Information Sciences 444 (2018) 172–186

Contents lists available at ScienceDirect

Information Sciences

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

Machine learning for Gravity Spy: Glitch classification and dataset

S. Bahaadini[a,*], V. Noroozi[b], N. Rohani[a], S. Coughlin[c,d], M. Zevin[c,d], J.R. Smith[e], V. Kalogera[c,d], A. Katsaggelos[a]

[a] *Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, USA*
[b] *Department of Computer Science, University of Illinois at Chicago, IL, USA*
[c] *Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Northwestern University, Evanston, IL, USA*
[d] *Department of Physics and Astronomy, Northwestern University, Evanston, IL, USA*
[e] *Department of Physics, California State University Fullerton, Fullerton, CA, USA*

ARTICLE INFO

Article history:
Received 23 December 2017
Revised 25 February 2018
Accepted 28 February 2018
Available online 6 March 2018

Keywords:
Deep learning
aLIGO
Dataset
Machine learning
Classification
Gravity Spy

ABSTRACT

The detection of gravitational waves with ground-based laser-interferometric detectors requires sensitivity to changes in distance much smaller than the diameter of atomic nuclei. Though sophisticated machinery and techniques have been developed over the past few decades to isolate such instruments from non-astrophysical noise, the detectors are still susceptible to instrumental and environmental noise transients known as “glitches,” which hinder searches for transient gravitational waves. The *Gravity Spy* project is an effort to comprehensively classify the glitches that afflict gravitational wave detectors into morphological families by combining the strengths of machine learning algorithms and citizen scientists.

This paper presents the initial Gravity Spy dataset used for citizen scientist and machine learning classification – a static, accessible, documented dataset for testing machine learning supervised classification. Previous versions of this dataset used in [8, 53] did not include all current classes and also for some of the classes, some samples were pruned and added. This set consists of time–frequency images of LIGO glitches and their associated metadata. These glitches are organized by time–frequency morphology into 22 classes for which descriptions and representative images are presented. Results from the application of state-of-the-art supervised classification methods to this dataset are presented in order to provide baselines for future glitch classification work. Standard splitting for training, validation, and testing sets are also presented to facilitate the comparison between different machine learning methods. The baseline methods are selected from both traditional and more recent deep learning approaches. An ensemble framework is developed that demonstrates that combining various classifiers can yield a more accurate model for classification. The ensemble classifier, trained with the standard training set, achieves 98.21% accuracy on the standard test set.

* Corresponding author.
E-mail address: SaraBahaadiniBeigyZarandi2019@u.northwestern.edu (S. Bahaadini).

https://doi.org/10.1016/j.ins.2018.02.068

1. Introduction

The recent observations of gravitational waves from binary black hole mergers [3–5,49] have inaugurated a new field of observational astronomy by providing a new method with which to explore the cosmos. These observations, made by the advanced Laser Interferometer Gravitational wave Observatory (LIGO, [28]), require sensitivity to fractional changes of distance on the order of 10^{-21}. The two LIGO detectors are in Hanford, Washington (LHO) and Livingston, Louisiana (LLO). To achieve this unprecedented sensitivity, all sensitive components of LIGO are exquisitely isolated from non-gravitational wave disturbances. Even with this isolation, the LIGO detectors are susceptible to disturbances that cause noise in the detectors and can afflict searches for gravitational waves.

Of particular concern are transient, non-Gaussian noise sources known colloquially as *glitches*. Glitches occur at a significant rate, come in many morphologies, and can mask or mimic gravitational wave signals. Work has been done in assessing whether or not an instance of excess noise is, in fact, a glitch [10], but a comprehensive classification and characterization of these noise features could allow their origin to be identified and their root cause to be removed from the instruments. Attempts to use machine learning algorithms have shown promise in glitch classification endeavors [35,36,40–42], however these techniques do not yet capture the full range of glitch morphologies present in LIGO data. In addition to the above methods, *Gravity Spy*[1] [53], a citizen science project hosted by the Zooniverse platform [12] that combines the classification power of machine learning and crowd-sourcing, provides a solution for addressing this problem. A critical component to the Gravity Spy method is the dataset used in training both the machine learning algorithm and the citizens.

In this paper, we present the Gravity Spy dataset, which is a collection of images of glitches and their associated metadata, in the context of machine learning tasks. We discuss the characteristics of the glitch classes within this data and provide an example for each class. To illustrate the complexity of the data and provide a better understanding of the relationship between various glitch classes, we visualize the feature space of Gravity Spy dataset. We further present a standard benchmark by defining the exact training, testing, and validation split sets that could be used to compare different machine learning algorithms. This dataset and standard benchmarks allow for further studies by the machine learning community, such as those performed in [21]. We apply state-of-the-art machine learning algorithms such as deep neural networks, support vector machines, and ensemble learning on this dataset to provide baselines for future works on this dataset. Although the problem considered in this paper is classification, the Gravity Spy dataset can be used for other machine learning tasks, such as clustering [51] and image retrieval [17].

Overall in this paper we have the following contributions:

- Introduction of the full specifications of the Gravity Spy dataset.
- Visualization of the Gravity Spy dataset.
- Determination of classification baseline accuracies for the dataset by developing classifiers based on neural networks, support vector machines, and ensemble learning, which are vital for establishing a control in testing future algorithms.

In the following sections, we describe the process of producing the images of glitches from raw gravitational wave detector data (Section 2.1), explain the specifications of glitch classes within this dataset (Section 2.2), investigate the feature space of all classes in the dataset (Section 2.3) and present the standard sets (Section 2.4). Finally, different machine learning baselines, with performance evaluation and data analysis for supervised classification tasks, are presented in Section 3. Concluding remarks are made in Section 4.

2. Gravity Spy data

Gravitational wave data, including transient noise in the detectors, is often visualized as time–frequency spectrograms. The images in the Gravity Spy dataset (used for both human classification and machine learning tasks) are a particular type of spectrogram based on decomposition using sine-Gaussian templates, a process known as the *Q-transform* [15]. The Gravity Spy dataset is composed of Q-transform images of any transients recorded by the gravitational wave channels of the detectors that exceed a certain threshold in loudness, specifically the signal-to-noise ratio (SNR), and pass the standard set of data quality criteria [2] used by LIGO's real-time gravitational wave searches.

Additionally, the Gravity Spy machine learning algorithms required a large training set of example images that belong to classes of morphologically distinct glitches to be constructed in order to allow machine learning pre-classification of the images that would be presented to citizen scientists [53]. To accomplish this, 22 different morphologically distinct classes of glitches were selected (the names and morphology of many of these classes had already been identified by the broader LIGO Scientific Collaboration [1,2,39]) and tens to hundreds of example images were hand selected (often with input from algorithms, such as the Hierachical Veto [48] that identify classes of glitches by their relationship with other types of disturbances, such as seismic noise). These classes are also the classification choices (buttons) that citizen scientists have to choose from the Gravity Spy project interface.

Over time, these training sets have expanded. When both the machine learning and volunteer classification of a given unlabeled image passes a certain confidence threshold [53], these images are "retired" and added to the training set. Furthermore, new glitch classes identified by citizen scientists, volunteers, or clustering algorithms are manually (following

[1] https://www.gravityspy.org.