Deep-Learning Inference of Rotational Core-Collapse Supernovae with Numerically-Generated Gravitational-Wave Signals

Solange Nunes

Universitat de València

Universidade do Minho

2nd Workshop on Compact Objects, Valencia, June 22, 2022

Outline

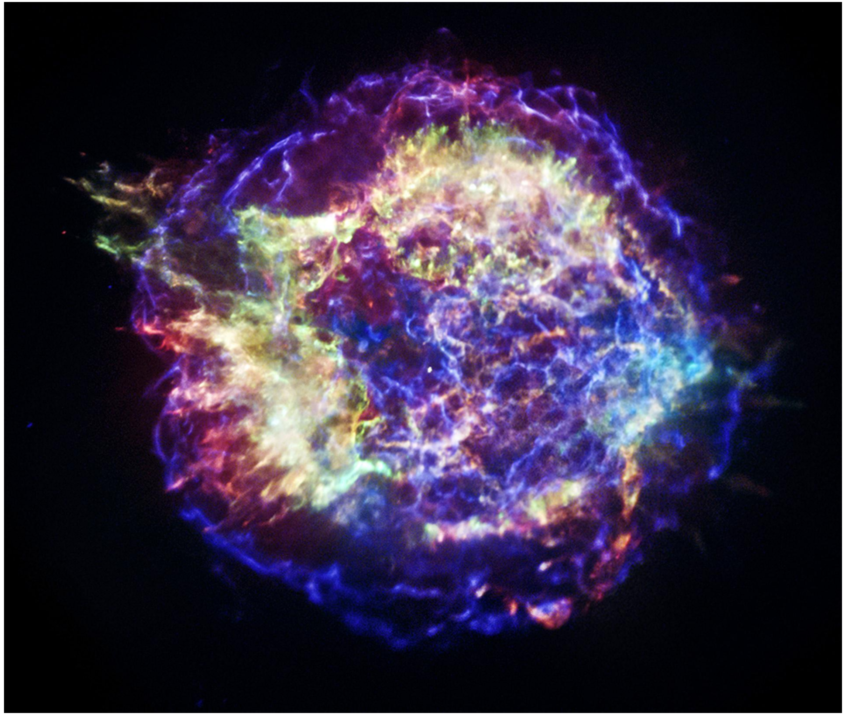
- Rotational Core-Collapse Supernovae
- Deep-Learning
- Spectrograms vs TimeSeries

- Dataset Construction
- Results:
  - Classification
  - Parameter Inference

Rotational Core-Collapse Supernovae (CCSN)

- Gravitational collapse of the core of massive stars and the subsequent explosion of such stars as supernovae.

- May provide valuable information about the physical processes operating during the gravitational collapse of the iron cores of massive stars.



Deep-Learning

For Classification and Regression:

Residual Convolutional Neural Networks (ResCNN)

Neurocomputing 367 (2019) 39–45

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

ELSEVIER

Integration of residual network and convolutional neural network along with various activation functions and global pooling for time series classification

Check for updates

Xiaowu Zou[a], Zidong Wang[b], Qi Li[a], Weiguo Sheng[a,*]

[a] Department of Computer Science, Hangzhou Normal University, Hangzhou, PR China
[b] Department of Computer Science, Brunel University London, Uxbridge, Middlesex UB8 3PH, UK

A R T I C L E I N F O

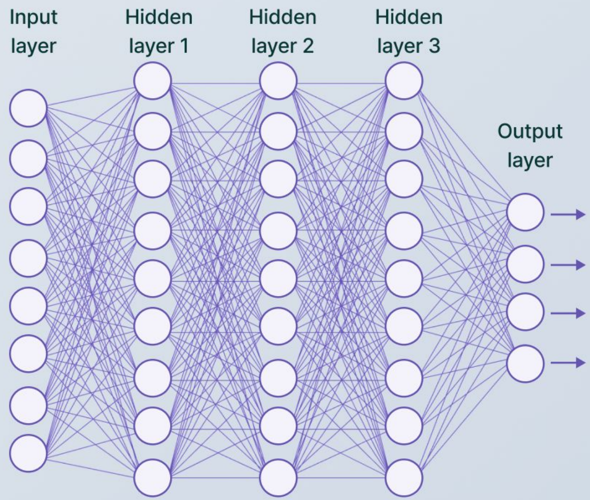
Article history:
Received 27 May 2019
Revised 21 July 2019
Accepted 8 August 2019
Available online 9 August 2019

Communicated by Steven Hoi

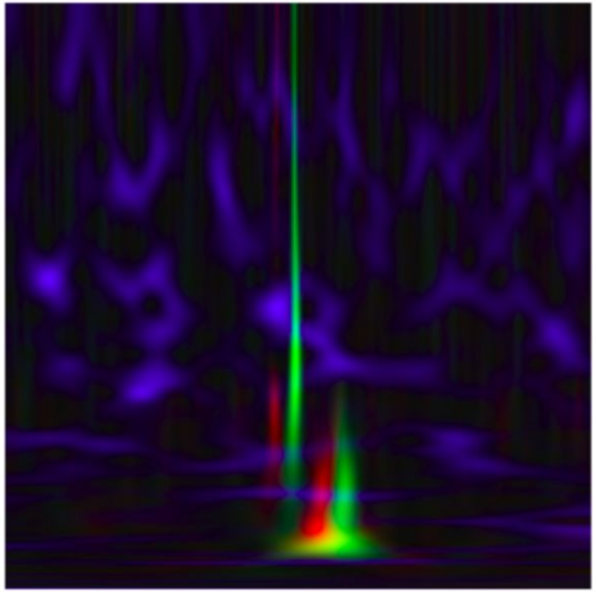
Keywords:
Time series classification
Residual network
Convolutional neural network
Deep learning

A B S T R A C T

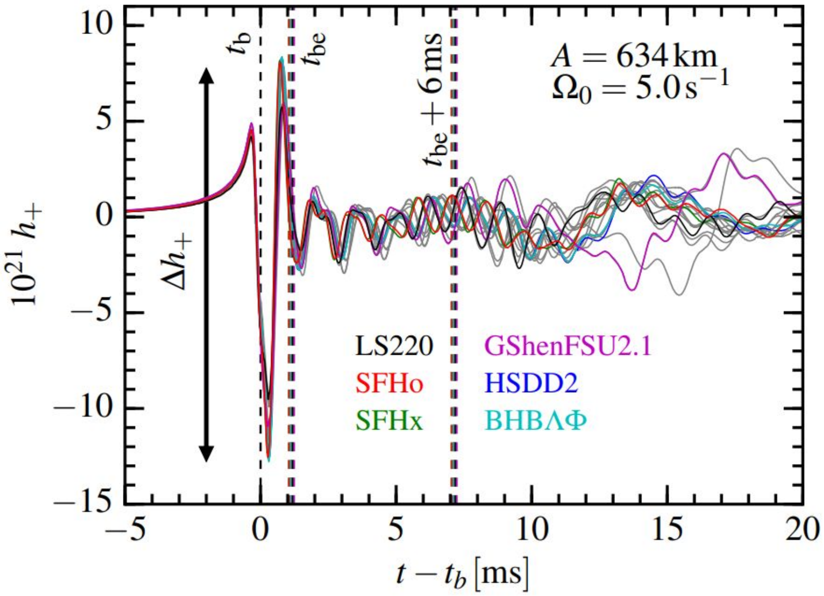
In this paper, we devise a hybrid scheme, which integrates residual network with convolutional neural network, for time series classification. In the devised method, the architecture of network is constructed by facilitating a residual learning block at the first three convolutional layers to combine the strength of both methods. Further, different activation functions are used in different layers to achieve a decent abstraction. Additionally, to alleviate overfitting, the pooling operation is removed and the features are fed into a global average pooling instead of a fully connected layer. The resulting scheme requires no heavy preprocessing of raw data or feature crafting, thus could be easily deployed. To evaluate our method, we test it on 44 benchmark datasets and compare its performance with related methods. The results show that our method can deliver competitive performance among state-of-the-art methods.



Spectrograms vs TimeSeries



Spectrogram of a CCSN signal.
[Figura 4 from **Gabriel Mas**, Trabajo de Fin de Grado]



Time-domain waveforms from CCSN
[Fig. 4 from **Richers et al** (1701.02752)]

Datasets Construction

- Selection of CCSN waveforms from the catalog developed by **Richers et al**:
  - ➢ $\omega_0 < 3.0$
  - ➢ $t_{collapse} < 1.0$ s
- Selection of parameter space

For each element of the Dataset

- Generation of a signal with random parameters;
- Projection of the signal into the detectores;
- Injection of the projected signals into the real noise of each detector;
- Whitening;

Equation of State Effects on Gravitational Waves from Rotating Core Collapse

Sherwood Richers,[1,2,3,4,*] Christian D. Ott,[1,5] Ernazar Abdikamalov,[6] Evan O'Connor,[7,8] and Chris Sullivan[9,10,11]

[1] *TAPIR, Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA, USA*
[2] *DOE Computational Science Graduate Fellow*
[3] *NSF Blue Waters Graduate Fellow*
[4] *Los Alamos National Lab, Los Alamos, NM, USA*
[5] *Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto, Japan*
[6] *Department of Physics, School of Science and Technology, Nazarbayev University, Astana 010000, Kazakhstan*
[7] *Department of Physics, North Carolina State University, Raleigh, NC, USA*
[8] *Hubble Fellow*
[9] *National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI, USA*
[10] *Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA*
[11] *Joint Institute for Nuclear Astrophysics: Center for the Evolution of the Elements, Michigan State University, East Lansing, MI, USA*

(Dated: January 10, 2017)

Gravitational waves (GWs) generated by axisymmetric rotating collapse, bounce, and early postbounce phases of a galactic core-collapse supernova will be detectable by current-generation gravitational wave observatories. Since these GWs are emitted from the quadrupole-deformed nuclear-density core, they may encode information on the uncertain nuclear equation of state (EOS). We examine the effects of the nuclear EOS on GWs from rotating core collapse and carry out 1824 axisymmetric general-relativistic hydrodynamic simulations that cover a parameter space of 98 different rotation profiles and 18 different EOS. We show that the bounce GW signal is largely independent of the EOS and sensitive primarily to the ratio of rotational to gravitational energy, $T/|W|$, and at high rotation rates, to the degree of differential rotation. The GW frequency ($f_{\rm peak} \sim 600-1000\,$Hz) of postbounce core oscillations shows stronger EOS dependence that can be parameterized by the core's EOS-dependent dynamical frequency $\sqrt{G\bar{\rho}_c}$. We find that the ratio of the peak frequency to the dynamical frequency $f_{\rm peak}/\sqrt{G\bar{\rho}_c}$ follows a universal trend that is obeyed by all EOS and

Classification Tests

GOAL: Separate signals from noise

Test 1:

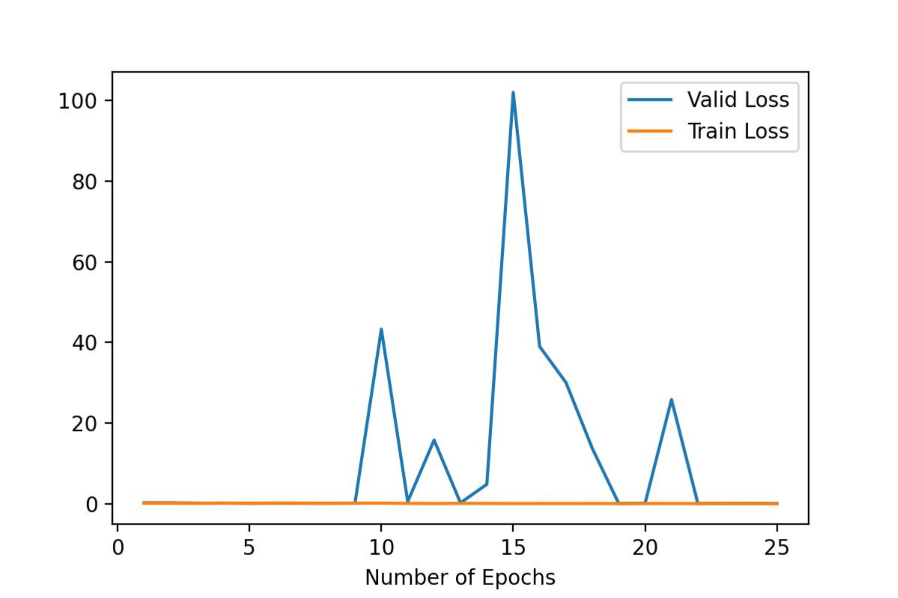
- 10k TimeSeries
- Fixed distance (20 kPc);
- Fixed sky position and polarization angle;
- Fixed inclination (π/2);
- Comparison with results from spectrograms;

Test 2:

- 10k TimeSeries
- distance between 5 and 20 kPc;
- Random sky position and polarization angle;
- Fixed inclination (π/2);

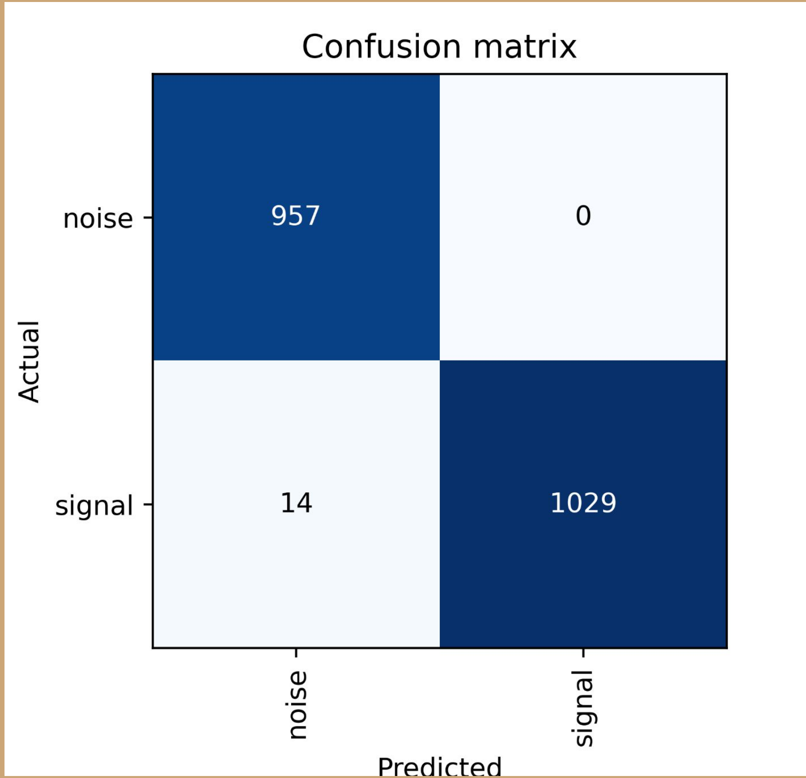
Classification Test 1

- Best model found at epoch 25 with valid loss value of 0.0295.
- Accuracy: 0.99



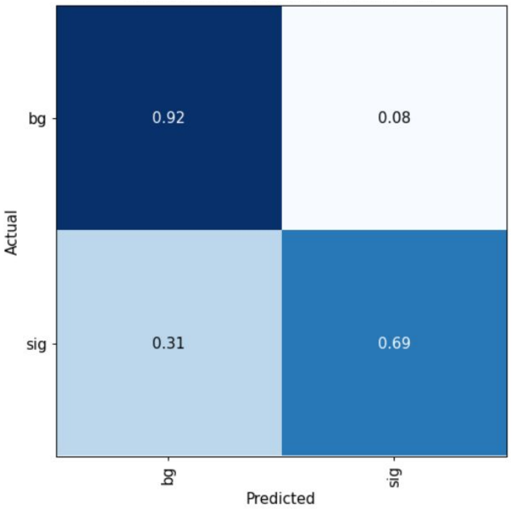
Classification Test 1

- No actual noise classified as signal;
- Only 1% of actual signals was predicted as noise;



Spectrograms vs Time Series

(a)



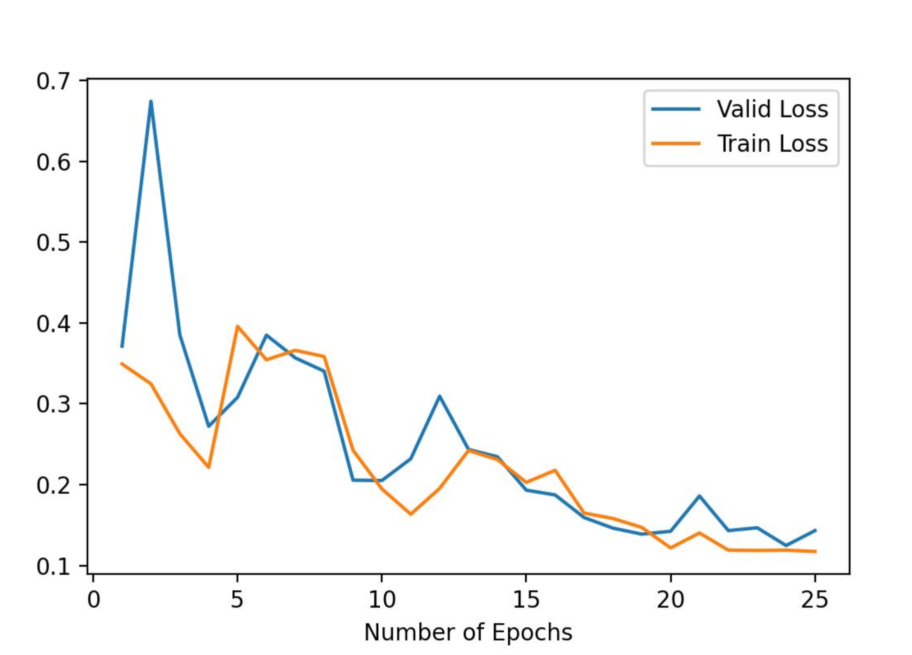
(b)

Actual Noise	1.00	0.00
Actual Signal	0.01	0.99
	Predicted Noise	Predicted Signal

(a) the confusion matrix obtain by Gabriel Mas with a similar dataset using Spectrograms and in (b) the confusion matrix obtain with a dataset of Time Series.
Differences: in (a) is used the noise from O2 and in (b) O3a; in (a) the signals have a window of 4 seconds and in (b) 1 second.

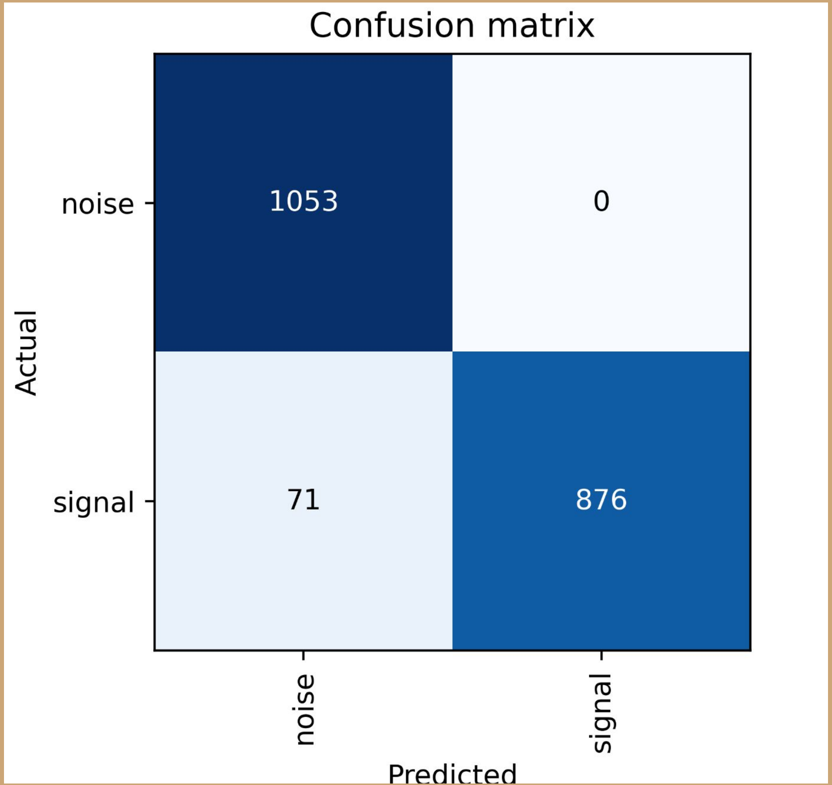
Classification Test 2

- Best model found at epoch 24 with valid loss value of 0.1247.
- Accuracy: 0.96



Classification Test 2

- No actual noise classified as signal;
- Only 7% of actual signals was predicted as noise;



Regression

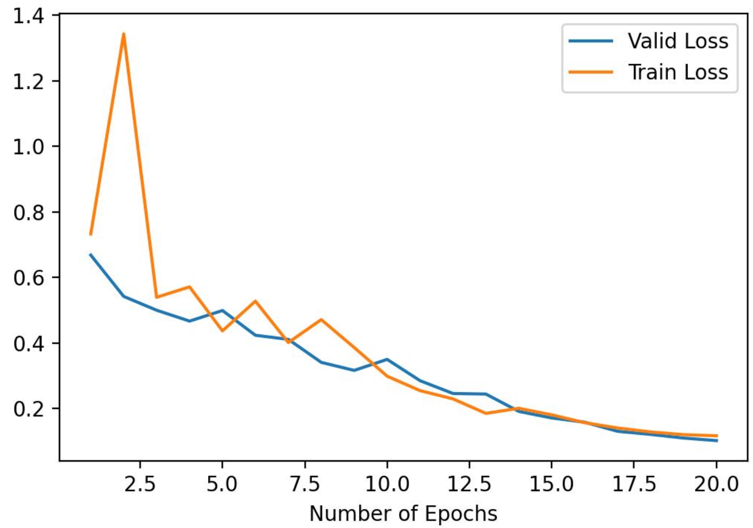
GOAL: Parameter Inference

- 10k TimeSeries;
- distance between 5 and 20 kPc;
- Random sky position;
- Fixed inclination (π/2);
- Inference:
  - ➢ Frequency at the peak of the signal, f_{peak}
  - ➢ Amplitude of the signal, Δh

Regression

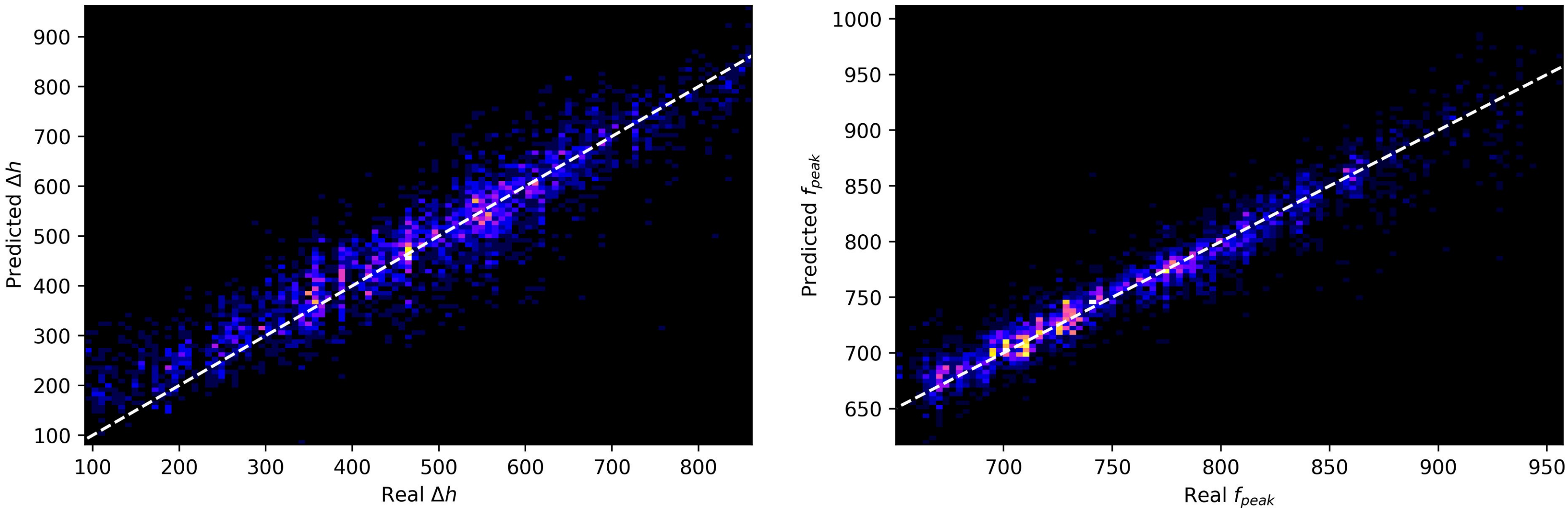
- Best model found at epoch 20 with valid loss value of 0.1164.

- RMSE: 0.34



Regression


Min
Max


Predicted Δh
Real Δh
100
200
300
400
500
600
700
800
900
Predicted f_{peak}
Real f_{peak}
650
700
750
800
850
900
950
1000

Conclusions

and

Next Steps

- Amazing results for Classifications
- Good results for Regression

What’s next?

- Larger Datasets;
- Inference on other parameters

Attachments

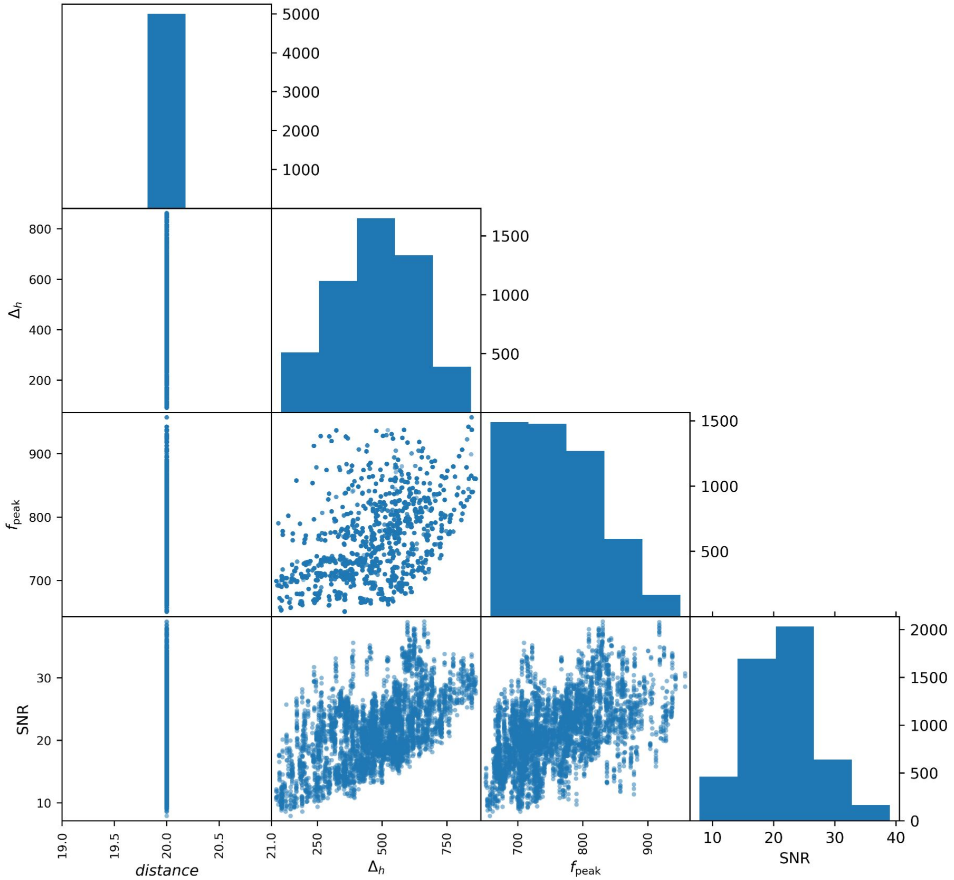
Classification Test 1

Dataset:

- 5k TimeSeries of noise;
- 5k TimeSeries of signals:
  - 1 second window
  - Sample rate: 4096Hz
  - 999 different waveforms with $w_0 < 3.0$
  - Distance: 20 kPc
  - Inclination = π/2
  - Declination, polarization and right ascension = 0

Network:

- Batch Size = 8
- Model: ResCNN(3,2)
- Weight decay: 1e-3
- Maximum learning rate: 0.5
- Monitoring: valid loss



Classification Test 2

Dataset:

- 5k TimeSeries of noise
- 5k TimeSeries of signals:
  - 1 second window
  - Sample rate: 4096Hz
  - 999 different waveforms with $w_0 < 3.0$
  - Distance: [5, 20] kPc
  - Inclination = π/2

Network:

- Batch size = 15
- Model: ResCNN(3,2)
- Weight decay: 1e-3
- Maximum learning rate: 0.05
- Monitoring: valid loss