Universidade do Minho

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

Solange Nunes

3rd Workshop on Compact Objects, Gravitational Waves and Deep Learning

Braga, September 23, 2022

Outline

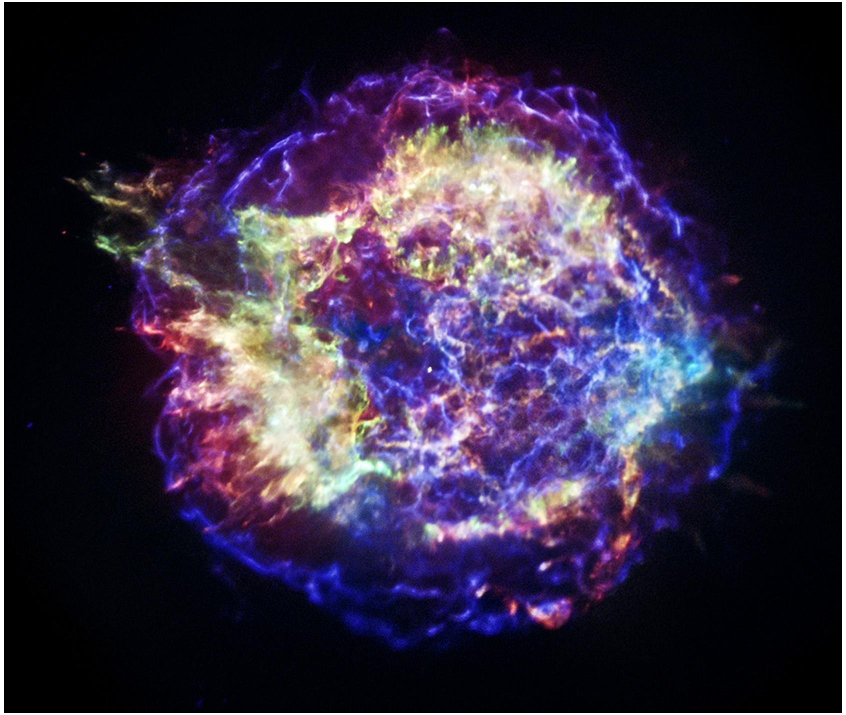
- Rotational Core-Collapse Supernovae
- Deep-Learning

- Dataset
- Results:
  - Classification
  - Regression

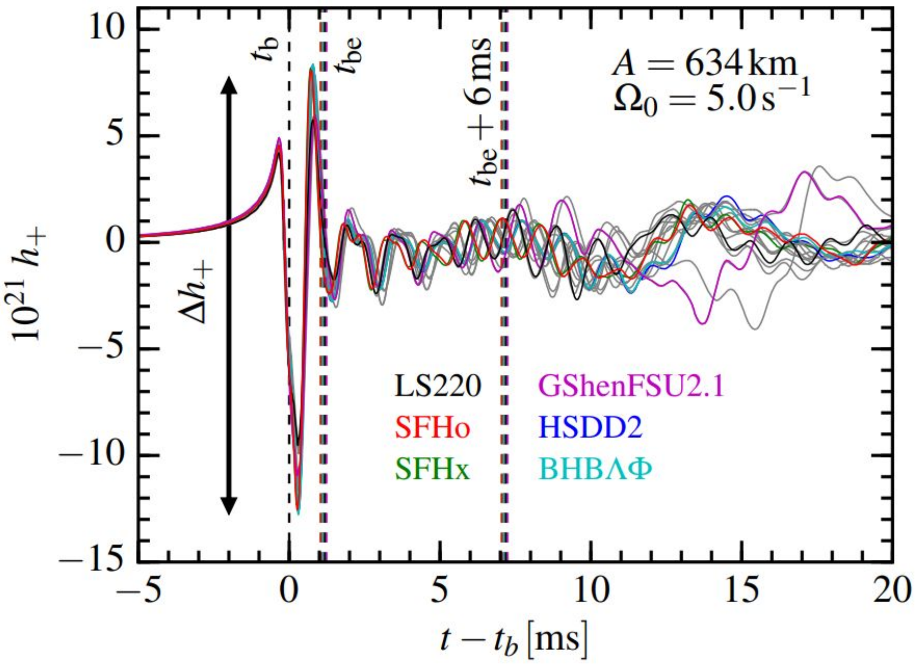
Rotational Core-Collapse Supernovae (CCSN)

- 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.



TimeSeries of CCSN



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

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

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

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

ARTICLE INFO

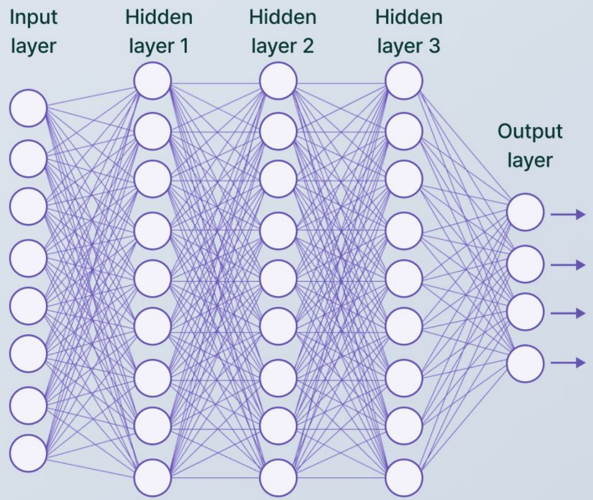
Article history:
Received 27 May 2019
Revised 21 July 2019
Accepted 8 August 2019
Available online 9 August 2019

Communicated by Steven Hoi

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

ABSTRACT

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.



Datasets

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 ($f_{\rm peak} \sim 600-1000\,\mathrm{Hz}$) 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

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

For each element of the Dataset

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

Classification

GOAL: Distinguish strains with detector background noise from strains with gravitational wave signals injected into the noise.

Dataset Properties:

- 50% background noise and 50% signal;
- Distance between 5 and 20 kPc;
- Random sky position and polarization angle;
- Fixed inclination ($\pi/2$ rad);
- All signals with SNR $\geq$ 5;

Classification Training

Training configurations:

- 80% training set and 20% validation set
- Training function: fit_one_cycle
- Maximum Learning Rate: 0.003
- Weight decay: 0.001
- Model: ResCNN(3,2)

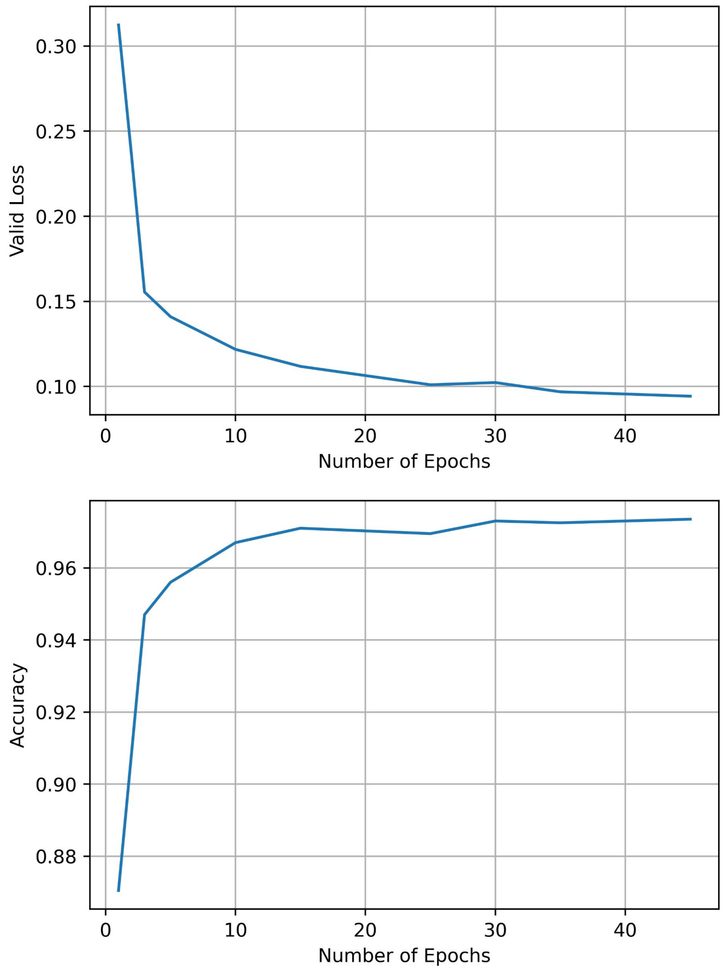
Classification Results: Dataset of 10k

Classification Results
Dataset of 10k

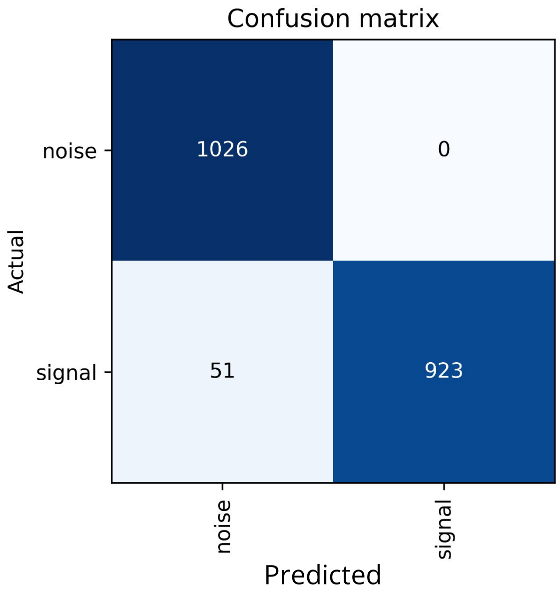
For the training with 25 epochs:

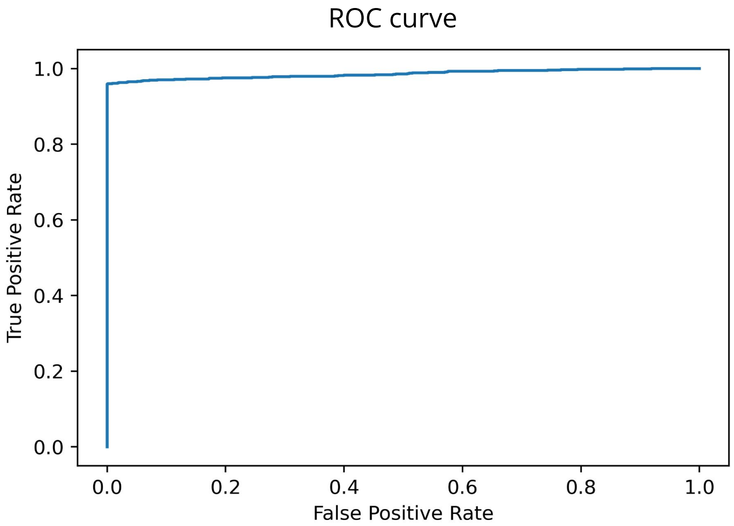
Valid Loss: 0.1010

Accuracy: 97.5%



Classification Results: Dataset of 10k (25 epochs)





- No actual noise classified as signal with score threshold of 0.5;
- Only 51 of actual signals was predicted as noise (5.24%);

- AUC of 0.985

Classification Results: Dataset of 10k (25 epoch)

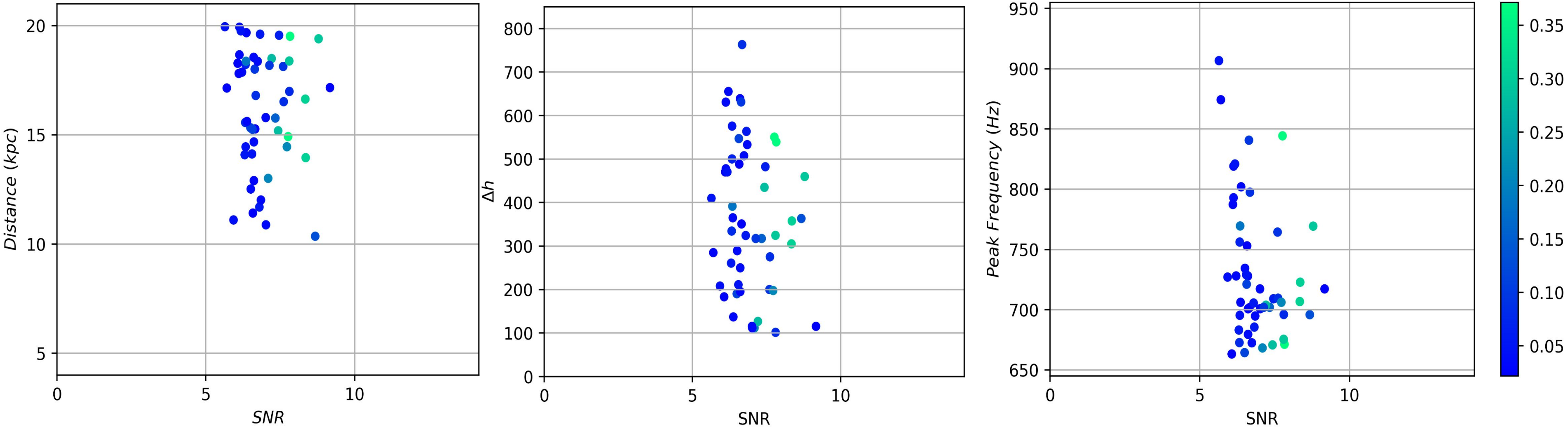
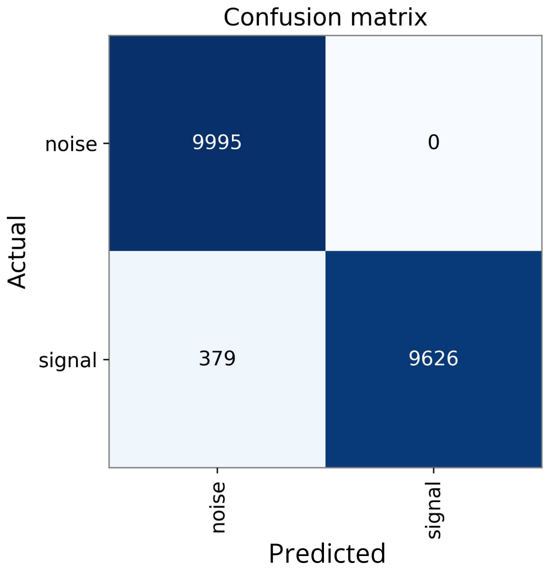
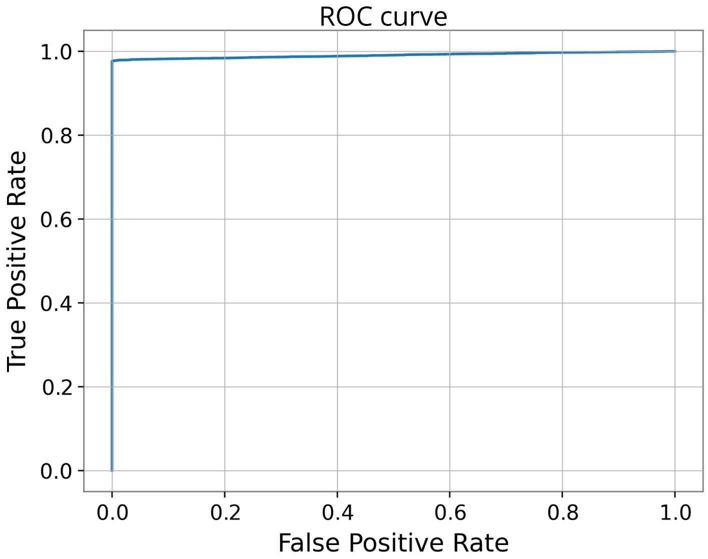


Fig. - Distribution of Distance, Dh and Peak Frequency as a function of the SNR for the wrongly classified real signals. The colors represents the score given by the model.

Classification Results: Dataset of 100k

Classification Results: Dataset of 100k (10 epoch)

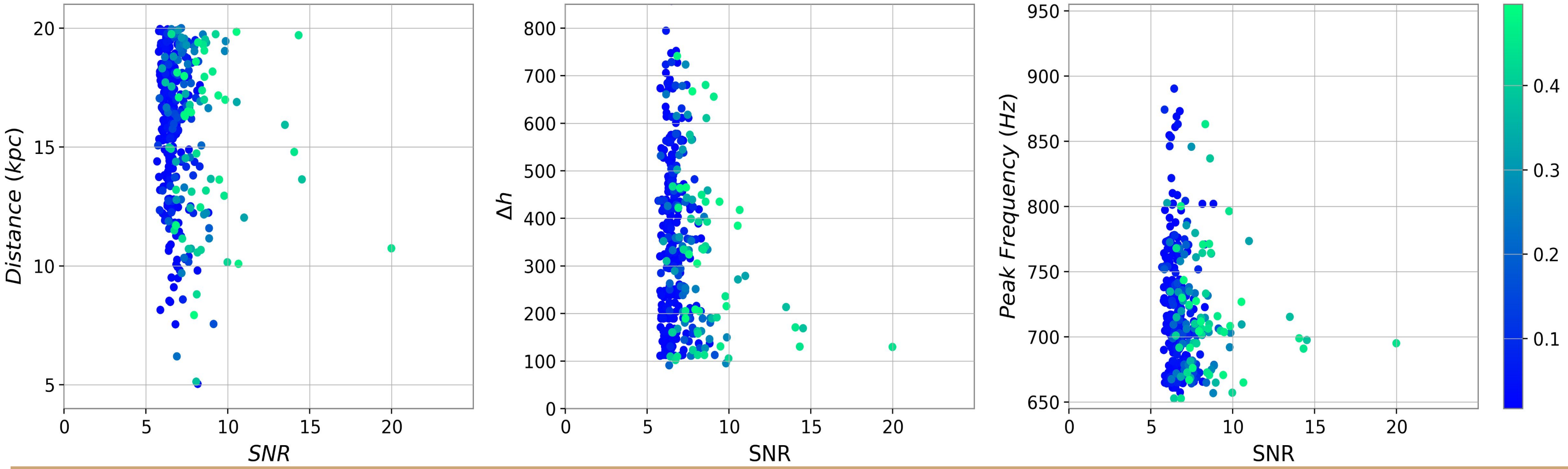




- No actual noise classified as signal with score threshold of 0.5;
- Only 379 of actual signals was predicted as noise (3.79%);

- Valid Loss: 0.07687
- Accuracy: 98.1%
- AUC: 0.991

Classification Results: Dataset of 100k



Distribution of Distance, Dh and Peak Frequency as a function of the SNR for the wrongly classified real signals. The colors represents the score given by the model.

Regression Results: Dataset of 10k

Regression

GOAL: Parameter Inference

Dataset Properties:

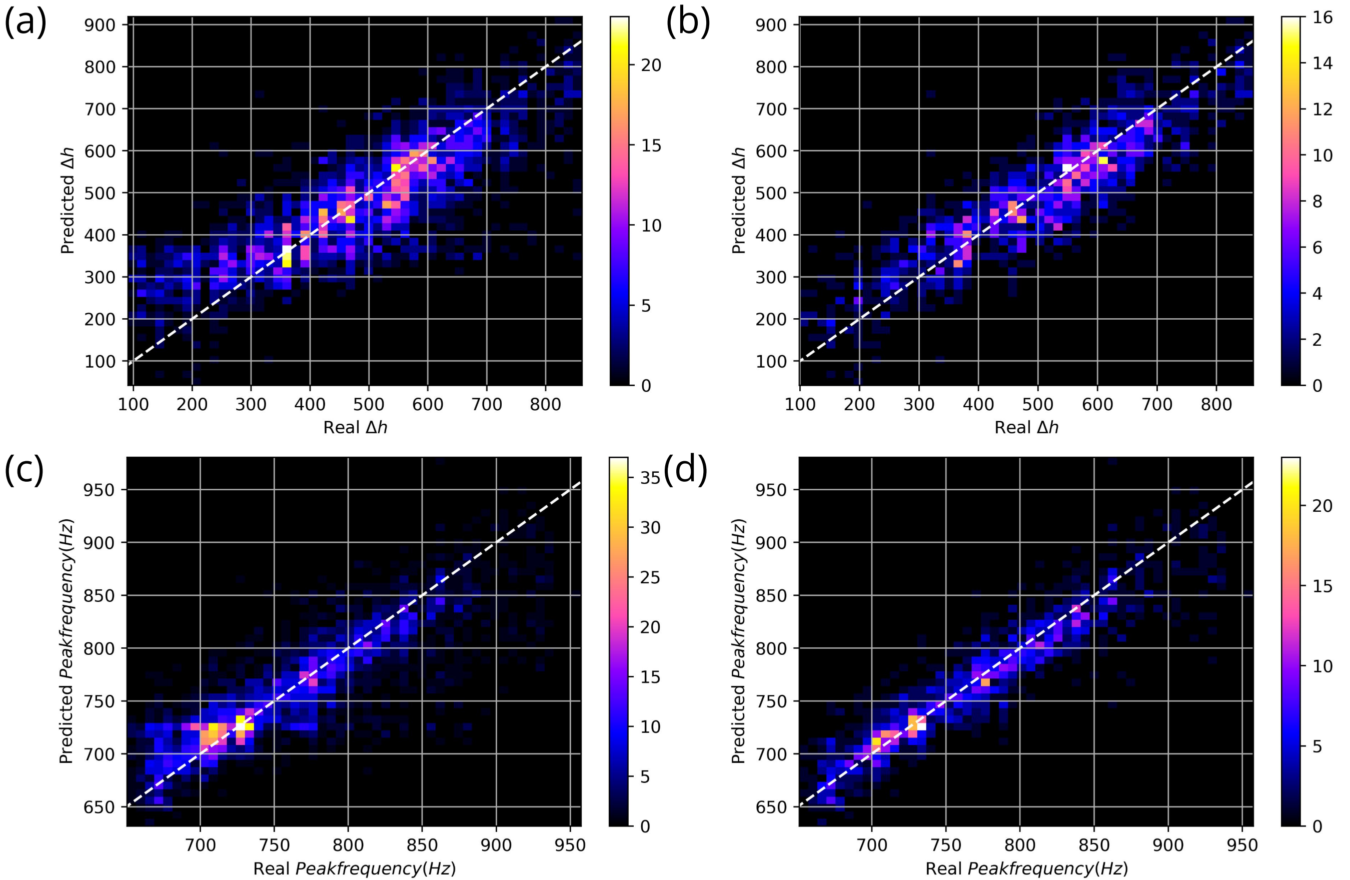
- 10k TimeSeries;
- Distance between 5 and 20 kPc;
- Random sky position;
- Fixed inclination (π/2 rad);
- All signals with SNR ≥ 5;
- Inference:
  - ➢ Frequency at the peak of the signal, f_{peak}
  - ➢ Amplitude of the signal, Δh

Regression Training

Training conditions:

- 70% training set and 30% validation set
- Training function: fit_one_cycle
- Maximum Learning Rate: 0.002
- Weight decay: 0.001
- Model: ResCNN(3,2)

Regression Results: Dataset of 10k (50 epochs)



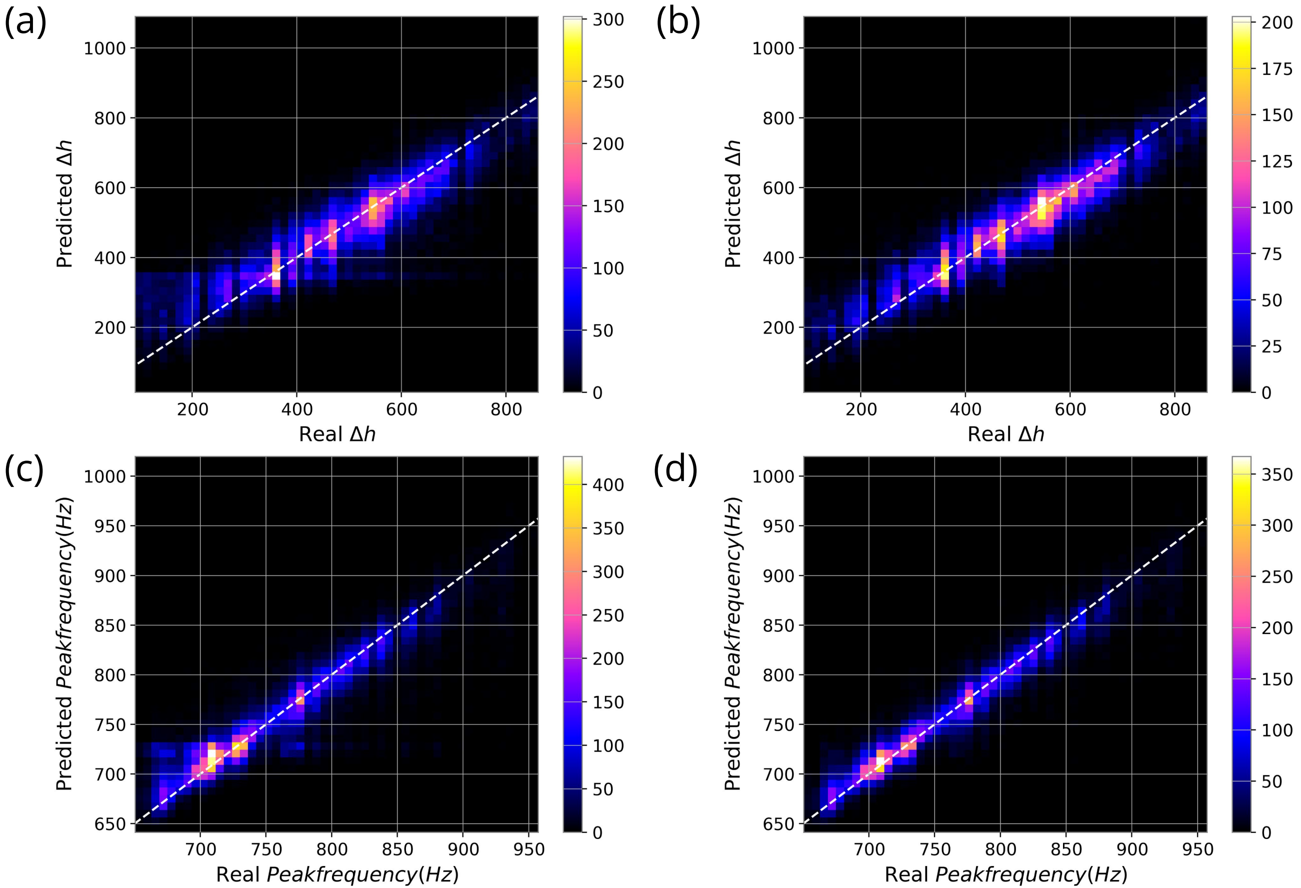
2D histogram of predicted values vs real values for:

(a) Δh with SNR ≥ 5
(b) Δh with SNR ≥ 15
(c) peak frequency with SNR ≥ 5
(d) peak frequency with SNR ≥ 15

Valid loss: 0.2975
Mean absolute error: 0.3986

Regression Results: Dataset of 100k

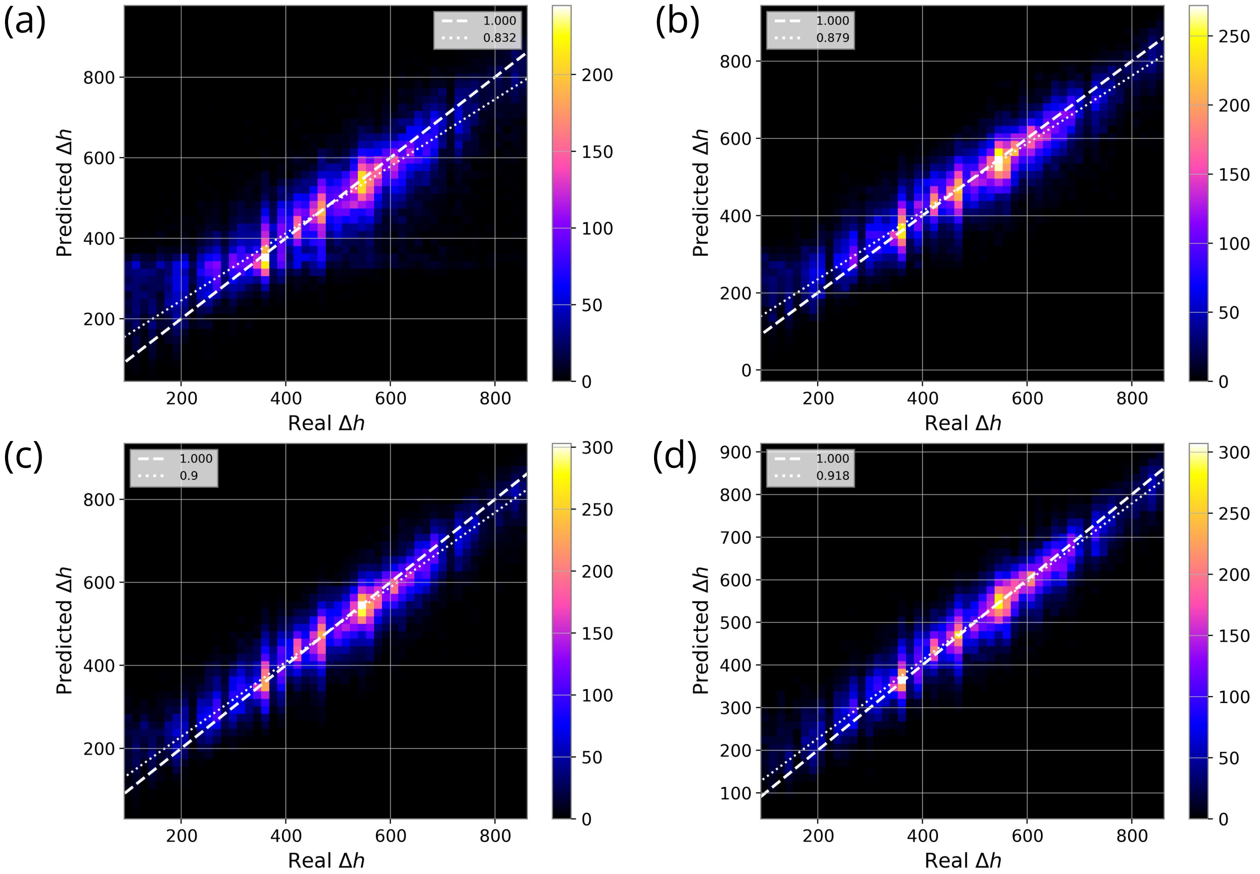
Regression Results: Dataset of 100k (50 epoch)



2D histogram of predicted values vs real values:
(a) Δh with SNR ≥ 5
(b) Δh with SNR ≥ 15
(c) peak frequency with SNR ≥ 5
(d) peak frequency with SNR ≥ 15

Valid loss: 0.2297
Mean absolute error: 0.3347

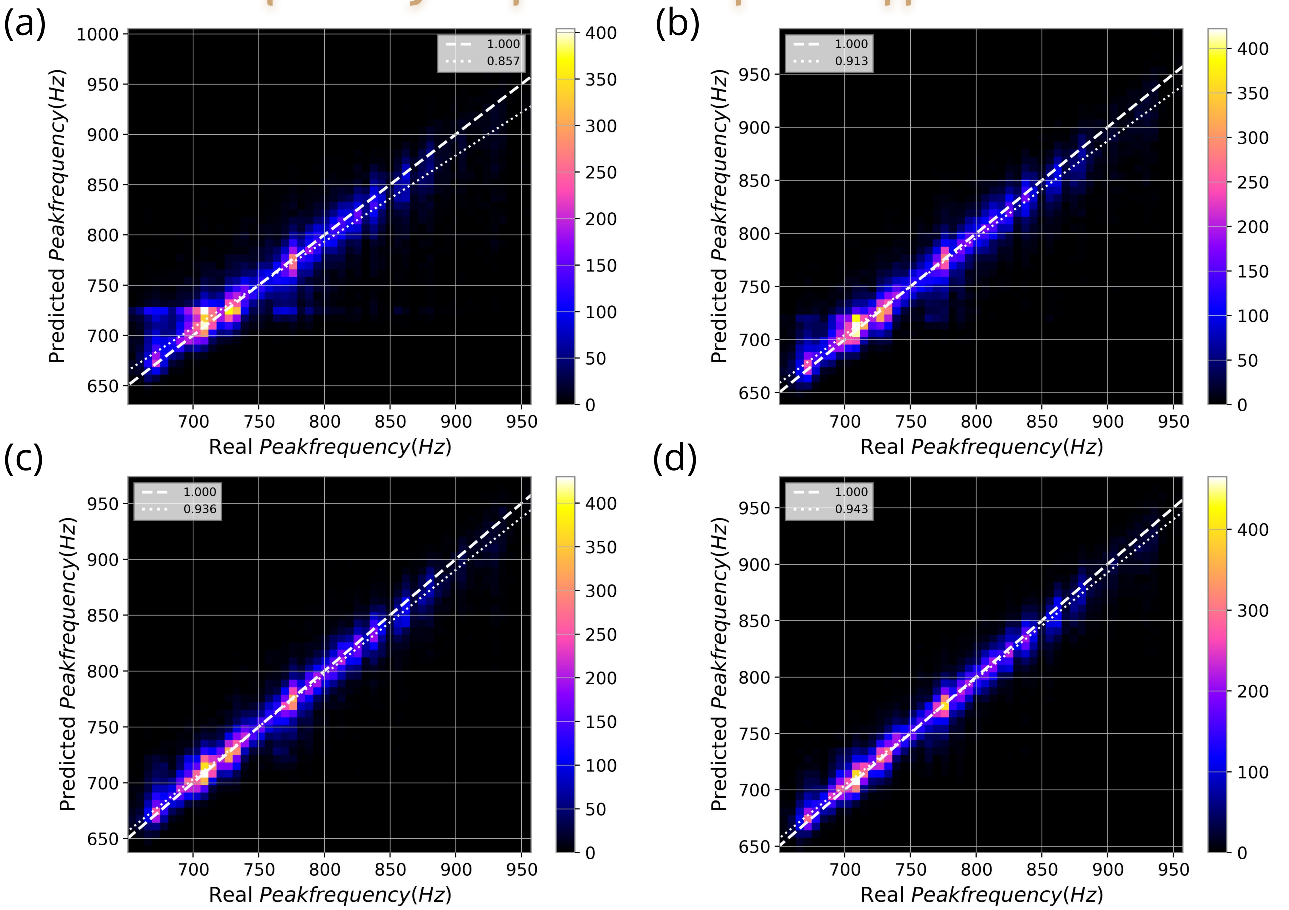
Δh inference for different minimum SNR (70 epochs)



2D histograms of predicted Δh vs real Δh for values of:
(a) SNR ≥ 5
(b) SNR ≥ 10
(c) SNR ≥ 15
(d) SNR ≥ 20

For SNR ≥ 20:
Valid loss: 0.09229
Mean absolute error: 0.2170

Peak Frequency inference for different minimum SNR (70 epochs)



2D histogram of predicted peak frequency vs real peak frequency for:
(a) SNR ≥ 5
(b) SNR ≥ 10
(c) SNR ≥ 15
(d) SNR ≥ 20

For SNR ≥ 20:
Valid loss: 0.09229
Mean absolute error: 0.2170

Conclusions

- This networks can perform classification with high accuracies
- No false positives
- False negatives appear only for lower values of SNR

- Regression performance is related to the SNR, giving the best results for SNR above 20

Attachments

Regression Results: Dataset of 100k for SNR ≥ 20

Valid loss: 0.09229
Mean absolute error: 0.2170