



GWskyNet: A Real-time Classifier for Public Gravitational-wave Candidates

Miriam Cabero¹ , Ashish Mahabal^{2,3} , and Jess McIver¹ ¹Department of Physics and Astronomy, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada; micamu@phas.ubc.ca²Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA³Center for Data Driven Discovery, California Institute of Technology, Pasadena, CA 91125, USA

Received 2020 September 22; revised 2020 October 19; accepted 2020 October 24; published 2020 November 19

Abstract

The rapid release of accurate sky localization for gravitational-wave (GW) candidates is crucial for multi-messenger observations. During the third observing run of Advanced LIGO and Advanced Virgo, automated GW alerts were publicly released within minutes of detection. Subsequent inspection and analysis resulted in the eventual retraction of a fraction of the candidates. Updates could be delayed by up to several days, sometimes issued during or after exhaustive multi-messenger follow-up campaigns. We introduce GWskyNet, a real-time framework to distinguish between astrophysical events and instrumental artifacts using only publicly available information from the LIGO-Virgo open public alerts. This framework consists of a non-sequential convolutional neural network involving sky maps and metadata. GWskyNet achieves a prediction accuracy of 93.5% on a testing data set.

Unified Astronomy Thesaurus concepts: [Convolutional neural networks \(1938\)](#); [Gravitational wave astronomy \(675\)](#); [Observational astronomy \(1145\)](#)

1. Introduction

The LIGO Scientific and Virgo Collaborations have so far reported 15 confirmed gravitational-wave (GW) signals from the mergers of compact objects (Abbott et al. 2019a, 2020b, 2020c, 2020d, 2020e). Depending on the nature of the objects and the environment surrounding the binary, merging compact binaries can also emit electromagnetic radiation and neutrinos. In the second observing run (O2) of Advanced LIGO and Advanced Virgo, low-latency GW candidates were distributed to partner astronomers for multi-messenger follow-up (Abbott et al. 2019b). This partnership resulted in the combined measurement of electromagnetic and GW emission from GW170817 (Abbott et al. 2017a, 2017b), which revolutionized our understanding of a wide range of physics related to merging neutron stars.

Accurate sky localization and the rapid release of GW candidates are crucial for multi-messenger observations of astrophysical events. The former is achieved through a global network of GW observatories: the two Advanced LIGO detectors in the United States (Hanford and Livingston) together with the Advanced Virgo detector in Italy have been able to pinpoint the sky location of nearby GW candidates to areas smaller than 50 square degrees. The latter is achieved through fast and efficient data analysis pipelines. During the third observing run (O3) of Advanced LIGO and Advanced Virgo, low-latency GW alerts were publicly available within minutes through the Gamma-ray Coordinates Network (GCN)⁴ and on the GW-candidate event DataBase (GraceDB).⁵ For compact binaries, the information released in Open Public Alerts (OPA) included⁶: the false-alarm rate (FAR) estimated by the search pipelines, the inferred sky position and distance of the source (Singer & Price 2016; Singer et al. 2016a), the probability of astrophysical origin with classification according to different mass regions (Kapadia et al. 2020), and the probability of neutron star matter content (Foucart et al. 2018;

Chatterjee et al. 2020). Rapid and public information release enabled every astronomer with access to telescopes or neutrino detectors to search for counterparts to the GW candidates (Andreoni et al. 2019; Antier et al. 2020; Coughlin 2020; Coughlin et al. 2020; Graham et al. 2020; Hussain et al. 2020; Kasliwal et al. 2020; Vieira et al. 2020).

Preliminary OPAs in O3 were generated automatically when low-latency GW search pipelines reported an effective FAR below either 1 per 2 months for the matched-filtering searches or 1 per year for the unmodeled burst searches.⁷ This FAR threshold targets an overall astrophysical purity of 90% (Abbott et al. 2020a), hence some candidates will likely originate from instrumental artifacts. GW candidates recognized as the result of an instrumental artifact after human inspection were retracted by LIGO and Virgo. Retractions usually appeared within an hour of the initial alert, although in some occasions it could take significantly longer, up to several days. Some surviving candidates were updated when further analyses provided more accurate information, sometimes changing the binary classification. LIGO and Virgo released a total of 80 OPAs during O3, of which 24 were subsequently retracted. Of the remaining 56 candidates, 26 were reported with a non-vanishing probability of being of terrestrial origin, and four have been confirmed so far as real astrophysical events by the LIGO Scientific and Virgo Collaborations (Abbott et al. 2020b, 2020c, 2020d, 2020e).

Telescope time is competitive and astronomers have to carefully select which candidates to pursue. Some of the retracted events in O3 had very accurate sky localization or very low FARs, qualities that make a candidate promising for follow-up. For instance, the binary neutron star candidate S190822c⁸ was reported with a FAR of 1 per 5×10^9 years. Within less than one hour, three different follow-ups with various instruments were reported (Cook et al. 2019;

⁴ <https://gcn.gsfc.nasa.gov/>⁵ <https://gracedb.ligo.org/>⁶ <https://emfollow.docs.ligo.org/userguide/index.html>⁷ <https://emfollow.docs.ligo.org/userguide/analysis/index.html#alert-threshold>⁸ <https://gracedb.ligo.org/superevents/S190822c/view/>

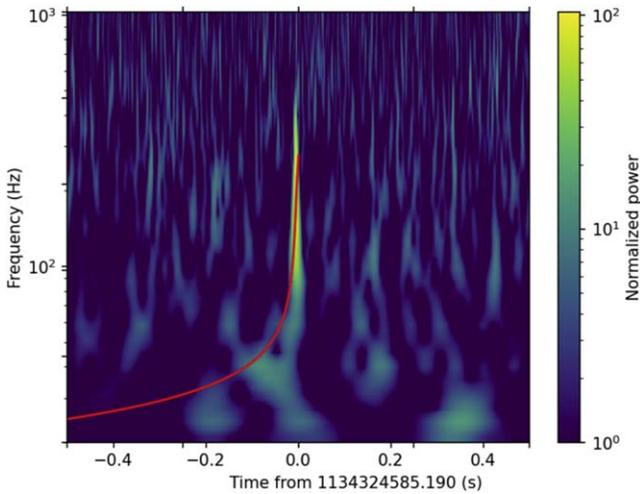


Figure 1. Time-frequency representation of a blip glitch included in the noise set. The red curve represents the GW template that triggered the 2-OGC search. This template corresponds to a binary black hole with masses $m_1 = 40M_\odot$ and $m_2 = 19M_\odot$. The network S/N reported by the search at the time of this glitch is $\rho_N = 11.3$.

Lipunov et al. 2019; Hussain 2019). The candidate was then retracted 1.5 hours after the preliminary alert. Depending on the time of the event and the available instruments, telescopes can be steered to point toward the sky location of GW candidates within minutes (though counterpart identification, if achieved, can take a much longer time depending on the area of uncertainty of the event). How can astronomers best decide which candidates to follow up based on the information released in OPAs?

In this work we introduce *GWSkyNet*, a low-latency classifier to separate real astrophysical events from terrestrial artifacts, supplementing information provided by the LIGO Scientific and Virgo Collaborations. Contrary to recent GW classifiers (Chatterjee et al. 2020; Kim et al. 2020), which rely on access to LIGO-Virgo data, *GWSkyNet* only requires publicly available data products and hence can be used by any astronomer. Built with state-of-the-art machine-learning algorithms such as non-sequential convolutional neural networks and multiple input machine-learning models, *GWSkyNet* achieves an accuracy of 93.5% on a testing data set.

This Letter is organized as follows. In Section 2 we describe the data set constructed to train and test the machine-learning algorithm. In Section 3 we describe the data preparation and the features selected for training, as well as the architecture chosen for the machine-learning model. Section 4 presents the results obtained with the fully trained model, including predictions for the O3 candidates in GraceDB. Finally, we summarize our findings in Section 5.

2. The Data Set

A data set with a large, balanced number of labeled examples is key to the success of supervised machine learning. With GW astronomy still in its infancy, the currently confirmed population of astrophysical GW events is insufficient to construct a robust training set. This section describes the methods used to simulate a population of real candidates and identify a population of noise candidates in real data.

2.1. Real Events

A population of simulated real events is constructed using gravitational waveform models for compact binaries, available in the *LALSuite* (LIGO Scientific Collaboration 2018) package. Masses are drawn from a uniform distribution in the total mass of the binary, with component masses in the range $m \in [1.2, 2.2]M_\odot$ for neutron stars and $m \in [3, 100]M_\odot$ for black holes. For simplicity, spins are restricted to be along the direction of orbital angular momentum. Spin magnitudes are constrained to $\chi \leq 0.05$ for neutron stars and $\chi \leq 0.99$ for black holes. With these conditions, we generate three types of binaries: binary neutron stars, neutron-star black-hole binaries, and binary black holes. Sources are distributed uniformly in co-moving volume with a maximum distance of up to 1.5 Gpc, depending on the binary type and the detectors sensitivities. Sky locations are randomly distributed, and the inclination angle, ι , is distributed uniformly over $\arccos \iota$.

To avoid contaminating the training set of real events with noise artifacts present in GW data, simulated signals are added into Gaussian noise colored with the power spectral densities (PSDs) of Advanced LIGO and Advanced Virgo. Three PSDs associated with three different GW events in the Gravitational-Wave Open Science Center (GWOSC; Abbott et al. 2019c) are chosen to represent different sensitivities of the detectors. The PSDs at the time of GW150914 and GW170104 represent the two Advanced LIGO detectors (Hanford and Livingston) in the first (O1) and second (O2) observing runs, respectively. To incorporate the Advanced Virgo detector, we additionally use the PSD at the time of GW170729, when Virgo had joined the observing run and the two LIGO detectors had significantly different sensitivities with respect to the rest of O2 (Abbott et al. 2019a). Simulated signals are distributed equally among these three different representative sensitivities.

To select only those signals that would be detectable by GW search pipelines, we impose the following conditions on the matched-filtering signal-to-noise ratio (S/N) of the signal. First, the signal must be detected in at least two observatories with $S/N \rho \geq 4.5$. Second, the network S/N, given by the quadrature sum of the individual S/Ns satisfying the first condition, must be $\rho_N \geq 7$. After applying these conditions, the final data set of simulated real events consists of 2857 candidates. While this population of real events can be easily increased, it is important to keep a good balance in the training set between real and noise candidates Johnson & Khoshgoftaar (2019).

2.2. Noise Events

Matched-filtering searches for GWs from compact binaries select candidates by requiring a temporal coincidence between detectors (Adams et al. 2016; Chu 2017; Messick et al. 2017; Nitz et al. 2018). The coincidence window is set by the maximum light-travel time between detectors plus the uncertainty in the inferred coalescence time in each detector. Noise transients (glitches) are independent between detectors and will rarely happen simultaneously. However, random noise fluctuations can have a sufficiently high S/N to trigger a coincidence in the search pipelines at the time of a glitch. If the resulting network S/N is greater than the pipeline's predetermined threshold (and the FAR is sufficiently low), the glitch can be identified as a GW candidate by the low-latency searches and generate an automatic OPA.

Two main ingredients are required to identify a population of noise events: a collection of known glitches in real detector data, and a collection of corresponding candidates (triggers) from the GW search pipelines. For the former, we use the collection of blip glitches from Cabero et al. 2019 and the O1 and O2 collections of glitches from the Gravity Spy classifier (Zevin et al. 2017; Bahaadini et al. 2018; Coughlin et al. 2019). For the latter, we use the search triggers from the second Open Gravitational-wave Catalog (2-OGC; Nitz et al. 2019). Each of these catalogs contain thousands of glitches and triggers, but only a small fraction of glitches will be associated with a coincident trigger and satisfy the S/N requirements imposed by the search pipelines.

While the temporal coincidence condition is consistent among GW searches, different search pipelines use different S/N requirements. Here, we apply the same threshold as for the simulated signals of the real events set: individual S/N $\rho \geq 4.5$ in at least two detectors, and network S/N $\rho_N \geq 7$. When a glitch is within ± 0.5 s of a 2-OGC coincident search trigger in the same detector, the candidate is selected for further evaluation. All the selected glitches are then visually inspected to ensure that the glitch and the waveform overlap. Figure 1 shows an example of a noise candidate, with the gravitational waveform of the coincident trigger overlaid on top of the glitch. The final data set of noise events consists of 1267 candidates, which makes 30% of the total training set. In the future, with the release of the O3 data sets, additional glitches will become available for the training set.

3. Machine-learning Classifier

Convolutional Neural Networks (CNNs) are state-of-the-art deep learning algorithms commonly used for image classification. We employ a non-sequential CNN architecture to design GWSkyNet, a real/noise classifier for real-time GW-candidate alerts. The main inputs to the classifier are the 3D sky map images released in the LIGO-Virgo OPAs. During O2, a method based on the mutual information distance of 2D GW sky maps was included in part of the LIGO-Virgo human vetting to distinguish between signal and noise (Essick 2017; Essick & Urban 2017). Here we introduce the first CNN classifier for GW sky maps, with a more diverse set of inputs. We find that additional numerical and categorical data, such as the distance to the source and the detector network, are important for sky map classification. Therefore, we construct a multiple input CNN model that accepts mixed data features. In this section, we first explain the features chosen for training and then describe the specifications of the network architecture.

3.1. Training Features

We use BAYESTAR (Singer & Price 2016; Singer et al. 2016a, 2016b) to reconstruct the 3D sky localization of the source for each candidate in the training set. The sky localization is stored as a Flexible Image Transport System (FITS) file. BAYESTAR FITS files contain four columns that represent four channels of a Hierarchical Equal Area isoLatitude Pixelization (HEALPIX) all-sky image. The first column (PROB) represents the 2D probability sky map. The other three columns (DISTMU, DISTSTD, and DISTNORM) are used to calculate the 3D probability volume map. The BAYESTAR FITS header provides additional metadata about the candidate. Of particular interest for the purpose of this work

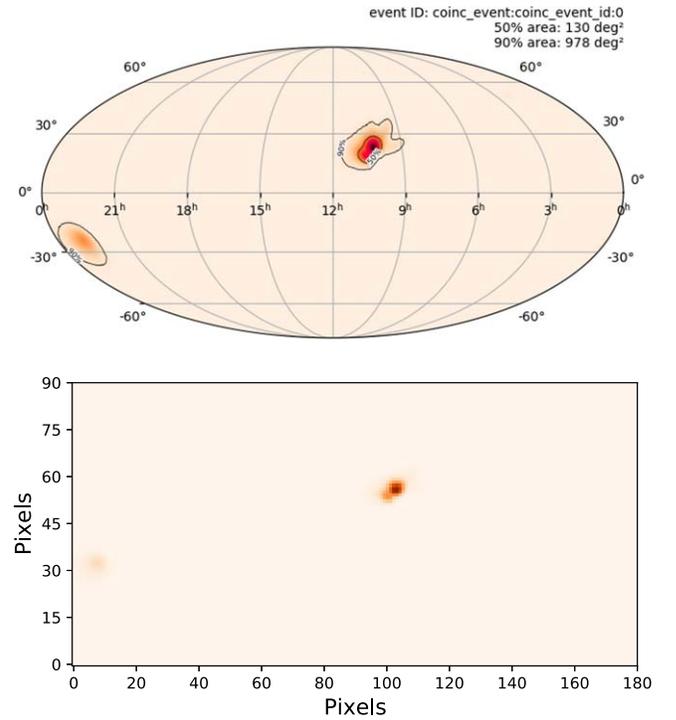


Figure 2. Sky map obtained with BAYESTAR for the blip glitch shown in Figure 1. Top: image as it would appear in an OPA on GraceDB. Bottom: same image after re-projection, normalization, and downsampling onto a 90×180 grid. This image is used for training the machine-learning algorithm.

are the posterior mean distance of the source and the list of GW instruments that triggered the candidate.

For training, we transform each column in the FITS file (the sky map images) to a rectangular projection with 180×360 pixels. Pixels with invalid values in the volume columns, as defined in Singer et al. (2016b), or with not-a-number (NaN) values originating from the re-projection, are replaced with zero. Each image is then normalized independently dividing by its maximum value. To reduce computations during training, images are downsampled to 90×180 pixels using the maximum pooling method. Figure 2 shows an example of a sky map image after re-projection, normalization, and downsampling. The three volume images are stacked together to form a three-channel cuboid that will be used as input for the classifier. The final set used as input for training the machine-learning algorithm consists of eight features:

1. Sky map image, shape (90, 180, 1),
2. Stacked volume images, shape (90, 180, 3),
3. Posterior mean distance in Mpc, normalized dividing by the maximum value in the training set,
4. Detector network in 3-bit multi-hot encoding format, with one bit reserved for each detector (LIGO Hanford, LIGO Livingston, and Virgo),
5. Four normalization factors (one for each image), normalized dividing by the corresponding maximum value in the training set.

3.2. GWSkyNet Architecture and Training

We use TensorFlow (Abadi et al. 2015) and the functional API in Keras to construct a multiple input model with eight branches. The first two branches are CNNs designed to operate

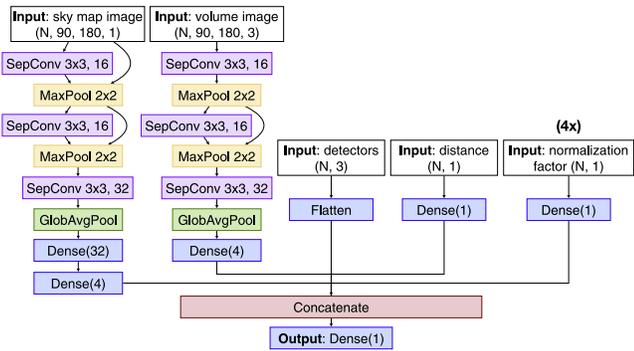


Figure 3. Architecture of the multiple input model constructed to classify GW sky maps. The first two branches are CNNs with residual connections for image data. The shape of the input data is indicated in parenthesis, with N being the number of examples in the training set. The numbers in the SeparableConv2D (SepConv) and MaxPool layers indicate the kernel size in pixels and the number of filters (when applicable). The number in the Dense layers indicates the number of units.

over the image data (sky map and three-channel volume images). To improve the performance of the classifier, residual connections are added to each CNN branch, with two residual blocks in the sky map branch and one block in the volume branch. The other six branches are simple multi-layer perceptrons for the categorical and numerical data. All branches are concatenated together into the final multiple input model. A complete description of the specifications of the network can be seen in Figure 3. Rectified Linear Unit (ReLU) activation functions are used for all hidden layers, and a sigmoid activation function is used for the output layer.

The training set introduced in Section 2, with 4124 candidates, is split into 81% for training, 9% for validation, and 10% for testing. The algorithm is trained on GPUs available through Google Colaboratory,⁹ and hyperparameter optimization is performed with KerasTuner (O’Malley et al. 2019). Training is enabled for 500 epochs, employing the binary cross-entropy loss function, the Adam optimizer with a learning rate of 0.002, and batch size of 32. Our use of GlobalAveragePooling ensured a small model (3687 free parameters in total), helping us avoid overfitting.

4. Performance and Predictions

The accuracy, which measures the fraction of correct predictions over the total number of classifications, is a performance metric commonly used for binary classifiers. With a Real-Noise (RN) score threshold of $RN \geq 0.5$ denoting a real source, the fully trained GWSkyNet achieves an accuracy of 93.5% on a test set with 413 candidates, independent from the training set. We also use precision and recall values as additional performance measures. The precision, defined as the fraction of true predicted positives over the total number of predicted positives, measures how often the model is correct when it predicts that a candidate is real. The recall or true-positive rate, defined as the fraction of true predicted positives over the total number of actual positives, measures how many real candidates are predicted correctly. On the test set, the classifier achieves a precision of 97.7% and a recall of 92.4% for $RN \geq 0.5$. Figure 4 shows the False Positive Rate (FPR) and False Negative Rate (FNR) on the test set for different RN score thresholds. The FPR varies very

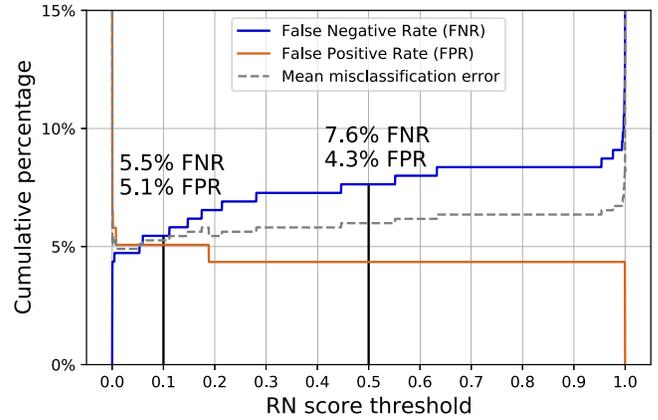


Figure 4. False Positive Rate (FPR) and False Negative Rate (FNR) as functions of the Real-Noise (RN) score threshold (based on Figure 9 in Duev et al. 2019). At a score threshold $RN \geq 0.5$, GWSkyNet yields 7.6% FNR and 4.3% FPR. Lowering the threshold to $RN \geq 0.1$ reduces the FNR to 5.5%, with small variation on the FPR (5.1%). The intersection is at $RN \simeq 0.06$, with $FPR = FNR = 5.1\%$.

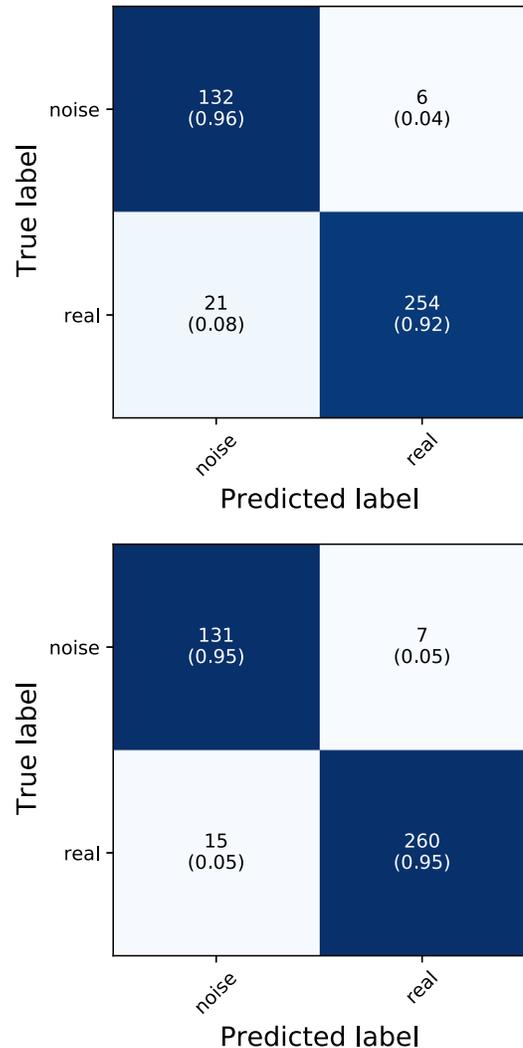


Figure 5. Confusion matrix, with normalized values in parenthesis, for thresholds $RN \geq 0.5$ (top) and $RN \geq 0.1$ (bottom) denoting a real source.

slowly with the RN threshold, so one can lower the FNR by lowering the threshold with only a small increase in the FPR. In Figure 5 we show the confusion matrices on the test set for two

⁹ <https://colab.research.google.com/notebooks/intro.ipynb>

different RN thresholds of 0.1 and 0.5. Based on one's preference for higher recall or higher precision, the threshold can be changed with the understanding that improving recall (precision) involves lowering precision (recall).

The O3 retracted candidates and confirmed astrophysical signals are a further test of the classifier performance. For each candidate, we use the BAYESTAR multi-order FITS files from GraceDB, because the training set was built using this pipeline. Of the four confirmed GW signals (Abbott et al. 2020b, 2020c, 2020d, 2020e), the classifier recognizes all of them as true astrophysical signals. Of the 22 retracted candidates identified by matched-filtering searches, 10 are correctly identified as instrumental artifacts and 12 are misclassified as astrophysical. Among the correct classifications is, for instance, the candidate S190822c mentioned in Section 1, which was retracted 1.5 hours after the preliminary alert. In future observing runs, GWSkyNet could provide immediate warning about the plausibility of a candidate, in some occasions sparing follow-up time invested on non-astrophysical candidates. The misclassifications on noise candidates can be explained from the data set used for training. On the one hand, most of the retracted candidates were identified by the GstLal and MBTAOnline search pipelines (as reported in GraceDB), while all the noise candidates in the training set originated from PyCBC search triggers. On the other hand, almost 70% of the retracted candidates included the Virgo detector, for which we could not include many noise events in the training set due to the small amount of available Virgo data. Incorporation of additional glitches in the training set, both by including results from different search pipelines as well as glitches from O3 data (when available), should lead to improved results. In addition, the visualization of layers of the trained network can provide an insight into the decisions reached. This is a nascent field and we will be exploring related methodology for building in more interpretability and explainability.

4.1. O3 GraceDB Predictions

We report predictions for the 51 unpublished candidates identified by at least one matched-filtering search for compact binaries. Using the initial BAYESTAR multi-order FITS files released in the preliminary alert, GWSkyNet predicts that 41 candidates are of astrophysical origin, and 10 candidates are due to terrestrial artifacts: S190510g, S190718y, S190901ap, S190910d, S190910h, S190930t, S191105e, S191205ah, S200213t, and S200302c. All 10 candidates have a non-zero probability of terrestrial origin, as estimated by LIGO and Virgo methods (Kapadia et al. 2020). However, two of these candidates are single-detector triggers (S190910h and S190930t). As there are no examples in the training set of GW candidates observed by only one detector, predictions on single-detector candidates should be taken with caution. The release of the O3 catalog by the LIGO Scientific and Virgo Collaborations will determine the accuracy of these predictions.

5. Conclusion

In this work we have introduced GWSkyNet, a new machine-learning binary classifier for low-latency GW candidates that complements the information released by the LIGO Scientific and Virgo Collaborations. This classifier, based solely on publicly available data products, can be used by the broader astronomy community for fast decision-making on

candidate follow-up. In future observing runs, real-time predictions from GWSkyNet on preliminary automated alerts could identify noise candidates without the delay of human-based retractions or analysis updates. More importantly, astronomers will receive additional information on the astrophysical probability of non-retracted candidates.

The classifier has been trained using LIGO and Virgo detector glitches as noise candidates and simulated signals of compact binary mergers as real candidates. With a threshold $RN \geq 0.5$ denoting a real source, the accuracy achieved on a test data set, which is independent of the training set, is 93.5%. Additionally, we reported predictions from GWSkyNet on O3 GraceDB candidates. Of the 51 unpublished non-retracted candidates, the classifier identified 10 as not astrophysical, with the caveat that two of those are single-detector candidates.

Increasing the size of the training set will be crucial to improve the classifier performance. In future work, we will incorporate additional candidates from different search pipelines, as well as from the O3 data set. With a larger training set in hand, we will construct a multi-class classifier capable of distinguishing between different classes of binary mergers. We provide the classifier and its data sources on GitHub¹⁰ and preserve the current version at the CaltechDATA repository: doi:10.22002/D1.1659.

We are thankful to the Gravity Spy team, and in particular Scott Coughlin, for sharing the collection of O1 and O2 glitches. We are also thankful to Deep Chatterjee, Tito Dal Canton, Derek Davis, Daryl Haggard, Ian Harry, Fergus Hayes, Alexander H. Nitz, Leo P. Singer, Nicholas Vieira, and the anonymous referee for useful comments and discussions.

M.C. and J.M. acknowledge funding from the Natural Sciences and Engineering Research Council of Canada (NSERC). A.M. acknowledges support from the NSF (1640818, AST-1815034). A.M. and J.M. also acknowledge support from IUSSTF (JC-001/2017).

This research has made use of data, software, and/or web tools obtained from the Gravitational Wave Open Science Center (<https://www.gw-openscience.org>), a service of LIGO Laboratory, the LIGO Scientific Collaboration, and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN), and the Dutch Nikhef, with contributions by Polish and Hungarian institutes.

ORCID iDs

Miriam Cabero  <https://orcid.org/0000-0003-4059-4512>
Ashish Mahabal  <https://orcid.org/0000-0003-2242-0244>
Jess McIver  <https://orcid.org/0000-0003-0316-1355>

References

- Abadi, M., Agarwal, A., Barham, P., et al. 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, <http://tensorflow.org/>
- Abbott, B., Abbott, R., Abbott, T. D., et al. 2017a, *PhRvL*, **119**, 161101
- Abbott, B., Abbott, R., Abbott, T. D., et al. 2017b, *ApJL*, **848**, L12
- Abbott, B., Abbott, R., Abbott, T. D., et al. 2019a, *PhRvX*, **9**, 031040
- Abbott, B., Abbott, R., Abbott, T. D., et al. 2019b, *ApJ*, **875**, 161
- Abbott, B., Abbott, R., Abbott, T. D., et al. 2020a, *LRR*, **21**, 3
- Abbott, B., Abbott, R., Abbott, T. D., et al. 2020b, *ApJL*, **892**, L3
- Abbott, R., Abbott, T. D., Abraham, S., et al. 2019c, arXiv:1912.11716

¹⁰ <https://github.com/GWML/GWSkyNet>

- Abbott, R., Abbott, T. D., Abraham, S., et al. 2020c, *PhRvD*, **102**, 043015
- Abbott, R., Abbott, T. D., Abraham, S., et al. 2020d, *ApJL*, **896**, L44
- Abbott, R., Abbott, T. D., Abraham, S., et al. 2020e, *PhRvL*, **125**, 101102
- Adams, T., Buskulic, D., Germain, V., et al. 2016, *CQGra*, **33**, 175012
- Andreoni, I., Goldstein, D. A., Anand, S., et al. 2019, *ApJL*, **881**, L16
- Antier, S., Agayeva, S., Almualla, M., et al. 2020, *MNRAS*, **497**, 5518
- Bahaadini, S., Noroozi, V., Rohani, N., et al. 2018, *Information Sciences*, **444**, 172
- Cabero, M., Lundgren, A., Nitz, A. H., et al. 2019, *CQGra*, **36**, 155010
- Chatterjee, D., Ghosh, S., Brady, P. R., et al. 2020, *ApJ*, **896**, 54
- Chu, Q. 2017, PhD thesis, The Univ. Western Australia
- Cook, D., van Sistine, A., Singer, L., et al. 2019, GCN, **25441**, 1
- Coughlin, M. W. 2020, *NatAs*, **4**, 550
- Coughlin, M. W., Dietrich, T., Antier, S., et al. 2020, *MNRAS*, **497**, 1181
- Coughlin, S., Bahaadini, S., Rohani, N., et al. 2019, *PhRvD*, **99**, 082002
- Duev, D. A., Mahabal, A., Masci, F. J., et al. 2019, *MNRAS*, **489**, 3582
- Essick, R. C. 2017, PhD thesis, Massachusetts Institute of Technology, <http://hdl.handle.net/1721.1/115024>
- Essick, R. C., & Urban, A. 2017, Skymap_Statistics, https://github.com/reedessick/skymap_statistics
- Foucart, F., Hinderer, T., & Nissanke, S. 2018, *PhRvD*, **98**, 081501
- Graham, M., Ford, K. E. S., McKernan, B., et al. 2020, *PhRvL*, **124**, 251102
- Hussain, R. 2019, GCN, 25439, 1
- Hussain, R., Vandenbroucke, J., & Wood, J. 2020, Proc. ICRC, **36**, 918
- Johnson, J. M., & Khoshgoftaar, T. M. 2019, *Journal of Big Data*, **6**, 27
- Kapadia, S. J., Caudill, S., Creighton, J. D. E., et al. 2020, *CQGra*, **37**, 045007
- Kasliwal, M. M., Anand, S., Ahumada, T., et al. 2020, arXiv:2006.11306
- Kim, K., Li, T. G., Lo, R. K., Sachdev, S., & Yuen, R. S. 2020, *PhRvD*, **101**, 083006
- LIGO Scientific Collaboration 2018, LIGO Algorithm Library - LALSuite Free Software (GPL) (doi:10.7935/GT1W-FZ16)
- Lipunov, V., Gorbvskoy, E., Kornilov, V., et al. 2019, GCN, **25440**, 1
- Messick, C., Blackburn, K., Brady, P., et al. 2017, *PhRvD*, **95**, 042001
- Nitz, A. H., Dal Canton, T., Davis, D., & Reyes, S. 2018, *PhRvD*, **98**, 024050
- Nitz, A. H., Dent, T., Davies, G. S., et al. 2019, *ApJ*, **891**, 123
- O'Malley, T., Bursztein, E., Long, J., et al. 2019, Keras Tuner, <https://github.com/keras-team/keras-tuner>
- Singer, L. P., Chen, H.-Y., Holz, D. E., et al. 2016a, *ApJL*, **829**, L15
- Singer, L. P., Chen, H.-Y., Holz, D. E., et al. 2016b, *ApJS*, **226**, 10
- Singer, L. P., & Price, L. R. 2016, *PhRvD*, **93**, 024013
- Vieira, N., Ruan, J. J., Haggard, D., et al. 2020, *ApJ*, **895**, 96
- Zevin, M., Coughlin, S., Bahaadini, S., et al. 2017, *CQGra*, **34**, 064003