PROCEEDINGS OF SCIENCE

# Convolutional Neural Network for Continuous Gravitational Waves Detection

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

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

*E-mail:* diego@gw.phys.titech.ac.jp, sasaoka@gw.phys.titech.ac.jp, gargsuyog@g.ecc.u-tokyo.ac.jp, elon@gw.phys.titech.ac.jp, f22l004a@mail.cc.niigata-u.ac.jp, somiya@phys.titech.ac.jp, hirotaka@tcu.ac.jp

The detection of gravitational waves (GWs) has opened up new avenues for studying the universe and testing fundamental physics. As LIGO, Virgo and KAGRA start another observation run (O4) this year with an improved sensitivity, non-axisymmetric neutron stars emitting quasi-monochromatic, long-standing GWs are expected to be within the detectors' frequency bands. However, their detection in the presence of noise is a challenging problem. In recent years, Convolutional Neural Networks (CNNs) have been proposed as a potential solution to this issue. This study explores the effectiveness of CNNs for detecting CWs embedded in simulated noise as well as in real noise obtained from the O3 observation run. The model is trained on a set of $10^4$ templates of Continuous Waves (CWs) signals immersed in Gaussian noise with time gaps. Moreover, we evaluated the CNN's capability to generalize in signal-to-noise-ratio (SNR) and signal frequency. The results show that the CNN model is successful in detecting signals in simulated noise for the frequency band ranging from 100 to 1000 Hz, achieving high detection accuracy and low false positive rates. Nevertheless, when evaluating the model in data with real noise, which contains non-stationary noise and instrumental artifacts, its performance deteriorates. Such limitations suggest that more advanced methods are needed to analyze gravitational wave data effectively.

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

PoS(ICRC2023)1519

*Speaker

## 1. Introduction

Unlike transient gravitational wave events from compact binary mergers [1–3], which manifest as brief bursts, Continuous Gravitational waves (CWs) are a class of GWs that endure over extended periods of time [4]. The emission mechanism of CWs arises from the intrinsic asymmetry in the mass distribution of rotating objects. As these entities spin, their non-spherical configurations induce periodic oscillations in the surrounding spacetime, giving rise to GWs characterized by a quasi-monochromatic frequency [5]. This frequency remains remarkably constant, as long as the rotational dynamics of the object remain stable.

With an improved sensitivity for Observation run O4 [6], ground-based interferometers are expected to detect CWs emitted within our galaxy. However, their detection and analysis present unique challenges, primarily due to their relatively low amplitudes, rendering them extremely susceptible to interfering background noise [7]. It has been estimated that the upper limit on the CWs strain amplitude $h_0$ is $\sim \mathcal{O}(10^{-25})$ [8]. Given that the signal-to-noise ratio (SNR) increases with longer observation time [9], one needs to integrate a substantial amount of data for optimal results. In the most favorable scenario, a signal with a SNR = 5, would require a period of $T \sim 300$ days comprising high-quality data, therefore yielding the computational costs associated with a coherent analysis of such signal practically unfeasible. Hence, alternative analysis methods are needed to circumvent the demand of the high amount of computational resources.

In recent years, Deep Learning (DL) techniques have emerged as powerful tools for addressing these kind of complexities [10]. It utilizes artificial neural networks with multiple layers to automatically learn and extract intricate patterns and representations from complex datasets. In particular, Convolutional Neural Networks (CNNs) are designed for processing grid-like data (such as images or spectrograms) to automatically learn and extract spatially local features, capturing important patterns and structures [11]. One of the key advantages of deep learning is its ability to automatically learn feature representations from raw data, reducing the need for manual feature engineering and allowing the model to discover complex and non-linear relationships.

In the context of GWs physics, CNNs have been used for a number of applications [12], namely parameter estimation of transient signals [13, 14], signal denoising [15], wave form modeling [16, 17] and detection and classification of glitches [18, 19] to name a few. For CWs physics, DL has the potential to significantly enhance the sensitivity and efficiency of searches. Neural Networks can be trained on simulated data sets, enabling them to recognize and distinguish subtle signals from background noise. Moreover, DL methods can facilitate real-time analysis and enable the rapid identification of potential gravitational wave candidates. They offer the advantage of adaptability and scalability, allowing the algorithms to improve over time as more data becomes available.

In this work, we explore the use of CNNs to develop a *binary classifier* capable of detecting CW signals immersed in Gaussian noise data with time gaps. Our goal is to determine how well a Neural Network can generalize its detection capabilities. The CNN under consideration is trained utilizing a curriculum learning approach and employing a transfer learning method rooted in the ResNet-50 architecture [20].

## 2. Method

Deep learning models use optimization algorithms that iteratively adjust the weights and biases of the network to minimize the difference between predicted outputs and the true labels. In this work, we use the amplitude of the Short Fourier Transform (SFT) spectrogram as the input data for the neural network. Such images contain Gaussian noise and a simulated CW signal. The model's output is a numerical value ranging from 0 to 1, signifying the level of confidence associated with the presence of a CW signal. Finally, after training and testing the model on synthetic data, we evaluate its performance when facing injected CW signals into real detector noise obtained during the observing run O3 [21].

### 2.1 Deep Learning approach and Neural Network architecture

An important advancement in the architecture of Convolutional Neural Networks (CNNs) is the integration of *Residual Networks* (ResNets), which employ skip connections to effectively bypass and shortcut specific layers mitigating the vanishing gradient problem. Another notable strategy for reducing the training time of CNNs is *Transfer Learning*, which capitalizes on the knowledge acquired from pre-training a CNN on a comprehensive dataset and applies it to a distinct yet related task. By leveraging the learned features of the pre-trained network, transfer learning enables the CNN to benefit from the broad representations and patterns extracted during the initial training phase. The process of transfer learning involves fine-tuning the pre-trained CNN on the target dataset by retaining the initial layers that capture low-level features, such as edges and textures, while the higher-level layers may be replaced to adapt to the specific requirements of the new task.

Here, we employ a transfer learning scheme making use of a 2D CNN called ResNet-50 [20] and we replace its final layers to adapt to our binary classification problem. Additionally, we adopt *Curriculum learning* as a training strategy that aims to enhance the learning process by gradually presenting the network with training samples in an increasing difficulty order. This approach mimics the way humans learn, starting from simpler concepts and progressively advancing to more complex ones. By designing a curriculum that exposes the CNN to easier examples first, the network can build a solid foundation before addressing more challenging instances.

### 2.2 Generation of Data

For training and validation of the CNN, we generate two classes of signals with different labels: *i*) monochromatic signals embedded in pure Gaussian noise with time gaps (label = 1) and *ii*) signals containing pure Gaussian noise only (label = 0). They were generated using the python package PyFstat [22]. To implement the curriculum learning scheme, we generated eight different sets of increasing signal depth (decreasing SNR), defined by $\mathcal{D} = \sqrt{S_n(f)}/h_0$, where $S_n$ is the power spectral density of the noise and $h_0$ is the signal amplitude. As for the frequency space, for each set of a given signal depth, we generate 10 000 spectrograms randomly sampled from 10 Hz to 1000 Hz following a uniform distribution (See Table 1).

### 2.3 Data injection

To test the generalization capabilities of our network, we use real noise from the H1 and L1 detectors during the observing run O3 [21] spanning a few weeks up to a couple of months. Then, by

**Table 1:** Definition of the parameters for synthetic data generation

| | |
|---|---|
| Data span | $T = 10^6$ s |
| Noise | Stationary, white, Gaussian |
| Sky-region | All-sky |
| Signal depth | $\mathcal{D} \in [10, 50]$ Hz$^{-1/2}$ |
| Frequency band | $f \in [10, 1000]$ Hz |
| Spin-down | $\dot{f} = -10^{-9}$ Hz/s |

using the PyFstat package, we inject a CW signal with random sky-location, $\mathcal{D} \in [10, 50]$ Hz$^{-1/2}$ and $f \in [10, 1000]$ Hz.

## 3. Results

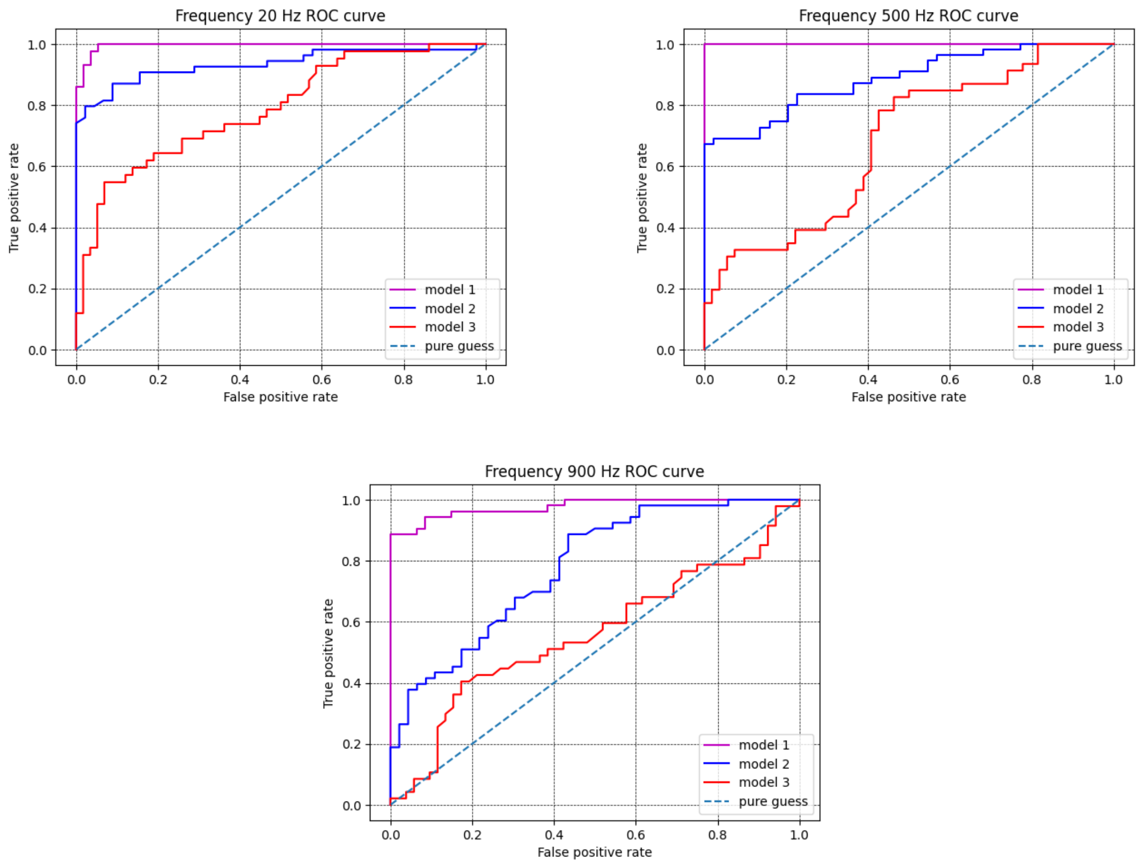


**Figure 1:** ROC curves of the Neural Network evaluated at three different frequency levels.

### 3.1 Test on generated data

For the trials on synthetic data, three different models are under consideration: *a*) *Model 1*, trained with signals whose $\mathcal{D} \in [10, 25]$ Hz$^{-1/2}$, *b*) *Model 2*, trained with signals having

$\mathcal{D} \in [10, 35]$ Hz$^{-1/2}$ and *c*) *Model 3*, with $\mathcal{D} \in [10, 50]$ Hz$^{-1/2}$. Later, the top performer will be chosen to be evaluated against injected data. To accomplish this, we employ a Receiver Operating Characteristic (ROC) curve, commonly used to assess the performance of binary classification models [23]. It plots the True Positive (TP) Rate or Sensitivity against the False Positive (FP) Rate as the discrimination threshold varies. The ROC curve illustrates the trade-off between sensitivity and specificity, providing valuable insights into the model's ability to correctly classify positive and negative instances at different thresholds.

Next, we test the models on spectrograms with varying frequencies and CWs having a signal depth $\mathcal{D} \sim 50$ (or SNR $\sim 60$). As we can see in Figure 1, *Model 1* shows a better performance over the other models. Even though it was trained only using the most intense spectrograms, it shows good generalization capabilities to less intense signals. Thus, *Model 1*, was chosen to be put to test against Injected data embedded in real detector noise.

### 3.2 Test on Injected data

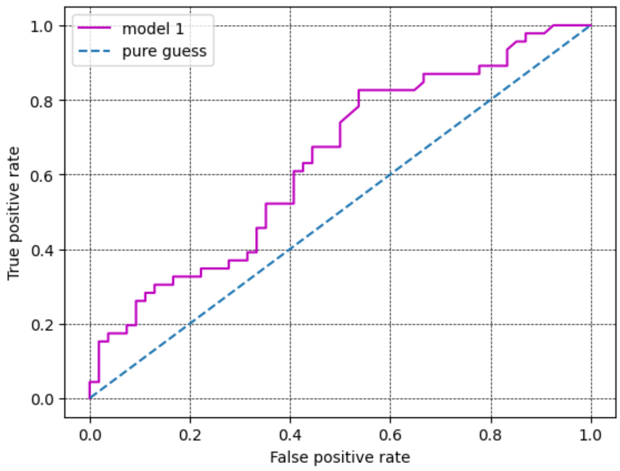


**Figure 2:** ROC curve for *Model 1* when evaluated on data containing real detector noise.

For this test, CWs with frequencies from all across the board and signal depths $\mathcal{D} \sim 50$ were injected in real interferometer noise. We took 10 different signals spanning from two weeks to a couple of months. Each sample was augmented via time shift, flipping and rotation to generate up to 500 different files. Approximately half of them had a CW signal embedded and the rest were only detector noise. Figure 2 shows that the performance of *Model 1* deteriorates notably when predicting the labels of signals in real noise data. However, predictions are still better than pure guess but not as consistent as they were when facing synthetic data. Since detectors' artifacts (included in the spectrograms of real noise) mimic the presence of a signal (See Figure 3), prediction becomes more difficult and proves that any realistic application needs CNNs to be trained with more realistic simulated noise. It's worth noticing there are two possible paths that may be followed: 1) training the CNN with CWs injected in real noise from the beginning and 2) finding a way of generating more realistic data using software, like the pyFstat package used in this study. It would be interesting to compare both methods and evaluate which one yields a better detection probability for future searches.

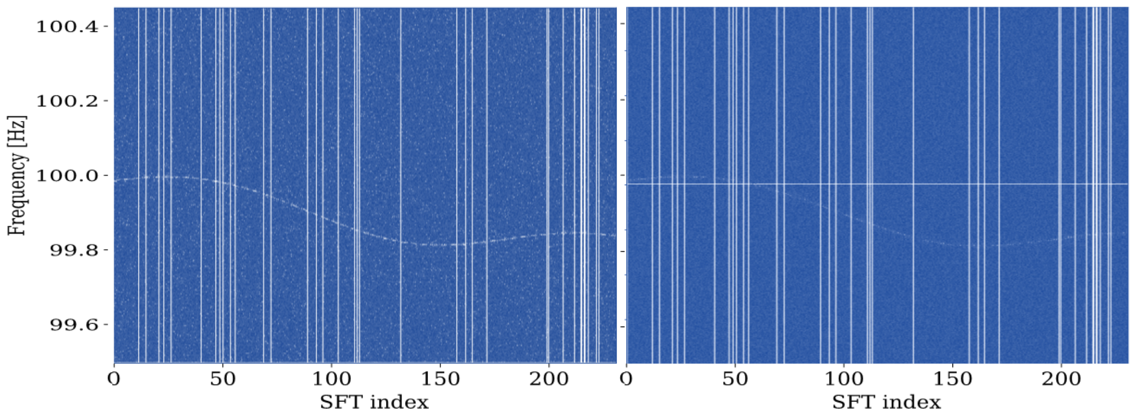


**Figure 3: Left**) Time-domain spectrogram with gaps showing a CW signal (note that the signal intensity was exaggerated for visualization purposes) and **right**) spectrogram including a detector artifact. We can see here how image-recognition-based ML falls short since both CW signal and artifact are visually similar.

## 4. Discussion

In this study, we explored the use of a Residual Neural Network trained with simulated data and its generalization capabilities when evaluated on CWs signals injected in real detector noise. As expected, performance decreases due to the presence of artifacts and other glitches. Therefore, in the future, it will be important to study the effect of such parasite signals and to develop a comprehensive training strategy for a robust CNN.

## Acknowledgments

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)]. Our work 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] LIGO Scientific Collaboration and Virgo Collaboration collaboration, *Observation of gravitational waves from a binary black hole merger*, *Phys. Rev. Lett.* **116** (2016) 061102.

[2] LIGO Scientific Collaboration and Virgo Collaboration collaboration, *GW170817: Observation of gravitational waves from a binary neutron star inspiral*, *Phys. Rev. Lett.* **119** (2017) 161101.

[3] LIGO Scientific Collaboration and Virgo Collaboration collaboration, *GWTC-1: A gravitational-wave transient catalog of compact binary mergers observed by LIGO and Virgo during the first and second observing runs*, *Phys. Rev. X* **9** (2019) 031040.

[4] K. Riles, *Recent searches for continuous gravitational waves*, *Modern Physics Letters A* **32** (2017) 1730035 [`https://doi.org/10.1142/S021773231730035X`].

[5] M. Sieniawska and M. Bejger, *Continuous gravitational waves from neutron stars: Current status and prospects*, *Universe* **5** (2019) .

[6] B.P. Abbott, R. Abbott, T.D. Abbott, S. Abraham, F. Acernese, K. Ackley et al., *Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA*, *Living Reviews in Relativity* **23** (2020) .

[7] P.D. Lasky, *Gravitational waves from neutron stars: A review*, *Publications of the Astronomical Society of Australia* **32** (2015) e034.

[8] LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration collaboration, *All-sky search for continuous gravitational waves from isolated neutron stars using Advanced LIGO and Advanced Virgo O3 data*, *Phys. Rev. D* **106** (2022) 102008.

[9] M. Maggiore, *Data analysis techniques*, in *Gravitational Waves: Volume 1: Theory and Experiments*, pp. 335–414, Oxford University Press (2007), DOI.

[10] L. Zdeborová, *Understanding deep learning is also a job for physicists*, *Nature Physics* **16** (2020) 602.

[11] D.A. Roberts, S. Yaida and B. Hanin, *Neural networks*, in *The Principles of Deep Learning Theory: An Effective Theory Approach to Understanding Neural Networks*, p. 37–52, Cambridge University Press (2022), DOI.

[12] E. Cuoco, J. Powell, M. Cavaglià, K. Ackley, M. Bejger, C. Chatterjee et al., *Enhancing gravitational-wave science with machine learning*, *Machine Learning: Science and Technology* **2** (2020) 011002.

[13] D. George and E. Huerta, *Deep learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data*, *Physics Letters B* **778** (2018) 64.

[14] H. Shen, E.A. Huerta, E. O'Shea, P. Kumar and Z. Zhao, *Statistically-informed deep learning for gravitational wave parameter estimation*, *Machine Learning: Science and Technology* **3** (2021) 015007.

[15] G. Vajente, Y. Huang, M. Isi, J.C. Driggers, J.S. Kissel, M.J. Szczepańczyk et al., *Machine-learning nonstationary noise out of gravitational-wave detectors*, *Phys. Rev. D* **101** (2020) 042003.

[16] Z. Doctor, B. Farr, D.E. Holz and M. Pürrer, *Statistical gravitational waveform models: What to simulate next?*, *Phys. Rev. D* **96** (2017) 123011.

[17] Y. Setyawati, M. Pürrer and F. Ohme, *Regression methods in waveform modeling: a comparative study*, *Classical and Quantum Gravity* **37** (2020) 075012.

[18] J. Powell, A. Torres-Forné, R. Lynch, D. Trifirò, E. Cuoco, M. Cavaglià et al., *Classification methods for noise transients in advanced gravitational-wave detectors II: performance tests on Advanced LIGO data*, *Classical and Quantum Gravity* **34** (2017) 034002.

[19] M. Cavaglià, S. Gaudio, T. Hansen, K. Staats, M. Szczepańczyk and M. Zanolin, *Improving the background of gravitational-wave searches for core collapse supernovae: a machine learning approach*, *Machine Learning: Science and Technology* **1** (2020) 015005.

[20] K. He, X. Zhang, S. Ren and J. Sun, *Deep residual learning for image recognition*, pp. 770–778, 06, 2016, DOI.

[21] LIGO Scientific, VIRGO, KAGRA collaboration, *Open data from the third observing run of LIGO, Virgo, KAGRA and GEO*, 2302.03676.

[22] G. Ashton, D. Keitel, R. Prix and R. Tenorio, *PyFstat*, July, 2020. 10.5281/zenodo.3967045.

[23] T. Fawcett, *An introduction to roc analysis*, *Pattern Recognition Letters* **27** (2006) 861.

PoS(ICRC2023)1519