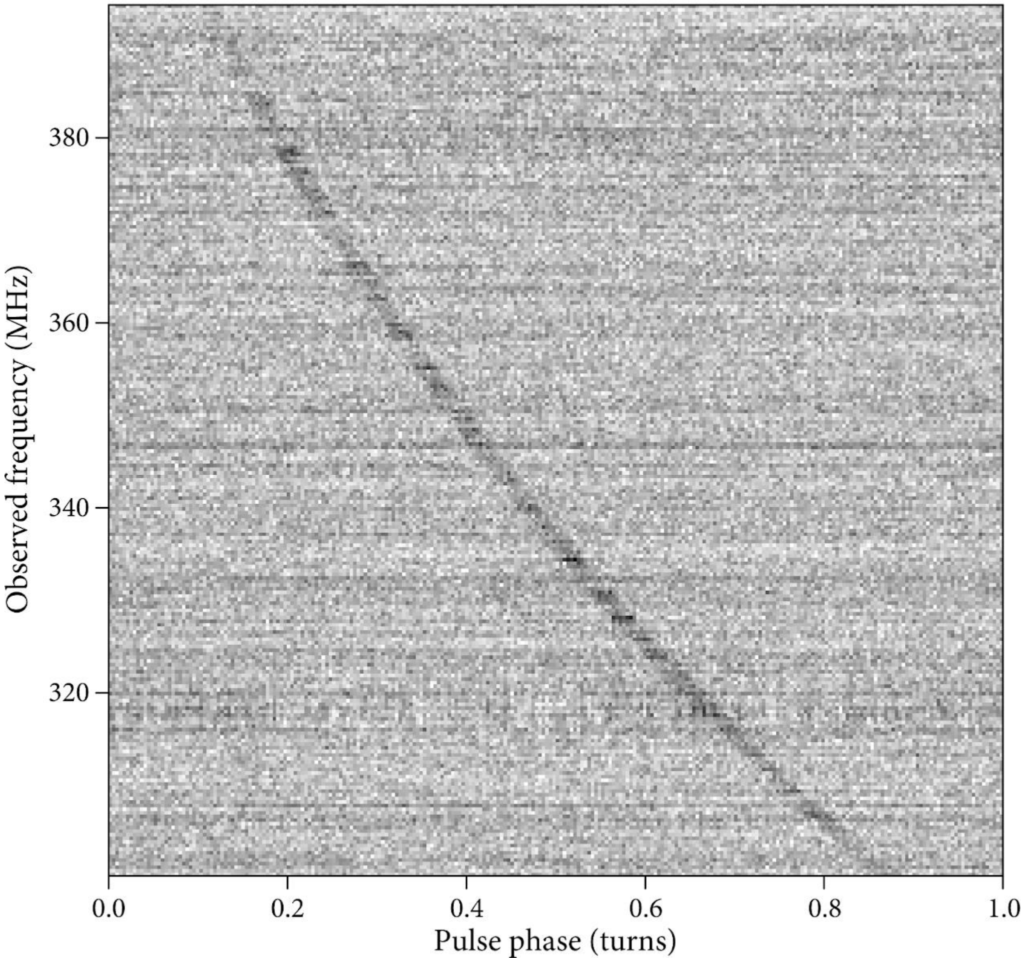
Information from the Signal



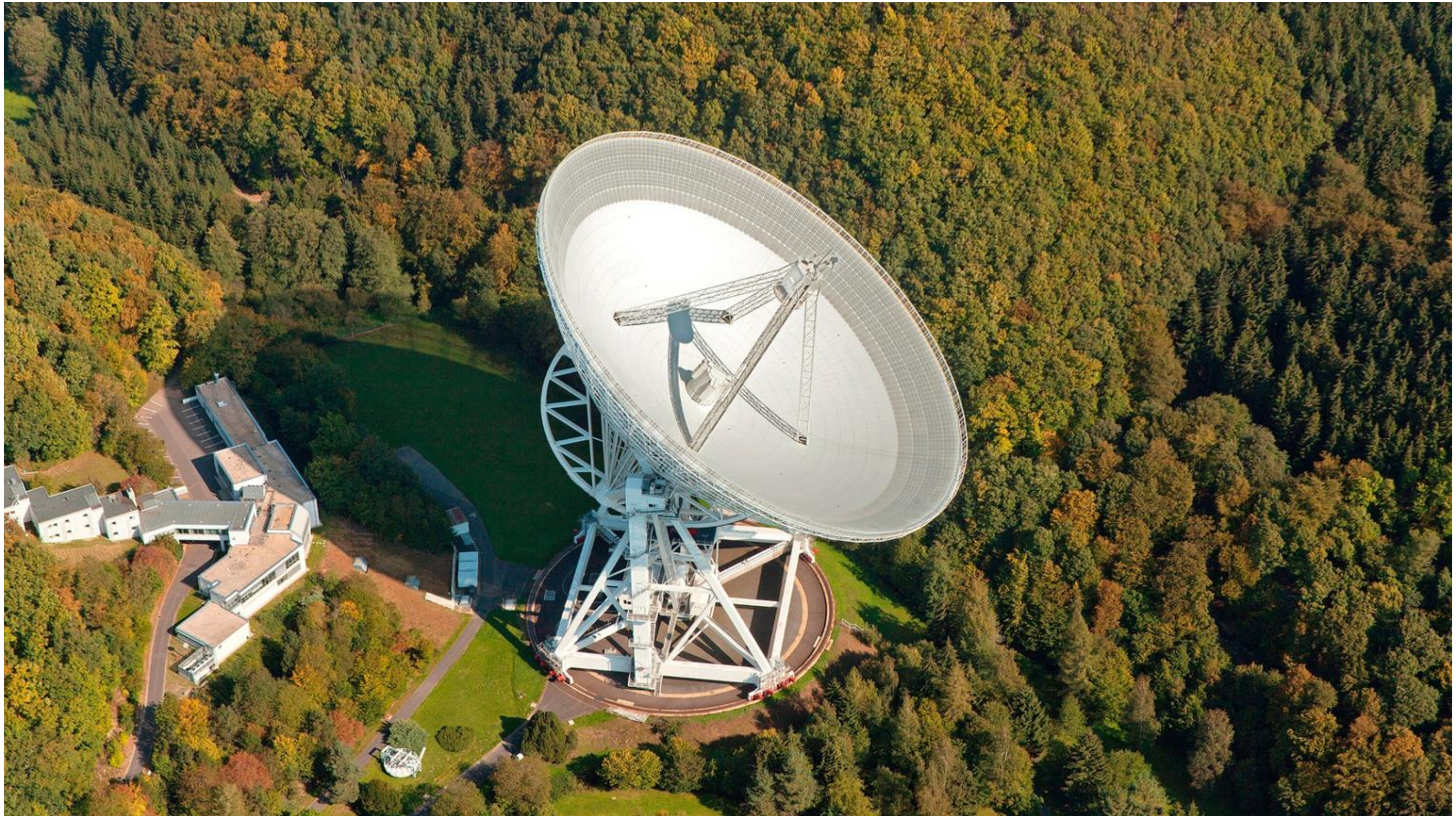
Information from the Signal



Roseta Stone for Pulsar Astrophysics

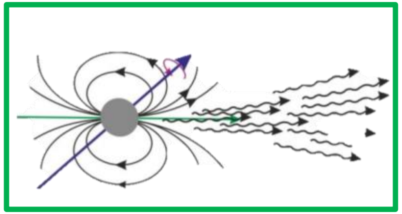


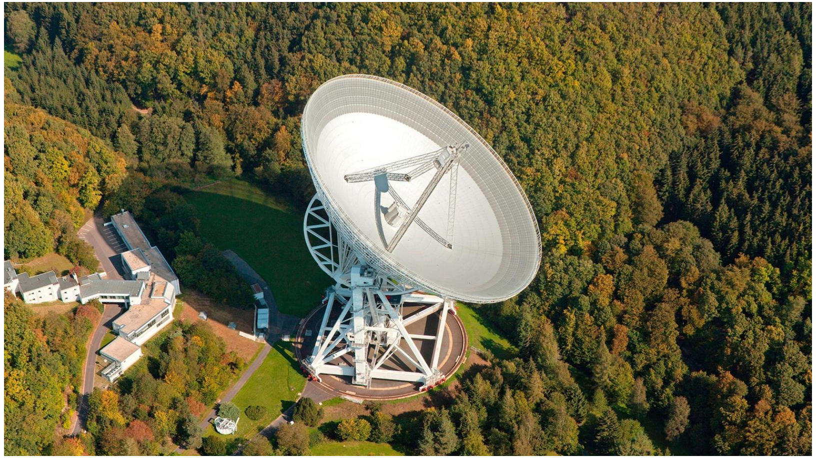
Detection of the Astronomical Radio Signals

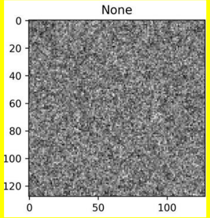


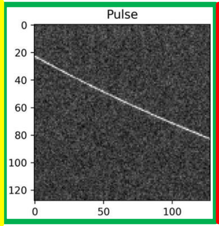
Radio Telescope Effelsberg, Bad Münstereifel, Germany © Hans Blossey

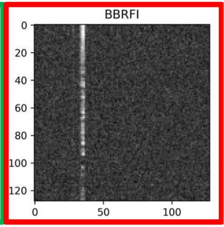
Astronomical and non- Astronomical Radio Signals

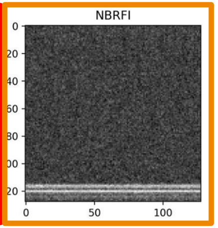










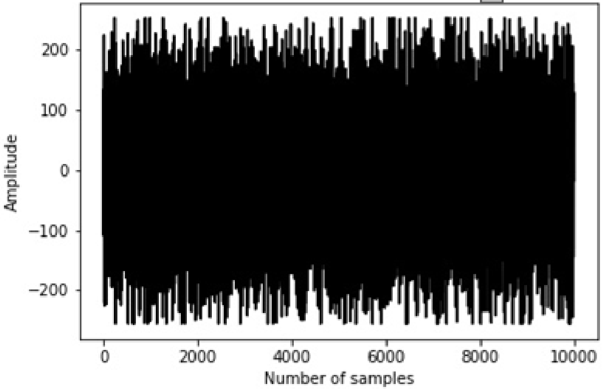


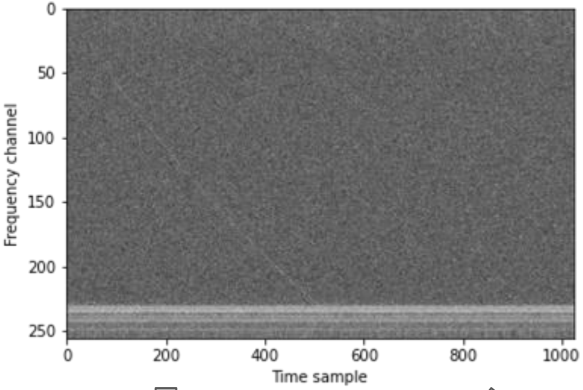
Data Formats in Radio Astronomy

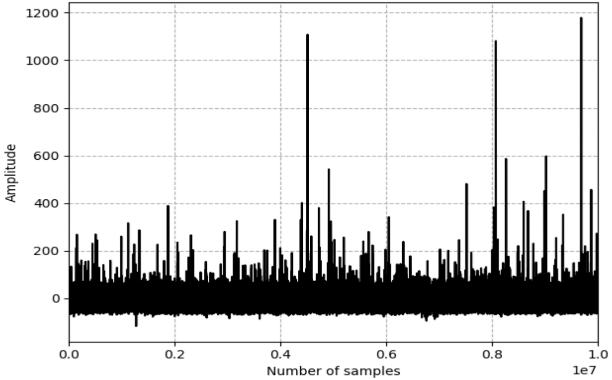
$$X(t)$$

$$S(f,t) = \left| \int_{-\infty}^{\infty} X(t) e^{-2\pi i f t} dt \right|$$

$$T(t) = \sum_{f} S(f,t)$$





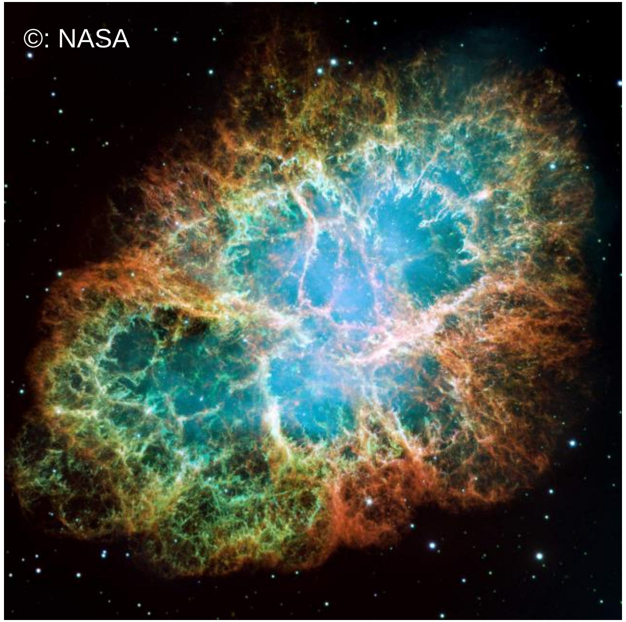


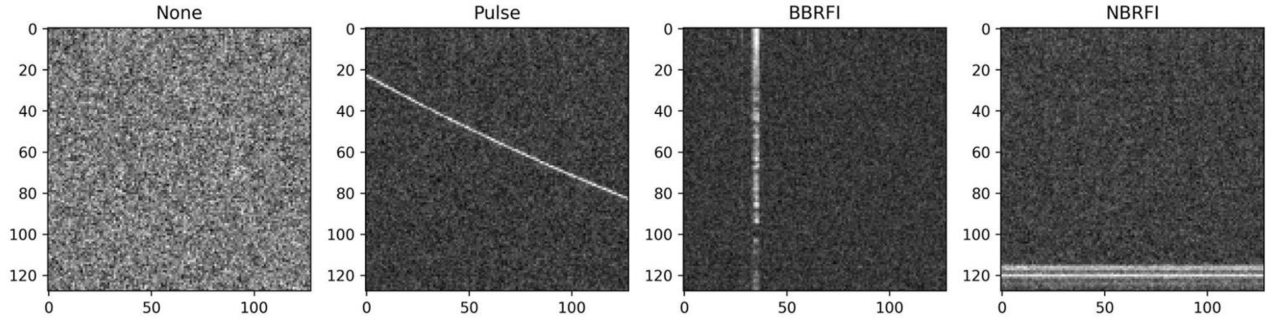
Format	Baseband data (dada files)	Spectrograms (filterbank file)	Time series
21 minutes	1.6 TB	3 GB	49 MB

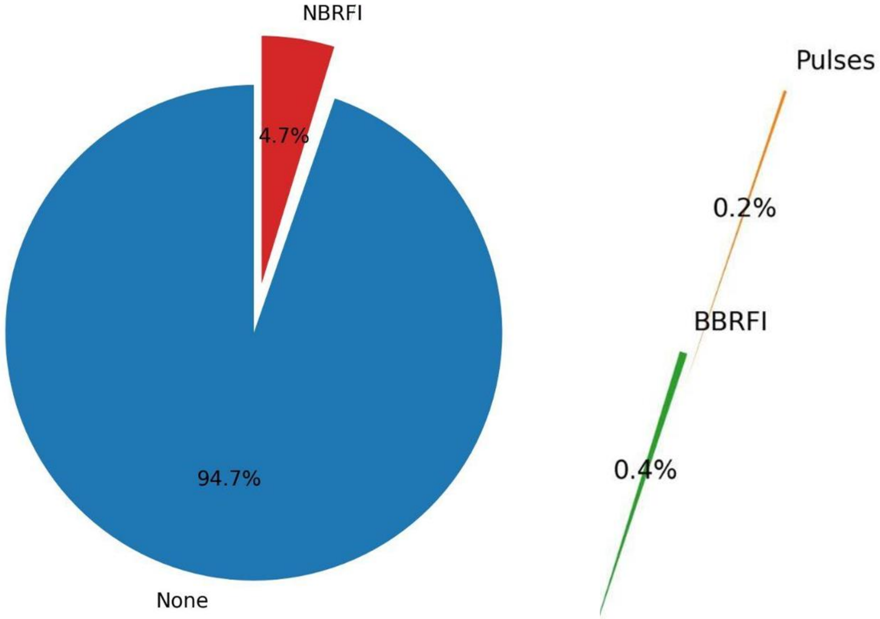
Data Rate in Radio Astronomy

Radio telescope name	Radio telescope exterior	Bitrate per beam	Total bitrate
Effelsberg		P210-7: 11.04 Gb /s	77 Gb / s (7)
		UWB: 290 Gb / s	290 Gb / s (1)
MeerKAT		107 Mb / s	1.7 Tb / s (~1024 beams)
Square Kilometer Array		~ 1 Gb / s	20 Tb / s (>2200 beams)

Statistics for One Observation

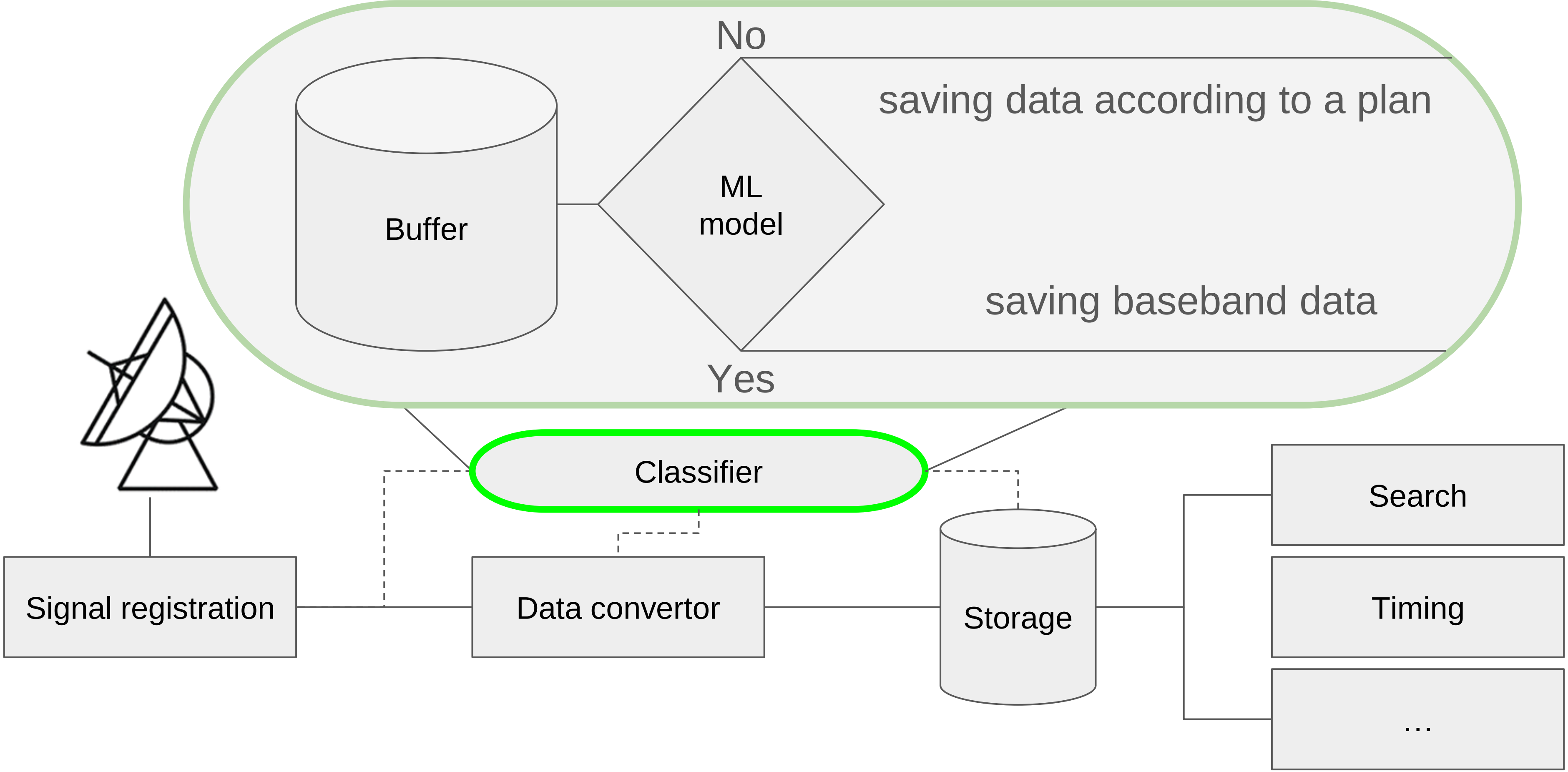




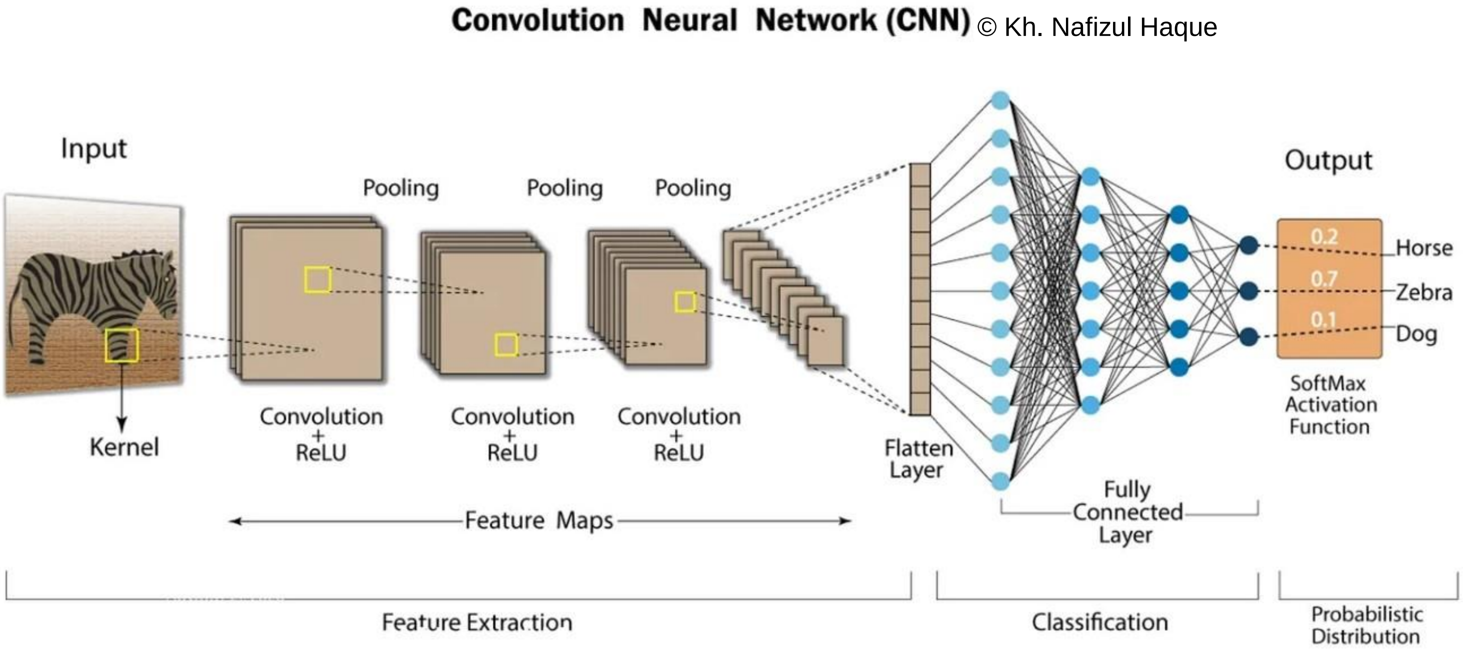


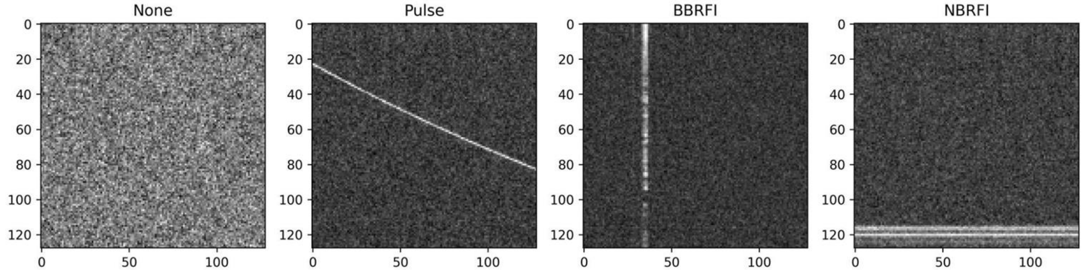
Object: **Crab pulsar (B0531+21)**
Data: **2020-05-31**
Time resolution: **102.4 μs**
Telescope: **Effelsberg**
Bandwight: **1210 - 1530 MHz**
Duration: **21 minutes**

Conception of the Task

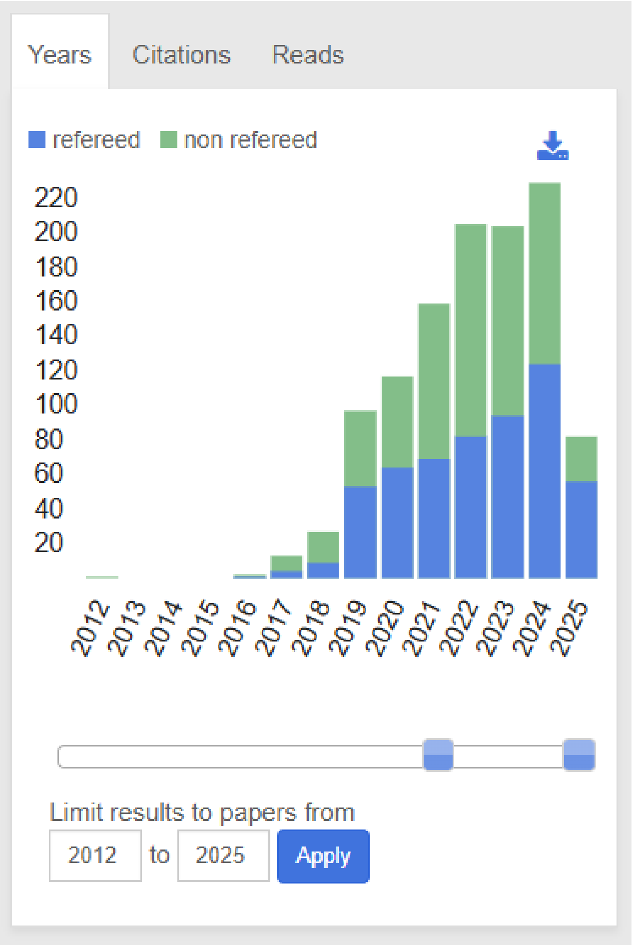


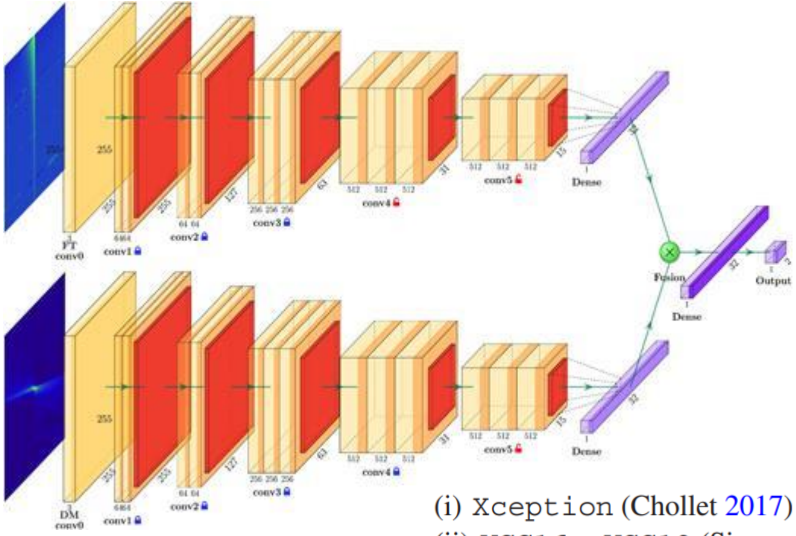
CNN Usage for Pattern Recognition





CNN Usage in Astronomy



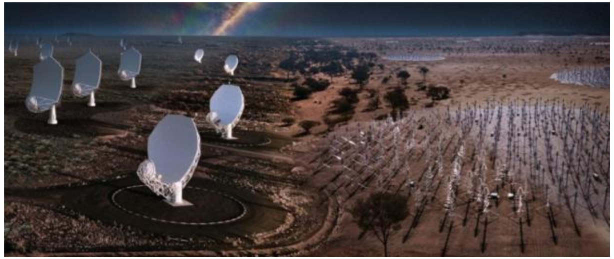


(i) `Xception` (Chollet 2017)
(ii) `VGG16, VGG19` (Simonyan & Zisserman 2014)
(iii) `ResNet50` (He et al. 2016)
(iv) `DenseNet121, DenseNet169, DenseNet201` (Huang et al. 2017)
(v) `InceptionV3` (Szegedy et al. 2016)
(vi) `InceptionResNetV2` (Szegedy et al. 2017)
(vii) `MobileNet` (Howard et al. 2017)
(viii) `MobileNetV2` (Sandler et al. 2018)

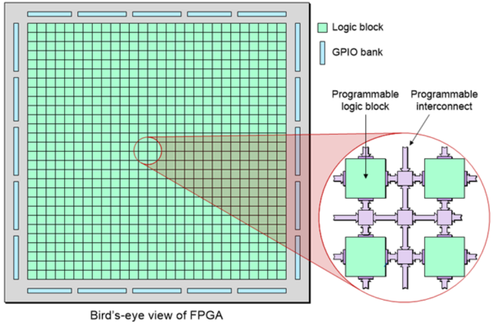
Devansh Agarwal, et al., MNRAS, 2020,
https://doi.org/10.1093/mnras/staa1856

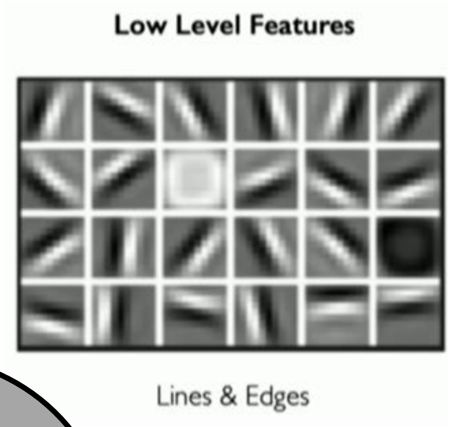
Usage of Extremely Minimalistic Model

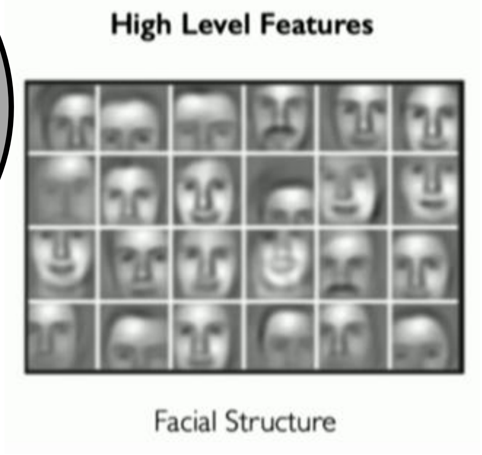
©: SKA



©: www.latticesemi.com

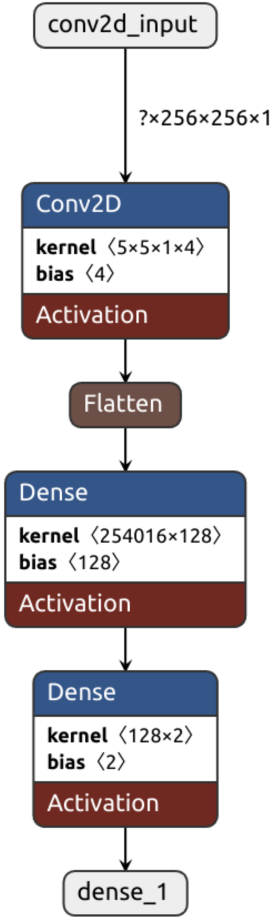


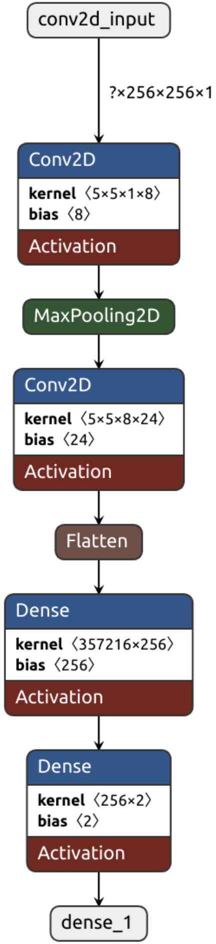


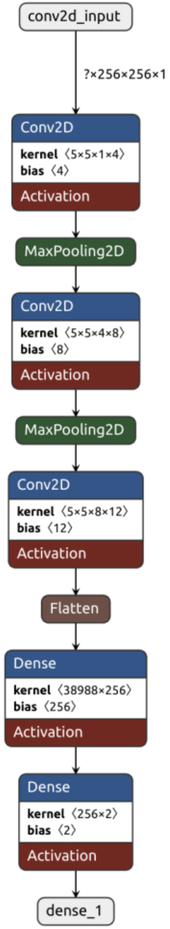


©:velog.io/@ktm1237/MIT-Introduction-to-Deep-Learning

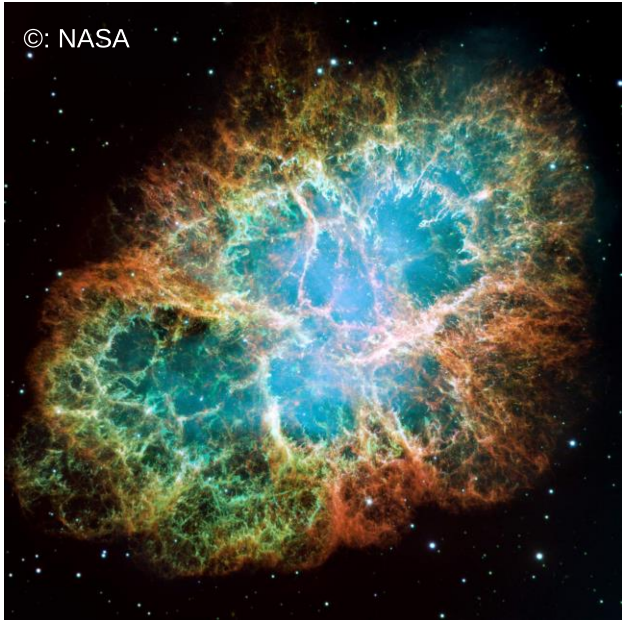
Extremely Minimalistic Classifier Architecture

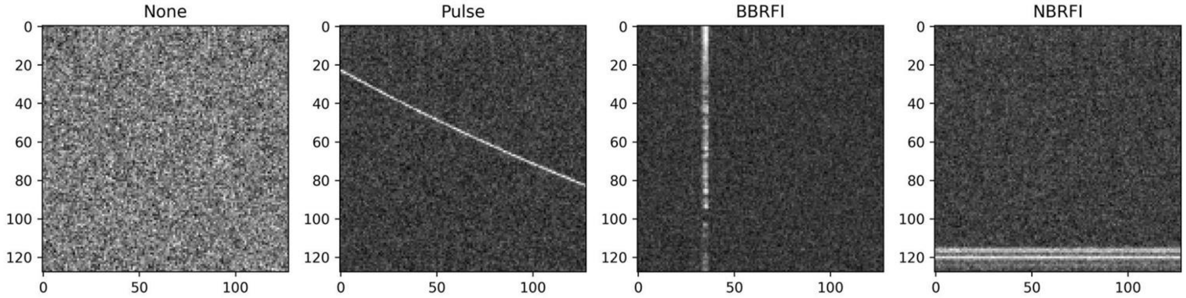




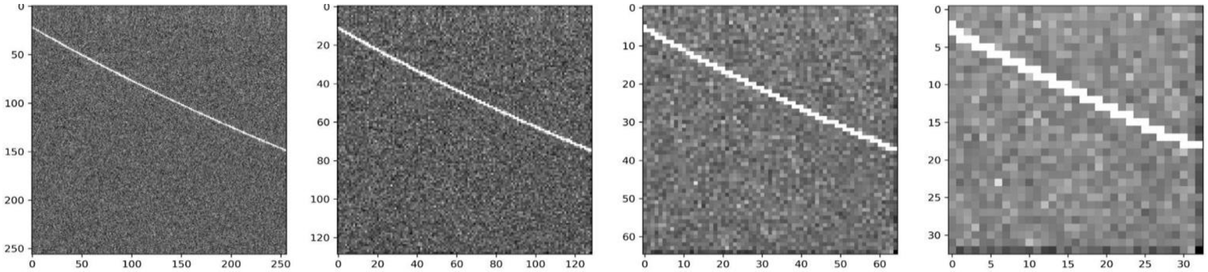


Spectrogram Classifier





Number of classes: 8

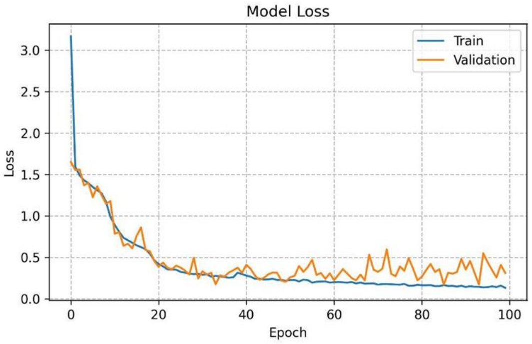


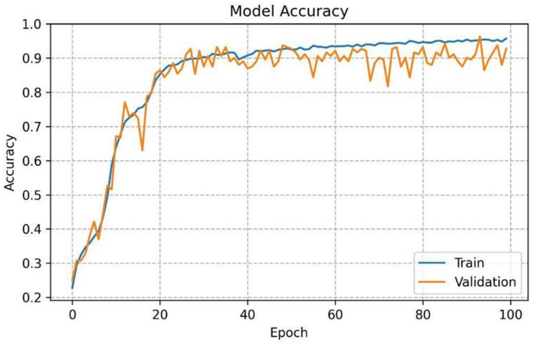
Test with four image resolutions

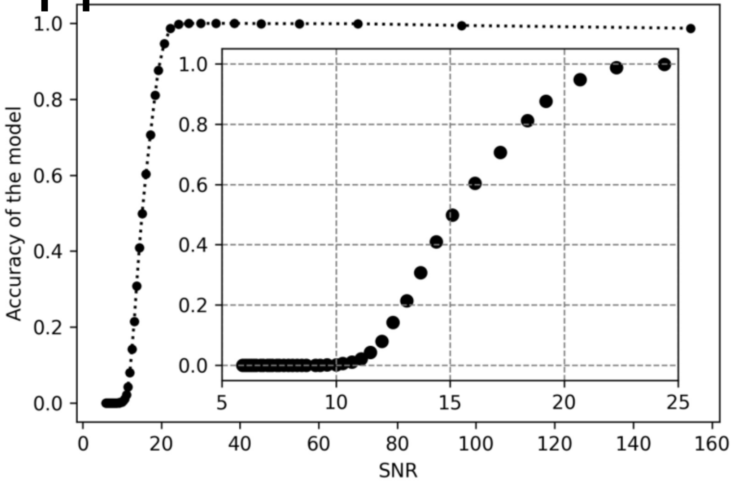
Data Preparation

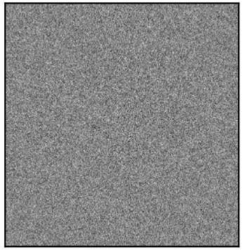
Model Selection → HP Tuning → Training → Validation → Validation with additional dataset

Limitations of the “Spectrogram” Approach

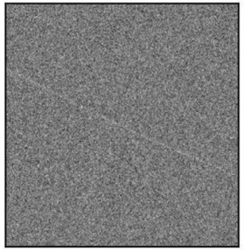






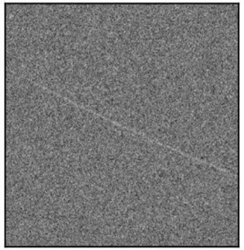


5.0

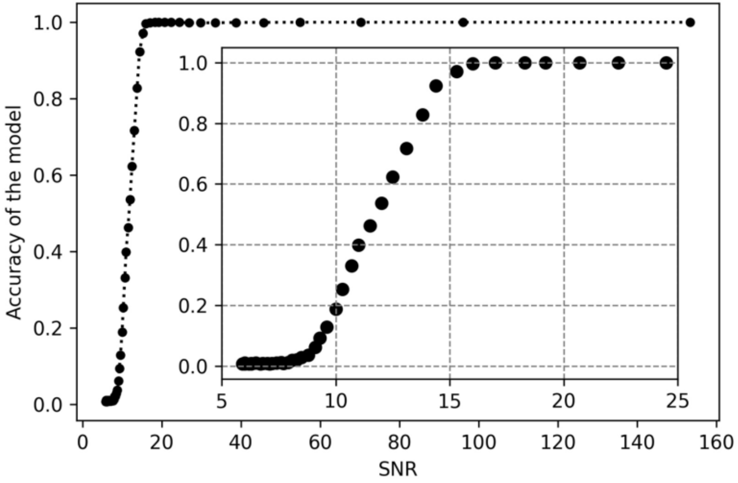


8.6

11.4

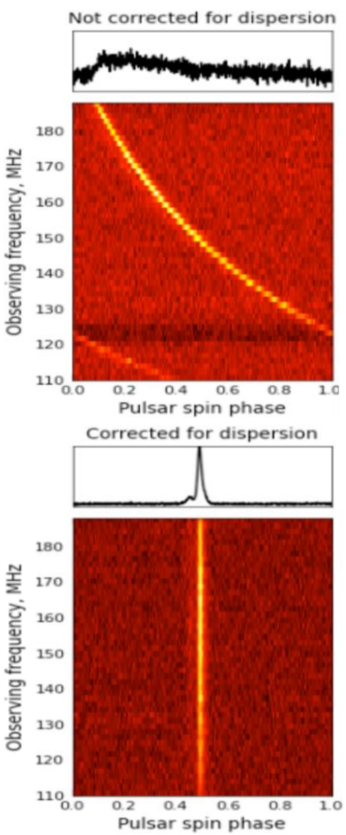


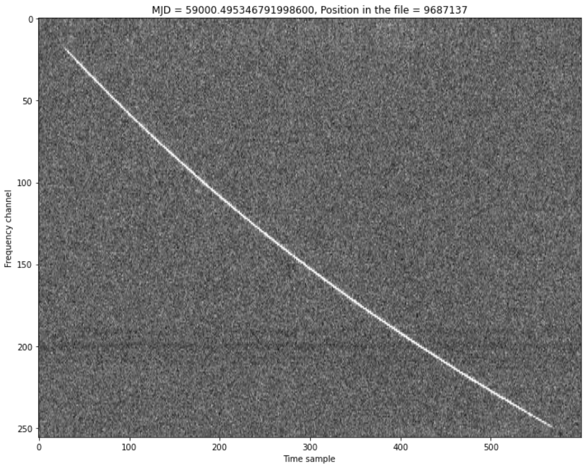
15.6

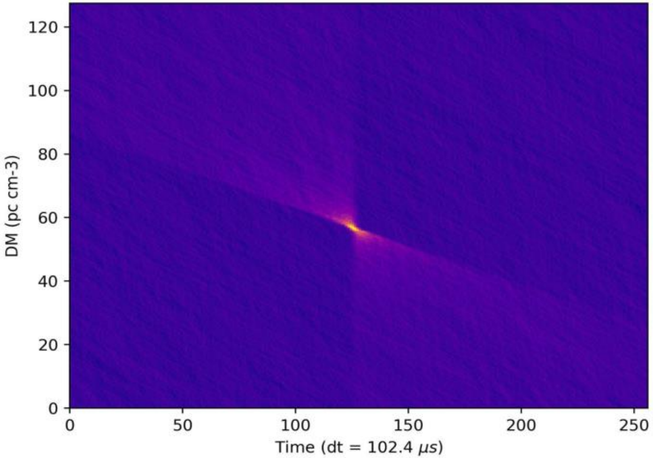


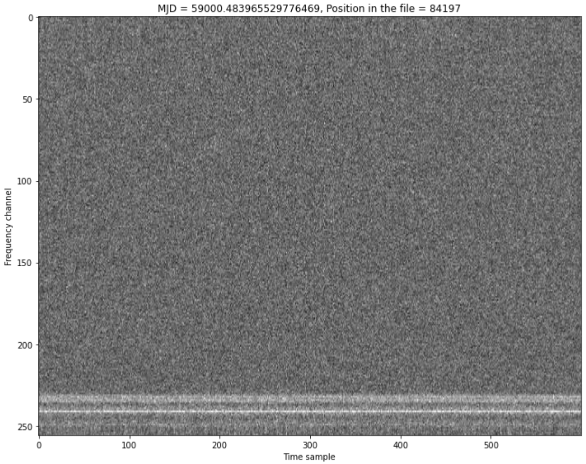
Spectrograms vs. DM-Images

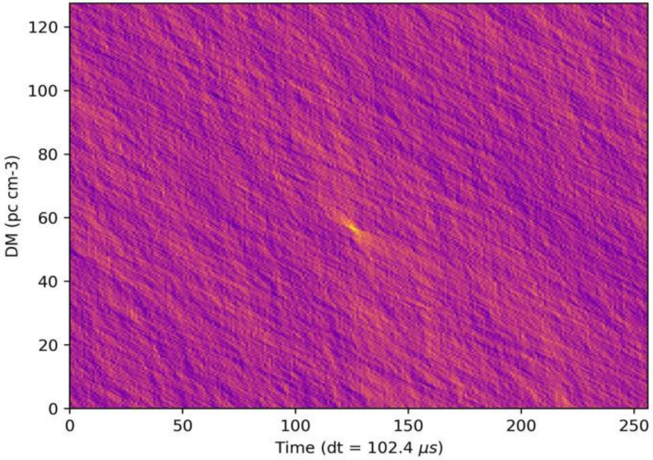
©: Anya Bilous



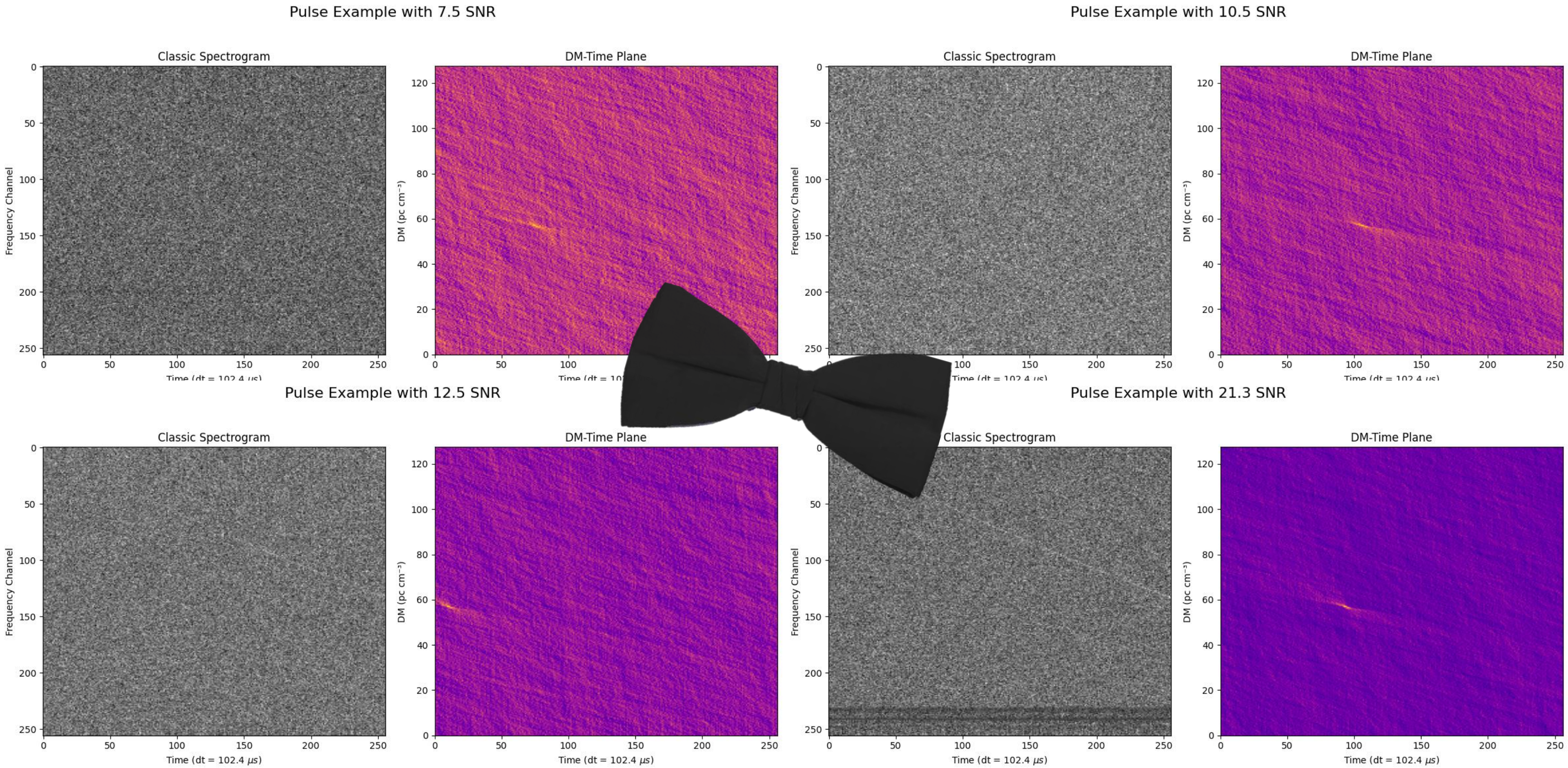






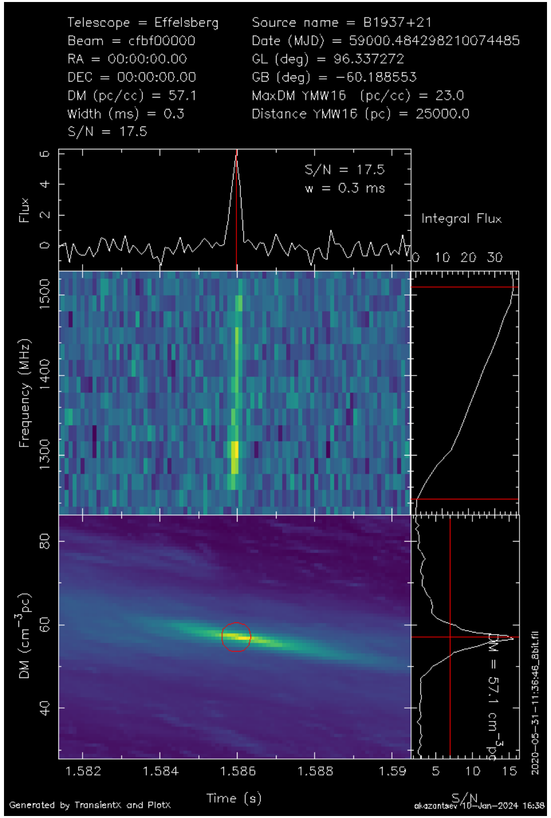


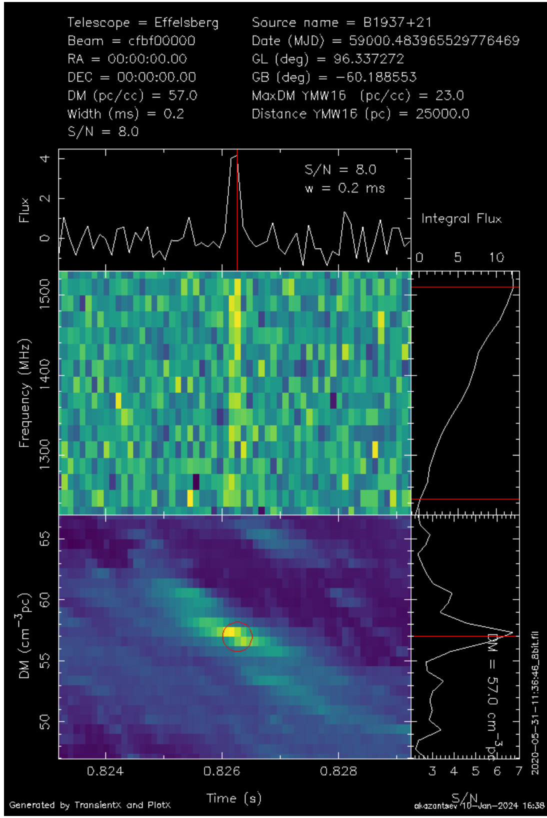
“Bow-tie” with Different SNR



TransientX as a Fast Solution for Dedispersion

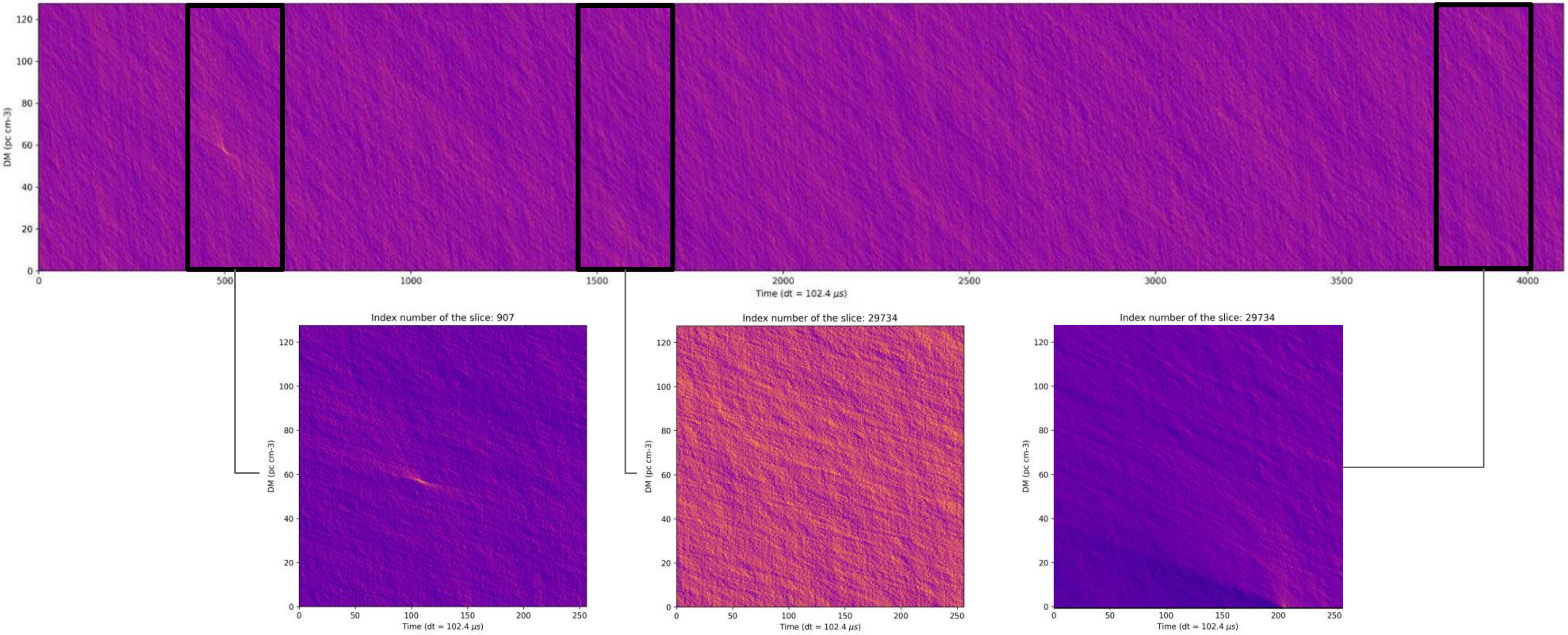
a one command line high performance transient search software.





https://github.com/ypmen/TransientX

DM-time Data from TransientX

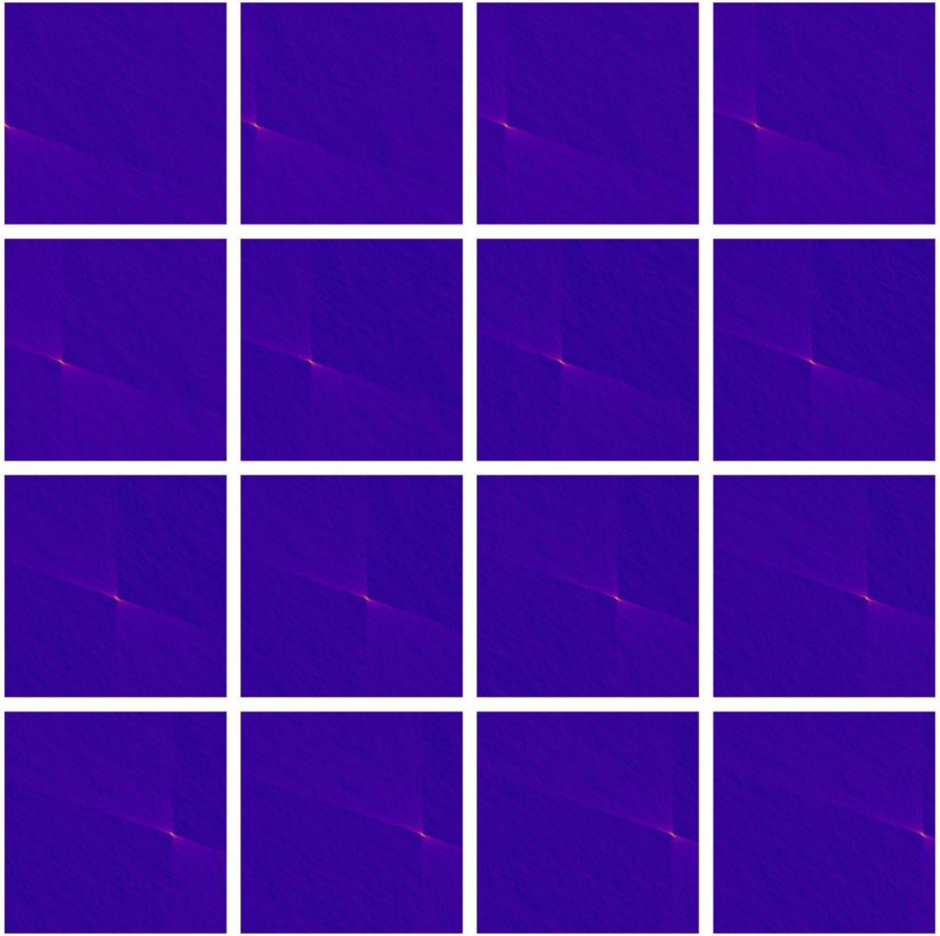


Pulse

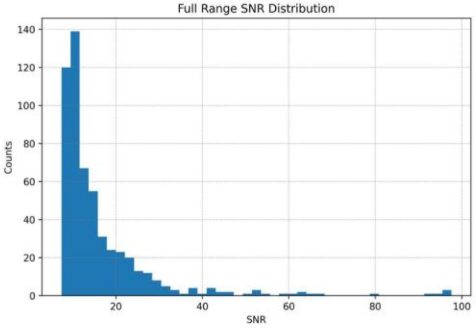
Noise

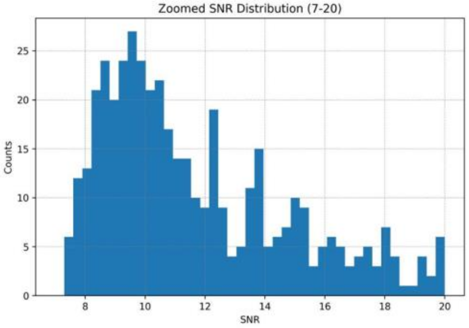
BBRFI

DM-time dataset

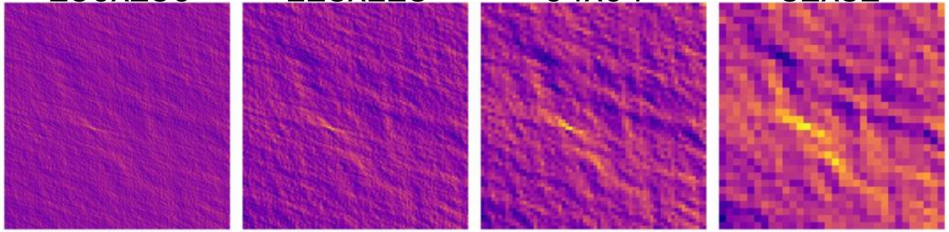


297 472 DM-time image in total. A big portion of this data (**80%, 237 977**) has been used for training sessions, The rest of the data (**20%, 59 495**) has been saved as an unseen dataset to calculate key metrics after training.

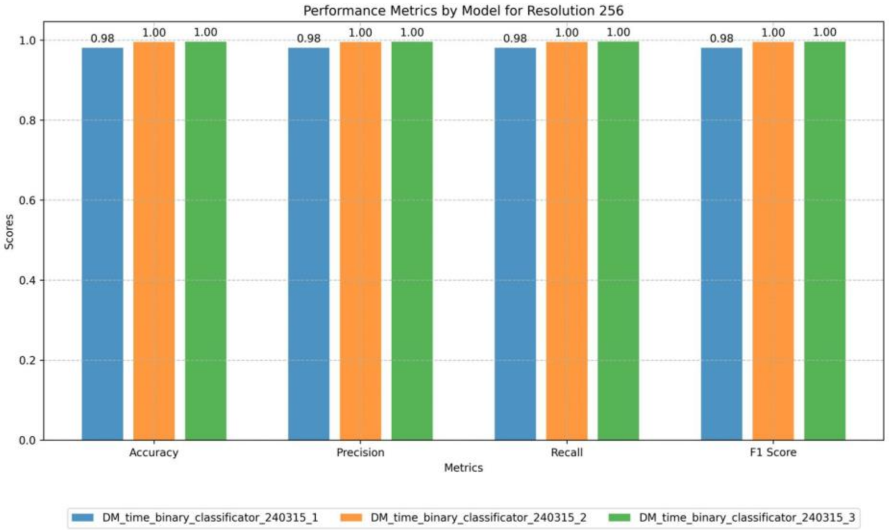


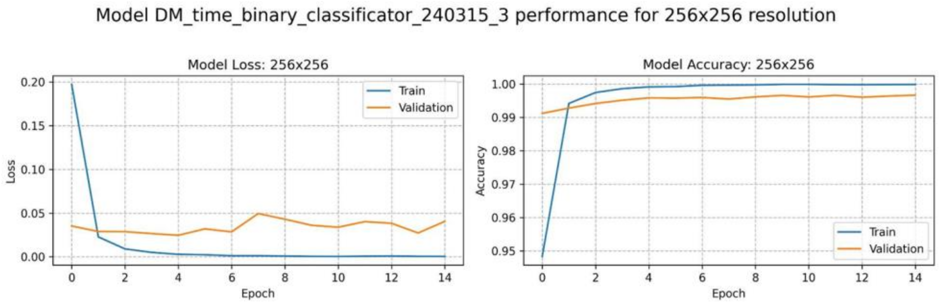


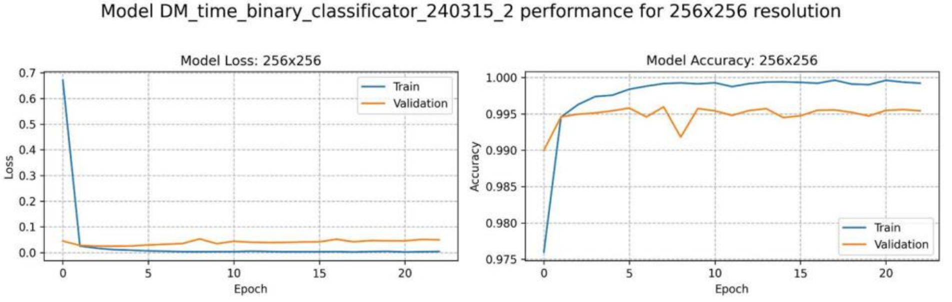
256x256 128x128 64x64 32x32

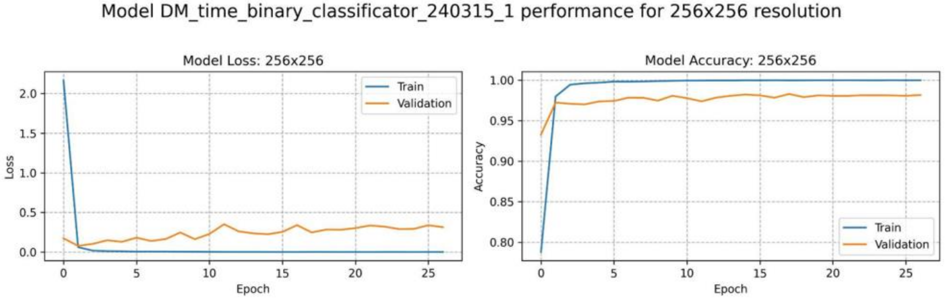


Results of Training and Validation

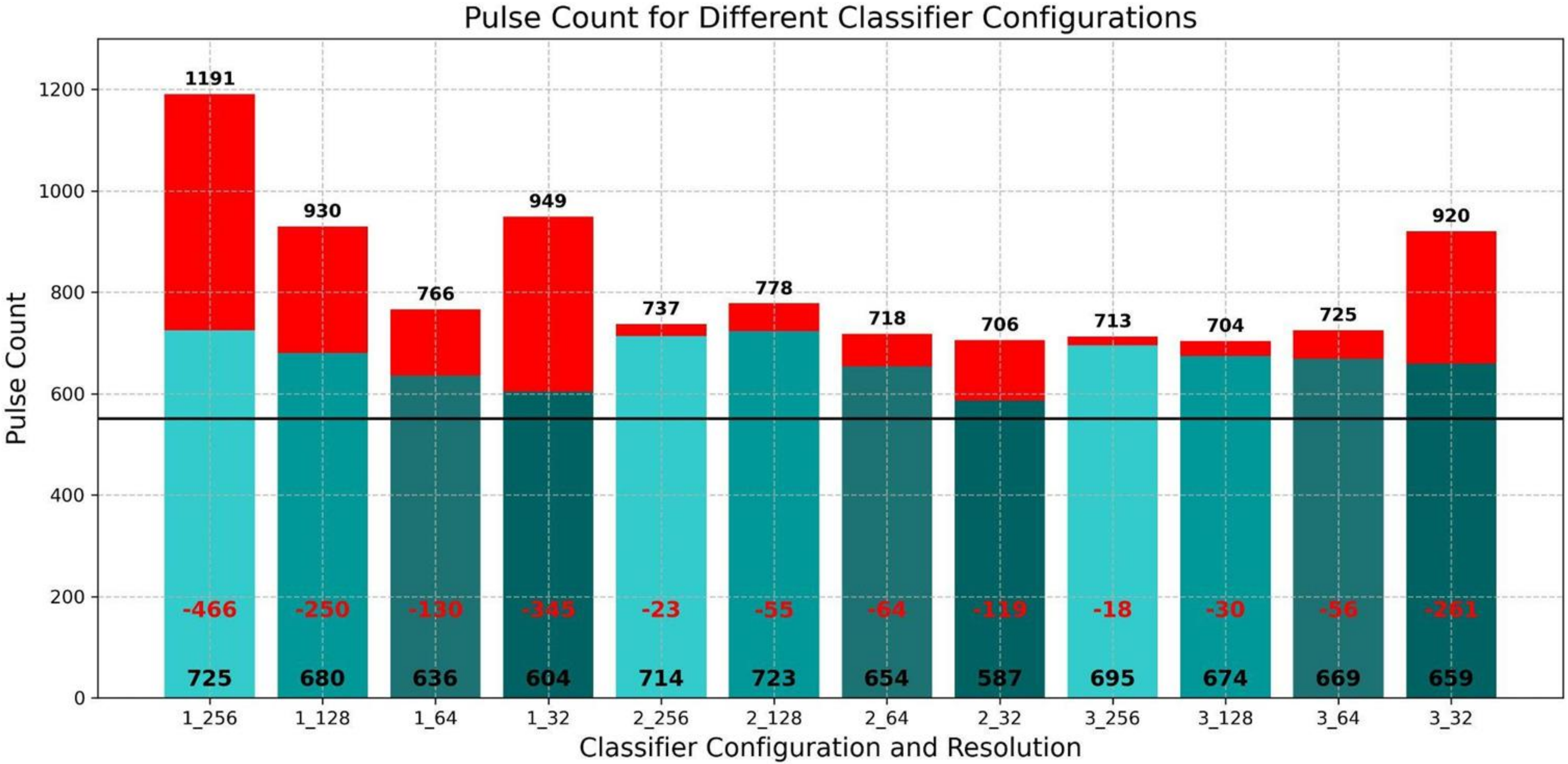








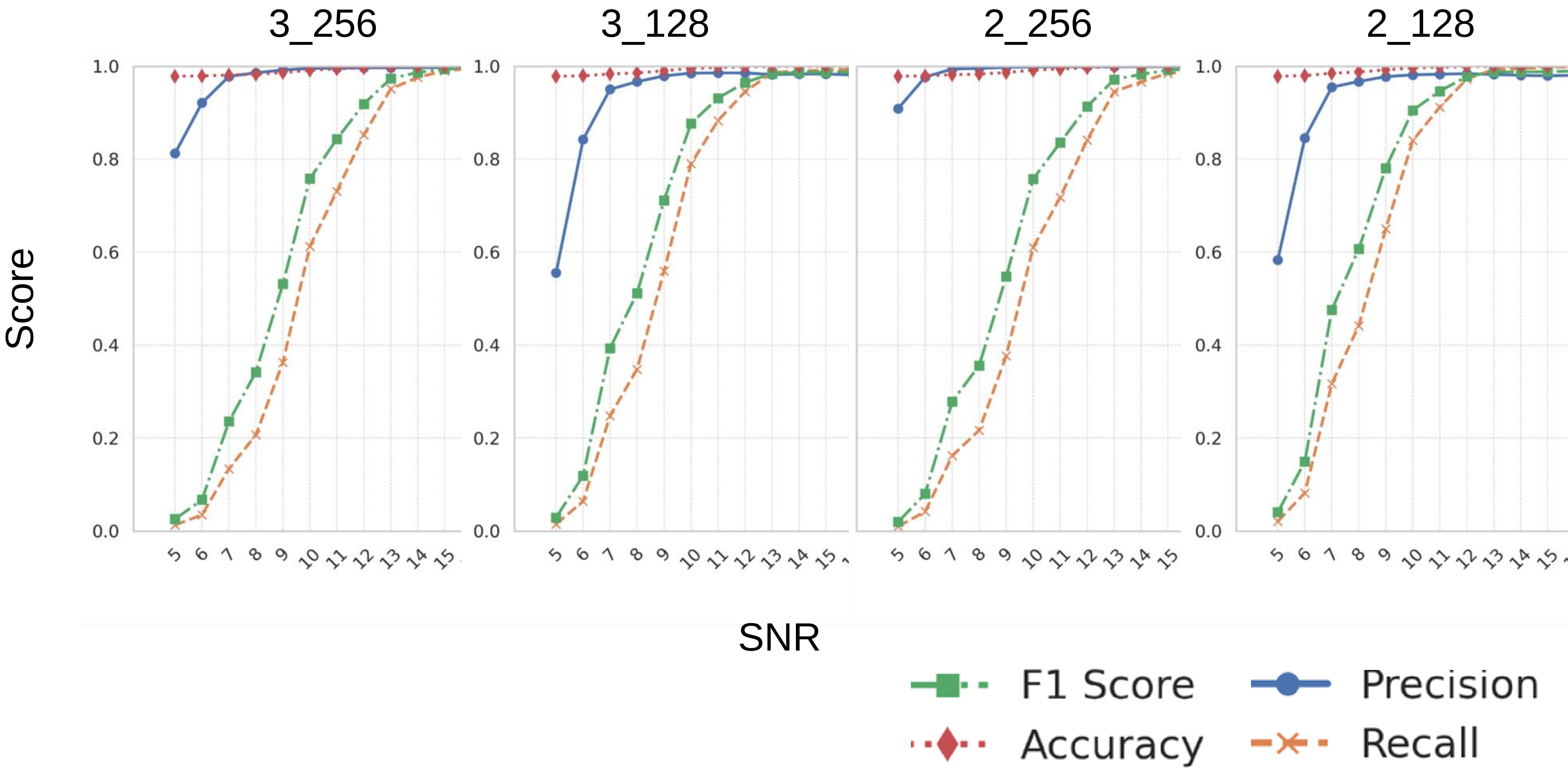
Comparison with TransientX Results



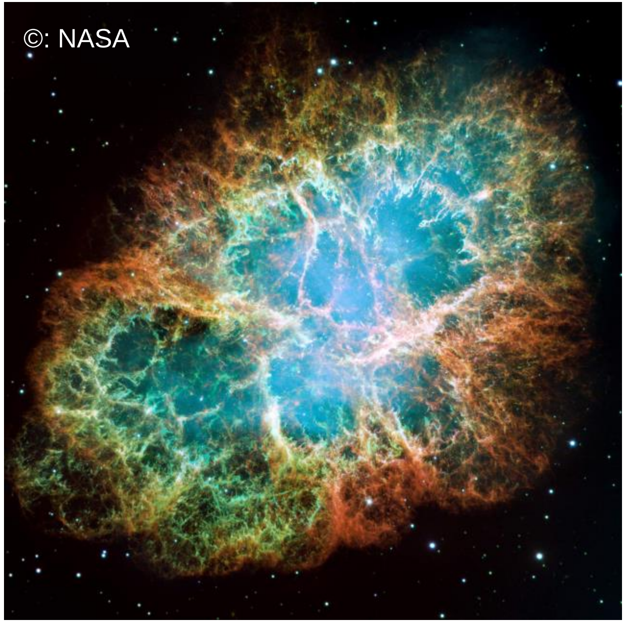
Results of Tests with Real Data

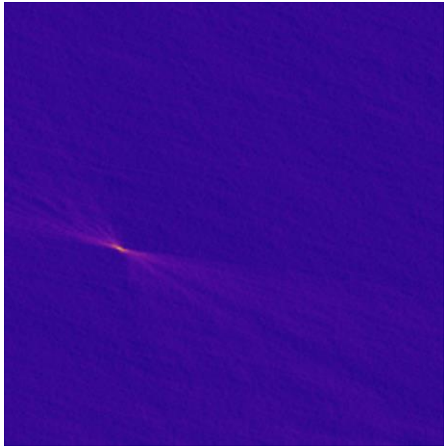
Model name	**%**	**Model did not find**	**Model found additionally**	**Unique for the model**
2_256	**93.5**	22	210 / **110**	25
2_128	**94.4**	19	212 / **90**	14
2_64	**94.4**	19	274 / **99**	11
2_32	**75.2**	87	164 / **33**	1
3_256	**94.6**	18	125 / **85**	11
3_128	**94.6**	18	97 / **66**	5
3_64	**94.4**	19	150 / **68**	4
3_32	**85.6**	50	477 / **58**	7

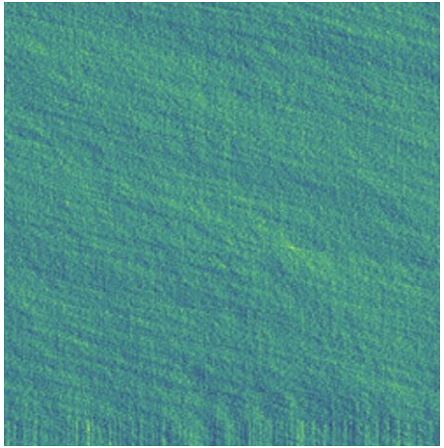
SNR Sensitivity

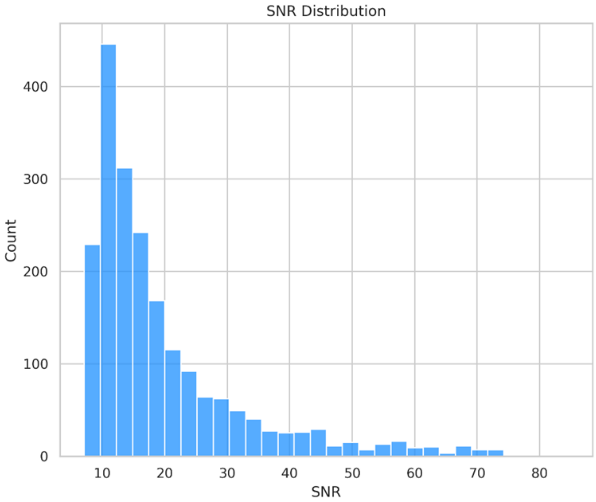


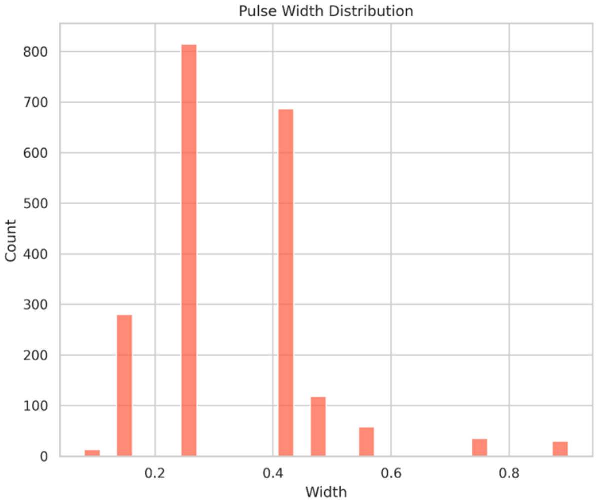
Testing with New Data





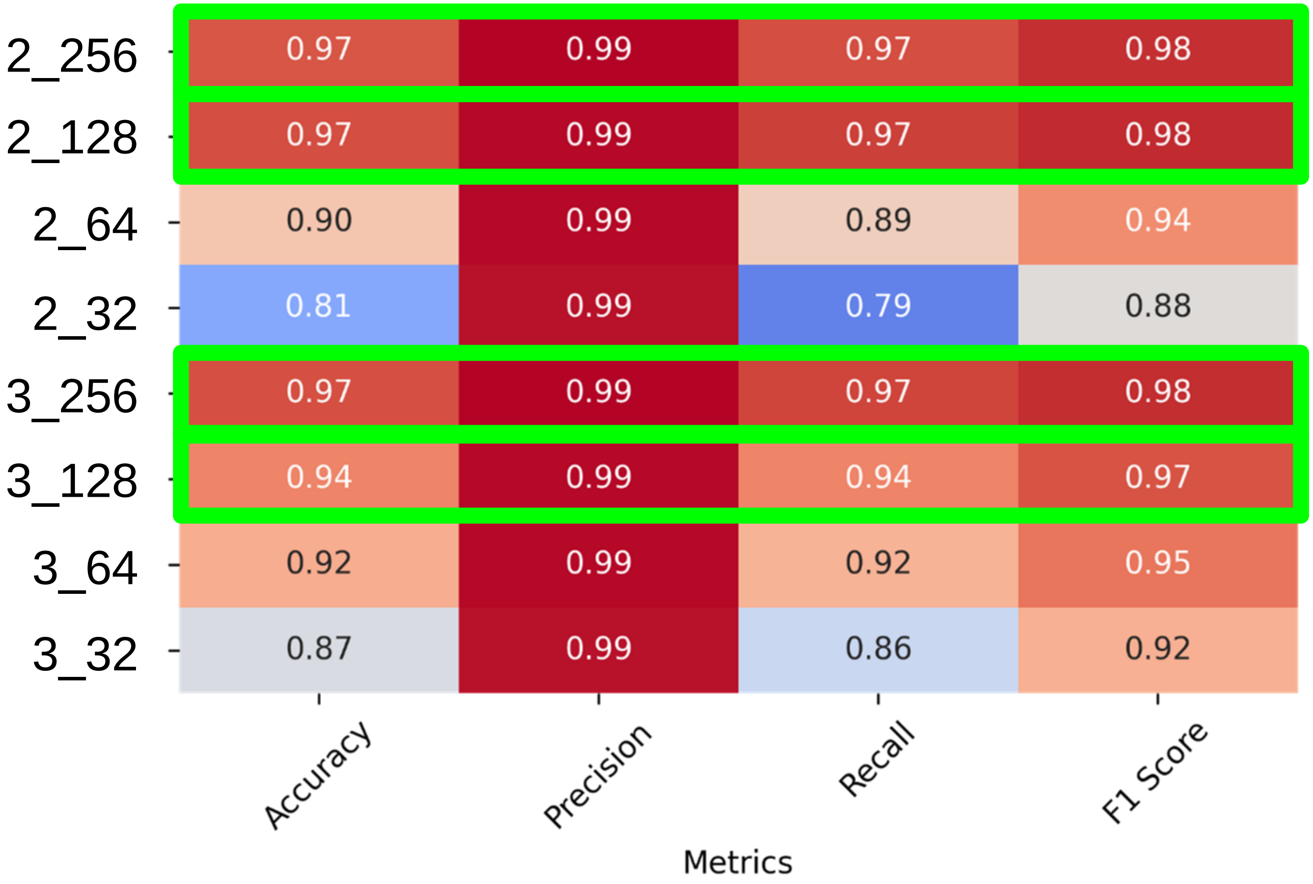




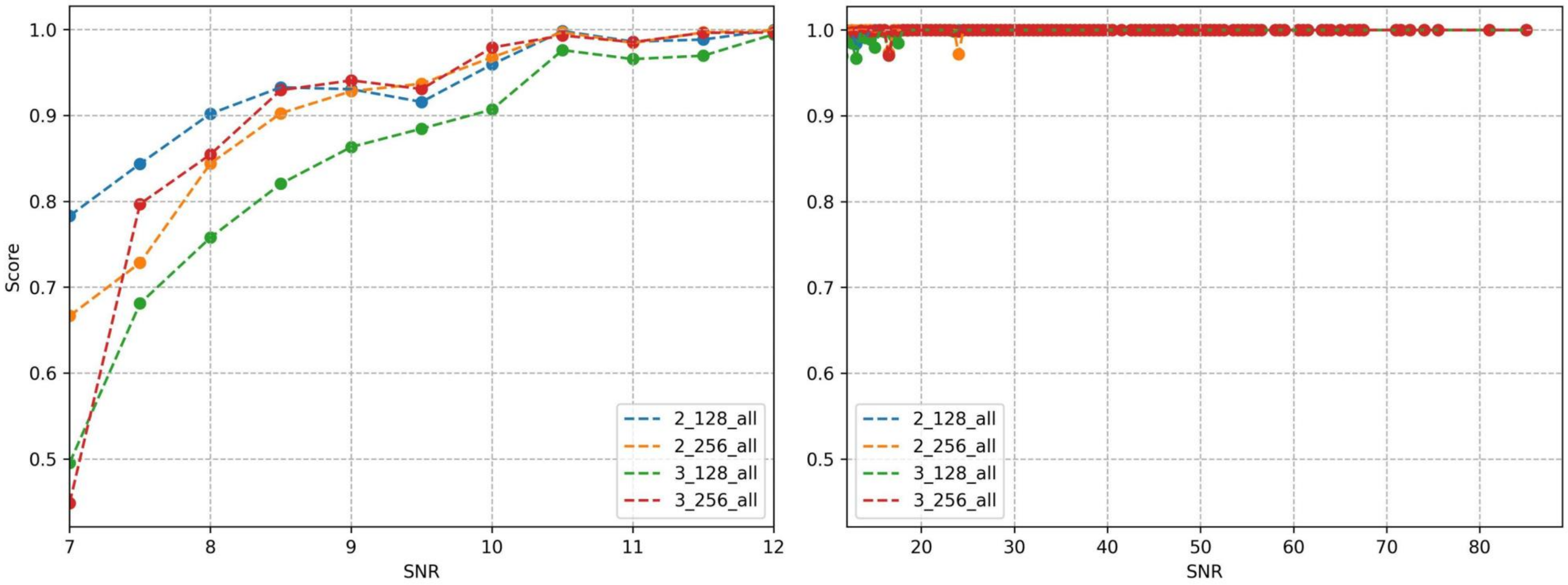


Object: **Crab pulsar (B0531+21)**
Data: **2024-06-21**
Time resolution: **81.92 μs**
Telescope: **Effelsberg**
Bandwight: **1281 - 1600 MHz**
Duration: **30 minutes**

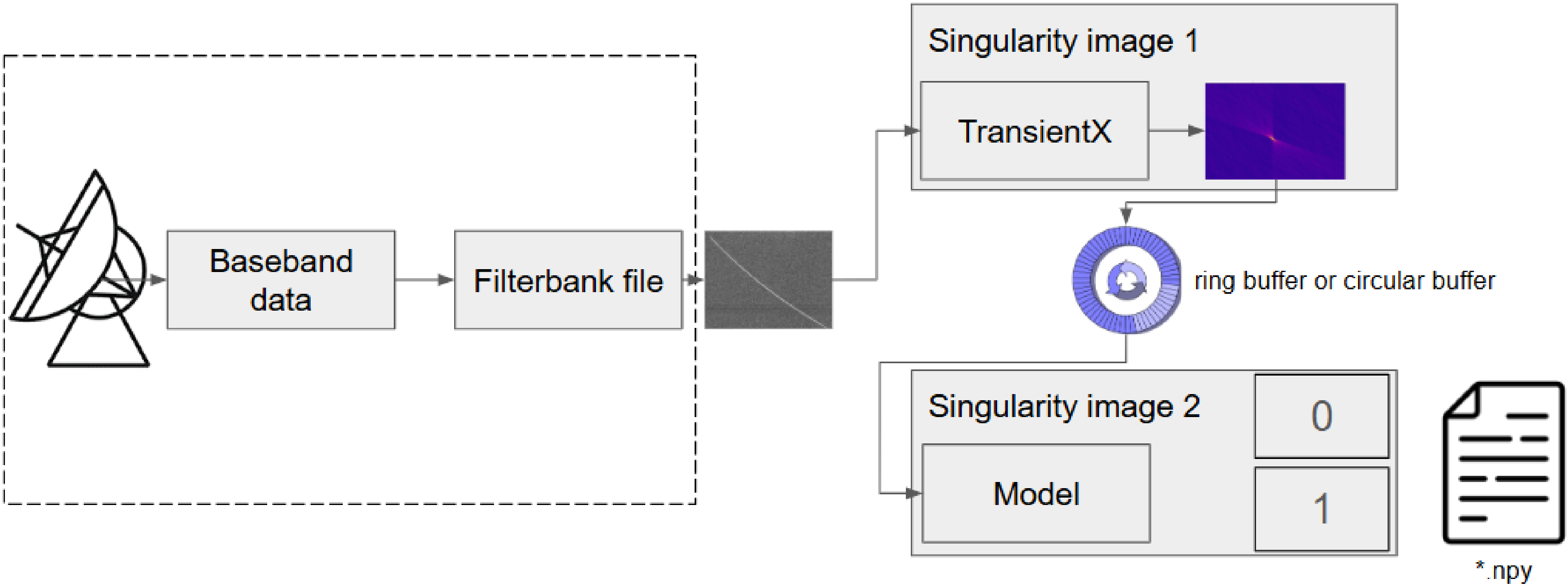
Key Metrics for the Models on New Data



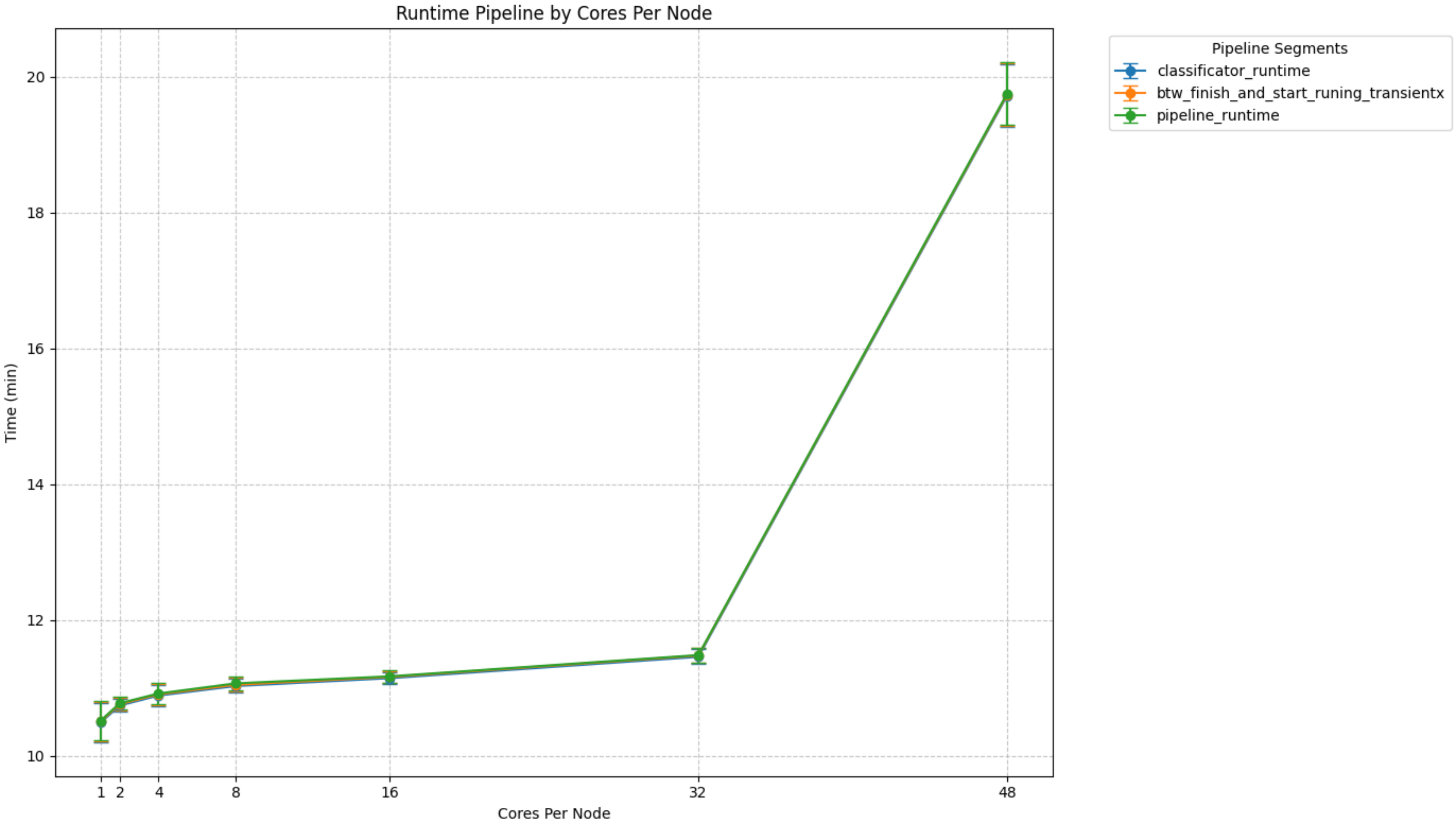
SNR Sensitivity on the New Data



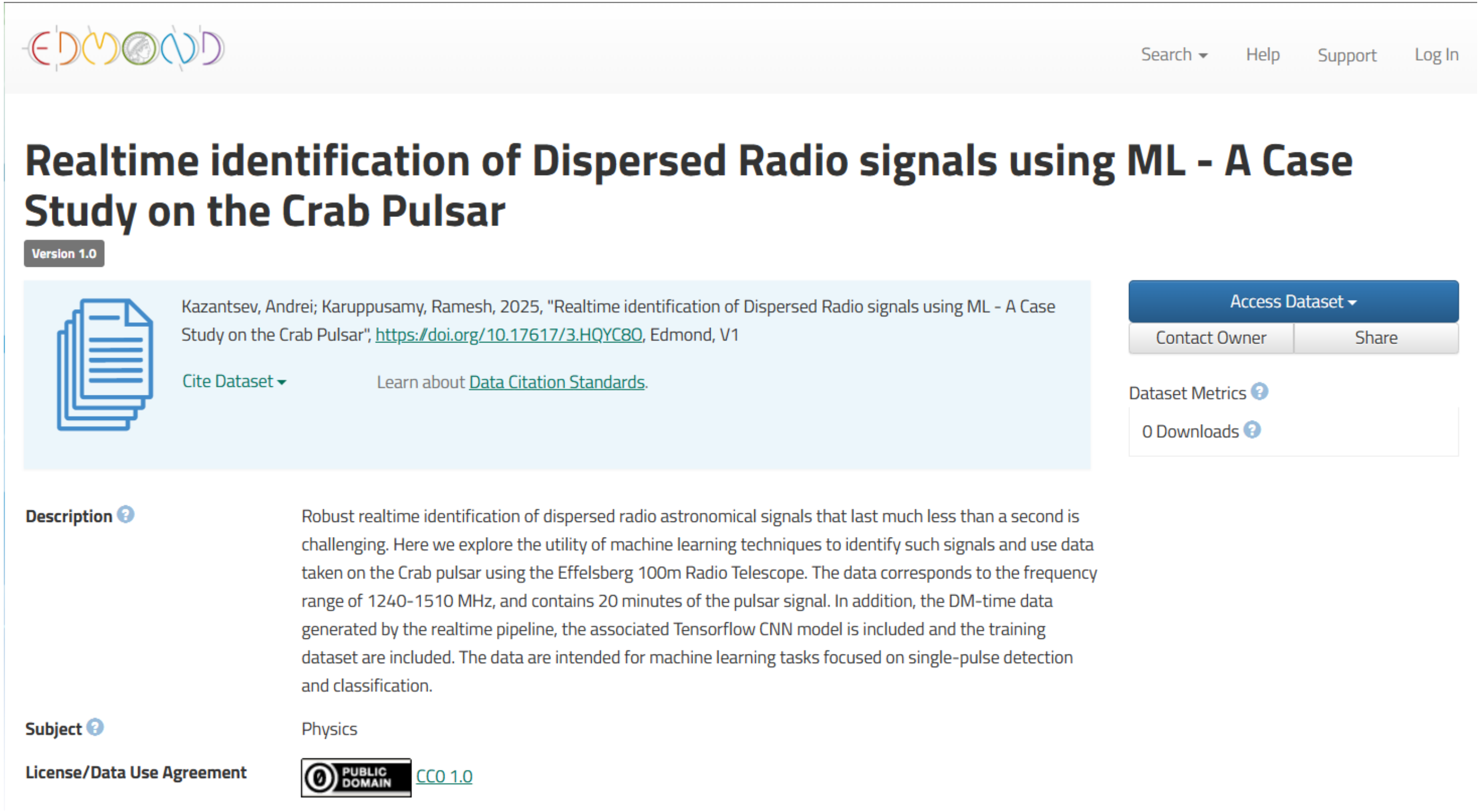
The Pipeline for Pulse Classification in a Filterbank File



Performance Metric of the Pipeline



Data Availability on the Edmond Platform

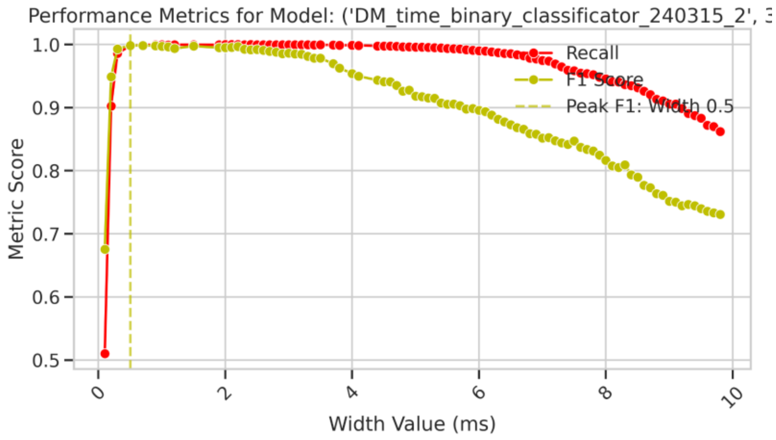


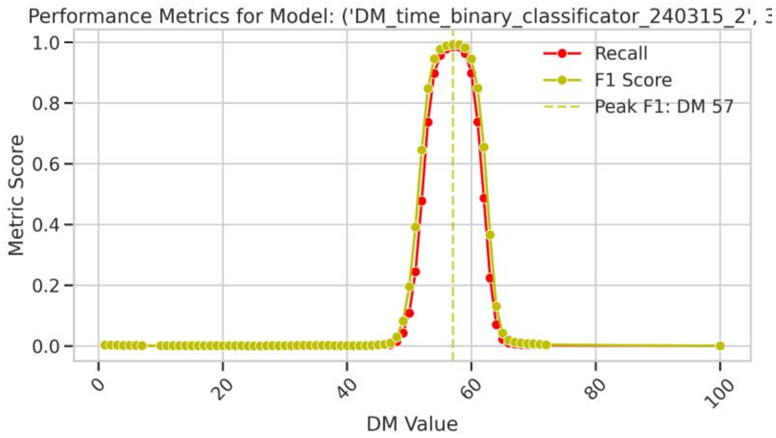
https://doi.org/10.17617/3.HQYC8O

Summary for the Current Stage

- The extremely minimalistic CNN model has been trained on DM-time images to detect individual pulses from the Crab pulsar with acceptable accuracy and sensitivity.
- The processing pipeline has been created to recognize individual pulses from Crab pulsar in filterbank files that includes TransientX and the trained model.
- The performance of the pipeline meets the set requirements.

Next Steps





Rishi Kumar

Questions? Suggestions?