Deep learning classification of the continuous gravitational-wave signal candidates from the time-domain $\mathcal{F}$-statistic search

Filip Morawski* and Paweł Ciecieląg

Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences, Bartycka 18, 00-716, Warsaw, Poland

Michał Bejger

APC, AstroParticule et Cosmologie, Université Paris Diderot,
CNRS/IN2P3, CEA/Irfu, Observatoire de Paris,
Sorbonne Paris Cité, F-75205 Paris Cedex 13, France and
Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences, Bartycka 18, 00-716, Warsaw, Poland

(Dated: July 17, 2019)

Many potential sources of gravitational waves still await for detection. Among them, particular attention is given to a non-axisymmetric neutron star. The emitted, almost monochromatic signal, is expected to be detected in the near future by LIGO and Virgo detectors. Although the gravitational waves waveform is well known, its small amplitude makes it extremely hard to detect. The accepted approach in searching for continuous gravitational waves is a matched filter technique, known as the $\mathcal{F}$-statistic method. The method consists in cross correlation of the collected data stream with signal templates in the frequency domain. Thus, for an all-sky search in which the parameters of the sources are not known, large number of templates have to be checked and therefore a large number of candidate gravitational-wave signals is produced and further analyzed. In this work, we propose deep learning as a fast method of classification for various types of candidates. We consider three types of signals: the Gaussian noise, the continuous gravitational wave, and the stationary line mimicking local artifacts in the detector. We demonstrate one and two-dimensional implementations of a convolutional neural network classifier. We present the limitations of our model with respect to the various signal-to-noise ratios and frequencies of the signal. The following work presents deep learning as a supporting method for the matched filtering detection pipeline.

I. INTRODUCTION

A. Gravitational waves

Gravitational waves (GWs) are distortions of the curvature of spacetime, propagating through it with the speed of light [1]. Experimental confirmation of their existence was recently provided by the LIGO and Virgo collaborations [2, 3] in the form of, till date, several binary black hole mergers [4–6], and one binary neutron star (NS) merger observations, the latter also electromagnetically bright [7]; see also the transient GW catalog for the summary of the LIGO and Virgo O1 and O2 runs [8].

In addition to merging binary systems, other promising sources of GWs are non-axisymmetric supernova explosions, as well as long-lived periodic radiation from rotating, non-axisymmetric NS emitting continuous GW radiation, sometimes called the GW 'pulsars'.

In this article we will focus on the latter type of the signal. The departure from axisymmetry in the mass distribution of a rotating NS can be caused by dense-matter instabilities, strong magnetic fields and/or elastic stresses in its interior (for a review see [9]). It depends on the largely unknown dense-matter equation of state, and therefore the time-varying mass quadrupole is not 'guaranteed' as in the case of binary system mergers. Till

date, the LIGO and Virgo collaborations performed several broadband searches for such signals, both targeted searches for pulsar sources of known parameters ([10, 11] and references therein), as well as all-sky searches for sources of unknown parameters ([12, 13] and references therein).

B. All-sky searches

The *all-sky searches* for continuous GW are blind searches which consist in probing parameter space to find the best matching template above certain threshold in terms of signal-to-noise ratio (SNR). The typical parameter space has four dimensions: frequency f, spindown $\dot{f}$, and two sky positions α, δ. There are various algorithms (for a recent review of the methodology of continuous GW searches with the Advanced LIGO O1 data see [14]) but in the core they rely on performing Fourier transforms of the input time series.

This type of search is particularly computationally demanding. In addition to unknown frequency parameters and sky location, the GW polarization, amplitude and phase of the signal is also not known *a priori*, so the grid or search parameters has to be sufficiently dense not to overlook the signal. Consequently, both the volume of the parameter space (number of evaluation of the Fourier transform), and the number of resulting candidate GW signals becomes very large, when the detection threshold is low. The threshold has to be low because the continu-

ous GW are expected to be very weak - currently only upper limits for their strength are set [12, 13, 15–17]. A natural way to improve the SNR is to analyze long stretches of data since the SNR increases as a square root of the data length. In practice, coherent analysis of the many-month long observations (typical length of LIGO/Virgo scientific run) is computationally prohibitive. Depending on a method the coherence time used ranges from minutes to days. Then additional methods are used to combine the results incoherently.

Many currently used algorithms to search for continuous periodic GW signals implement the $\mathcal{F}$-statistic methodology [18], a data analysis algorithm based on matched filtering by maximizing the likelihood function with respect to the set of unknown parameters. Recent examples include searches in the LIGO and Virgo data [15–17], as well as mock data challenge [19]. The method uses few-days long data segments for coherent analysis, and in the second step the candidate signals from the coherent analysis are searched for coincidences. In the following we will explore alternative approach to the second step.

C. The aim of this research

The aim of this work is to analyse, with the use of deep artificial neural networks, the results of a first stage of the all-sky time-domain (TD) $\mathcal{F}$-statistic pipeline [16, 20], `TD-Fstat search`, implemented in [21] (see the documentation in [22]) i.e., to classify the output of `TD-Fstat search`, the clusters of candidate GW signals crossing a pre-defined threshold value of the $\mathcal{F}$-statistic.

Specifically, we study the application of deep convolutional neural network (DCNN) on the distribution of candidate signals obtained by evaluating the `TD-Fstat search` algorithm on a pre-defined grid of parameters ('optimal' grid defined in [23]). The data contains either pure Gaussian noise, Gaussian noise with injected astrophysical-like signals, or Gaussian noise with injected purely monochromatic signals, mimicking local artifacts of the detector (so-called stationary lines).

D. Previous works

The convolutional neural network (CNNs) [24] have already proved to be useful in the field of gravitational wave physics. In particular in the domain of image processing. Razzano and Cuoco [25] has been using convolutional neural networks for classification of noise transients in gravitational wave detectors. George and Huerta [26] have been developing the *Deep Filtering* method for the signal processing, based on a system of two deep CNNs, designed to detect and estimate parameters of compact binary coalescence signal in highly noisy time-series data streams. Dreissigacker *et al.* [27] have been using DL as a search method for the CWs from unknown spinning

neutron stars over broad range of frequencies. Whereas, Gebhard *et al.* [28] have been studying the general limitations of CNNs as a tool to search for merging black holes.

The last three papers discuss the deep learning (DL) as an alternative to matched filtering. However, it seems that the DL has too many limitations for the application in the classification of GW based on raw data from interferometer (see discussion in [28]). For this reason we have decided to study the application of DL differently. In this paper we have considered DL as a tool complementary to matched filtering. A tool that allows to effectively classify large number of signal candidates obtained with the matched filter method. Instead of studying only binary classification, we have covered the multi-label classification assessing the case of artifacts resembling the CW signal. Finally our work compares two different types of convolutional neural networks implementations: one-dimensional (1D) and two-dimensional (2D).

E. The article's structure

The article is organized as follows. In Sect. II we introduce the deep learning algorithms with particular emphasis on the convolutional neural networks and their application in astrophysics. Section III describes data processing we used to develop accurate model for the `TD-Fstat search` candidate classification. Section IV summarizes our results which are further discussed. Summary and a description of future plans are provided in Sect. V.

II. DEEP LEARNING

DL [29] has commenced a new area of machine learning, a field of computer science based on special algorithms that can learn from examples in order to solve problems and make predictions, without the need of being explicitly programmed [30]. DL stands out as a highly scalable method that can process raw data without any manual feature engineering. By stacking multiple layers of artificial neurons (called neural networks) combined with learning algorithms based on backpropagation and stochastic gradient descent ([24] and references therein), it is possible to build advanced models able to capture complicated non-linear relationships in the data by composing hierarchical internal representations. The deeper the algorithm is, the more abstract concepts it can learn from the data, based on the outputs of the previous layers.

The DL is commonly used in commercial applications associated with computer vision [31], image processing [32], speech recognition [33] and natural language processing [34]. What is more, it is also becoming more popular in science. The deep learning algorithms for image analysis and recognition have been successfully tested in many fields of astrophysics like galaxy classification [35]

and asteroseismology [36]. Among many deep learning algorithms there is one that might especially be useful in the domain of gravitational wave physics – the convolutional neural networks.

A. Convolutional Neural Network

CNN is a deep, feed-forward artificial neural network (network that process the information only from the input to the output) whose structure is inspired by the studies of visual cortex in mammals, the part of the brain which specializes in processing visual information. The crucial element of CNNs is called a convolution layer. It detects local conjunctions of features from the input data and maps their appearances to a feature map. As a result the input data is split into parts, creating local receptive fields and compressed into feature maps. The size of the receptive field corresponds to the scale of the details to be looked in the data.

CNNs are faster than typical fully-connected [37], deep artificial neural networks because sharing weights significantly decreases the number of neurons required to analyze data. They are also less prone to the overfitting (the model learned the data *by heart* preventing the correct generalization).

To further reduce the computational cost *pooling layers* (subsampling layers) might be used, coupled to the convolutional layers. They constrain the size of the CNN and make it more resilient to the noise and translations which enhances their ability to handle new inputs.

III. METHOD

A. Generation of data

To obtain a sufficiently large, labeled training set, we generate an artificial `TD-Fstat search` results containing signals with known parameters, as well as local detector disturbances and pure noise. We adopt three different classes of the candidate signals: 1) a GW signal, modeled here by injecting an astrophysical-like signal that matches the $\mathcal{F}$-statistic filter, corresponding to spinning triaxial NS ellipsoid [20], 2) an injected purely monochromatic signal, mimicking local artifacts of the detector (so-called stationary lines) [38], for which the $\mathcal{F}$-statistic is not an optimal filter, or 3) the pure Gaussian noise. These three classes are henceforth denoted by the "signal", "line" and "noise" labels, respectively.

By analyzing sufficiently long segments of data (of the order of days) one produces the GW candidates with the matched-filter SNR ρ depending on both the amplitude h and the time-series duration T_0: $\rho \propto h\sqrt{T_0}$ [39].

To generate the candidate signals for the classification, the `TD-Fstat search` uses narrow-banded time series data as an input. In this work we focus on stationary white Gaussian time series, into which we inject
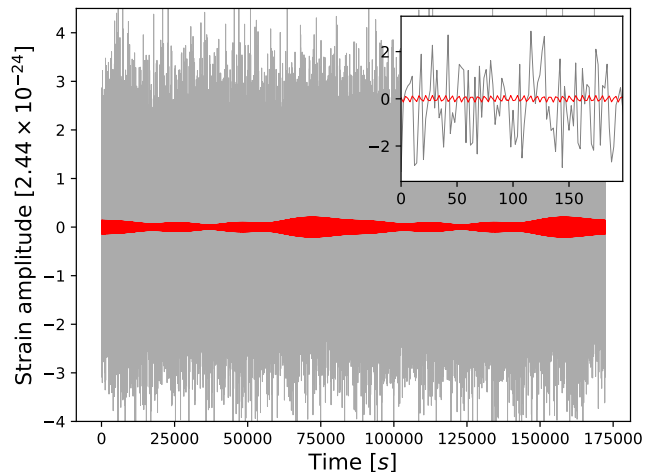

FIG. 1. An example of a continuous GW time-domain data, input of `TD-Fstat search`. Grey time series of $T_0 = 2$ sidereal days length mimics the downsampled, narrow banded data produced from the raw interferometer data [16, 20]. The data contains an almost-monochromatic astrophysical GW signal (red curve) of the SNR $\rho = 20$ and the following parameters (see also Table I for the parameters of the search and the text for more details): frequency $f = 2.16$ (in the units of the narrow band, between 0 and π), spindown $\dot{f} = -3.81 \times 10^{-8}$ (in dimensionless units of the pipeline, corresponding to $\dot{f}_{astro} = -3.03 \times 10^{-9}$ Hz/s; $\dot{f} = \pi\dot{f}_{astro}dt^2$ [20]), $\delta = 0.474$ (range between $-\pi/2$ and $\pi/2$) and $\alpha = 5.84$ (range between 0 and 2π). The reference frequency of the narrow band equals 100 Hz. Visible modulation is the result of the daily movement of the detector with respect to the astrophysical source, as well as of their relative positions, reflecting the quadrupolar nature of the detector's antenna pattern; in the case of a stationary line local to the detector such modulation is absent.

astrophysical-like signals, or monochromatic 'lines' imitating local detector's disturbances. An example of such input data is presented in Fig. 1. It mimics the raw data taken from the detector, downsampled from the original 16384 Hz (20000 Hz) in the case of the LIGO and Virgo detectors, respectively, and is divided into narrow frequency bands, as the frequency of an astrophysical almost-periodic GW signal is not expected to vary substantially. Details of the input data are gathered in Table I. Additional `TD-Fstat search` input include the ephemeris of the detector (the position of the detector with respect to the Solar System Barycenter and the direction to the source of the signal, for each time of the input data), as well as the pre-defined grid parameter space, on which the search is performed [23].

By using the $\mathcal{F}$-statistic one does not need to directly search for the polarization, amplitude and phase of the GW signal (after a signal is found these values are recovered from the distribution of the $\mathcal{F}$-statistic; see [18, 39] and references therein). Instead, in the simplest realistic case of a rotating triaxial ellipsoid one explores a 4-dimensional space parameterized by the GW frequency

f, frequency derivative $\dot{f}$ (also called the *spindown*), to describe the fact that the source is most likely decreasing or increasing its spin frequency, and the position of the source in the sky (e.g., declination δ and right ascension α of the sky equatorial coordinate system). Each point in this 5-dimensional space of $(f, \dot{f}, \delta, \alpha, \mathcal{F})$ that crosses a pre-defined SNR threshold is called a *candidate signal*. Signal-to-noise of each candidate signal ρ is related to the value of $\mathcal{F}$-statistic as $\rho = \sqrt{2(\mathcal{F} - 2)}$ (see [39] for detailed description).

In the signal-injection mode, the `TD-Fstat search` implementation adds an artificial signal to the data at the specified frequency $f, \dot{f}, \delta, \alpha$ and SNR; in the case of pure noise class no additional signal is added to the original Gaussian data, but the data is evaluated in pre-described range of $f, \dot{f}, \delta, \alpha$.

Subsequently, the code performs a search around the injection parameters $(f, \dot{f}, \delta, \alpha)_{inj}$ for a given time segment in the range of a few grid points, and all the candidate signals crossing the pre-defined threshold are recorded. We generate the training data set using the parameters specified in the Table I. For each configuration of the SNR and reference frequency of the narrow band, we have produced 2500 signals per class (292500 in total). We limited our studies to signals with an SNR $\geqslant 8$, which is a robust threshold for the detection of continuous GW signal.

To summarize, each candidate GW signal instance consist of the SNR ρ, and values of the frequency f (in fiducial units of the narrow-band, from 0 to π), spindown $\dot{f}$ (in Hz/s), and two angles describing its sky position in equatorial coordinates, right ascension α (values from 0 to 2π) and declination δ (values from $-\pi/2$ to $\pi/2$); see Fig. 2 for an exemplary output distribution of the candidate signals.

Detector	LIGO Hanford
Reference frequency	50, 100, 200, 300, 500, 1000 Hz
Segment length T_0	2 days
Bandwidth	0.25 Hz
Sampling time dt	2 s
Grid range	± 5 points
$\mathcal{F}$-statistic threshold	14.5 (corresponding to $\rho = 5$)
Signal-to-noise ρ of injections	from 8 to 20

TABLE I. Parameters of the input to the `TD-Fstat search` code. Time series consist initially of random instances of white Gaussian noise, into which signal or stationary lines were injected. T_0 equal to 2 sidereal days with 2 s sampling time results in 86164 data points.

The CNN required the input matrix of the fixed size. However, the number of points on distributions shown in Fig. 2 may vary for each simulation. Depending on the frequency (see Table I) it could increase a few times. To address this issue, we transformed point-based distributions into two different representations: set of four 2D images (four distributions) and set of five 1D vectors (five
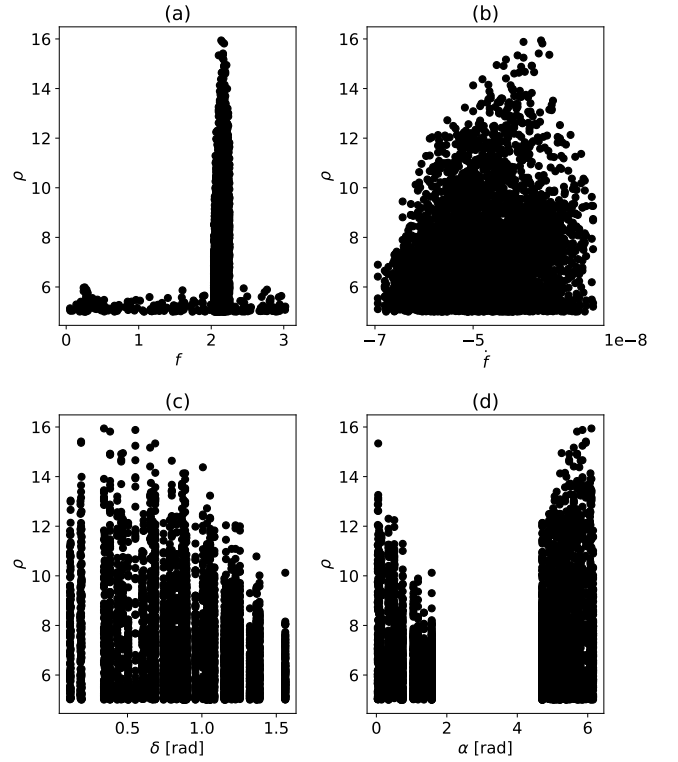

FIG. 2. Exemplary `TD-Fstat search` output for GW signal generated for the SNR = 20 and frequency $f = 100$ Hz and parameters of the injected signal from Fig. 1 ($f = 2.16$, $\dot{f} = -3.81 \times 10^{-8}$, $\delta = 0.474$ and $\alpha = 5.84$). The distributions represent the relation between ρ ($\mathcal{F}$-statistic SNR) with respect to: the frequency (a), the derivative of frequency (b), the declination (c) and the right ascension (d).

$\mathcal{F}$-statistic parameters).

The image-based representation was created via conversion to the two-dimensional histogram (see Fig. 3) of the corresponding point-based distributions. Their sizes were fixed to 64×64 pixels. We chose this value empirically; smaller images lost some information after the transformation, whereas bigger images led to the significantly extended training time of the DCNN we used.

The vector-based representation was created through the selection of the 50 greatest values of the ρ distribution and their corresponding values from the other parameters $(f, \dot{f}, \alpha$ and $\delta)$. The length of the vector was chosen empirically. The main limitation was related to the density of the point-like distributions which changed proportionally to the frequency. For the 50 Hz signal candidates, the noise class had sparse distributions of slightly more than 50 points. Furthermore, the vectors were sorted with respect to the ρ values, since this step allowed to reach slightly higher values of classification accuracy; see Fig. 4.

The created datasets were then split into three separate subsets: training set (60% of signals from the total dataset), validation set (20% of signals from the total
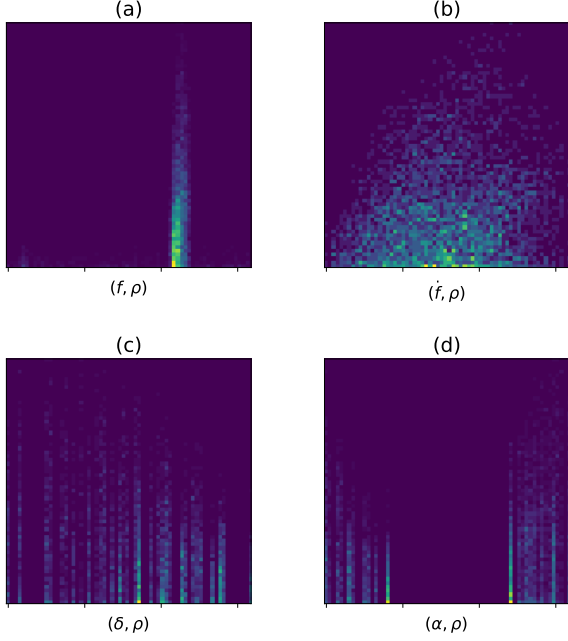
FIG. 3. The 2D representation of `TD-Fstat search` used as an input to the 2D DCNN. Images presented here have same size equal to 64x64 pixels. They correspond to the distributions shown on Fig. 2: (a) frequency $- \rho$, (b) spin-down $- \rho$, (c) declination $- \rho$, (d) right-ascension $- \rho$. The colours correspond to the density of the distribution – the brighter it is, the more points contributed to the pixel. These four images describe single signal candidate.

FIG. 4. The 1D representation of `TD-Fstat search` used as an input to the 1D DCNN. The plots present five parameters of $\mathcal{F}$-statistic describing single signal candidate limited to the 50 maximum values of ρ: (a) frequency, (b) spin-down, (c) declination, (d) right ascension and (e) ρ. The vector of ρ was sorted since it allowed to reach higher accuracy during training.

dataset) and the testing set (20% of signals from the total dataset). Validation part was used during training to monitor the performance of the network (whether it overfits). Testing data was used after training to check how the DCNN performs with unknown samples.

B. Neural network architecture

The generated datasets required two different implementations of the DCNN. Overall we tested more than 50 architectures ranging $2 - 6$ convolutional layers and $1-4$ fully connected layers for both models. Final layouts are shown in Figs. 5 and 6.

In case of 1D DCNN, the classifier containing three convolutional layers and two fully connected layers yielded the highest accuracy (more than 93% for whole validation/test datasets). Whereas the 2D DCNN required four convolutional neural networks and two fully connected layers to reach the highest accuracy (84% over whole validation/test datasets). Although both models were considered as trained after 50 epochs, the 2D version needed 3 times more time to process the data than 1D. The training for the first model took around 4 hours whereas the latter one was done after 80 minutes (on the same machine).
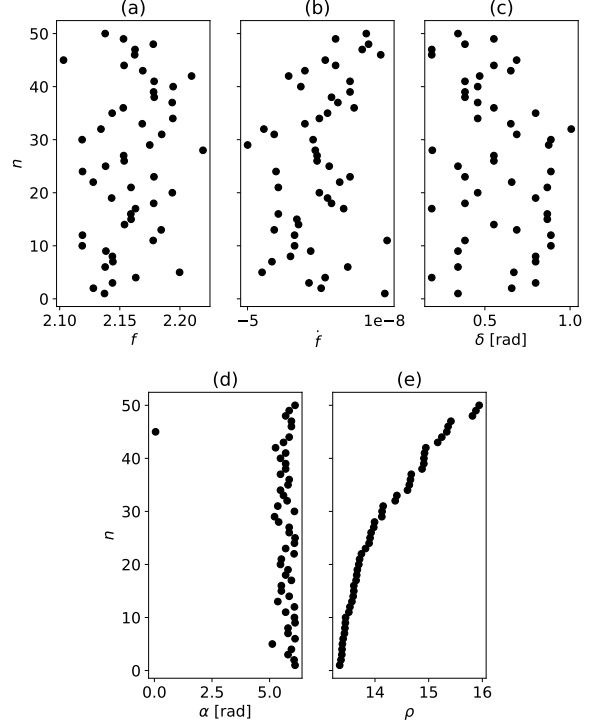
To avoid the overfitting we included dropout [40] in the architecture of both models. The final set of hyperparameters used for the training was the following (for both implementations; definitions of all parameters specified here can be found in [24]):

- activation function for hidden layers: ReLU,

- activation function for output later: softmax function,

- loss function: cross-entropy,

- optimizer: ADAM [41],

- batch size: 128,

- learning rate: 0.001.

The DCNN architectures were built using the Python library Keras [42] on top of the Tensorflow [43], with support for the GPU. We developed the model on Nvidia-sponsored Quadro P6000 [44] and performed the production runs on the Cyfronet Prometheus cluster [45] equipped with Tesla K40 GPU, running CUDA 9.0 [46] and the cuDNN 7.3.0 [47].
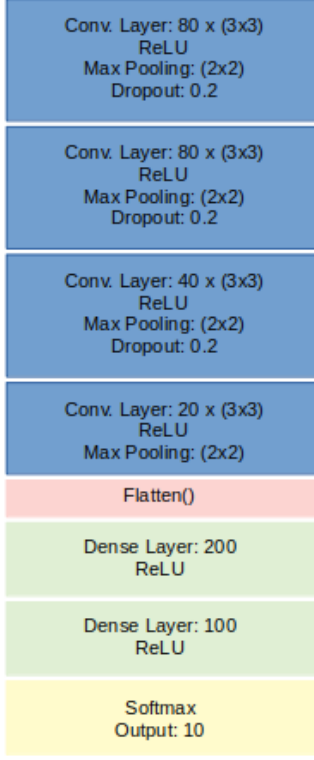
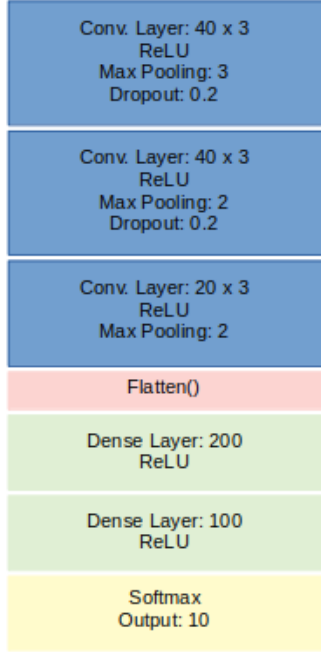FIG. 5. The diagram showing the architecture of the designed 2D DCNN.

FIG. 6. The diagram showing the architecture of the designed 1D DCNN.

IV. RESULTS AND DISCUSSION

The designed DCNNs (described in Section III B, Fig. 5 and Fig. 6) were trained on the generated datasets. During the training model implementing 1D architecture was able to correctly classify 93% of all candidate signals. Whereas the model implementing 2D architecture reached 84% of accuracy (see the comparison between trainings in Fig. 7). Since the very first epoch, the first model indicated a better ability to generalize candidate signals over large range of frequencies and SNRs. To further confirm this results, both DCNNs were tested against unknown before samples (test dataset), after the training. The results were shown in Fig. 8 in the form of confusion matrix. Both models were able to correctly classify majority of GW signals (93% for the 1D model and 95% for 2D model) as well as the noise (89% and 93% respectively). However the difference in the classification of the line was significant. 1D DCNN was able to correctly classify 97% of line candidates whereas 2D DCNN only 66%. Although the 2D model seemed to be more suited for binary classification task (detection of GW signal from the noise), the 1D DCNN outperformed 2D version in the multi-label classification.
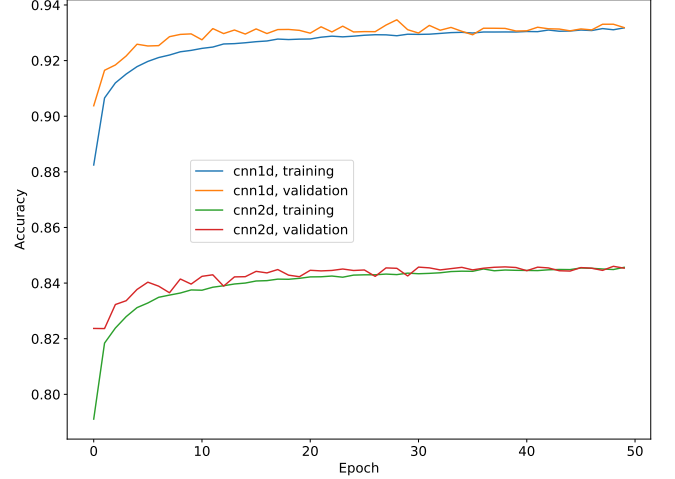

FIG. 7. The evolution of the accuracy as a function of the training epoch for the three-label classification for the 1D DCNN and 2D DCNN. Both models reached the maximum accuracy after 40 epochs (based on the results on the validation set). We prolonged training to 50 epochs to check the appearance of the overfitting. The 1D DCNN was still properly learning (although without increase of validation accuracy), however 2D version seemed to start ovefitting.

Knowing the general limitations of designed DCNNs, we performed additional tests trying to understand the response of our models against signal candidates of the specific parametrization. We generated additional datasets for particular values of the SNR and the frequency (see Table I). We expanded the ρ range to the values up to 4 which corresponded to the $\mathcal{F}$-statistic
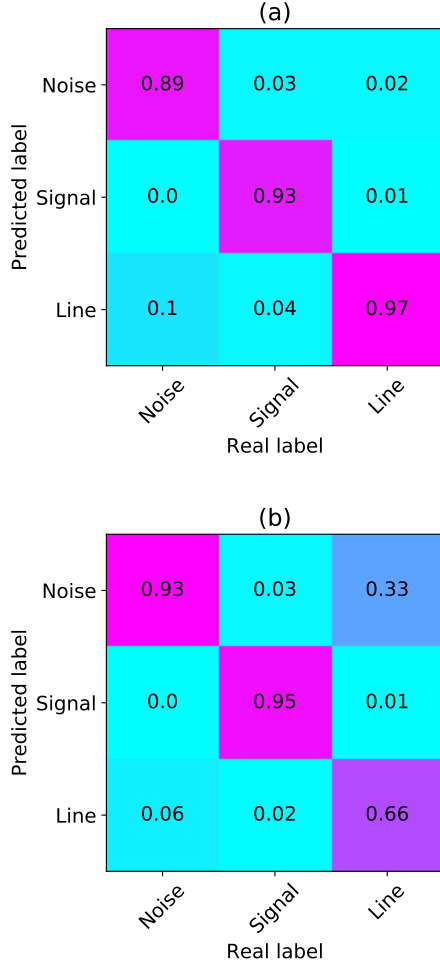
FIG. 8. Confusion matrix for the three-label classification evaluated on the test set for the 1D DCNN (a) and 2D DCNN (b) after the training. Although the GW signal and the noise were classified on the similar level, the line caused significant problem for the 2D model. The image representation of this type seemed to resemble the noise for the significant amount of signal candidates.

tributing to the peak (see Fig. 2 (a) for comparison) with respect to the background noise, made these candidates hardly distinguishable from the noise class.

On the other hand, 2D DCNN varied significantly in relation to the frequency. It reached the highest accuracy for the 100 Hz (99% for the SNR>10). For the other frequencies, the maximum accuracy was gradually shifted toward increasing SNR. Interestingly, the accuracy for the 50 Hz reached the maximum for the SNR=10; then it gradually decreased. The 2D DCNN seemed to outperform 1D model only for the narrow band of the frequency. Nevertheless, the general performance of this implementation was much worse.
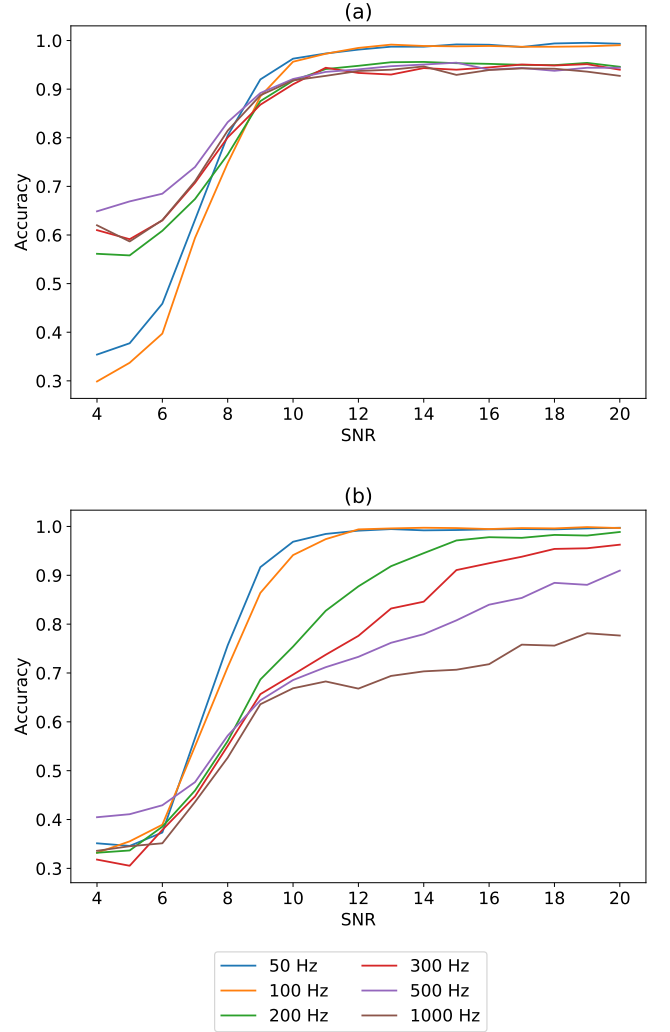

FIG. 9. The evolution of accuracy as the function of the SNR for 1D DCNN (a) and 2D DCNN (b). The first model achieved maximum level of accuracy for SNR=10-12 and maintained its value for whole range of frequencies. The 2D version varied dramatically in relation to the frequency with the maximum accuracy being gradually shifted toward larger values of SNR.

threshold for the signal candidate. This step allowed us to test the response of the DCNN against unknown during training very weak signals that seemed to be nearly indistinguishable from the noise.

The results were presented in Fig. 9 (a) and (b) (for 1D and 2D DCNNs, respectively). The 1D model presented significantly more stable behaviour toward the candidates over whole range of considered frequencies. It also maintained nearly stable accuracy for the SNR $\geq$ 10 (reaching the value of more than 90% for all of them). Interestingly, candidates with SNR < 8 were correctly classified in $60-70\%$ of samples for frequency $\geq$ 200 Hz. This was relatively high value, taking into consideration their noise-like pattern (for the GW and the line). This pattern had the biggest influence on the classification of the signal candidates generated for frequencies: 50, 100 Hz and the SNR < 8. The small number of points con-

Since the 1D DCNN proved to be more accurate over broad range of frequencies, we chose it as a more useful

model in the classification of the $\mathcal{F}$-statistic signal candidates. Below we present the results of additional tests we performed to better understand its usability.

We conducted the analysis of the true positives rates to test the model response in relation to the type of the considered signal class. Figure 10 presented the accuracy as the function of SNR for the three studied types of signal candidate.

The results for the noise class for the given frequency maintained at relatively constant value. This was expected since the noise did not depend on the SNR (it had fixed amplitude equal to the threshold of $\mathcal{F}$-statistic, see Section III A). The values of SNR on the Fig. 10 (a) were used for the comparison with the other types of signal candidates. The differences between accuracies for various frequencies were caused by variability of distribution densities of the initially generated data (compare with Fig. 2). The density increased with increasing frequency. Chosen 50 points for the 1D representation instead of having random character (as for low frequencies) would cover area around some fluctuation in the parametric space that could resemble other types of signal candidates. However this was the case for not more than 20% of samples.

The GW signal had the strongest effect on the total accuracy (compare Fig. 10 (b) with Fig. 9 (a)). The DCNN was expected to reach around 0% accuracy for SNR=4. The GW signal should not be distinguishable from the noise at this level. Nevertheless, for majority of considered configurations of SNR/frequency, this class was recognized in more than 95% of cases.

The line for majority of frequencies also maintained at the constant level of accuracy. The significant decrease for the 50 and 100 Hz datasets was the reason behind the poor performance in the Fig. 9 (a) below SNR=8 (orange and blue curves). Responsible for the decrease of accuracy in case of low frequencies was the size of grid range used by `TD-Fstat search` (see Tab. I). The generated distributions of $\mathcal{F}$-statistic parameters had sparse character for 50 and 100 Hz making both, the line and GW signal, hardly distinguishible from the noise for lower SNRs.

We additionally performed tests on the signal candidates generated for different frequencies than specified in the Table I. We chose five new frequencies to test the model on: 20, 250, 400, 700, 900 Hz. The results were presented in Fig. 11. The 20 Hz case is missing since the number of available points (from initial distributions) to create set of five 1D vectors was much smaller than the chosen length (some distributions for the noise class contained less than 10 points). Nevertheless, the DCNN for the other frequencies reached similar accuracies as those presented in Fig. 9 (a). This result proved the generalization ability of the 1D DCNN toward unknown frequencies. However the limitation of the model was the minimum number of points available to create input data. Since this number was proportional to the frequency of the searched signal, our DCNN was not suited to search
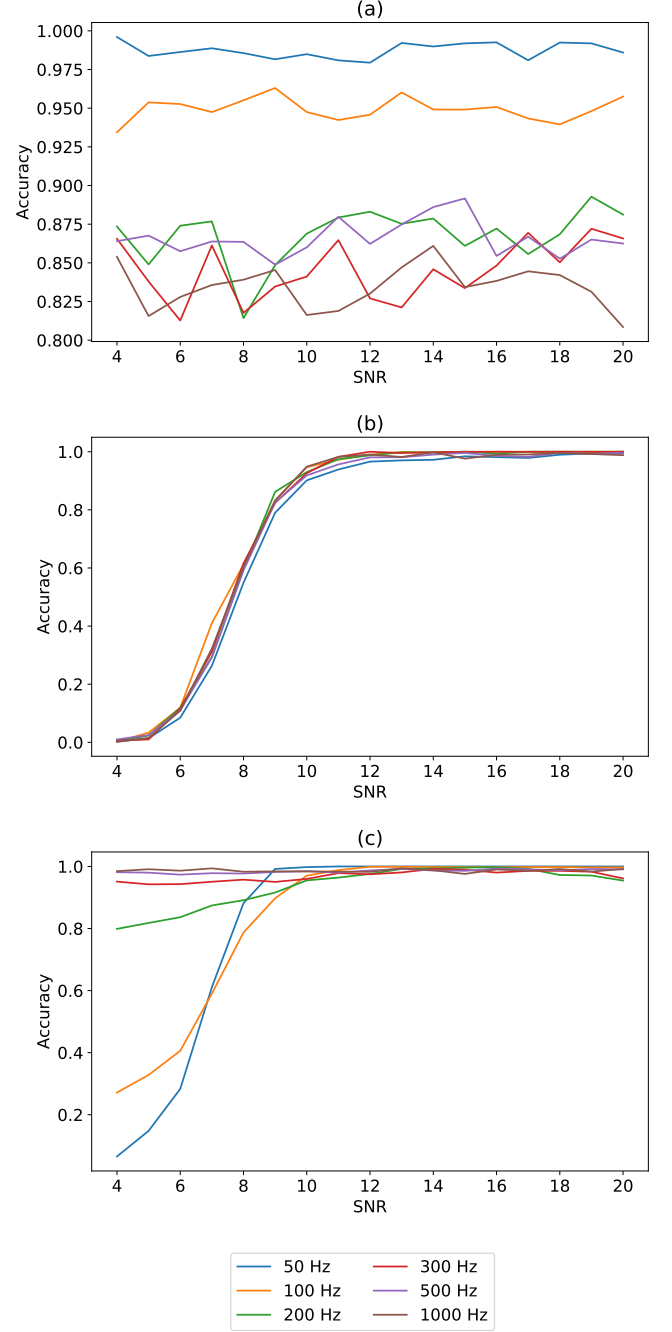

FIG. 10. The true positive rates of the 1D DCNN for the three types of signal candidates: the noise (a), the GW (b) and the line (c). Values of SNR on the (a) plot are used for comparison with the other types of signal candidates (SNR for the noise is fixed and equal to the threshold of $\mathcal{F}$-statistic).

for candidates below 50 Hz.

Our work, as one of the few, covered the application of DL as a supplementary component to MF. Analysing signal candidates instead of raw data allowed us to avoid problems that other researchers encountered. From the one hand, this approach limits the number of signals
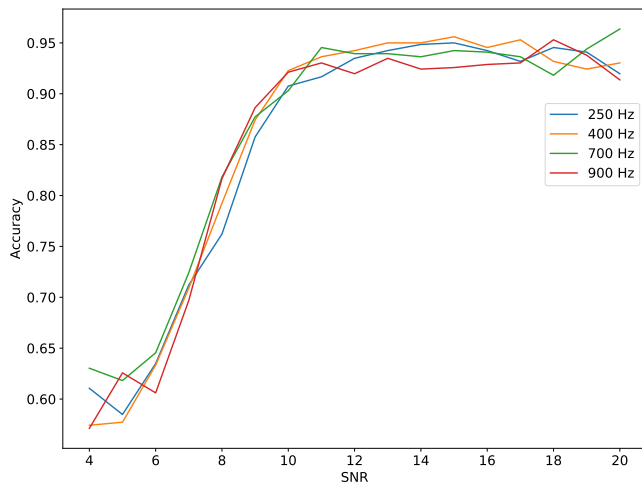
FIG. 11. The evolution of accuracy as the function of the SNR for 1D DCNN for the signal candidates of unknown during training frequency. The 20 Hz case is missing since the generated samples did not have enough points to create an input data (vectors of length 50, see Section III A).

to those that exceeded the $\mathcal{F}$-statistic threshold. From the other hand, the generated data are characterized by known significance. As Gebhard *et al.* [28]) described, application of DL on raw data provides signal candidates of unknown or hard to define significance. Before DL could be used as a safe alternative to MF for the detection of GW, it has to be studied further. However, it can already be considered in terms of supporting role to MF. For example, it could be used in the preparation of the candidates for the follow-up via fast classification and limiting the parameter space to be processed further. As our project showed, the DCNN can also be used in the classification of glitches. In case of short-time scale artifacts, these methodology has already reached satisfactory results as Razzano and Cuoco [25] proved.

V. CONCLUSIONS

We proved that the DCNN can be successfully applied in the classification of `TD-Fstat search` results, multi-

dimensional vector distributions corresponding to three signal types: the GW signal, the stationary line and the noise. We compared 2D and 1D implementations of DCNN. The latter achieved much higher accuracy (93% with respect to 84%) over candidate signals generated for broad range of frequencies and SNR. For majority of signals (SNR ≥ 10) 1D DCNN maintained more than 90% of accuracy. This level of accuracy was preserved for the signal candidates of unknown frequency during the training.

2D DCNN represented a different character. Although, the overall accuracy was worse than 1D model, the 2D version seemed to achieve better results as a binary classifier (between the GW signal and the noise). Representation of the input data in the form of the image seemed to cause significant problems for the proper classification of the line. Even though the 2D DCNN had worse generalization ability, it was able to outperform 1D implementation for the targeted search for the signal candidates of 100 Hz.

Nevertheless, 1D DCNN with its ability to generalize unknown samples (in particular with respect to the frequency) seemed to be better choice for the realistic applications.

Our work opens many possibilities for further development within CW searches. Among the projects we are considering, is the application of DL in the follow-up of signal candidates as well as analysis of data from many detectors.

VI. ACKNOWLEDGMENTS

[1] A. Einstein, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 688-696. (1916).

[2] J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, *et al.*, Class. Quant. Grav. **32**, 074001 (2015), arXiv:1411.4547.

[3] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, *et al.*, Class. Quant. Grav. **32**, 024001 (2015), arXiv:1408.3978.

[4] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al., Physical Review Letters **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].

[5] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al., Physical Review Letters **118**, 221101 (2017), arXiv:1706.01812 [gr-qc].

[6] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al., Physical Review Letters **119**, 141101 (2017), arXiv:1709.09660 [gr-qc].

[7] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al., Physical Review Letters **119**, 161101 (2017), arXiv:1710.05832 [gr-qc].

[8] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, and et al., arXiv e-prints (2018), arXiv:1811.12907 [astro-ph.HE].

[9] P. D. Lasky, Publications of the Astronomical Society of Australia **32** (2015), 10.1017/pasa.2015.35.

[10] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al., ApJ **839**, 12 (2017), arXiv:1701.07709 [astro-ph.HE].

[11] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al., Phys. Rev. D **96**, 122006 (2017), arXiv:1710.02327 [gr-qc].

[12] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al., Phys. Rev. D **96**, 062002 (2017), arXiv:1707.02667 [gr-qc].

[13] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al., Phys. Rev. D **97**, 102003 (2018), arXiv:1802.05241 [gr-qc].

[14] M. Bejger, Rencontres de Moriond (2017), arXiv:1710.06607 [gr-qc].

[15] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al., Phys. Rev. D **96**, 122004 (2017), arXiv:1707.02669 [gr-qc].

[16] J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, T. Accadia, F. Acernese, K. Ackley, C. Adams, T. Adams, and et al., Classical and Quantum Gravity **31**, 165014 (2014), arXiv:arXiv:1402.4974 [gr-qc].

[17] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, and et al., arXiv e-prints , arXiv:1903.01901 (2019), arXiv:1903.01901 [astro-ph.HE].

[18] P. Jaranowski, A. Królak, and B. F. Schutz, Phys. Rev. D **58**, 063001 (1998).

[19] S. Walsh, M. Pitkin, M. Oliver, S. D'Antonio, V. Dergachev, A. Królak, P. Astone, M. Bejger, M. Di Giovanni, and O. Dorosh, Phys. Rev. D **94**, 124010 (2016), arXiv:1606.00660 [gr-qc].

[20] P. Astone, K. M. Borkowski, P. Jaranowski, M. Pietka, and A. Królak, Phys. Rev. D **82**, 022005 (2010), arXiv:arXiv:1003.0844 [gr-qc].

[21] "Time-domain $\mathcal{F}$-statistic pipeline repository," https://github.com/mbejger/polgraw-allsky (), accessed: 2018-09-04.

[22] "Time-domain $\mathcal{F}$-statistic pipeline documentation," http://mbejger.github.io/polgraw-allsky (), accessed: 2018-09-04.

[23] A. Pisarski and P. Jaranowski, Classical and Quantum Gravity **32**, 145014 (2015), arXiv:arXiv:1302.0509 [gr-qc].

[24] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (The MIT Press, 2016).

[25] M. Razzano and E. Cuoco, Class. Quant. Grav. **35**, 095016 (2018), arXiv:1803.09933 [gr-qc].

[26] D. George and E. A. Huerta, Physics Letters B **778**, 64 (2018), arXiv:1711.03121 [gr-qc].

[27] C. Dreissigacker, R. Sharma, C. Messenger, and R. Prix, arXiv e-prints , arXiv:1904.13291 (2019), arXiv:1904.13291 [gr-qc].

[28] T. D. Gebhard, N. Kilbertus, I. Harry, and B. Schölkopf, arXiv e-prints , arXiv:1904.08693 (2019), arXiv:1904.08693 [astro-ph.IM].

[29] Y. Lecun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).

[30] A. L. Samuel, IBM J. Res. Dev. **3**, 210 (1959).

[31] A. Maronidis, E. Chatzilari, S. Nikolopoulos, and I. Kompatsiaris, Digital Signal Processing **74**, 14 (2018).

[32] P. N. Druzhkov and V. D. Kustikova, Pattern Recognition and Image Analysis **26**, 9 (2016).

[33] Z. Zhang, J. Geiger, J. Pohjalainen, A. E.-D. Mousa, W. Jin, and B. Schuller, ACM Trans. Intell. Syst. Technol. **9**, 49:1 (2018).

[34] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, J. Mach. Learn. Res. **12**, 2493 (2011).

[35] V. Lukic and M. Brüggen, Proceedings of the International Astronomical Union **12**, 217–220 (2016).

[36] M. Hon, D. Stello, and J. Yu, Royal Astronomical Society. Monthly Notices **469**, 4578 (2017).

[37] B. Research, *DeepBench* (2018 (accessed February 9, 2018)).

[38] P. B. Covas, A. Effler, E. Goetz, P. M. Meyers, A. Neunzert, M. Oliver, B. L. Pearlstone, V. J. Roma, R. M. S. Schofield, V. B. Adya, and et al., Phys. Rev. D **97**, 082002 (2018), arXiv:1801.07204 [astro-ph.IM].

[39] P. Jaranowski and A. Krolak, *Analysis of Gravitational-Wave Data* (Cambridge, UK: Cambridge University Press, 2009).

[40] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, Journal of Machine Learning Research **15**, 1929 (2014).

[41] D. P. Kingma and J. Ba, arXiv e-prints , arXiv:1412.6980 (2014), arXiv:1412.6980 [cs.LG].

[42] F. Chollet *et al.*, "Keras," https://keras.io (2015).

[43] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," (2015), software available from tensorflow.org.

[44] Benefiting from the donation via the NVidia GPU seeding grant.

[45] Prometheus, Academic Computer Centre CYFRONET AGH, Kraków, Poland.

[46] J. Nickolls, I. Buck, M. Garland, and K. Skadron, Queue **6**, 40 (2008).

[47] S. Chetlur, C. Woolley, P. Vandermersch, J. Cohen, J. Tran, B. Catanzaro, and E. Shelhamer, ArXiv e-prints (2014), arXiv:1410.0759.