Gravity Spy: Humans, Machines and The Future of Citizen Science

Kevin Crowston
Syracuse University
Syracuse, NY, USA
crowston@syr.edu

The Gravity Spy Team
https://www.zooniverse.org
/projects/zooniverse
/gravity-spy/about/team

Abstract
Gravity Spy is a citizen science project that draws on the contributions of both humans and machines to achieve its scientific goals. The system supports the Laser Interferometer Gravitational Observatory (LIGO) by classifying "glitches" that interfere with observations. The system makes three advances on the current state of the art: explicit training for new volunteers, synergy between machine and human classification and support for discovery of new classes of glitch. As well, it provides a platform for human-centred computing research on motivation, learning and collaboration. The system has been launched and is currently in operation.

CSCW '17 Companion, February 25 - March 01, 2017, Portland, OR, USA
ACM 978-1-4503-4688-7/17/02.
http://dx.doi.org/10.1145/3022198.3026329.

Author Keywords
Citizen science; machine learning; learning; training.

ACM Classification Keywords
H.1.2 [MODELS AND PRINCIPLES]: User/Machine Systems—Human factors; I.2.1 [ARTIFICIAL INTELLIGENCE]: Applications and Expert Systems; J.2 [PHYSICAL SCIENCES AND ENGINEERING]: Astronomy

Introduction
Citizen science describes scientific projects that rely on contributions to scientific research from volunteers from the general public (i.e., citizens in the broadest sense of the term). Automation of data collection has led to a dramatic increase in the volume of available data, making it infeasible for science teams to analyze them without some assistance. To address this problem, citizen science projects recruit volunteers to help analyze data. These systems work via the Internet and so are examples of crowdsourcing and of interest to social computing researchers. However, with better data collection, data volumes continue to grow, and soon will be too much even for large numbers of volunteers. Machine processing holds out the promise of handing such big data, but these techniques alone cannot handle the diversity of scientific data. In this poster, we present ways to create partnerships

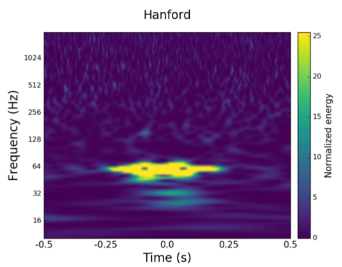

Figure 1: Example glitch represented as a spectrogram (time on the x axis, frequency on the y axis and intensity represented by colour). All figures are from the Gravity Spy system.

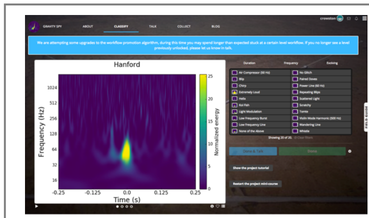

Figure 2: Gravity Spy classification interface, showing the spectrogram of a glitch on the left and choices for classification on the right.

between human and machine processing, in the context of a novel citizen science project called Gravity Spy [1].

Citizen Science to Support LIGO

The Gravity Spy citizen science project has been developed to support the Laser Interferometer Gravitational Observatory (LIGO). The purpose of LIGO is to detect gravitational waves generated by astrophysical events. It made the first ever detection of gravitational waves, from a pair of merging black holes, on 14 September 2015 [2]. Even though they are created by incredible violent events, gravitational waves reaching the earth are incredibly weak. LIGO can detect gravitational-wave-created length fluctuations in the 4-kilometer arms of the detector that are only one-thousandth the diameter of a proton.

The extreme sensitivity that enables LIGO to detect astrophysical events also makes it susceptible to many other disturbances, either environmental (e.g., small ground motions) or within the instrument (e.g., fluctuations in the laser). These disturbances create "glitches", noise in the detector that interfere with the search for gravitational waves (e.g., Figure 1). To improve the performance of the detector, glitches must be identified to remove them from the data stream and ideally to identify their causes to eliminate them.

There are currently about twenty known classes of glitches with different underlying causes. However, it is assumed that there are further classes of glitch that have not yet been identified and that glitches of different classes will come and go as the detector is worked on. However, the volume of glitches makes it challenging for the LIGO scientists to examine them all. Glitches occur every few seconds (up to 20,000 per day depending on the cutoff selected for signal strength). In comparison, detectable astrophysical events are predicted to occur only about once a month.

Gravity Spy System Design

The goal of the Gravity Spy project is to classify glitches by recruiting volunteers to assist in the process. The classification interface is shown in Figure 2. As with other citizen science projects, multiple volunteer classifications are aggregated to determine the consensus for a glitch.

The main advance in the system is that it incorporates machine learning (ML) techniques to work in conjunction with the volunteers. The current system uses deep learning approaches to classify images [3]. The algorithms were trained initially on a small collection of glitches classified by experts working with the LIGO project. A key point is that the ML provides an assessment of the confidence of its classification.

Most applications of ML to image classification aim at automation, i.e., using human-classified data to train algorithms to replace the humans. Pure automation would be difficult in this case though because of the diversity of the appearance of glitches, the presence of glitches of unknown classes and because the classes of glitches are expected to be dynamic. Therefore, we are exploring ways that humans and machines can work together on the citizen science task, thus taking advantage of the distinct skills of both.

In addition to automation, we have identified three ways that humans and machines can support each other's work that we are implementing: learning, task support and discovery.

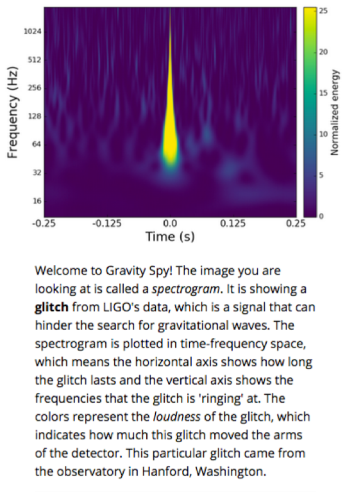

Figure 3: Sample slide from the project tutorial.

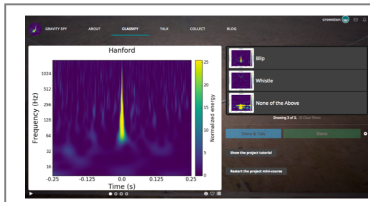

Figure 4: Reduced Gravity Spy classification interface for beginners, with just three options.

First, most online citizen science projects provide at most minimal training, and mostly separate from the task of the project. While we do provide such training (e.g., a project tutorial as shown in Figure 3), we are also providing explicit in-task training, guided by the ML. Specifically, new volunteers are presented with images classified by ML models as highly likely to be of just a few distinctive classes. Because the ML confidence is high, it is likely that these are exemplary images that will help the volunteer to learn to identify glitches of those classes. As well, the interface is reduced to just those choices plus “none of the above”, as shown in Figure 4.

Once volunteers can successfully classify the initial classes, they advance to successive training levels, in which they see ML-selected images of additional classes, and eventual graduate as fully-qualified volunteers. Conversely, glitches that have been classified by the volunteers can be added back to the training set for the ML, thus improving its performance, as well as being provided to the LIGO project to improve the performance of the detector.

Second, the system creates a synergy between machines and volunteers for work on the task of glitch classification. The ML provides an initial guess at the classification, but beginners check the ML classifications as part of their training. Images with low ML certainty—or even images for which the ML has no classification at all—are routed for classification to more experienced users, i.e., those who have completed the training. An advantage of this approach is that even the classifications done by beginners provide useful work for the project by verifying ML results.

Third, the system allows volunteers to take part in discovering new knowledge, not just applying knowledge from the project science team. Specifically, experienced users can identify new classes of glitches by examining those that do not fit in any existing class. Of course, discoveries have been made in many citizen science projects, most famously the Voorwerp found in Galaxy Zoo [4]. However, these discoveries have been byproducts of the project, the discovery process supported only informally. In contrast, supporting such discoveries is a primary goal of the Gravity Spy project, albeit still a work in progress.

Current Status

The first phase of the Gravity Spy system has been implemented using the Zooniverse project builder, with custom coding to manage the learning process. Classification of images and tracking of user performance is handled by a separate server that communicates with the Zooniverse backend. The current implementation provides three training levels for beginners (with two, six and ten classes of glitch) and two levels for classification by experienced users (both with all classes of glitch, but at different levels of ML confidence).

After an extensive beta test, the system launched in October 2016. It can be seen at http://gravityspy.org/. The system currently has more than 4,000 volunteers who have done more than 750,000 classifications on more than 70,000 images. It currently averages about 6,000 classifications per day. Volunteers are successfully proceeding through the training and advancing in levels.

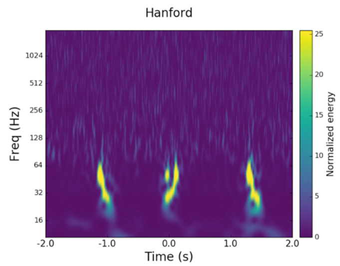

Figure 5: Spectogram of a "paired doves" glitch discovered by a Gravity Spy volunteer.

The system currently provides only minimal support for discovering new glitches. Nevertheless, a few advanced volunteers have already undertaken this task. A few new classes have been identified that are of interest to the LIGO scientists. An example is shown in Figure 5, a "paired dove" glitch first identified by volunteer EcceruElme. This class of glitch was of particular interest to LIGO scientist because its morphology is similar to a gravitational wave event.

In addition to being a system for classifying glitches, Glitch Zoo is also platform for conducting human-centred computing research. For example, we have already run one experiment comparing different appeals for new volunteers to identify which motives are most salient and are currently undertaking another experiment to test the effectiveness of the training.

Future Work

The current development work on the project is to more explicitly support volunteers identifying new classes of images. Specifically, we plan to explore the use of computerized clustering of images to support advanced users. For instance, volunteers currently create collections of similar glitches that they have encountered. We plan to develop algorithms to enable them to search for additional glitches that resemble those in the collection. Further work is planned to enable volunteers to explore the detector status data recorded by LIGO to track down the cause of glitches. In addition, we are seeking ways to support teams of volunteers collaborating on the project. Finally, we continue to seek additional areas for collaboration. For example, a commonly studied topic in citizen science is volunteer motivation. Could ML be used to identify tasks that volunteers will find particularly interesting?

In summary, through our work on the Gravity Spy system, we seek to test and develop ways to develop human-machine partnerships for complex work, going beyond a focus on simple automation. While our focus has been on citizen science, we expect our learnings from this project to be valuable as machines learn to support more kinds of work.

Acknowledgements

We thank the many citizen scientists whose engagement have made Gravity Spy possible and our collaborators on the project, including M. Zevin, S. Coughlin, S. Bahaadini, E. Besler, N. Rohani, S. Allen, M. Cabero, A. Katsaggelos, S. Larson, T. Littenberg, A. Lundgren, J. Smith and L. Trouille. Gravity Spy is partly supported by the National Science Foundation award INSPIRE 15-47880.

References

[1] Zevin, M., Coughlin, S., Bahaadini, S., Besler, E., Cabero, M., Rohani, N., . . . Kalogera, V. Gravity Spy: Methods for integrating aLIGO detector characterization, machine learning, and citizen science (Under review).

[2] LIGO Scientific Collaboration, VIRGO Collaboration, Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., . . . Zweizig, J. Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, 116, 6 (2016), 061102.

[3] Bahaadini, S., Rohani, N., Coughlin, S., Zevin, M. and Katsaggelos, A. K. Deep Multi-View Models For Glitch Classification (under review).

[4] Lintott, C. J., Schawinski, K., Keel, W., Van Arkel, H., Bennert, N., Edmondson, E., . . . Vandenberg, J. Galaxy Zoo: 'Hanny's Voorwerp', a quasar light echo? *Monthly Notices of the Royal Astronomical Society*, 399, 1 (October 11, 2009 2009), 129-140.