5th International Conference on Advances in Signal Processing and Artificial Intelligence (ASPAI' 2023),
7-9 June 2023, Tenerife (Canary Islands), Spain

Type of Presentation:

Oral: ☒ In-person: ☒

Poster: ☐ Virtual in Zoom: ☐

The same: ☐

Topic: Artificial Intelligence

Classification of time series as images using deep convolutional neural networks: application to glitches in gravitational wave data

Shuzhu Jin [1], Soumya D. Mohanty [2], Qunying Xie [1], Hanzhi Wang [3], and Xue-Hao Zhang [3]

[1] School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China

[2] The University of Texas Rio Grande Valley, One West University Blvd, Brownsville, TX 78526, USA

Tel.: + 19568826680 E-mail: soumya.mohanty@utrgv.edu

[3] Institute of Theoretical Physics, Lanzhou University, Lanzhou 730000, China

Summary: The classification of frequently occurring terrestrial-origin transient signals, called *glitches*, in the time series data from gravitational wave detectors is important for mitigating their adverse effects on searches for rare and valuable astrophysical signals. While formally a time series classification problem, recent successes in glitch classification have all come from using their time-frequency image representations. Using transfer learning with the VGG16 deep convolutional neural network for image classification, we compare the efficacy of different types of image representations for classifying simulated glitches. We find the novel result that training the network with 2D plots of the noisy glitch time series provides better classification accuracy than their time-frequency images.

Keywords: Time series classification, CNN, gravitational waves

1. Introduction

With the detection of the first signal from two merging black holes in 2015 [1], the birth of gravitational wave (GW) astronomy has opened a new window for observing the Universe. Data analysis plays a crucial role in this field because astrophysical GW signals must be extracted from the noise dominated time series output of a GW detector. A challenge here is the high rate of diverse types of transient terrestrial-origin signals, called *glitches*, that elevate the false alarm rate of GW searches. The classification of glitches, therefore, is of great importance to improving GW search sensitivity.

The Gravity Spy [2] project has produced a dataset of ≈ 8000 glitches that have been divided into 22 manually labeled classes based on their constant Q-transform [3] (CQT) time-frequency images. Although formally a time series classification (TSC) problem [4,5], the most successful methods to date for automated glitch classification (e.g., [6]) have used convolutional neural networks trained on CQT images. Here, we propose a novel approach that uses 2D plots of noisy glitch time series and find better classification performance than CQT images. This result may have wider implications for the TSC problem in general.

2. Methodology

We used transfer learning with VGG16 [7], a pre-trained deep convolutional neural network with 16 layers. The weights of the base model layers were frozen up to the 5th convolutional block and a new fully connected head model was created to train on our dataset. We used the adaptive moment estimation (Adam) optimizer with the categorical cross-entropy loss function and trained the model over 10 epochs.

We simulated an ensemble of glitch time series, each 2.0 sec long and sampled at 4096 Hz with one glitch signal occupying the central T sec, where $T \leq 0.04$ sec. The strength of a glitch relative to unit variance white Gaussian noise (WGN) is quantified by its signal to noise ratio (SNR) defined as the Euclidean norm of the noise-less glitch signal. We considered the following different types of glitch time series. (1) **Noisy:** glitch added to WGN and SNR drawn from a uniform distribution over $[10, 100]$. (2) **Noise-less:** glitch SNR fixed at 60 and no WGN added. (3) **Denoised**: time series obtained by passing an SNR=60 glitch added to WGN through an adaptive spline fitting method called SHAPES [8]. Fig.1 (left panel) illustrates the noisy and denoised time series. The noise-less time series is an idealization that serves as a benchmark. For real data, the denoised time series is the closest that one can get to the noise-less case.

For each type of time series, we compared two image representations for training VGG16: (1) **CQT** computed with the full 2.0 sec, and (2) the **2D plot** of a 0.2 sec section of the time series centered on the glitch. We also used a multi-view approach **(2D+CQT)** in which the two types of images are paired with a common class label. The time series used for generating the test data have glitches added to WGN with SNRs having a uniform distribution over (1) a **broad** ([10, 100]) and (2) a **narrow** ([5, 20]) range. The test dataset presented to VGG16 always has the same image representation as the corresponding training data.

To make the simulations realistic, we took representative glitches from three of the most frequently occuring Gravity Spy classes (called Blip, Koi Fish, and Tomte) and fitted each with a model consisting of a superposition of Gaussians. Each best fit model was then used as the centroid of a cluster in $\mathbb{R}^N$, where N is the number of samples in the fitted model. For each centroid, a cluster of glitch signals was generated by independently drawing the parameters of the Gaussians in the glitch model from uniform distributions centered on their respective values in the centroid. Of these, only the glitches that fell within a given Euclidean distance of the centroid were retained. Similarly, we generated 6 additional clusters with arbitrarily chosen centroids (also modeled as a superposition of Gaussians). Finally, the training and testing datasets had 600 and 450 images, respectively, of each type of time series for each of the 9 classes.

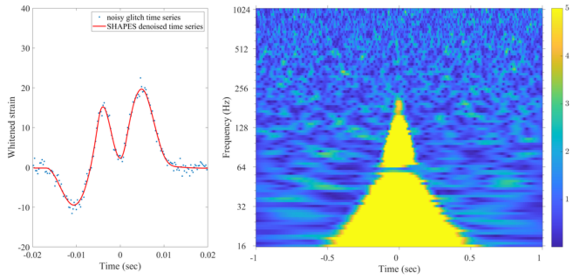

Fig. 1. (Left) Noisy (dots) and denoised (red) glitch time series example. (Right) CQT of the noisy time series.

3. Results

Tables 1 and 2 summarize our results. Each numerical entry corresponds to a given combination, labeled by the column and row headings, of the type of time series and image representation used for the training data. The numerical value is the accuracy of classification given by the fraction of test images that were classified correctly.

We see that training with 2D plots of the noisy glitch time series outperforms all other combinations. Furthermore, training using the denoised time series is comparable in performance to the ideal case of noise-less glitches. Training with variable SNR shows the best performance in all cases. 2D+CQT does not confer an advantage over 2D alone. As expected, performance of all methods worsens at lower SNRs.

4. Conclusions

We conclude that a 2D plot of time series as the image representation for training deep convolutional networks offers a promising approach to glitch classification. Its use for general TSC merits further investigation.

Acknowledgements

S.D.M is supported by U.S. National Science Foundation (NSF) grant PHY-2207935.

Table 1. Classification accuracies: broad SNR range.

SNR ∈ [10,100]		Type of time series		
		Noise-less	**Noisy**	**Denoised**
Image type	**2D plot**	0.7573	0.9552	0.7287
	CQT	0.8382	0.9162	0.8225
	2D+CQT	0.7725	0.9315	0.7531

Table 2. Classification accuracies: narrow SNR range.

SNR ∈ [5,20]		Type of time series		
		Noise-less	**Noisy**	**Denoised**
Image type	**2D plot**	0.5843	0.7324	0.5329
	CQT	0.6007	0.7065	0.5934
	2D+CQT	0.5820	0.7216	0.5411

References

[1]. B. P. Abbott et al. Observation of gravitational waves from a binary black hole merger, *Physical Review Letters*, Vol. 116, 2016, pp. 061102.

[2]. Michael Zevin et al. Gravity spy: integrating advanced ligo detector characterization, machine learning, and citizen science, *Classical and Quantum Gravity*, Vol. 34, 2017, pp. 064003.

[3]. Judith C Brown. Calculation of a constant Q spectral transform, *The Journal of the Acoustical Society of America*, Vol. 89, 1991, pp. 425–434.

[4]. Ismail Fawaz et al. Deep learning for time series classification: a review, *Data Mining and Knowledge Discovery*, Vol. 33, 2019, pp. 917–963.

[5]. S Mukherjee et al. in *Proceedings of the 14th Gravitational Wave Data Analysis Workshop*, Rome, Italy, 26-29 January 2010, *Journal of Physics: Conference Series*, Vol. 243, 2010, pp. 012006.

[6]. Tiago S. Fernandes et al. Convolutional Neural Networks for the classification of glitches in gravitational-wave data streams, *arXiv:2303.13917 [gr-qc]*, 2023.

[7]. Karen Simonyan, Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition, in *Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)*, San Diego, USA, 7-9 May 2015, *arXiv:1409.1556 [cs.CV]*.

[8]. Soumya D Mohanty and Ethan Fahnestock. Adaptive spline fitting with particle swarm optimization. *Computational Statistics*, Vol. 36, 2021, pp. 155–191.