PROCEEDINGS OF SCIENCE

Deep Learning Application for Detecting Gravitational Waves from Core-Collapse Supernovae

Seiya Sasaoka,[a,*] Yilun Hou,[a] Diego Dominguez,[a] Suyog Garg,[b,c] Naoki Koyama,[d] Yusuke Sakai,[e] Yuto Omae,[f] Kentaro Somiya[a] and Hirotaka Takahashi[b,e]

[a] *Department of Physics, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8551, Japan*

[b] *Institute of Cosmic Ray Research (ICRR), The University of Tokyo, 5-1-5 Kashiwa-no-Ha, Kashiwa City, Chiba 277-8582, Japan*

[c] *Department of Physics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

[d] *Graduate School of Science and Technology, Niigata University, 8050 Ikarashi-2-no-cho, Nishi-ku, Niigata City, Niigata 950-2181, Japan*

[e] *Research Center for Space Science, Advanced Research Laboratories, Tokyo City University, 3-3-1 Ushikubo-Nishi, Tsuzuki-Ku, Yokohama, Kanagawa 224-8551, Japan*

[f] *Department of Industrial Engineering and Management, Nihon University, 1-2-1 Izumi-cho, Narashino, Chiba 275-8575, Japan*

E-mail: sasaoka@gw.phys.titech.ac.jp, elon@gw.phys.titech.ac.jp, diego@gw.phys.titech.ac.jp, gargsuyog@g.ecc.u-tokyo.ac.jp, f221004a@mail.cc.niigata-u.ac.jp, g2191402@tcu.ac.jp, oomae.yuuto@nihon-u.ac.jp, somiya@phys.titech.ac.jp, hirotaka@tcu.ac.jp

Core-collapse supernovae are potential sources of gravitational waves that could be detected by current and future detectors, and their detection and analysis are of great importance for understanding the explosion mechanism. Since matched filtering cannot be used for these signals due to the stochastic nature of the waveforms, detection methods based on time-frequency representation have been developed. Recently, deep learning has been applied to the analysis of gravitational wave data and has the potential to greatly improve our ability to detect and analyze these signals. In this study, we apply a convolutional neural network to detect and classify gravitational waves from core-collapse supernovae. The model is trained on waveforms obtained from 3D numerical simulations, injected in real noise of O3 observing run. We also apply class activation mapping technique to visualize from which part of the input the model predicted the result. The results show that our model is able to classify 9 different waveforms and noise with 96.9% accuracy at 1 kpc. The maps visualized by class activation mapping technique show that the model's predictions are based on g-mode shapes of input spectrograms.

38th International Cosmic Ray Conference (ICRC2023)
26 July - 3 August, 2023
Nagoya, Japan

*Speaker

PoS(ICRC2023)1499

1. Introduction

The first detection of gravitational waves (GWs) from a binary black hole merger in 2015 marked the beginning of GW astronomy [1]. Subsequently, the first joint observation of GWs from a binary neutron star merger and the related electromagnetic signals opened the door to multi-messenger astronomy [2]. With numerous binary mergers detected, the field is anticipating the detection of short-duration GW bursts, with core-collapse supernova (CCSN) being a prominent target. Despite the detection of neutrinos from the SN1987A event [3, 4], the details of the explosion mechanism are still an open question, and the direct probe of its internal dynamics by neutrinos and GWs is crucial for the study of the supernova engine. Since GW signals from CCSNe have stochasticity in nature, traditional detection techniques such as matched filtering, which relies on specific waveform templates, are not applicable. As an alternative, detection methods based on time-frequency representation have been developed [5].

In recent years, machine learning techniques, especially deep learning, have demonstrated remarkable success in various scientific domains. Deep learning algorithms excel at recognizing complex patterns and extracting meaningful features. Their ability to learn from large data sets has led to breakthroughs in fields such as computer vision and natural language processing. In the context of CCSNe, Astone *et al.* [6] first proposed to apply convolutional neural network (CNN) to detect them, and they showed that CNN is a promising approach to identify CCSN signals from background noise.

In this study, we take a similar approach as Iess *et al.* [7] to detect and classify CCSNe using two-dimensional CNN model, but we add more simulated signals and consider signals from various distances. In addition, we apply gradient-weighted class activation mapping (Grad-CAM) technique [8] to visualize the regions in the input which affects the model's predictions.

The remainder of this paper is organized as follows. Section 2 describes our datasets, CNN model and Grad-CAM technique. The results and the visualization of the model are presented in Sec. 3. We summarize and conclude the paper in Sec. 4.

2. Method

Our CNN model is trained to classify strains at three detector LIGO H1, LIGO L1 and Virgo into 10 class: noise and 9 different CCSN waveforms. In this section, we describe the data and the model used in this analysis, and briefly explain the CAM technique.

2.1 Data set

9 different CCSNe waveforms obtained by 3D numerical simulation [9–13] are used to train and test our CNN model. The characteristics of each waveform are summarized in Table 1. The directions of radiation (θ, ϕ) are uniformly sampled and the plus and cross polarizations of each GW are calculated using the formulae

$$h_+ = \frac{1}{D}\frac{2G}{c^4}(\ddot{Q}_{\theta\theta} - \ddot{Q}_{\phi\phi}), \tag{1}$$

$$h_\times = \frac{1}{D}\frac{G}{c^4}\ddot{Q}_{\theta\phi}, \tag{2}$$

Table 1: CCSN waveforms used in this study.

Paper	Equation of State	Waveform Identifier	M_{star} $[M_\odot]$
Powell and Müller 2019 [9]	LS220	s3.5_pns	3.5
		s18	18
Radice *et al.* 2019 [10]	SFHo	s13	13
		s25	25
Mezzacappa *et al.* 2020 [11]	LS220	c15	15
Powell and Müller 2020 [12]	LS220	s18np	18
		m39	39
		y20	20
Powell *et al.* 2021 [13]	SFHo	z85	85

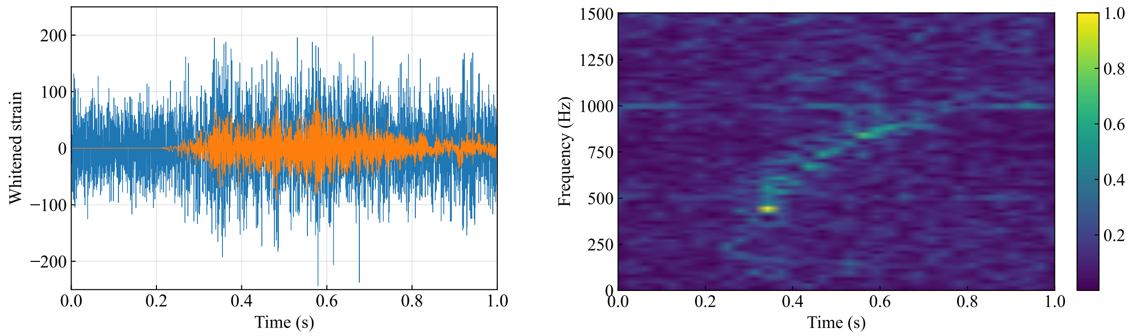


Figure 1: Sample whitened strain of H1 detector and raw s18 signal at 1kpc (left) and corresponding spectrogram (right).

where Q is the traceless quadrupole moment, and D is the distance between a source and Earth. These waveforms are resampled at a sampling rate of 4096 Hz, and a high pass filter with a cutoff frequency of 11 Hz and a Tukey window with $\alpha = 0.1$ are applied prior to the zero padding to make the length of each sample to one second. Each sample is then randomly time shifted and rescaled so that the time of core bounce is between 0 and 0.15 seconds and the distance is between 1 and 10 kpc. Sky location is also randomly selected and gravitational wave amplitude $h(t)$ is computed, taking into account the antenna pattern function and the delay in arrival time of each detector.

Noise used in this study are O3 real data of Advanced LIGO and Advanced Virgo, obtained from Gravitational Wave Open Science Center [14]. Data from GPS time 1238265720 to 1238252308 was used for training set, 1238265720 to 1238354855 was used for validation set, and 1238404064 to 1238457121 was used for test set. After a signal is injected in noise, each sample is whitened in frequency domain and short-time Fourier transformed with window size of 0.0625 seconds to produce a spectrogram. We generated 200,000 samples each of training, validation and test data. One of the training samples is shown in Fig. 1.

2.2 Model

Our CNN model consists of two convolutional layers of kernel size 3, each followed by a max-pooling layer of size 2. The outputs of these layers are fed into two fully connected layers, which output a size 10 vector whose elements represent a probability of each class: noise, c15, m39, s3.5_pns, s13, s18np, s18, s25, y20 and z85. The model is trained using cross entropy as the loss function and Adam optimizer with a learning rate of 10^{-3} to update the weights. We trained the model for 100 epochs with a mini-batch size of 2048.

2.3 Visualization

After training the model, we apply Grad-CAM to visualize the regions in the input that influenced the model's prediction. It is computed using the feature maps A^k at the last convolutional layer and the gradients of the predicted score of the class of interest with respect to the feature maps as weighted parameters w_c^k. The ReLU function is applied to extract only features that have a positive influence on the predictions. The resulting map of class c is expressed as

$$w_k^c = \sum_{i,j} \frac{\partial Y^c}{\partial A_{ij}^k}, \quad M^c = \mathrm{ReLU}\left(\sum_k w_k^c A^k\right). \tag{3}$$

To obtain higher resolution maps, we use Guided Grad-CAM, which is a combination of Grad-CAM and Guided Backpropagation [15]. Guided Backpropagation modifies the standard backpropagation algorithm to only propagate positive gradients, highlighting the significant input features that influence the model's predictions while ignoring negative gradients.

3. Results

This section describes the results of the test set applied to the trained model and discusses the performance of our model. We then show the class activation mapping images for interpreting the model.

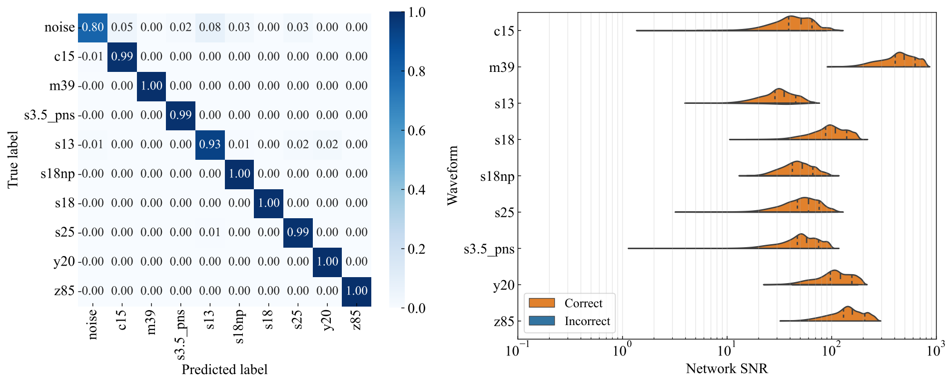
3.1 Classification result

Performance of a multi-class classification model is usually expressed by a confusion matrix, which shows the number of samples classified into each class. In Fig. 2, confusion matrices normalized for each class and the distribution of matched filter signal-to-noise ratio (SNR) for signals at 1, 5 and 10 kpc are plotted. Our classification model shows 96.9% accuracy for signals at 1 kpc and 59.3% for those at 5 kpc. We can see from the confusion matrices that as the distance increases, the amplitude of signal becomes smaller and the number of samples misclassified as noise increases. The accuracy for signals at 10 kpc is 30.2%, and our model cannot identify most of those signals except for m39 waveforms, whose SNR is much higher than others.

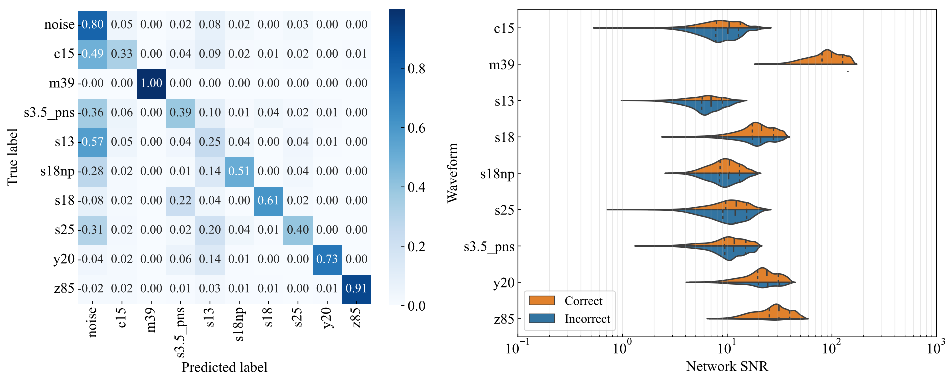
3.2 Visualization

Figure 3 shows some of the correctly classified samples and corresponding class-activation mapping images produced by Guided Grad-CAM. The CAM image of the noise sample show nothing particularly significant. The CAM image of the signal samples are shaped like the g-mode

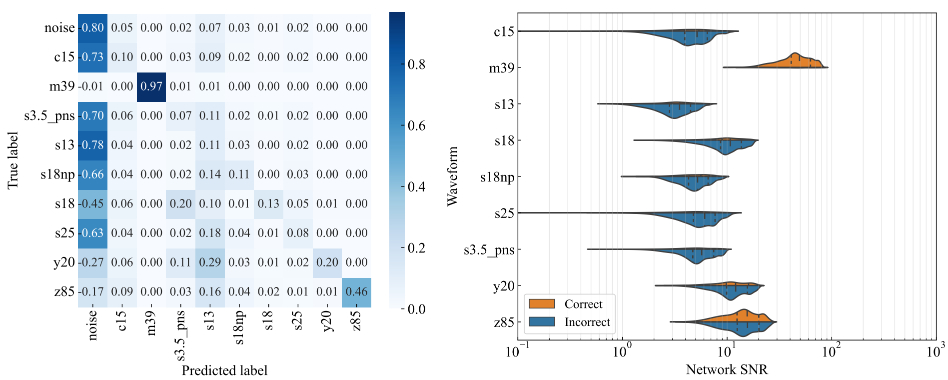
PoS(ICRC2023)1499



(a) 1 kpc



(b) 5 kpc



(c) 10 kpc

Figure 2: Confusion matrices of the test set (left) and violin plots of network SNR of each waveform (right) at distance 1, 5 and 10 kpc.

of the input spectrograms. Thus we can conclude that the model's predictions are based on the shape of the g-mode in the input spectrograms.

Waveforms such as c15, s3.5_pns, s13, s18np, s25, and z85, reported in their respective papers, have standing accretion shock instability (SASI) induced GW mode. Despite the presence of these distinctive features in the waveforms, the CAM images show that the model does not use them to make predictions. This appears to be because they are not adequately captured in the short-time

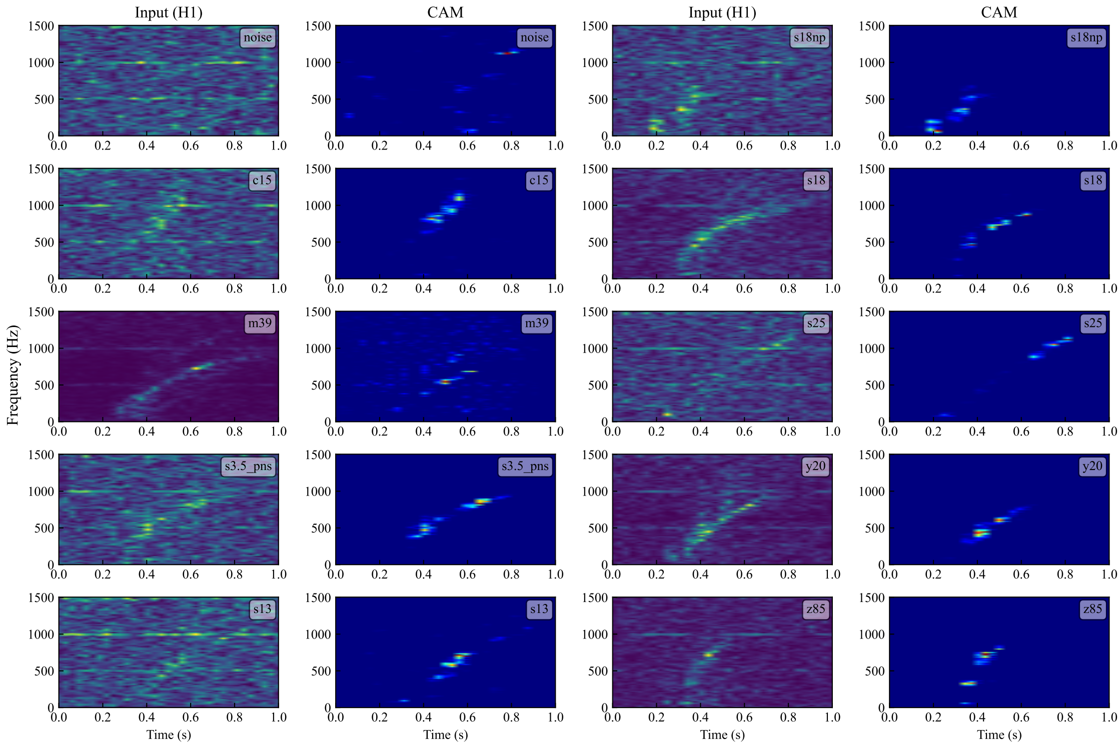


Figure 3: Input spectrograms and CAM visualizations of 9 waveform samples at 1 kpc and a noise sample.

Fourie transformed spectrograms. As a result, the CNN model likely could not utilize such features for its predictions.

4. Conclusions

In this study, we applied a two-dimensional CNN model to detect and classify CCSN signals immersed in noise. Our model achieved a high accuracy of 96.9% for signals at 1 kpc distance, but the model struggled to correctly identify most of the signals at 10 kpc.

To gain insights into the decision-making process of the model, we applied Grad-CAM technique to visualize the regions in the inputs that were influencial to the predictions. The CAM images of correctly classified signal samples revealed that the model's predictions were heavily affected by the shape of the g-mode oscillation appeard in the spectrograms.

Time-frequency maps used in this analysis were produced from the short-time Fourier transform, but it is expected that by using methods such as the Hilbert-Huang transform [16], which can produce higher resolution time-frequency maps, CNN models identify modes other than g-mode to make predictions.

Acknowledgments

The authors would like to thank J. Powell for providing us gravitational wave simulation data. This research was supported in part by JSPS Grant-in-Aid for Scientific Research [No. 22H01228

(K. Somiya), and Nos. 19H01901, 23H01176 and 23H04520 (H. Takahashi)]. This research was also supported by the Joint Research Program of the Institute for Cosmic Ray Research, University of Tokyo and Tokyo City University Prioritized Studies. This research has made use of data or software obtained from the Gravitational Wave Open Science Center (gwosc.org), a service of the LIGO Scientific Collaboration, the Virgo Collaboration, and KAGRA. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation, as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain. KAGRA is supported by Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan Society for the Promotion of Science (JSPS) in Japan; National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea; Academia Sinica (AS) and National Science and Technology Council (NSTC) in Taiwan.

References

[1] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116**, 061101 (2016) [gr-qc/1602.03837].

[2] B. P. Abbott *et al.*, *Multi-messenger Observations of a Binary Neutron Star Merger*, *ApJL* **848**, L12 (2017) [astro-ph/1710.05833].

[3] K. Hirata *et al.*, *Observation of a neutrino burst from the supernova SN1987A*, *Phys. Rev. Lett.* **58**, 1490 (1987).

[4] R. M. Bionta *et al.*, *Observation of a neutrino burst in coincidence with supernova 1987A in the Large Magellanic Cloud*, *Phys. Rev. Lett.* **58**, 1494 (1987).

[5] S. Klimenko *et al.*, *Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors*, *Phys. Rev. D* **93**, 042004 (2016) [gr-qc/1511.05999].

[6] P. Astone, P. Cerdá-Durán, I. Di Palma, M. Drago, F. Muciaccia, C. Palomba and F. Ricci, *New method to observe gravitational waves emitted by core collapse supernovae*, *Phys. Rev. D* **98**, 122002 (2018) [astro-ph/1812.05363].

[7] A. Iess, E. Cuoco, F. Morawski, C. Nicolaou and O. Lahav, *LSTM and CNN application for core-collapse supernova search in gravitational wave real data*, *A&A* **669**, A42 (2023) [astro-ph/2301.09387].

[8] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, *Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization*, in *2017 IEEE International Conference on Computer Vision (ICCV)* , pp. 618–626 [cs/1610.02391].

[9] J. Powell and B. Müller, *Gravitational wave emission from 3D explosion models of core-collapse supernovae with low and normal explosion energies*, *Monthly Notices of the Royal Astronomical Society* **487**, 1178 (2019) [astro-ph/1812.05738].

[10] D. Radice, V. Morozova, A. Burrows, D. Vartanyan and H. Nagakura, *Characterizing the Gravitational Wave Signal from Core-collapse Supernovae*, *ApJL* **876**, L9 (2019) [astro-ph/1812.07703].

[11] A. Mezzacappa *et al.*, *Gravitational-wave signal of a core-collapse supernova explosion of a 15 $M_\odot$ star*, *Phys. Rev. D* **102**, 023027 (2020) [astro-ph/2007.15099].

[12] J. Powell and B. Müller, *Three-dimensional core-collapse supernova simulations of massive and rotating progenitors*, *Monthly Notices of the Royal Astronomical Society* **494**, 4665 (2020) [astro-ph/2002.10115].

[13] J. Powell, B. Müller and A. Heger, *The final core collapse of pulsational pair instability supernovae*, *Monthly Notices of the Royal Astronomical Society* **503**, 2108 (2021) [astro-ph/2101.06889].

[14] R. Abbott *et al.* (LIGO Scientific Collaboration, Virgo Collaboration and KAGRA Collaboration), *Open data from the third observing run of LIGO, Virgo, KAGRA and GEO*, arXiv:2302.03676 (2023).

[15] J. T. Springenberg, A. Dosovitskiy, T. Brox and M. A. Riedmiller, *Striving for Simplicity: The All Convolutional Net*, arXiv:1412.6806 (2014).

[16] M. Takeda *et al.*, *Application of the Hilbert-Huang transform for analyzing standing-accretion-shock-instability induced gravitational waves in a core-collapse supernova*, *Phys. Rev. D* **104**, 084063 (2021) [astro-ph/2107.05213].