ALBUS: a machine learning algorithm for gravitational wave burst searches

1st Vincent Boudart
STAR Institute
Bâtiment B5, Université de Liège
Sart Tilman B4000 Liège, Belgium
vboudart@uliege.be

2nd Maxime Fays
STAR Institute
Bâtiment B5, Université de Liège
Sart Tilman B4000 Liège, Belgium
Maxime.Fays@uliege.be

***Abstract*—Minute-long gravitational-wave bursts are transient events that arise from a wide variety of astrophysical phenomena. In opposition to the already-detected compact binary mergers, minute-long bursts are poorly modeled, preventing the use of matched filtering techniques. These events are thus probed through the excess-power method, consisting in searching for a local excess of power in the time-frequency space correlated between detectors. The search for minute-long GWs can then be viewed as a search for high-value clustered pixels within an image, which has already been tackled extensively in the machine learning realm. In this paper, we use a convolutional neural network as an anomaly detection tool. We show that our algorithm can reach a pixel-wise detection despite trained with minimal assumptions.**

***Index Terms*—gravitational waves, bursts, machine learning, CNN, anomaly detection, time-frequency maps**

I. Introduction

On September 14, 2015, the collision of two black holes was revealed through their gravitational-wave (GW) signal for the first time by the Advanced LIGO [1] detectors. Since then, the Advanced Virgo detector [2] has joined the efforts of LIGO to unravel more than 90 compact binary coalescence (CBC) events [3], among which the newly detected black hole-neutron star [4] and binary neutron collisions [5]. As both Virgo and LIGO have planned new sensitivity improvements in the coming years, new gravitational wave sources are expected to be observed. Among these expected candidates, unmodeled GW transients, also known as bursts, are a prime target for the next observing run. Bursts include a variety of astrophysical phenomena, such as supernova [6], nonaxisymmetric deformations in magnetars [7] and gamma-ray bursts [8]. They include both short (< 10 seconds) and long (from 10 to a few hundreds of seconds) events. The uncertainties in their physical models force us to make minimal assumptions on GW waveform characteristics. The existing models cannot consequently be taken as accurate patterns to be recognized and are thus used as tests for pipelines rather than actual targets of the search. A template-free approach has thus been developed, known as excess-power method, to search for GW transients with minimal assumptions. It consists in cross-correlating the data from two or more detectors into time-frequency maps (TF maps) or spectrograms and search for high-value clusters of pixels. The current search algorithms that implement this method do not use neural networks as their main detection engine, leading to hours of tuning and lack of speed. This works thus aims at providing a new tool to the search for minute-long GW transients that takes advantage of the speed and robustness of convolutional neural networks (CNNs).

V. B. is supported by the Gravitational Wave Science (GWAS) grant funded by the French Community of Belgium. M. F. acknowledges the support of the Fonds de la Recherche Scientifique-FNRS, Belgium, under grant No. 4.4501.19.

II. Data

The time-frequency maps used in the excess-power method are produced thanks to the cross-correlation, also known as coherence, of the data from at least two detectors. In this work, we use the data from the two Advanced LIGO interferometers, located at Hanford Livingston in the US. The coherence between two signals x and y is expressed as:

$$C_{x\,y}(f) = \frac{|G_{x\,y}(f)|^2}{G_{x\,x}(f)\,G_{y\,y}(f)}, \tag{1}$$

where $G_{xy}(f)$ is the cross-spectral density (CSD) between signals x and y, while $G_{xx}(f)$ and $G_{yy}(f)$ are the power spectral densities (PSD) of x and y respectively. The cross-spectral density is defined as:

$$G_{x\,y}(f) = \int_{-\infty}^{\infty} \left[\lim_{T\to\infty} \frac{1}{T} \int_{-\infty}^{\infty} x_T(t-\tau)\, y_T(t) dt \right] e^{-i2\pi f\tau} d\tau \tag{2}$$

A vector of coherence values versus frequency bins can be generated by evaluating expression (1) at several sampled frequencies. Then, we apply Welch's method [9] to small overlapping subsets of our original signal to produce a full time-frequency array. This is equivalent to repeatedly updating the coherence vector and compiling its time evolution as a single map. The resulting array is called a spectrogram or a TF map. The last pre-processing step consists in whitening [10] the spectrogram. To constitute a sufficient number of spectrograms containing only detector noise, we use time-slides [11]. It consists in shifting the detector data by time delays larger than the time of flight of GWs between detectors to guarantee our cross-correlated data to contain only detector noise. We use data from the third LIGO-Virgo

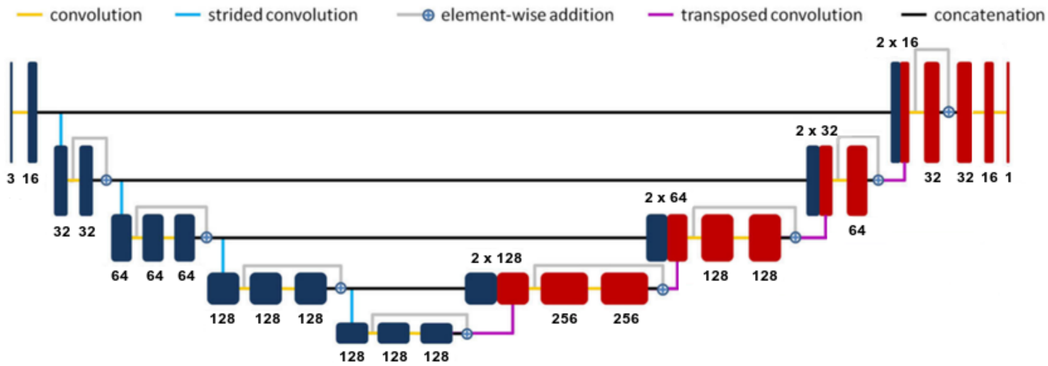

Fig. 1. Architecture of ALBUS

observing run to produce 4500 TF maps, constituting our background dataset. We use a time resolution of 6 seconds and a frequency bin of 2 Hz. Taking a 1000s data stream from the spanning frequencies up to 2048 Hz, results in spectrograms with dimensions of *166*x*1025* pixels.

Neural networks are particularly good at recognizing shapes and patterns in images (YOLO [12], AlexNet [13], etc) and are therefore well suited for the detection of minute-long GW transients in TF maps. However, we cannot use the existing models as accurate patterns to recognize and our algorithm should also detect unexpected GW signals. That is why we make use of the Scipy library [14] to generate chirp signals sufficient close to the actual waveform models. This allows us to generate random chirps to cover the whole time-frequency plane with varying parameters such as the duration, the frequency bandwidth, and the frequency and energy evolution. The chirping signals are then injected into noise-only spectrograms with 9 levels of visibility, defined as:

$$V = \sum_{i,j} \left(S_{ij} - N_{ij} \right) \tag{3}$$

where N_{ij} is a noise-only spectrogram and S_{ij} refers to the same spectrogram in which a chirp has been injected. The sum is carried over all the pixels (i, j) in the map. The visibility has been introduced to ensure chirps to be visible in the TF maps, which is not guaranteed with existing injection methods relying on intensity criteria based solely on signal characteristics. We choose 9 intensity levels in order to cover a quite large intensity range. We use this intensity criterion to build our second dataset, containing 4500 TF maps. Our final dataset thus contains 9000 spectrograms.

III. Methodology

In order to detect high-value clusters of pixels in our TF maps, we will make use of CNNs. Most of the CNNs that detects objects also involve a classification task [12] [13]. However, we want to highlight the pixels of the burst signals rather than assigning a label to the whole spectrogram. Therefore, inspired by the authors in [15], we build a network that returns a pixel-by-pixel localization map. The network, shown in Fig. 1, is made up of two parts, a downscaling part that keeps the useful information through its different layers, and an upscaling part that aims at localizing precisely this information in a map with the same dimensions as the input. The connections between the downscaling and upscaling parts help both the gradients to flow in the network and the network to learn the precise position of the patterns detected.

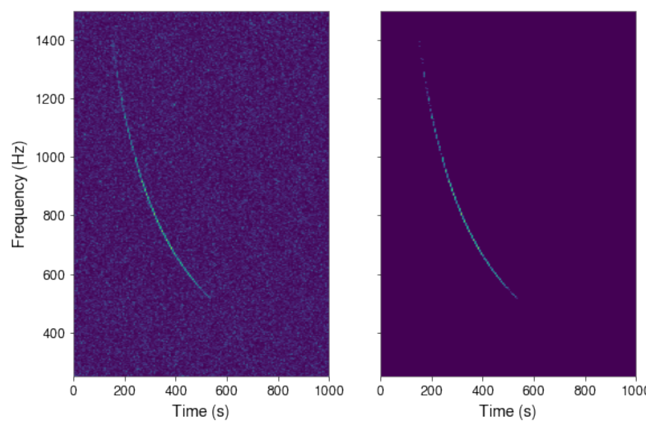

Fig. 2. Background spectrogram in which a chirp signal is injected (left) and its associated target map used for the training phase (right).

The training method then consists in minimizing a loss between the output of the network and a target map, so that the former keeps approaching the latter as the training progresses. To form the target maps corresponding to the TF maps in our dataset, we set a threshold on the spectrogram pixels corresponding to the 99th percentile of the values. This is equivalent to keeping the top 1% pixels showing the highest values. We then normalize the target map. This procedure leads to a target map that follows the intensity evolution of the signals through the input map. An example of a spectrogram containing a chirp signal and its corresponding target map can be seen in Fig. 2.

To train our network, the ADAM optimizer [16] has been chosen with a learning rate of 10^{-4}. We use the Mean Squared Error loss as minimization loss, guaranteeing a well-behaved

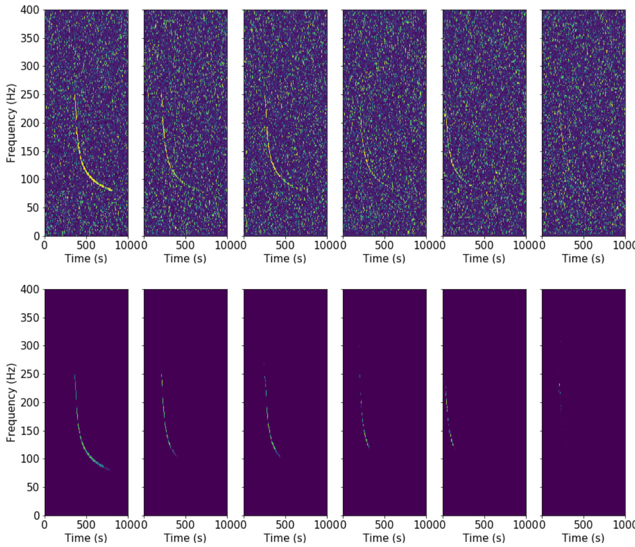

Fig. 3. Detection performance of ALBUS on 6 different visibility levels (from left to right : 60, 40, 30, 20, 16 and 12) for the waveform model *GRBplateau* [8]. The top panel shows the input images and the lower panel shows the output of ALBUS.

loss surface. The batch size is set to 20 where one half is taken from the background images and the other half from the chirp images. The validation is made of 10% of the original dataset. We decided to stop the training after 30 epochs because both losses started to reach a plateau, indicating that the network do not improve anymore.

IV. Results

Once the network is trained on random chirp signals, we need to evaluate how good it is at detecting minute-long models. Fig. 3 shows the output of ALBUS for one of the selected models in [17] across 6 different intensities. The signals are well recognized even at low intensity and the variation of intensity in the input TF map is also seen in the localization map. Our network is not only looking at the pixels having a high value but also at the connectivity between these pixels. It then naturally looks prolongs the main structure to catch pixels following the general trend of the signal. Such a propriety can be a relevant tool to reject background images showing isolated hot pixels.

V. Discussion and conclusion

We have shown that convolutional neural networks can be applied to the search for minute-long gravitational wave transients in the time-frequency space of the cross-correlated LIGO noise. Our approach allows a fast and pixel-precise identification of the long-duration signals with no training on the latter.

The threshold for the detection of burst signals is determined by the highest background candidates, i.e. the background candidates that show the highest detection score as defined by a particular pipeline. Usually, the highest candidates are identified after analyzing at least 50 years of background data, making more than 1 million spectrograms to process. In order to rank these candidates and automate the detection, a detection statistics needs to be defined in follow-up works.

Acknowledgment

This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. The authors are grateful for computational resources provided by the LIGO Laboratory and supported by the National Science Foundation Grants No. PHY-0757058 and No. PHY-0823459.

References

[1] J. Aasi et al., "Advanced LIGO", Class. Quant. Grav., vol. 32, num. 7, pp. 074001, Mars 2015.

[2] F. Acernese et al., "Advanced Virgo: a second-generation interferometric gravitational wave detector", Class. Quant. Grav., vol. 32, num. 2, pp. 024001, 2015.

[3] R. Abbott et al., The LIGO Scientific, the Virgo and the KAGRA Collaborations, "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run", arXiv:2111.03606, 2021.

[4] R. Abbott et al., The LIGO Scientific, the Virgo and the KAGRA Collaborations, "Observation of Gravitational Waves from Two Neutron Star–Black Hole Coalescences", Astrophys. J. Lett., vol. 915, num. 1, pp. L5, 2021.

[5] B.P. Abbott et al., "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral", Phys. Rev. Lett., vol. 119, pp. 161101, 2017.

[6] M. H. P. M. van Putten et al., "Prospects for multi-messenger extended emission from core-collapse supernovae in the Local Universe", Eur. Phys. J. Plus, vol. 134, num. 10, pp. 537, 2019.

[7] S. Dall'Osso, B. Giacomazzo, R. Perna and L. Stella, "Gravitational Waves from Massive Magnetars Formed in Binary Neutron Star Mergers", Astrophys. J., vol. 798, num. 1, pp. 25, 2015.

[8] A. Corsi and P. Mészaros, "Gamma-Ray Burst Afterglow plateaus and gravitational waves: multi-messenger signature of a millisecond magnetar?", Astrophys. J., vol. 702, num. 2, pp. 1171-1178, 2009.

[9] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms", IEEE Transactions on Audio and Electroacoustics, vol. 15, num. 2, pp. 70-73, 1967.

[10] E. Cuoco et al., "On-line power spectra identification and whitening for the noise in interferometric gravitational wave detectors", Class. Quant. Grav., vol. 18, num. 9, pp. 1727-1751, 2001.

[11] M. Was et al., "On the background estimation by time slides in a network of gravitational wave detectors", Class. Quant. Grav., vol. 27, num. 1, pp. 015005, 2010.

[12] J. Redmon, S. Divvala, R. Girshick and A. Farhadi., "You Only Look Once: Unified, Real-Time Object Detection", in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE Computer Society, pp. 779-788, June 2016.

[13] A. Krizhevsky, I. Sutskever and G. E. Hinton., "ImageNet Classification with Deep Convolutional Neural Networks", in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, Curran Associates Inc., pp. 1097-1105, 2012.

[14] P. Virtanen et al., "SciPy 1.0: fundamental algorithms for scientific computing in Python", Nature Methods, Springer Science and Business Media LLC, vol. 17, num. 3, pp. 261-272, 2020.

[15] F. Xing et al., "Towards pixel-to-pixel deep nucleus detection in microscopy images", BMC Bioinformatics, vol. 20, pp. 472, 2019.

[16] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", in *3rd International Conference on Learning Representations ICLR Conference Track Proceedings*, 2015.

[17] R. Abbott et al., The LIGO Scientific and the Virgo and the KAGRA Collaborations, "All-sky search for long-duration gravitational-wave bursts in the third Advanced LIGO and Advanced Virgo run", Phys. Rev. D., vol. 104, pp.102001, 2021.