

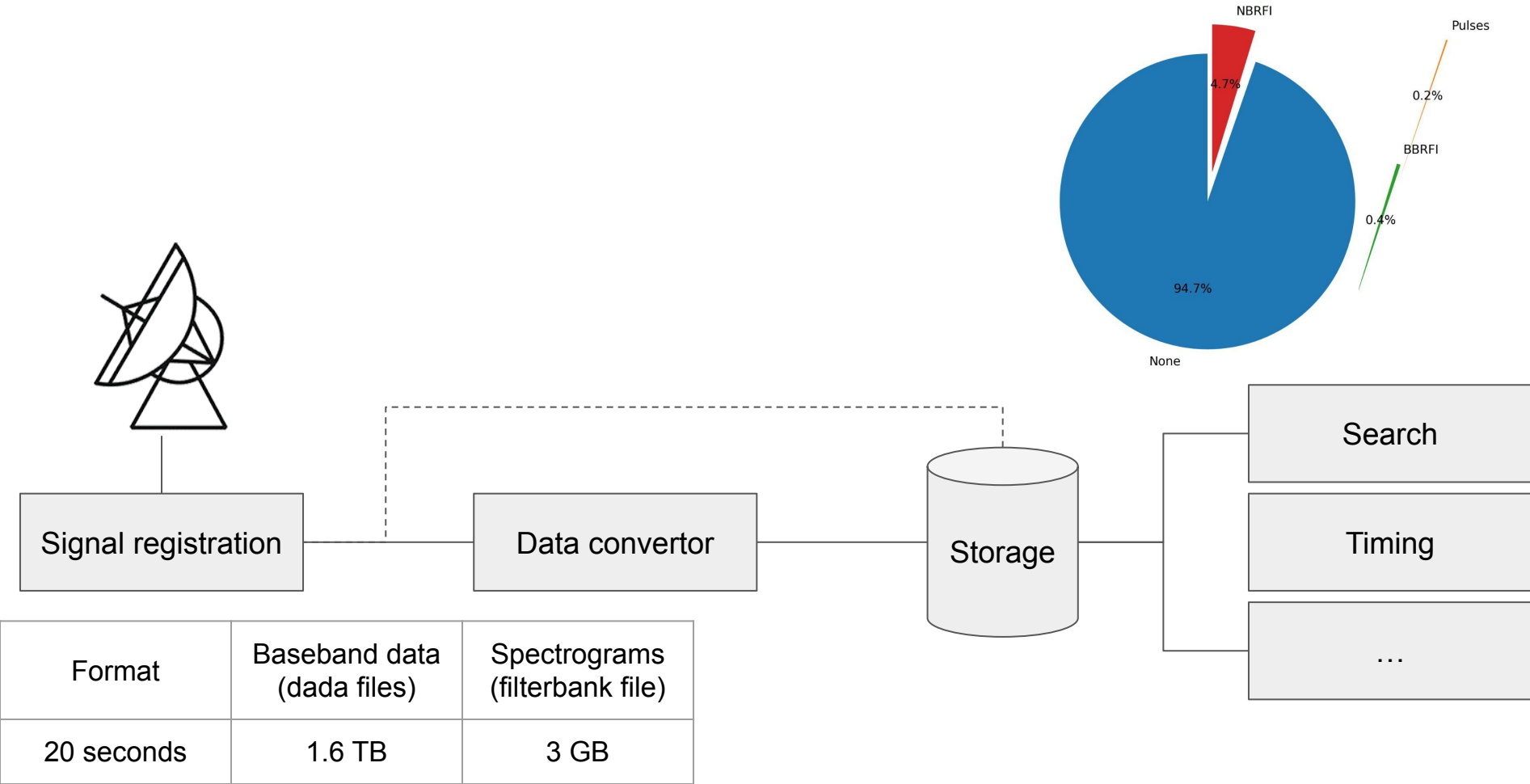
Deep Learning for real-time classification of astronomical radio signals

Andrei Kazantsev, Ramesh Karuppusamy, Yunpeng Men, Rishi
Kumar, Michael Kramer

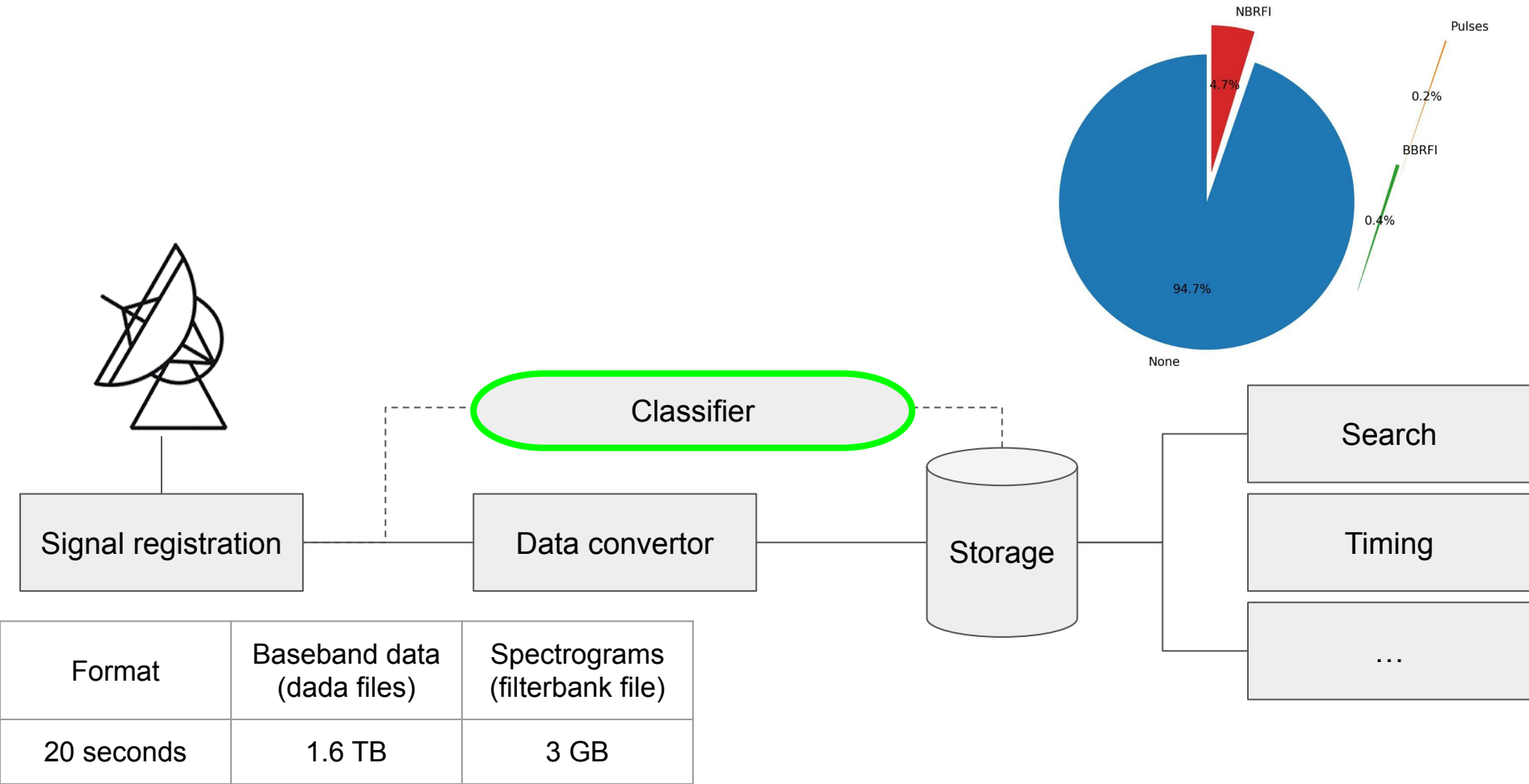
Max Planck Institute for Radio Astronomy
Bonn, Germany



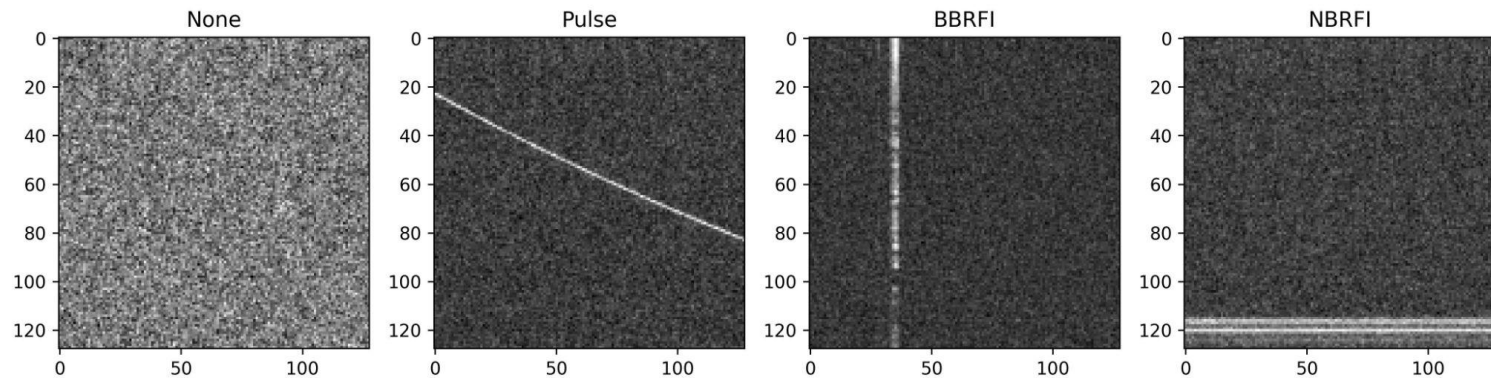
The main concept of the real-time classification



The main concept of the real-time classification



ML-PPA v.0.1



M ML-PPA

master ml-ppa History Find file Code

publicly available
Harry Enke authored 4 weeks ago cc433a57

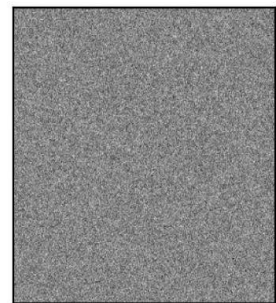
Name	Last commit	Last update
gitlab-profile	publicly available	4 weeks ago
pulsardt	publicly available	4 weeks ago
pulsardtpp	publicly available	4 weeks ago
pulsarrfi_gen	publicly available	4 weeks ago
pulsarrfi_nn	publicly available	4 weeks ago
tutorial_project	publicly available	4 weeks ago

Synthetic and real-based
data generation

Training and testing

https://gitlab-p4n.aip.de/punch_public/ml-ppa

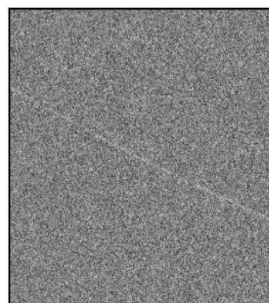
Limits of “spectrogram” approach



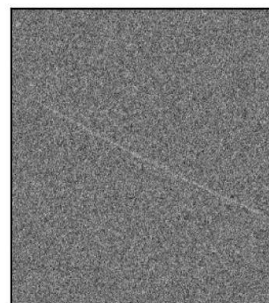
5.0



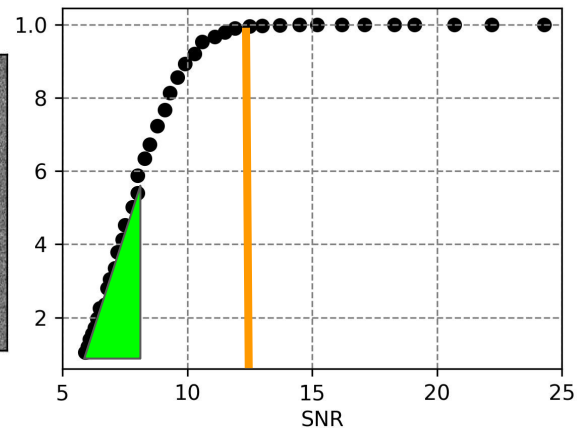
8.6



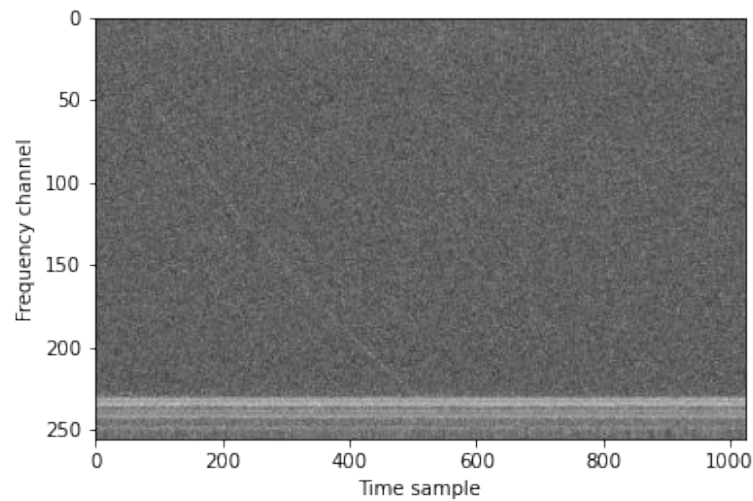
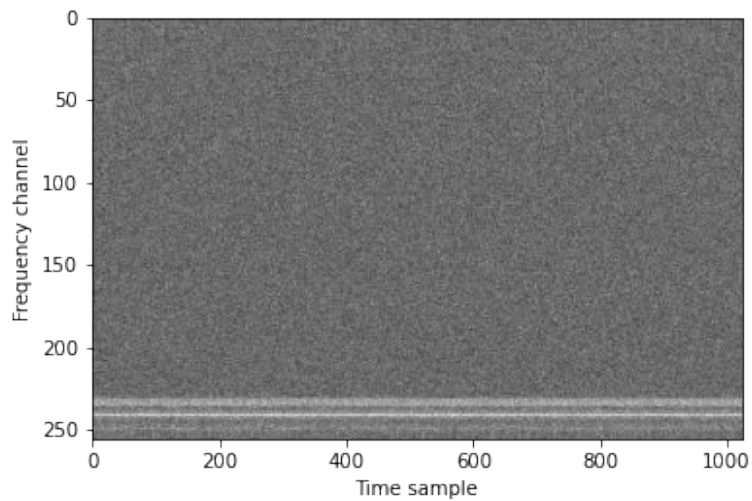
11.4



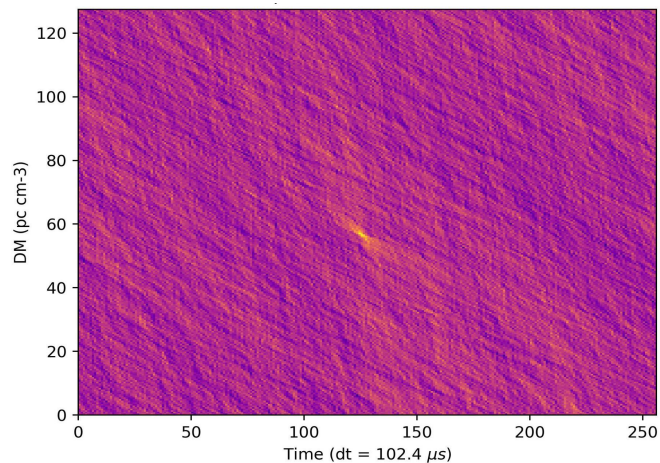
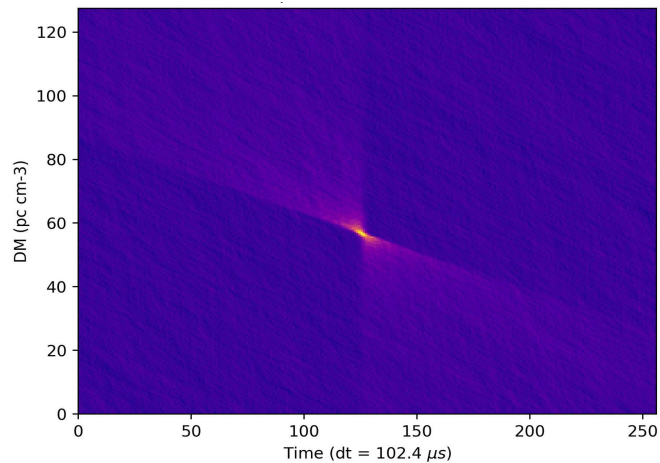
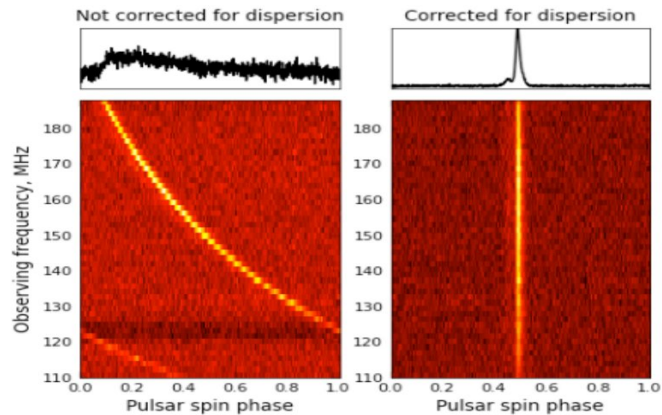
15.6



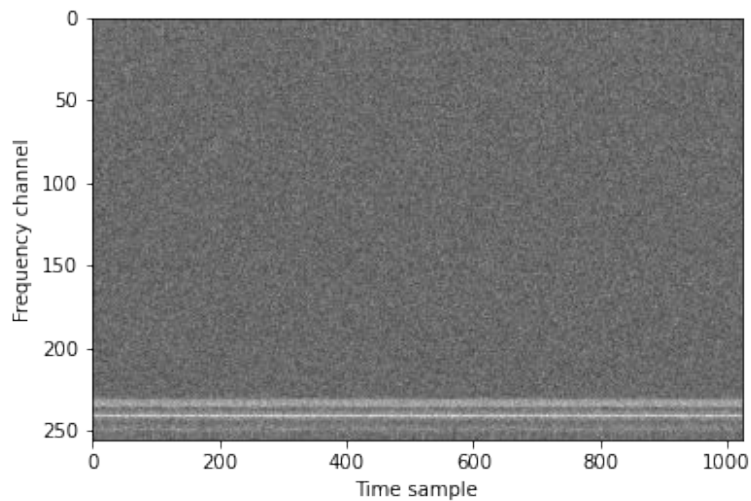
SNR



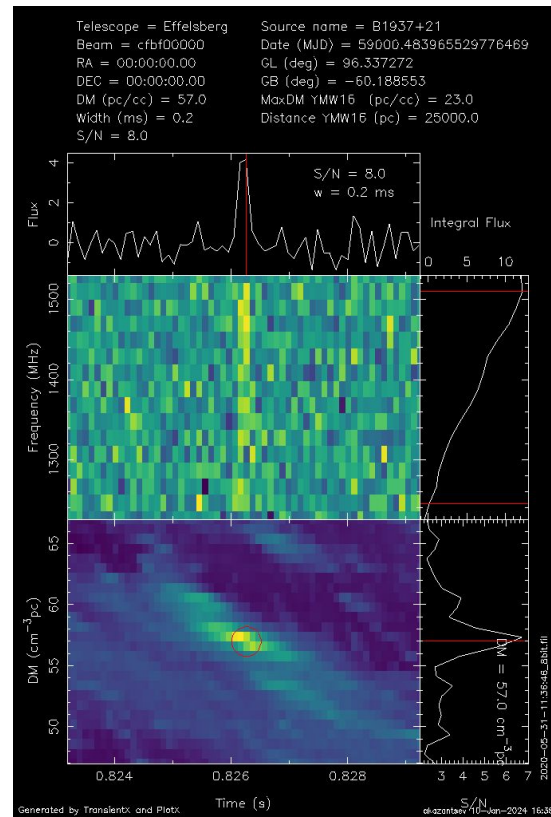
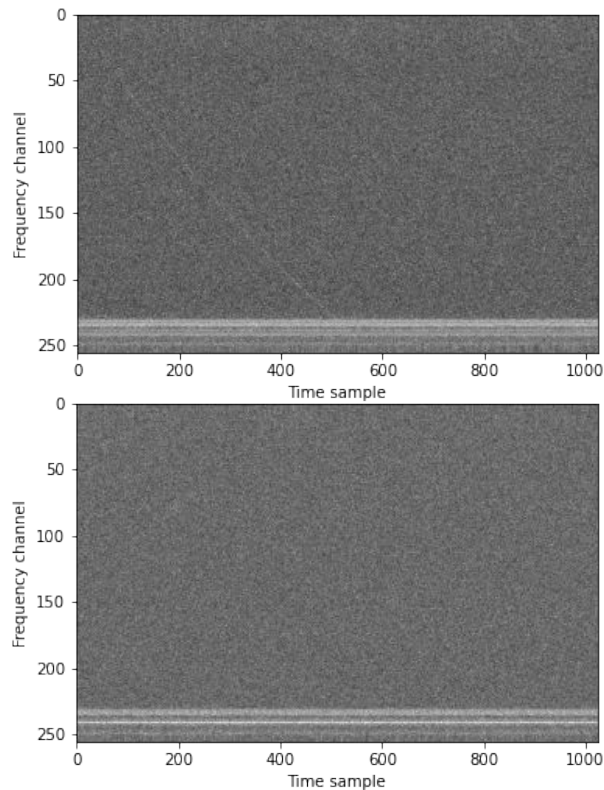
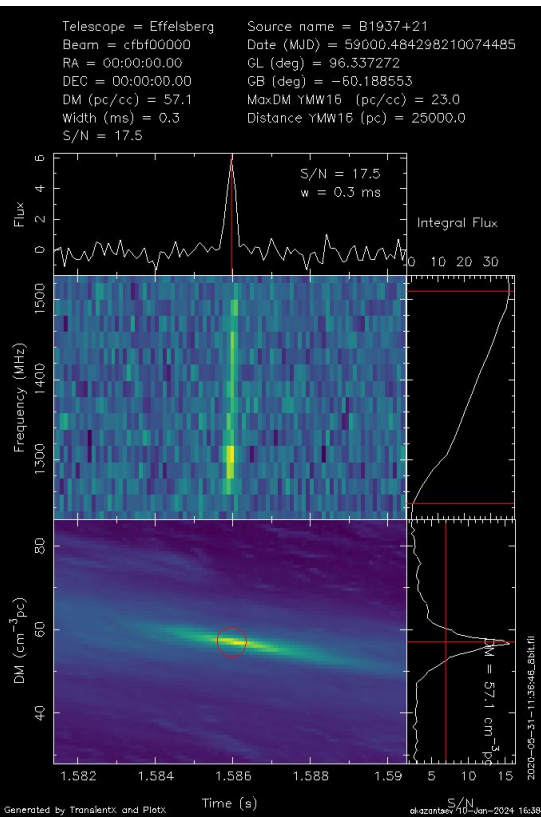
Spectrogram vs. DM-images



Credit: Anya Bilous



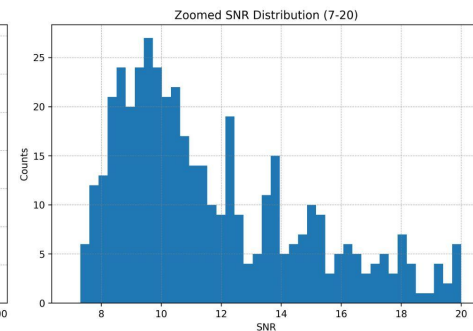
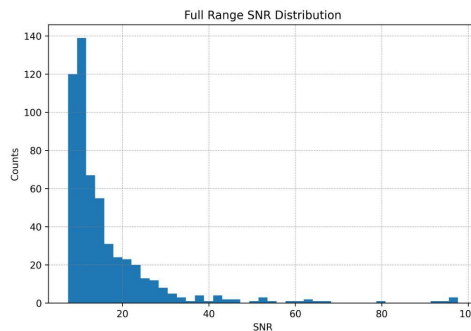
TransientX by Yunpeng Men



<https://github.com/ypmen/TransientX>

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.

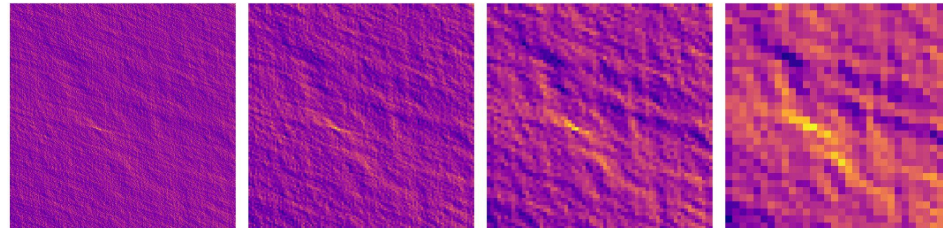


256x256

128x128

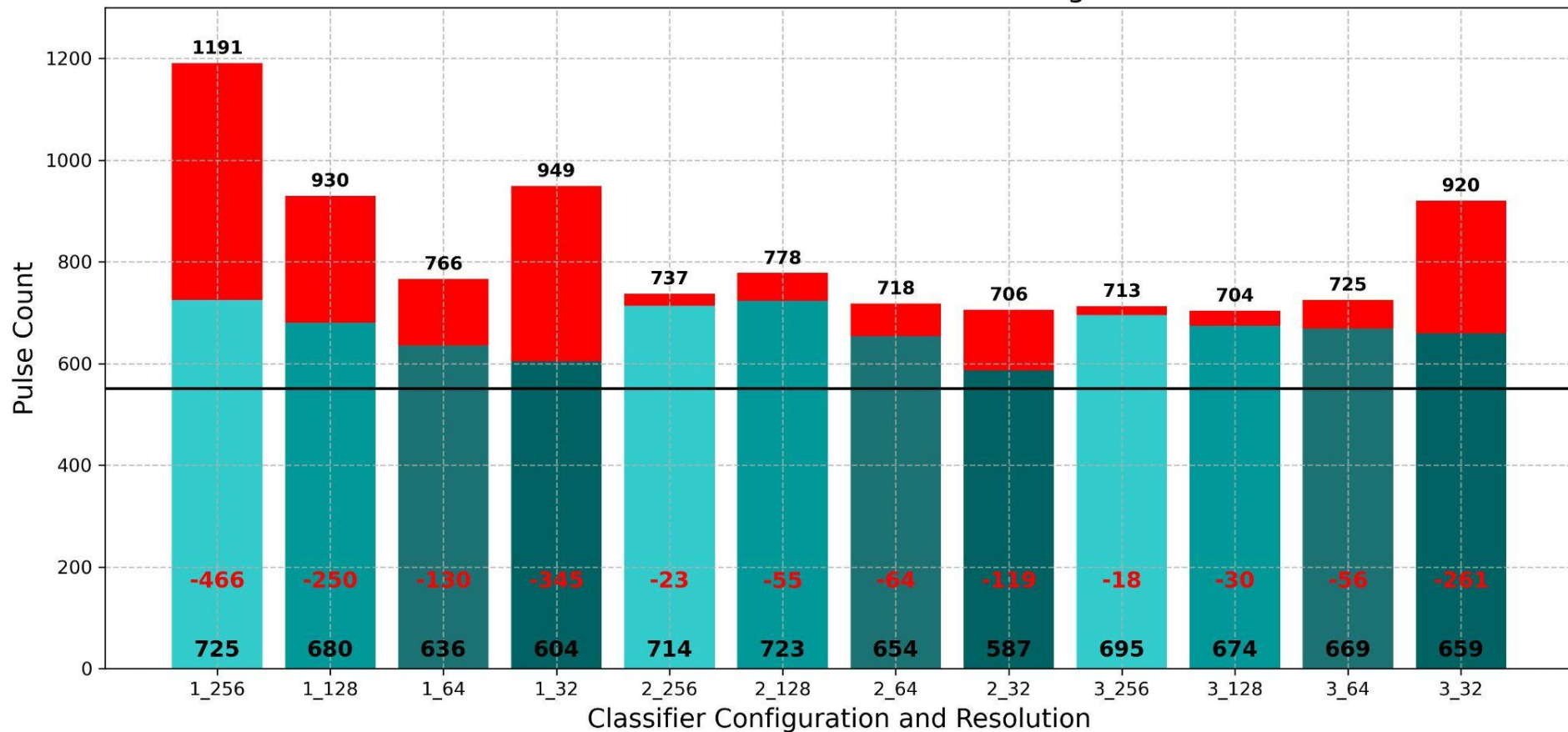
64x64

32x32



Results for test with real data

Pulse Count for Different Classifier Configurations



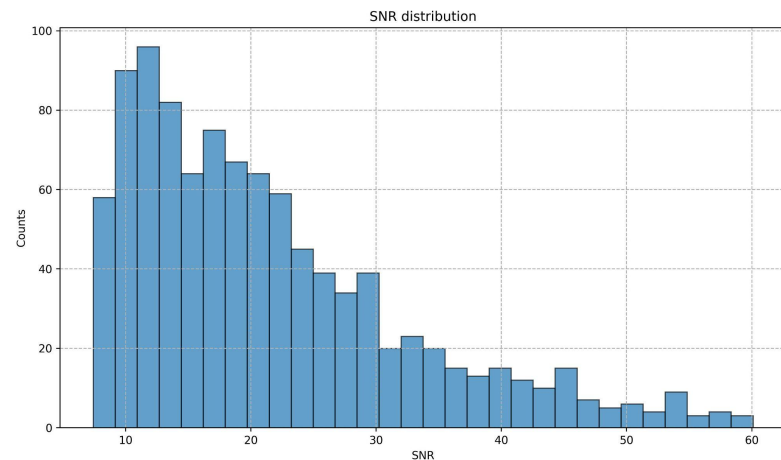
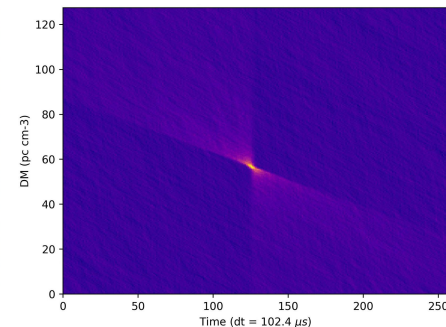
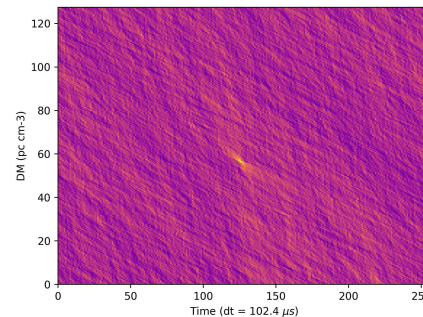
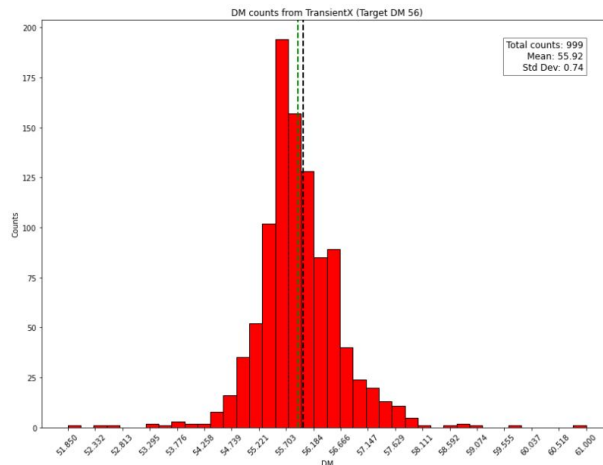
Results for test with real data

Model name	%	Model did not find	Model found additionally	Unique for the model
256_2	93.5	22	210 / 110	25
128_2	94.4	19	212 / 90	14
64_2	94.4	19	274 / 99	11
32_2	75.2	87	164 / 33	1
256_3	94.6	18	125 / 85	11
128_3	94.6	18	97 / 66	5
64_3	94.4	19	150 / 68	4
32_3	85.6	50	477 / 58	7

Totally synthetic filterbank file



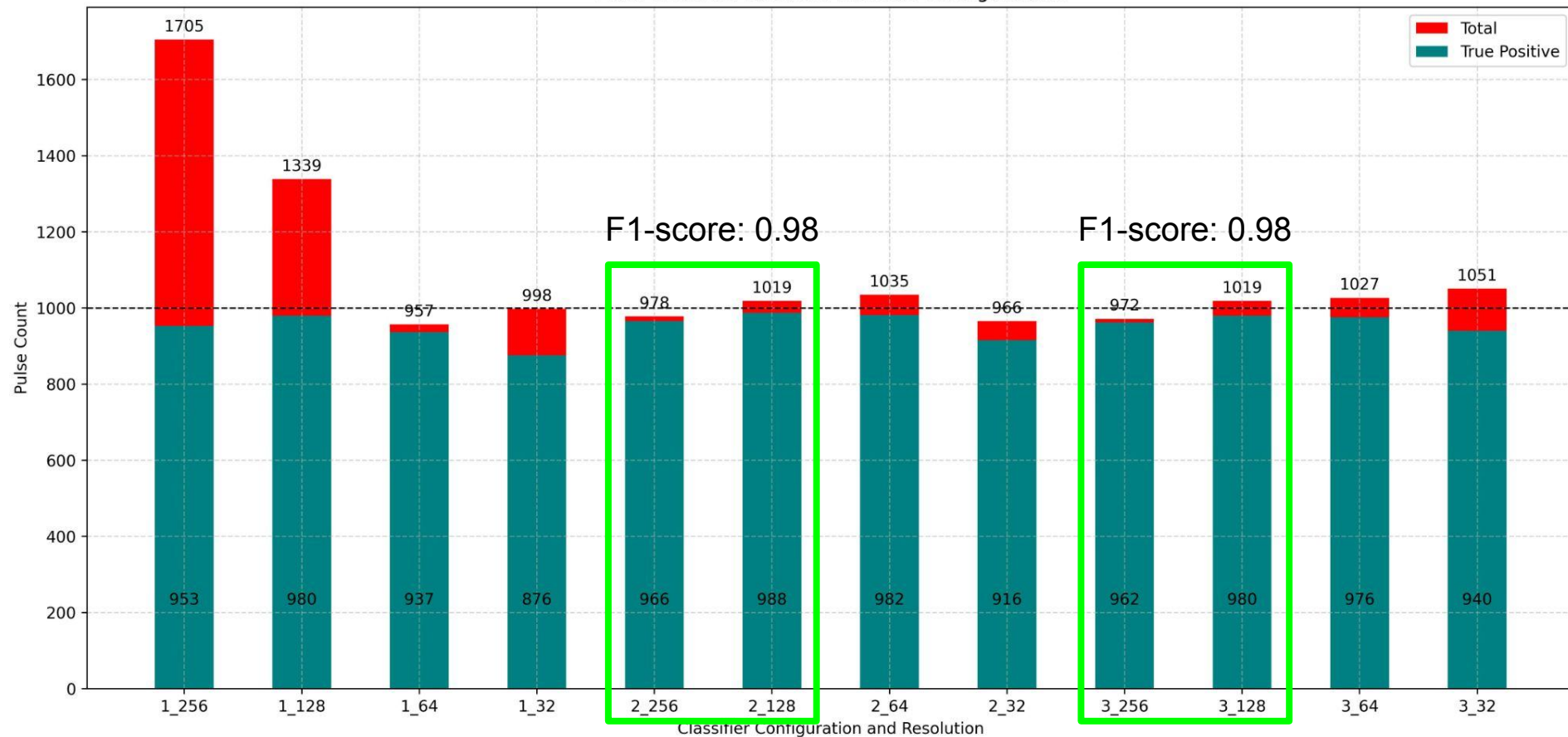
Rishi Kumar



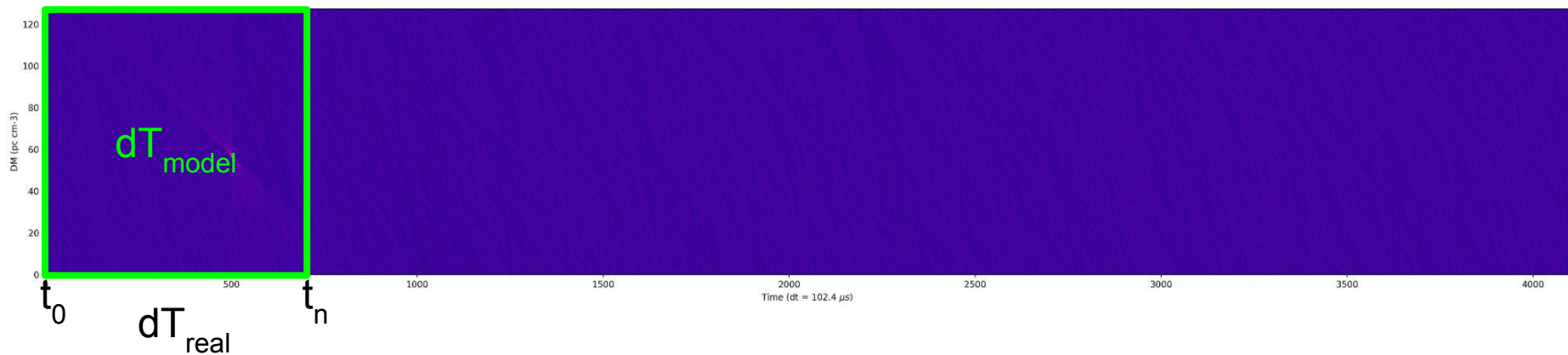
- Script to create a filterbank file with synthetic pulses has been developed (based on FRB Faker).
- Filterbank file with 1000 synthetic pulses has been created (pulses parameters - Crab pulsar, observation parameters - B0531+21_59000_48386.fil)
- First analysis of the filterbank file with TransientX has been made.

Results of search in fake filterbank

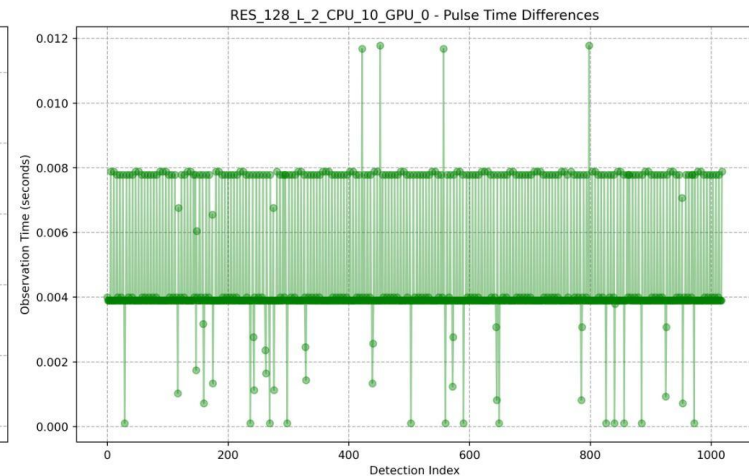
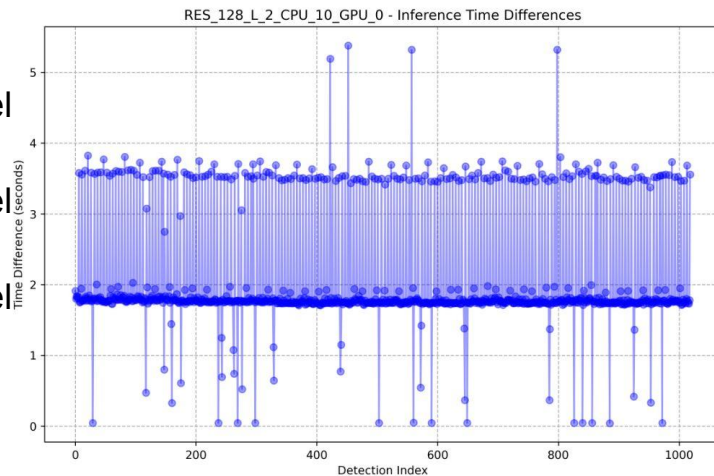
Pulse Count for Different Classifier Configurations



Inference time vs. Real-time

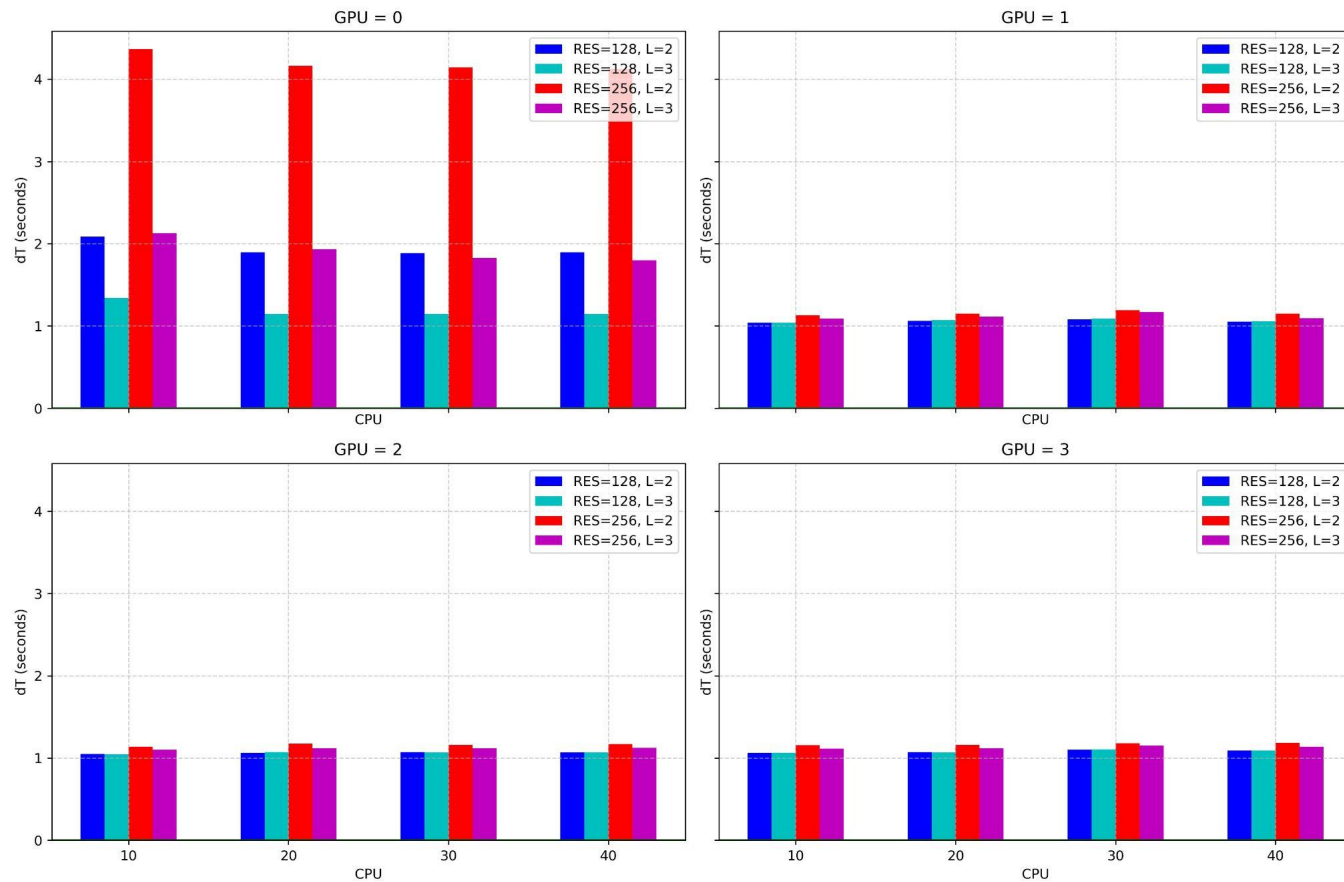


1. $dT_{\text{real}} < dT_{\text{model}}$
2. $dT_{\text{real}} = dT_{\text{model}}$
3. $dT_{\text{real}} > dT_{\text{model}}$



Inference time vs. Real-time

dT vs CPU for different GPU, RES, and L values (Histogram)



$$t_1 = 0.0001024 \text{ s}$$

$$t_{256} = 0.026 \text{ s}$$

tensorflow version: 2.4.3

CPU: Intel(R) Xeon(R) Platinum 8268
2.90GHz

GPU: Quadro RTX 6000

Our models are minimalistic

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 252, 252, 4)	104
max_pooling2d (MaxPooling2D)	(None, 126, 126, 4)	0
conv2d_1 (Conv2D)	(None, 122, 122, 8)	808
max_pooling2d_1 (MaxPooling2D)	(None, 61, 61, 8)	0
conv2d_2 (Conv2D)	(None, 57, 57, 12)	2412
flatten (Flatten)	(None, 38988)	0
dense (Dense)	(None, 256)	9,981,184
dense_1 (Dense)	(None, 2)	514
Total params:		9,985,022
Trainable params:		9,985,022
Non-trainable params:		0

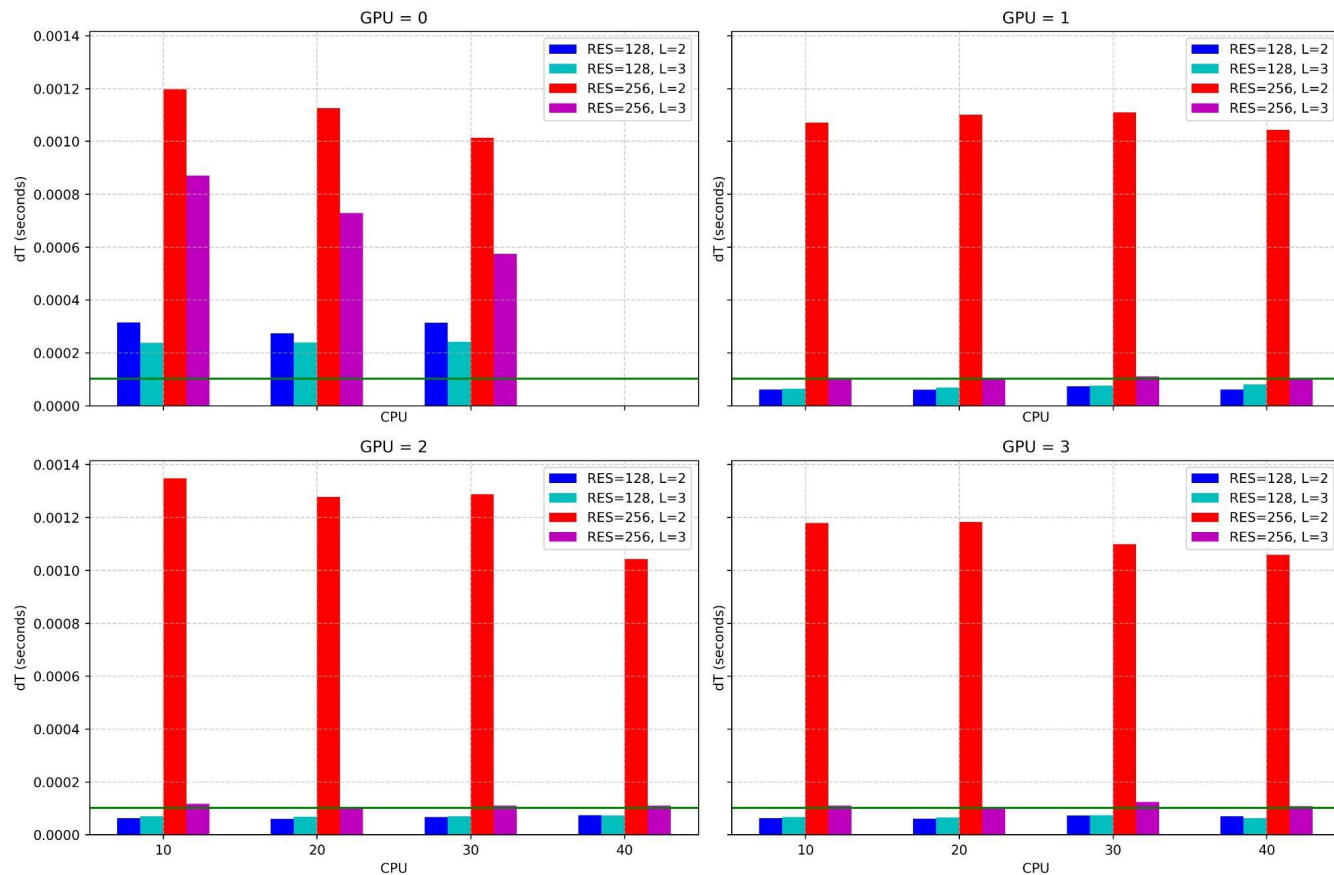
predict(element) vs predict(datasets)

```
model = load_model(...)
file = open(..., 'a')
file.write(f"{datetime.now()}\t{0}\n")
for i, element in enumerate(datasets):
    prediction = model.predict(element)
    if prediction == 1:
        file.write(f"{datetime.now()}\t{i + 1}\n")
file.close()
```

```
model = load_model(...)
with open(..., 'w') as file:
    start_time = datetime.now()
    predictions_for_the_models = model.predict(datasets)
    finish_time = datetime.now()
    time_difference = (finish_time - start_time).total_seconds()
    file.write(f"{time_difference/datasets.shape[0]}\n")
```

Inference time vs. Real-time

dT vs CPU for different GPU, RES, and L values (Histogram)



$$t_1 = 0.0001024 \text{ s}$$

$$t_{256} = 0.026 \text{ s}$$

tensorflow version: 2.4.3

CPU: Intel(R) Xeon(R) Platinum 8268
2.90GHz

GPU: Quadro RTX 6000