



Imperial College London
Department of Physics

Searching for Gravitational Waves Using Convolutional Neural Networks

Project Code: THEO-Contaldi-1

Word Count: 9153

Date: April 13, 2020

CID Number:
01218411

Supervisor:
Carlo Contaldi

Assessor:
Toby Wiseman

Contents

1	Introduction	2
2	Theory	3
2.1	Detectors and Search Methods	3
2.1.1	Detectors	3
2.1.2	Search Methods Overview	4
2.1.3	Binary Coalescence Search	4
2.1.4	Noise, Data Quality, and Detection Confidence	5
2.1.5	Parameter Estimation	5
2.1.6	Limitations of the Current Search Methods	6
2.2	Neural Network	7
2.2.1	Neural Network Overview	7
2.2.2	Activation Function	8
2.2.3	Optimiser and Loss Function	9
2.2.4	Convolutional Neural Network	10
3	Methods	11
3.1	Noise Simulation	11
3.1.1	Sampling of Real LIGO Noise	11
3.1.2	Characterisation of Noise Using Power Spectrum	11
3.1.3	Gaussian Noise Generation in Frequency	12
3.1.4	band-pass Filtering	13
3.1.5	Inverse Fourier Transform	14
3.2	Waveform Simulation	15
3.2.1	Application of PyCBC	15
3.2.2	Parameters, Approximations, and Assumptions	16
3.2.3	Completion of Waveform	17
3.2.4	Strain Data Simulation	18
3.3	Detection and Parameter Estimation	20
3.3.1	Tensorflow and Keras	20
3.3.2	Training Data	20
3.3.3	Detection Model	21
3.3.4	Parameter Estimation Model	22
4	Results and Discussion	23
4.1	Detection	23
4.2	Parameter Estimation	25
4.3	Limitations and Future Work	28
5	Conclusion	29

Abstract

The detection of gravitational waves from compact binary mergers has inaugurated a new era of astrophysics. To expedite the analysis of gravitational-wave signals, deep learning convolutional neural networks were leveraged to attain more confident detection and accurate parameter estimation. In this project, both detector noise and gravitational waveform were characterised and simulated to generate large, extensive banks of strain data. Convolutional neural network models were then constructed and trained with simulated data. The detection model achieved an overall accuracy of 97.07% and reached 100% sensitivity for signals with SNR greater than 6. Also, it is able to perform single-detector analysis and detect weak signals in noisy strain data. On the other hand, the parameter estimation model obtained average percentage errors of 13.57% and 15.39% for estimating the chirp mass and luminosity distance respectively. These results suggest that the application of convolutional neural networks is an ideal alternative solution in the search of gravitational waves.

1 Introduction

The existence of gravitational waves was postulated by Albert Einstein in 1916 on the basis of the Field Equations in General Relativity. In essence, gravitational waves are ripples of spacetime caused by massive accelerating objects. Such waves propagate through the universe at the speed of light and are not easily dissipated or altered by interactions with matter, therefore preserving the information about their cosmic origins. As such, gravitational waves are excellent probes for astrophysical phenomena that cannot be investigated using electromagnetic radiation. Detecting and analysing gravitational waves provides a new perspective for observing the universe, especially the parts that are otherwise impossible to perceive. The effect of gravitational waves is measured with gravitational-wave strain, the fractional change in distance due to the distortion of spacetime. This dimensionless quantity is typically on the order of less than 10^{-20} , indicating that the amplitudes of gravitational waves are remarkably small. This is a characteristic of gravitational waves due to the intrinsic strength of gravitational interaction, as well as the distance between the source and the Earth. As a result, unprecedented levels of instrumental sensitivity and analytical capability are required to detect minuscule gravitational waves.

On September 14th 2015, the advanced LIGO reported the first direct detection of gravitational waves GW150914, originating from a binary black hole merger. This groundbreaking discovery was promptly awarded a Nobel Prize in Physics in 2017, marking the significance of this accomplishment to both the scientific community and mankind. In O1 and O2 of operation from 2015 to 2017, 11 gravitational wave events were successfully observed by LIGO and Virgo, with 10 being binary black hole coalescences and 1 being a binary neutron star collision.

The current search conducted by LIGO and Virgo relies on first identifying large coincident signals at multiple detectors to trigger further analysis, meaning that successful detection can only be achieved when gravitational waves are adequately strong. This limitation motivated the development of a more sensitive detection method capable of extracting weak gravitational-wave signals from noisy strain data.

In this project, the application of convolutional neural networks (CNN) was proposed as an alternative solution for gravitational-wave detection. Simulation algorithms of detector noise and gravitational waveform were developed to generate comprehensive strain datasets in order to inform the analysis. Analytical CNN models for detection and parameter estimation were then devised and trained to obtain both high detection sensitivity and prediction accuracy.

2 Theory

2.1 Detectors and Search Methods

2.1.1 Detectors

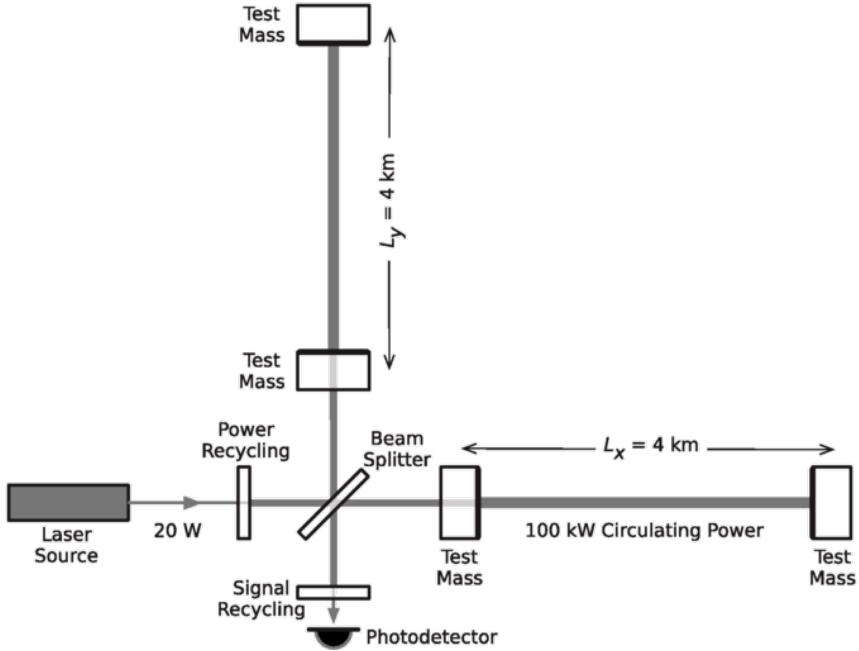


Figure 1: Simplified diagram of an Advanced LIGO detector (not to scale). The two orthogonal arms stretch and compress alternately when gravitational waves pass through, creating phase difference between the two light beams to measure the gravitational-wave strain.

A gravitational-wave detector is a large-scale modified Michelson Interferometer consisting of two orthogonal arms with an equal length of 4km arranged in an L shape [1], as shown in Figure 1. A laser beam is projected along each arm and reflected by the mirrors (test masses) suspended at both ends of the arm. When undisturbed by gravitational waves, two perpendicular laser beams are calibrated to interfere destructively and produce zero light signal at the beam splitter.

In contrast, passing gravitational waves causes the arms to stretch and compress alternately, changing the travel time of the laser beams and shifting the phase alignment. This creates optical signals that can be sensed by the photodetector and converted into strain data. In mathematical terms, a gravitational wave propagating through an interferometer changes the arm lengths by $\delta L_x(t)$ and $\delta L_y(t)$ respectively, such that the measured difference is

$$\Delta L(t) = \delta L_x(t) - \delta L_y(t) = h(t)L \quad (1)$$

where $h(t)$ is the gravitational-wave strain amplitude projected onto the detector, and L is the unperturbed arm length ($L = L_x = L_y$).

To attain the high sensitivity necessary for observing gravitational waves, several technologies were integrated into the detector infrastructure to amplify the optical responses [2]. Namely, its optical cavity that multiplies the phase difference, power recycling mirror for increasing the laser power, and signal recycling mirror which optimises signal extraction.

On the other hand, the reduction of noise is also crucial for obtaining quality data [3]. Considering that the detectors are extremely sensitive to vibrations, seismic noise is the source that most

notably invalidates the measurements. For the isolation of seismic noise, each heavy test mass is suspended at the final stage of a stable quadruple pendulum, as well as supported by an active platform that performs counter-movements to eliminate the recorded ground vibration noise. Other noise sources, such as quantum (photon counting) noise, thermal noise, and gas noise are also reduced by the implementation of various advanced technologies and methods.

Moreover, a vast array of sensors is equipped for each detector to monitor its instrumental behaviors and nearby environmental disturbances in auxiliary data channels, in order to characterise the noise and determine the quality of data. These auxiliary channels are essential for pinpointing and addressing unexpected bursts of noise to enable the confident detection of gravitational waves.

2.1.2 Search Methods Overview

Currently, the search for gravitational-wave candidates is achieved by two types of analytical methods. The first one is a preliminary investigation with an online, low latency generic transient search, which identifies potential gravitational-wave signals, almost in real-time. This search method promptly looks for coincident bursts of excess power from multiple detectors without modeling gravitational waveforms, facilitating the detection of un-modeled and unforeseen astrophysical events. The second one is an offline binary coalescence search that targets the gravitational waves emitted from compact binary mergers. This method leverages the modeling of gravitational waveforms and the matched filtering technique to acquire more accurate parameter estimation. These search methods calculate the signal to noise ratio for each candidate event and construct ranking statistics to determine the likelihood of the observed signal truly being gravitational waves.

2.1.3 Binary Coalescence Search

A compact binary merger is the coalescing process of an orbiting pair of massive, dense objects, such as black holes or neutron stars [4]. As two components in a binary system spiral in towards each other, they revolve at an increasing orbital angular frequency and radiate gravitational-wave energy. This orbital inspiral eventually leads to the cataclysmic merger and the subsequent ring-down, where a new single massive object is formed and oscillates in shape before stabilising into its final configuration.

Figure 2 displays the evolution of the binary black hole merger GW150914 and its corresponding gravitational waveform. In the inspiral phase, the waveform appears to be moderately stable and consistent, with its frequency gradually increases as the black holes spiral closer to each other. The waveform has its peaks with rapidly increased frequency when the black holes merge, then it is damped and fades away quickly during the ringdown phase.

The binary coalescence search method employs the matched filtering technique to compare the detected gravitational-wave signals with an extensive bank of modeled waveforms [?]. The template bank built for the search includes a large number of theoretically predicted waveforms for merging compact binary systems, taking into account various combinations of parameters such as the component masses, component spins, and luminosity distance.

The waveform modeling of compact binary mergers involves the application of many analytical and numerical techniques, namely the effective-one-body model, post-Newtonian approximation, black hole perturbation theory, and numerical relativity [7] [8]. This analytical approach leads to an accurate, computationally feasible algorithm of waveform modeling which provides great insights into the properties of gravitational-wave sources.

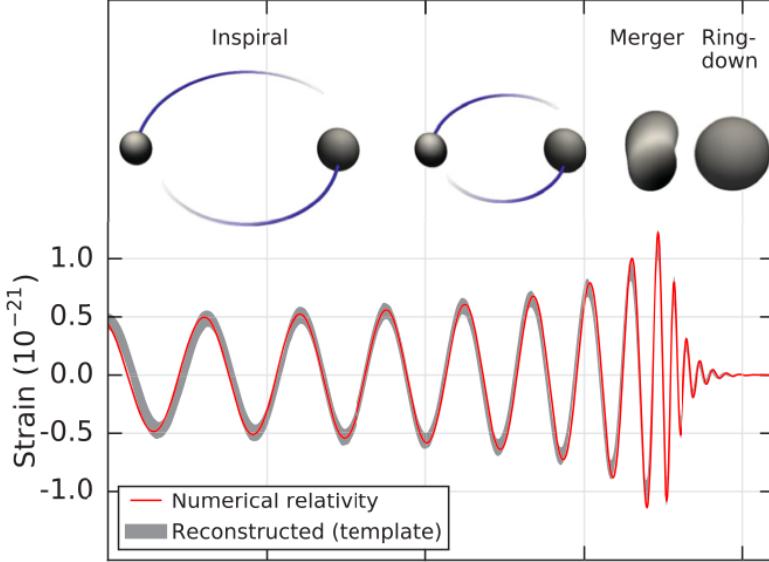


Figure 2: Evolution of the binary black hole merger GW150914 and its waveform, for the inspiral, merger, and ringdown phases. A relatively stable and consistent waveform was maintained in the inspiral phase, followed by peaks with high frequency for the merger and an immediate decay in the final ringdown phase.

2.1.4 Noise, Data Quality, and Detection Confidence

The LIGO-Virgo data can be challenging to analyse because the noise is a dynamic superposition of numerous technical and environmental noise of non-astrophysical origins. Also, the noise has non-stationary and non-Gaussian properties due to the presence of glitches.

Glitches are short bursts of transient noise caused by environmental disturbances and system instabilities. Considering that the matched filtering technique is only optimal when the noise is stationary and Gaussian, the handling of glitches is especially important [9]. By exploiting the information from auxiliary data monitoring channels, recognisable glitches can be eliminated or reduced, whereas the remaining unknown ones have to be examined by comparing data collected from detectors across several observatories. Data of poor quality due to glitches or other operational issues can be vetoed and removed from the data processing pipelines, in order to increase the confidence in observation of candidate events which pass the checks.

After cleansing the glitches, the probability of unrecognised transient noise occurring simultaneously in multiple detectors and producing false signals mimicking gravitational waves should be remarkably small. As a result, the statistical significance of a candidate event observed in two or more detectors can be measured with the false alarm rate, the expected rate of false detection due to coincidental noise fluctuation.

2.1.5 Parameter Estimation

The set of parameters characterising a binary compact merger can be estimated by performing a Bayesian inference analysis to derive the posterior distributions that pinpoint the most fitting values [10]. The parameters necessary to properly describe a gravitational wave include the intrinsic ones which determine the time-frequency morphology and amplitude-phase evolution of gravitational waves, such as the primary and secondary component masses and spins, as well as the extrinsic parameters such as luminosity distance and sky position.

In the analysis, the Bayes Theorem was applied:

$$P(\vec{\theta} | \{d\}, M) = \frac{P(\vec{\theta} | M)P(\{d\} | \vec{\theta}, M)}{P(\{d\} | M)} \quad (2)$$

where $\vec{\theta}$ is the compact binary merger parameters, $\{d\}$ is the strain data, and M is the waveform and noise models. The posterior $P(\vec{\theta} | \{d\}, M)$ gives the probability density of the degree that a modeled waveform M with parameter values $\vec{\theta}$ describes the data $\{d\}$. Having completed the calculation, the resulting posterior probability distribution can be applied then estimate the parameters such as the component masses and luminosity distance. As an exception, the sky localisation which positions gravitational-wave sources is done by timing triangulation for multiple detectors in the operating network.

2.1.6 Limitations of the Current Search Methods

The current search methods developed by LIGO and Virgo relies on first identifying coincident signals with excess power across multiple detectors to trigger further analysis. However, the detection sensitivity and accuracy of this approach is low for small signals deeply embedded in noise, meaning that oftentimes only a small fraction of gravitational-wave events can be successfully observed. Inevitably, the method is ineffective when observing events with low signal to noise ratios because of the difficulty distinguishing the bursts of true signals from background noise.

The current method also requires having more than one detector in operation at a given time. No valid analysis can be conducted with a single detector when other detectors undergo planned or emergency maintenance, or when the data collected by other detectors is severely contaminated by local or internal disturbances.

Furthermore, the matched filtering technique adopted in the current approach depends on comparing the input strain data to a template bank of modeled waveforms to estimate the parameters, rather than establishing some more general non-linear regression solutions. The need to filter the data against a bank with approximately 250000 template waveforms and evaluate the probability for each one of them might not be the most ideal method for parameter estimation.

To overcome these limitations, a rigorous and robust search method needs to be developed to provide high detection accuracy for small gravitational-wave signals, enable independent analysis on data from only one detector, and achieve more efficient and extensive parameter estimation. In this project, the application of deep learning convolutional neural networks was proposed to be a potential solution for the search of gravitational waves.

2.2 Neural Network

A neural network is an architecture structured conceptually based on these properties of the visual cortex in the human brain. With this framework, a neural network can learn from data and recognise underlying relationships in a dataset. In this project, a search method powered by deep learning convolutional neural networks was developed as an alternative solution for the gravitational-wave search.

The search of gravitational-wave signals in noisy data is a highly challenging task given the complex nature of both gravitational waves and the detector noise. Gravitational waveforms can only be computationally modelled by capitalising on numerous analytical and numerical techniques, and the detector noise is non-stationary and non-Gaussian. Consequently, the search for gravitational waves requires a non-linear solution that can extract important features from data and attain accurate predictions.

2.2.1 Neural Network Overview

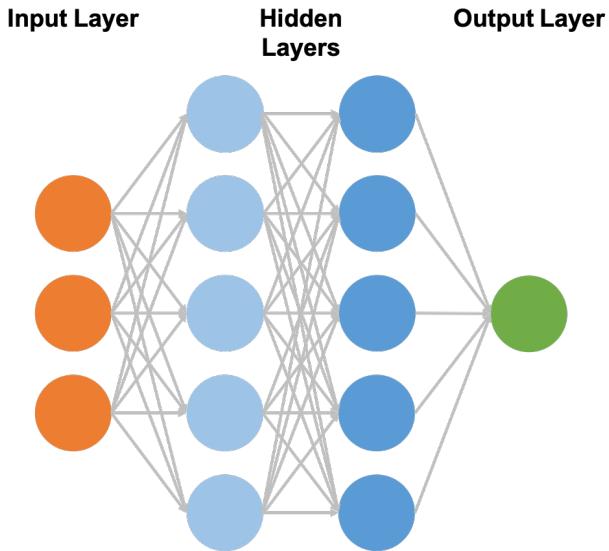


Figure 3: Example of a neural network with an input layer, two hidden layers, and an output layer. The nodes in the adjacent layers are interconnected to each other by the edges, such that the output of one layer becomes an input for the next one.

As illustrated in Figure 3, a neural network consists of multiple **layers** that are analogous to the columns in the human brain cortex, and a layer is made of multiple **nodes** that are equivalent to neurons [11]. Nodes in adjacent layers are interconnected by **edges**, which are functionally identical to synapses that train mitt electrical or chemical signals between neurons. That being said, an edge is responsible for passing input to a node, and each edge has an assigned **weight** that signifies the relative importance of the input [11]. On a fundamental level, a node receives inputs and computes an output by applying its associated **activation function** to the weighted sum of all inputs. This output then becomes an input in the next layer and continues to propagate.

Overall, a deep neural network comprises an input layer, many hidden layers, and an output layer. The functions of input and output layers are self-evident, and hidden layers perform the function of transforming raw inputs to outputs. The layer structure enables a deep neural network model to compute from low-level to high-level features incrementally, therefore eliminates the need for

domain expertise and feature extraction for conventional machine learning, meaning larger, more complicated data can be analysed.

With this architecture, a neural network model has the ability to learn and predict. More precisely, a neural network learns by finding the optimal weights that lead to the most accurate predictions. When training a neural network, the weight values are adjusted to highlight the most correlated, important features and disregard the irrelevant ones. On the other hand, a neural network predicts by simply applying the optimal weight set obtained from the prior training to produce outputs accordingly.

In more technical terms, a neural network model learns from **backpropagation** and predicts by forward propagation [12]. For an untrained model, all weights are initially randomly assigned. When a labeled data sample is passed forward through the network, an output is calculated by the network. By comparing this prediction with the expected labeled outcome of the data sample, the error can be evaluated using a **loss function**. This error value is subsequently passed backward to modify the weights by an **optimisation** algorithm for minimising the loss. The updated weight set acquired after this process straightforwardly results in a reduced error, hence higher precision. All data samples in a dataset are passed forward and backward through the network once over an epoch. The training of a neural network model can be considered completed after it has trained with a sufficiently large dataset over many epochs to achieve a high standard of accuracy.

2.2.2 Activation Function

A neural network can be considered as a universal function approximators, meaning that it can compute and learn any functions. Complicated problems can then be solved by establishing non-linear, high-degree functional mappings between inputs and outputs. This crucial nonlinearity is introduced to a neural network by activation functions [13].

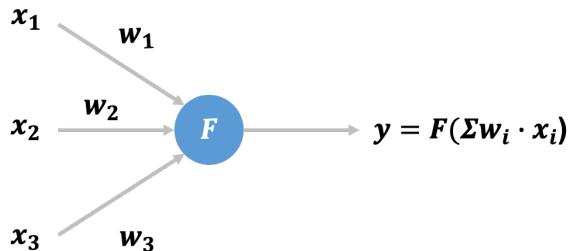


Figure 4: Example of the computational operation at a node in a neural network. The inputs x_i are assigned a set of weights w_i , such that the output y can be calculated from the weighted sum of the inputs $\sum w_i \cdot x_i$ using the activation function $F(x)$.

As shown in Figure 4, an activation function converts the weighted sum of inputs to an output, which then becomes an input to the next layer. Mathematically, this operation can be described as $y = F(\sum w_i \cdot x_i)$, where y is the output, $F(x)$ is the activation function, w_i is the weights, and x_i is the inputs. An activation function should be differentiable so that the weights can be optimised by a gradient-based method. There exist several prominent non-linear activation functions that are commonly utilised in neural networks, such as sigmoid, hyperbolic tangent (\tanh) and **rectified linear unit (ReLU)** functions [14]. The shapes of these three activation functions are compared in Figure 5.

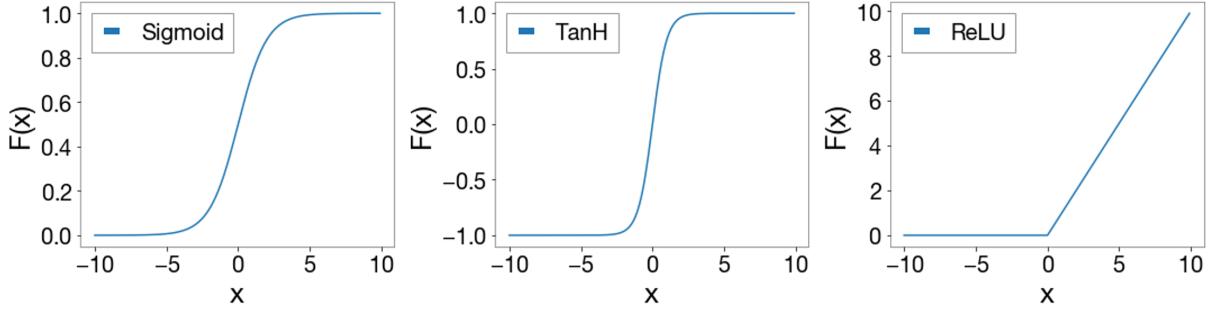


Figure 5: Sigmoid, hyperbolic tangent (tanh) and rectified linear unit (ReLU) activation functions. These non-linear activation functions enhance the complexity of a neural network and improve its ability to find accurate functional mappings of data.

In particular, the ReLU is the most prevalent activation function integrated into almost all neural networks [15]. The ReLU activation function can be expressed as $F(x) = \max(0, x)$, such that if $x < 0$, $F(x) = 0$ and else if $x \leq 0$, $F(x) = x$. It is linear for positive input values and non-linear for negative ones. As a result, the simplicity of linear models in the positive regime makes it computationally efficient, and the non-linear properties required for complex functional mapping are preserved by the negative domain. Moreover, its constant gradient of 1 implies that it can be optimised easily and the vanishing gradient problem can be avoided. It should be noted that ReLU is only used within hidden layers, whereas functions such as sigmoid and linear are employed in the output layer.

2.2.3 Optimiser and Loss Function

As previously described, a neural network model is trained by adjusting its weights with a gradient-based optimisation method to minimise a loss function. When a model with a given weight set predicts an output, a loss function is applied to calculate the error of the prediction. Following that, an optimiser then updates the weights to reduce the error in the next evaluation by navigating down the error gradient. The choices of optimiser and loss function partially define the training strategy for a neural network model.

Outlining different types of optimiser and loss functions is beyond the scope of the report, therefore only the ones relevant to the project are discussed here. As an example, **adaptive moment estimation (Adam)** is an optimiser that is broadly adopted in neural network applications [16]. This method is robust for most deep learning problems and provides both high learning speed and fast convergence.

On the other hand, the selection of loss function depends on the framing of the predictive task, such as classification or regression. A loss function is essentially a statistical representation that quantifies the effectiveness of a neural network model [17]. In the case of regression, **mean squared error (MSE)** loss is the default choice. It calculates the average of the squared differences between the predicted and actual values across all data samples. The squaring indicates that the model is heavily penalised for making deviated predictions. In contrast, the default loss function for binary classification is **binary cross-entropy loss**. This loss function evaluates a logarithmic score summarising the average differences between the predicted and actual probabilities across all data samples.

2.2.4 Convolutional Neural Network

A convolutional neural network (CNN) is a specialised model for high-level feature extraction, composed of a series of convolutional, max pooling, and fully-connected layers. Compared with a regular neural network, CNN generally has fewer weights to learn, resulting in faster training time and less computational cost.

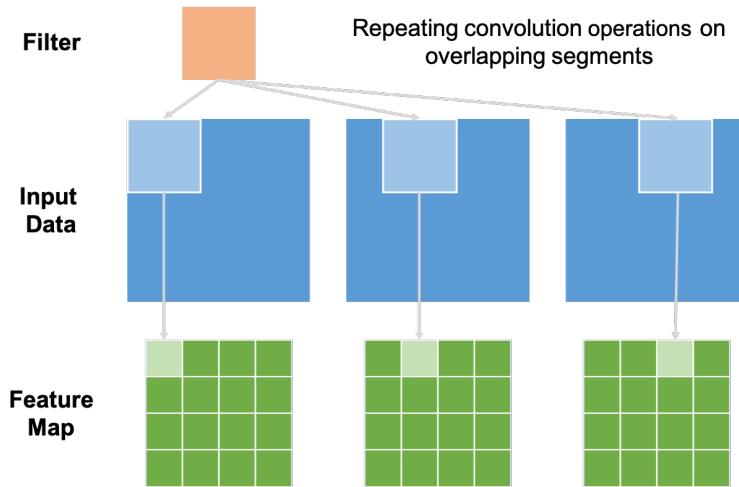


Figure 6: Convolution operation on a two-dimensional input data. A filter is repeatedly applied to overlapping segments of input data to generate an overall feature map.

Functionally, a **convolutional** layer in a network carries out a convolution operation on the input data to create a feature map [18]. As illustrated in Figure 6, a filter can be multiplied with an input data segment to obtain the product which summarises the information within that specific segment. By systematically repeating this operation across overlapping input data segments, a feature map can be produced to provide an overview of the features extracted from the input data.

A **max-pooling** layer is often applied after the convolution to downsample the feature map, lowering the resolution of the map to accentuate the dominant features. That being said, the output from the convolutional and max-pooling layers represents the prominent high-level extracted features. This output can then be flattened and propagated through some non-linear **fully-connected** layers, where the extracted features are used to acquire the optimal functional mapping between the input and the final output.

In recent years, CNN has been capitalised on extensively in a wide range of fields such as computer vision and natural language processing [19]. It is arguably the most powerful and efficient type of deep learning models, especially for image interpretation. Nevertheless, an abundance of research conducted in the past has suggested that CNN is also highly applicable to time series analysis.

Some preliminary study was conducted in 2017 to examine whether CNN can be leveraged to search for gravitational waves [20]. The results reported confirmed that the CNN model *Deep Filtering* developed can detect the true gravitational-wave events observed by LIGO and Virgo. This promising outlook thus motivated the development of an innovative CNN-based gravitational-wave search method which overcomes the limitations discussed in Section 2.1.6.

3 Methods

The objective of this project is to establish an alternative solution for gravitational-wave search using deep learning convolutional neural networks. Simulation algorithms of noise and gravitational waveform were developed to generate comprehensive strain datasets for informing the analysis. Analytical CNN models for detection and parameter estimation were then devised and trained to attain both high detection sensitivity and prediction accuracy.

3.1 Noise Simulation

In general, the noise in the LIGO detectors can be considered approximately stationary and Gaussian. To authentically simulate the detector noise, we derived the power spectral density from sampled real LIGO noise, generated simulated noise using a Gaussian model in the frequency domain, then converted the noise to a time series through the inverse Fourier transform.

3.1.1 Sampling of Real LIGO Noise

Real LIGO detector noise samples were collected from the LIGO Open Science Center (L OSC) around the 11 detected events in O1 and O2. By taking samples around the observed event to ensure the detectors were in steady-state operation, the noise behaviour captured in the samples can be considered representative of the full analysis period. Strain data can be accessed remotely by utilising GWpy, a package for analysing and characterising gravitational-wave data in python. In this project, 15 second time series were fetched from both detectors 30 seconds before the recorded events, allowing the sampled data to encapsulate the noise characteristics and exclude gravitational-wave signals at the same time.

3.1.2 Characterisation of Noise Using Power Spectrum

The LIGO detector noise is approximately stationary in the absence of glitches, meaning that its statistical properties do not change over time and it can be characterised by its **power spectral density (PSD)** $S_n(f)$. Such an approach in the frequency domain involves complex Fourier transform, hence it is essential to conduct the analysis cautiously to accurately select and preserve information of interest.

To prevent any non-stationary glitches contaminating the sampled data and resulting in atypical PSD, the PSD was estimated from the samples by applying Welch's method. This method divides a sampled time series into successive overlapped segments, computes the PSD for the individual segment using the fast Fourier transform (FFT), and take all segments into account to give an averaged PSD. With a default sampling frequency of $f_s = 4096$, the number of data points in each segment was set to be $4f_s$, the number of points of overlap between segments was set to be $2f_s$, and a Tukey window with its shape parameter $\alpha = 0.25$ was applied to each segment before the FFT. Applying this transformation method, the average PSD of all samples was calculated to provide a quintessential description of the detector noise in O1 and O2.

In Figure 7, the average noise PSD of LIGO Livingston (L1) and Hanford (H1) are presented. Both detectors are devised to observe gravitational waves in the frequency range from 10Hz to 10kHz while having their most sensitive band around $100 - 300\text{Hz}$. It can also be seen that although the two detectors have similar power spectra, there are still visible deviations in low and high-frequency regimes.

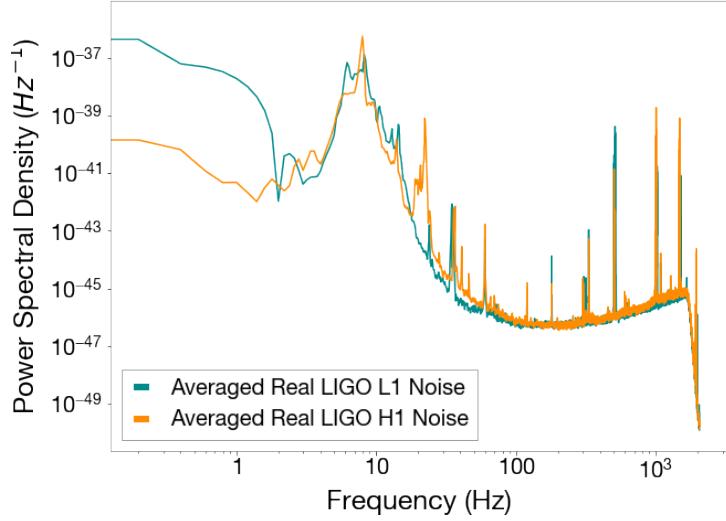


Figure 7: Average power spectral density of the detector noise for LIGO Livingston (L1) and Hanford (H1) in O1 and O2, acquired from 11 samples of 15s of real LIGO data respectively.

For the remaining part of the analysis, the LIGO H1 detector was chosen to be the detector of interest to ensure that the differences between detectors are considered, so that noise simulation is detector specific and accurate.

3.1.3 Gaussian Noise Generation in Frequency

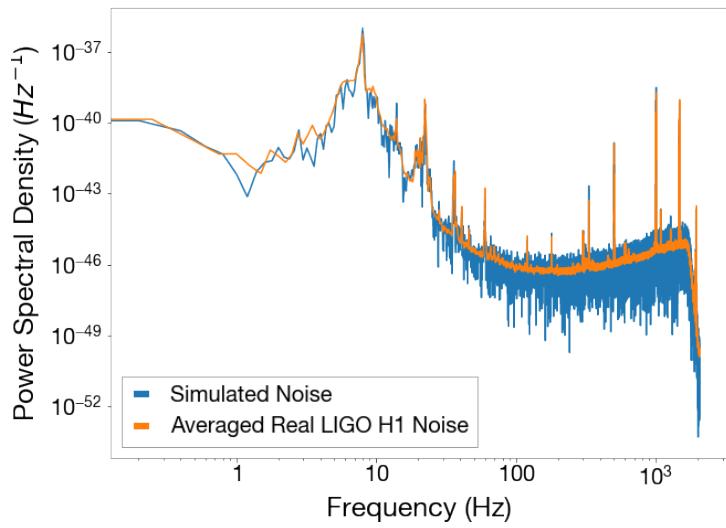


Figure 8: Power spectral density of an example of simulated LIGO H1 noise, compared with the averaged real LIGO H1 noise. The simulated noise was computed using Gaussian functions regarding the average power spectral density as described in Equation 3.

Having obtained the average PSD which succinctly characterises the detector noise, the noise was then simulated by exploiting its Gaussian property. That being said, the noise is independent at

each frequency and follows Gaussian distribution with amplitude proportional to $S_n^{1/2}(f)$ and random phase. Simulated noise $\tilde{n}(f)$ was computed in the frequency domain using Gaussian functions $G(\mu, \sigma)$ with expected value $\mu = 0$ and variance $\sigma^2 = S_n(f)/2$:

$$\tilde{n}(f) = G(0, \frac{S_n^{1/2}(f)}{\sqrt{2}}) + iG(0, \frac{S_n^{1/2}(f)}{\sqrt{2}}) \quad (3)$$

where the $1/\sqrt{2}$ factor was introduced to scale the modulus of the complex noise to $S_n(f)$. Figure 8 shows the PSD of a simulated noise in the LIGO H1 detector. The simulated noise agrees with the general shape of the average detector PSD while having greater irregularity for a real noise. Depending on the desired time length of simulated noise, sometimes it is necessary to perform interpolation on the average PSD to increase the number of data points prior to the noise generation. This is equivalent to applying a higher sampling rate to avoid undersampling and aliasing for the inverse FFT. More specifically, generating T second of simulated noise by the inverse FFT requires $t_s = T$ data points per Hertz for the average PSD, where t_s is the sampling rate in frequency. In other words, this is a consequence of the Nyquist theorem when sampling in frequency.

3.1.4 band-pass Filtering

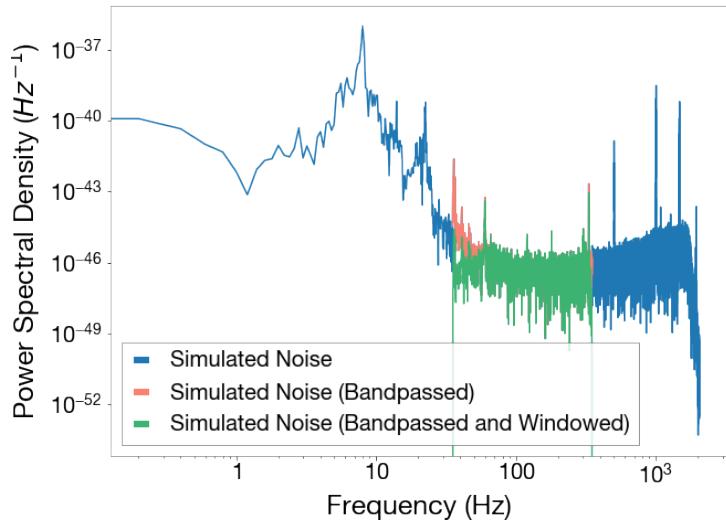


Figure 9: Band-pass filtered and windowed power spectral density of simulated LIGO H1 noise. A passband of $[35, 350] \text{Hz}$ and a Tukey window with shape parameter $\alpha = 0.25$ were imposed. The band-pass filtering operation is crucial for highlighting the frequency range of interest and disregarding inconsequential information. Applying the window function further facilitates the inverse FFT.

Considering the large noise amplitude in both low and high-frequency regimes, band-pass filtering was needed to enhance the visibility of features of interest by removing irrelevant noise. A passband of $[35, 350] \text{Hz}$ was implemented to accentuate the most sensitive region on the power spectrum for observing gravitational waves. Furthermore, window functions had to be applied to suppress spectral leakage due to discontinuous boundaries, which could potentially lead to false features in transformation. In this case, a Tukey window with shape parameter $\alpha = 0.25$ was convolved with the band-pass filtered PSD. The window was chosen because of its flat central region which preserves the original form of the spectrum.

Figure 9 visualises how the simulated noise PSD was band-pass filtered and windowed. The noise in low and high-frequency regimes was discarded and the region of interest was appropriately windowed for the inverse FFT operation.

3.1.5 Inverse Fourier Transform

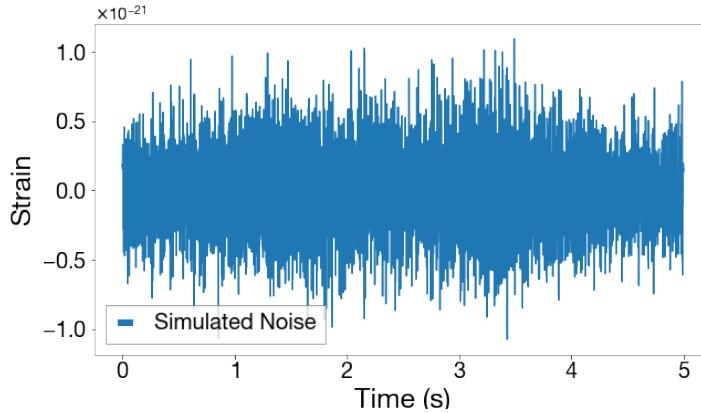


Figure 10: Example of a simulated LIGO H1 noise in the time domain with $T = 5\text{s}$ and $f_s = 4096$. It was band-pass filtered with a $[35, 350]\text{Hz}$ passband.

Subsequently, the simulated noise PSD was transformed into a real-time series through the real inverse FFT method to enable analysis in time. The normalisation method was set to "ortho" and a factor of $f_s^{1/2}$ was introduced such that the output noise was scaled to the correct amplitude in the time domain. The final output of the noise simulation pipeline is a time series of T second with sampling frequency f_s which can be characterised by its power spectrum in the frequency domain, as shown in the example in Figure 10.

Figure 11 exhibits the comparison between time series of simulated noise and band-pass filtered real LIGO H1 noise in time frames of 5s and 0.1s . The simulated noise appears to be consistent with the real one in terms of amplitude and patterns, therefore this visual comparison provides a preliminary validation for the noise simulation method discussed in this section.

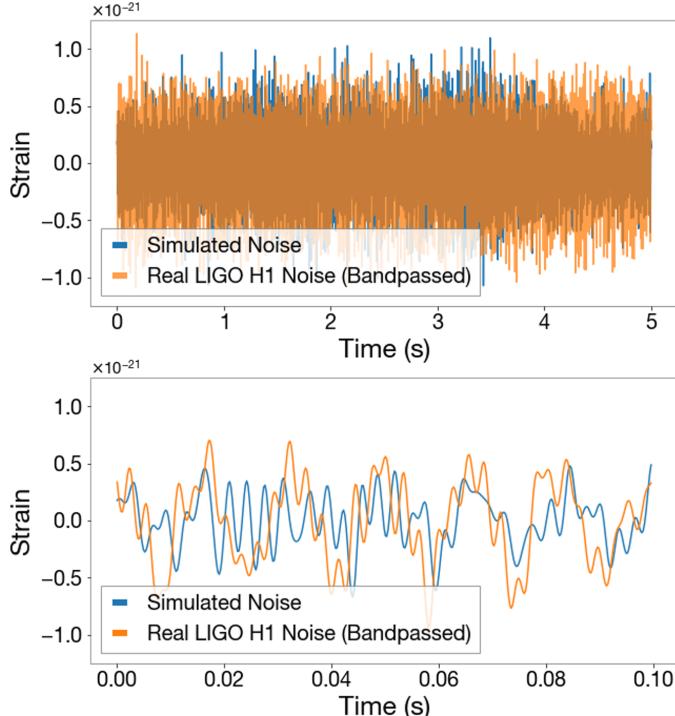


Figure 11: Comparisons between a simulated LIGO H1 noise and a band-pass filtered real LIGO H1 noise in the time domain, for $\Delta T = 5\text{s}$ (top) and $\Delta T = 0.1\text{s}$ (bottom) with $f_s = 4096$. Both noise are band-pass filtered with a $[35, 350]\text{Hz}$ passband. The similarity in both the waveforms and amplitudes verifies that the noise simulation method developed is effective.

3.2 Waveform Simulation

We first established the key parameters, approximations, and assumptions considered in gravitational waveform modeling. Waveforms were initially generated using PyCBC, then completed by concatenating a basic sine wave. Simulation of gravitational-wave strain data can be accomplished by simply injecting gravitational waveforms into noise.

3.2.1 Application of PyCBC

In the project, simulated gravitational waveforms were generated by PyCBC, a software package specialised in exploring astrophysical sources of gravitational waves. PyCBC is the foundation of the binary coalescence search, and it is heavily deployed in the ongoing gravitational-wave analysis at LIGO and Virgo. In the PyCBC analysis, compact binary systems with a total mass larger than $4M_\odot$ are modeled on the effective-one-body formalism, which leverages the post-Newtonian approximation, black hole perturbation theory, and numerical relativity.

In practice, stimulated waveforms are generated in the time domain with the function `get_td_waveform()`. This model can take into account numerous parameters for the construction of waveform, including the masses, spins on the x, y, z directions, luminosity distance, tidal deformability, angular frequencies, coalescence phase, inclination, etc. Although all these factors collaboratively govern the waveform of a compact binary merger, the masses and luminosity distance are the two leading parameters which predominantly determine the waveform.

3.2.2 Parameters, Approximations, and Assumptions

The **chirp mass** \mathcal{M} of a compact binary system, which is a function of the primary mass m_1 and the secondary mass m_2 :

$$\mathcal{M} = \frac{m_1 m_2^{3/5}}{(m_1 + m_2)^{1/5}} \quad (4)$$

The chirp mass determines the orbital evolution of the system as a result of energy loss from emitting gravitational waves. In other words, it guides the time-frequency morphology of gravitational waves, which evolves with the inspiral, merger, and ringdown phases in a coalescence event.

On the other hand, the amplitude of gravitational waves obeys an inverse law with respect to the distance from its source. The **luminosity distance** d_L can be defined by the relationship between the flux F measured by an observer and the luminosity L of a source:

$$F = \frac{L}{4\pi d_L^2} \quad (5)$$

The luminosity of a source locating at a certain physical distance is affected by the curvature of spacetime, meaning that the true physical distance can only be derived from the luminosity distance if the spacetime geometry is known. For this reason, the luminosity distance was chosen to be the distance measure considered in the study of gravitational waves since it does not require assumed knowledge on the cosmology.

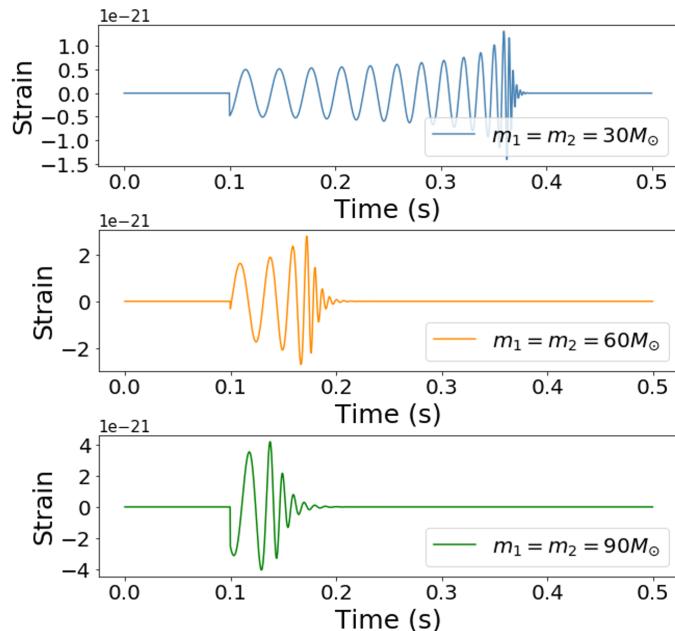


Figure 12: Simulated gravitational waves generated by PyCBC for equal, for varying masses $m_1 = m_2 = 30, 60, 90 M_\odot$ with a constant luminosity distance $d_L = 500 Mpc$. Greater component masses are related to shorter duration and larger amplitude.

As shown in Figure 12, the waveform changes with the component masses, hence the chirp mass of a compact binary system. More massive binary systems have shorter process durations, whereas less massive ones tend to have longer, more prominent inspiral phases. Conversely, Figure 13 demonstrates that varying the luminosity distance only changes the amplitude of gravitational

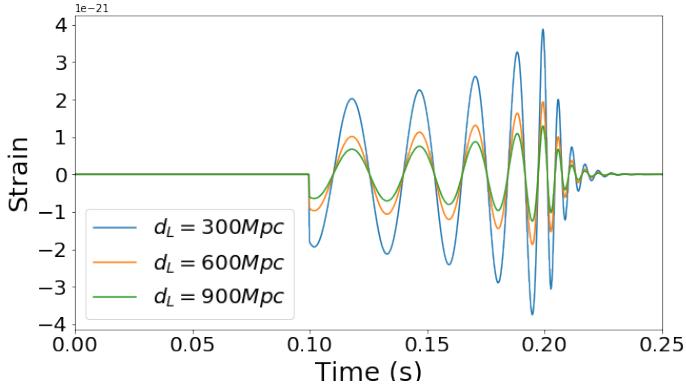


Figure 13: Simulated gravitational waves generated by PyCBC for varying luminosity distance $d_L = 300, 600, 900 Mpc$ with constant component masses $m_1 = m_2 = 50 M_\odot$. Larger luminosity distance is related to smaller amplitude.

waves, and it has no impact on the overall waveform pattern. Intuitively, longer distance is associated with smaller amplitude as a consequence of the inverse law.

It needs to be noted that the magnitudes of component spins and their misalignment with respect to the orbital angular momentum also have a considerable effect on the waveform. If the spins are aligned, the separation before merging will be small, resulting in a longer inspiral process before the objects get sufficiently close to merging. In contrast, the misalignment of spins implies that objects will merge at a greater separation and give out shorter signals. However, the modeling of component spins and orbital angular momentum entails the computation of 9 additional parameters. This approach will significantly complicate the problem so that it is computationally infeasible to incorporate these factors. As a result, the decision of neglecting spins in waveform simulation was made.

In summary, only the primary and secondary masses m_1, m_2 with a unit of solar mass M_\odot , as well as the luminosity distance d_L with a unit of Mpc were considered for gravitational-wave simulation using the PyCBC function `get_td_waveform()`. Nevertheless, there are two supplementary parameters to be specified for the PyCbC waveform generation: the approximant and low frequency bound. The approximant specifies the mathematical model for waveform simulation. In this case, the approximant chosen was “SEOBNRv2” [21], which computes the dominant mode of the gravitational waves emitted by binary black holes. And the low frequency bound was set to $30 Hz$ to ensure the waveforms created to have a frequency range compatible with the simulated noise.

Furthermore, `get_td_waveform()` returns both the plus and cross polarisations of a waveform in the time domain. To simplify the modeling of waveform, it was assumed that the incident gravitational waves passing through the detector are always perfectly aligned with the orthogonal arms, so the projected amplitude measured is the maximum possible. That is to say, only the plus polarisation was considered for waveform simulation.

3.2.3 Completion of Waveform

A waveform generated by `get_td_waveform()` has an abrupt, arbitrary start and only includes a fraction of the inspiral phase. To construct a complete, realistic wave within the sampled time frame, a sine wave was concatenated prior to the waveform. This sine wave was computed based on the instantaneous amplitude $A(t)$, phase $\phi(t)$, and angular frequency $\omega(t)$ around the beginning of the waveform to facilitate a smooth, natural transition from stable trigonometric wave to gravitational waveform.

First, the waveform was converted into an analytical signal by applying the Hilbert transform. Following this step, the instantaneous amplitude $A(t)$ was attained from the amplitude envelope, which is the modulus of the analytical signal. Similarly, the instantaneous phase $\phi(t)$ was obtained from the argument of the analytical signal, with the values unwrapped to their 2π compliments. The instantaneous angular frequency $\omega(t)$ was then calculated by finding the first derivative of the instantaneous phase. Subsequently, the sine wave to be concatenated can simply be modeled as $A(t_0)\sin[\omega(t_0)t + \phi(t_0)]$, where t_0 was set to be 0.025s from the start of the waveform to eliminate edge effects in the Hilbert transform.

It should be noted that in practice, a zero time series with a specified size was created first. A PyCBC waveform was introduced into the time series at a random time and the sine wave was concatenated to fill the gap between the $t = 0$ and the original waveform.

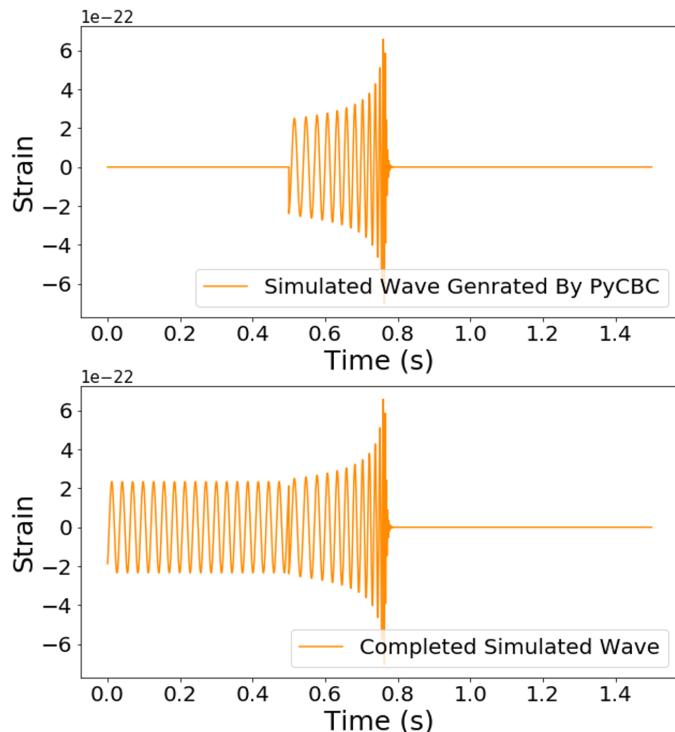


Figure 14: (Top) Example of a simulated gravitational waveform generated by PyCBC for $m_1 = m_2 = 50M_{\odot}$ and $d_L = 1000Mpc$. (Bottom) Example of a complete simulated gravitational waveform. A sine wave with amplitude $A = 2.34 \times 10^{-22}$, phase $\phi = 5.36rad$, and angular frequency $\omega = 5.3 \times 10^{-2}rads^{-1}$ was concatenated prior to the original waveform.

In Figure 14, the top graph shows a gravitational waveform generated by PyCBC with an abrupt, discontinuous start. And the bottom graph illustrates that a complete gravitational waveform can be created by concatenating a sine wave prior to the original waveform. This gives an authentic, realistic simulation of overall gravitational waves contained in strain data.

3.2.4 Strain Data Simulation

The final step towards simulating strain data is simply injecting waves into noise. In other words, a simulated strain data $d(t)$ can be written as the sum of a gravitational wave $h(t)$ and a background noise $n(t)$ such that $d(t) = h(t) + n(t)$. In Figure 15, it can be seen that a wave is embedded

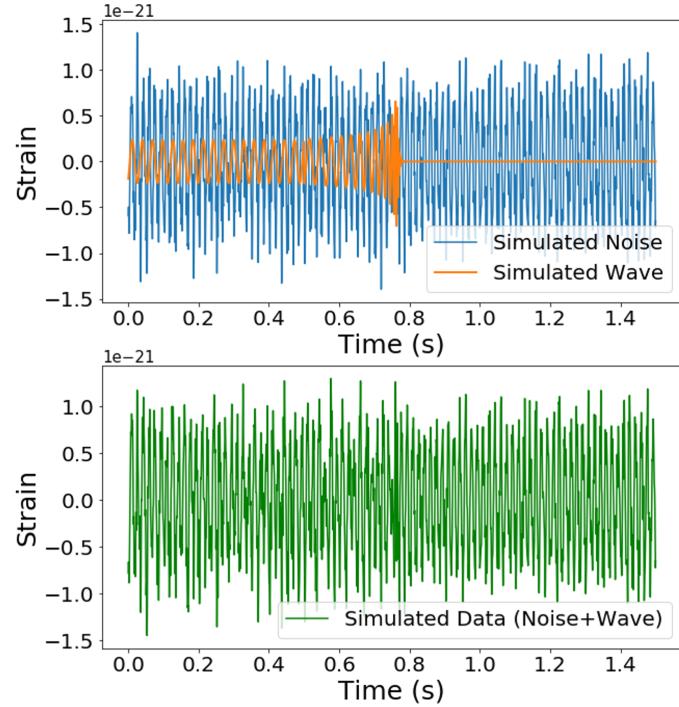


Figure 15: (Top) Illustration of injecting a gravitational wave into noise. (Bottom) Example of a simulated strain data as a sum of gravitational wave and noise.

into noise to produce strain data, highlighting that it is difficult to determine whether or not a gravitational wave is present in the strain data.

3.3 Detection and Parameter Estimation

The objective of this project is to develop the convolutional neural network (CNN) models for detection and parameter estimation of gravitational waves. In order to attain them, we first generated training datasets by employing the strain simulation method described in the previous section. Two CNNs were then constructed and trained with some simulated strain datasets to obtain the models for detection and parameter estimation.

3.3.1 Tensorflow and Keras

A machine learning framework is an interface, library or organised tool which facilitates the development of machine learning solutions. For this project, Tensorflow was selected as the framework for devising the CNN models. It is an open-source platform with a comprehensive, flexible ecosystem of tools, libraries, and community resources to support the implementation of machine learning applications. Additionally, Keras, a high-level application programming interface, was used in conjunction with Tensorflow to provide a more user-friendly experience.

3.3.2 Training Data

The training dataset for the detection model includes both data samples with and without signals of gravitational waves, whereas the training dataset for the parameter estimation model is simply composed of data samples with gravitational-wave signals of random component masses and luminosity distances.

In both cases, the training datasets comprise 10000 time series with a length of 5 seconds and a sampling frequency of 4096. Moreover, the gravitational-wave signals present in the data samples correspond to binary coalescence events with primary masses between 10 to $100M_{\odot}$ and luminosity distance between 10 and $1000Mpc$.

Data Attribute	Training Dataset
Dataset Size	10000
Strain Data Length	5 s
Primary Mass m_1 range	$10 - 100 M_{\odot}$
Secondary Mass m_2 range	$10 - m_1 M_{\odot}$
Luminosity Distance d_L range	$10 - 1000 Mpc$

Table 1: Key data attributes defining the training datasets for the detection and parameter estimation models.

In principle, a non-linear model only generalises well and achieves high accuracy when trained with an adequately large dataset. This means that it is often favourable to have a training dataset as large as possible. However, this is not feasible because the hardware deployed in the project lacked the computational capacity and storage space needed for managing a large amount of data. Therefore, it was decided that the optimal size of the training datasets is 10000.

Considering that gravitational-wave signals typically have durations less than a second, it is preferred to have slightly longer strain data samples with a length of a few seconds. This introduces further uncertainties for realistic physical data simulation since in this case, a signal has greater flexibility to appear at a random time within the data time frame. Consequently, a data length of 5

second was chosen alongside a default sampling frequency of 4096.

The simulated gravitational waveforms were generated by using uniform random number generators for both the component masses and luminosity distance. The primary mass m_1 was set to range from 10 to $100M_\odot$, and the secondary mass m_2 was set to range from 10 to m_1 to avoid mass degeneracy. This arbitrary mass range was designed to be comparable to the masses of the binary black holes observed in O1 and O2 ($7.6 - 50.6M_\odot$) [22]. Similarly, the range of the luminosity distance d_L was set between 10 and $1000Mpc$ such it has a magnitude comparable to the luminosity distances of the binary black holes discovered in O1 and O2 ($320 - 2750Mpc$).

For the detection training dataset, a uniform random number generator was utilised to ensure that approximately half of the strain data samples contain gravitational wave signals. Each of the data samples was then assigned a label indicating the presence or absence of gravitational-wave signals. More specifically, the data samples with signals were labeled 1 and the ones without were labeled 0. In other words, detection is a binary classification problem.

For the parameter estimation training dataset, each data sample was associated with two labels: chirp mass \mathcal{M} and luminosity distance d_L . This suggests that parameter estimation is a two-dimensional regression problem. The chirp mass was considered here because it governs the orbital evolution of a compact binary system. Based on the component mass range used, the chirp mass has a range from 8.7 to $87M_\odot$.

The datasets were normalised prior to the training to improve training efficiency. Also, the signal-to-noise ratio (SNR) for each strain data sample was calculated to assess the relative strength of the gravitational-wave signal to background noise in terms of power. The amplitudes for both the signal and noise were squared and integrated over the data segment where the signal is present to measure the powers.

3.3.3 Detection Model

	Layer Type	Output Shape
1	Reshape	(20480,1)
2	Convolution	(20478,8)
3	MaxPooling	(5119,8)
4	ReLU	(5117,8)
5	Flatten	40960
6	ReLU	128
7	ReLU	64
8	ReLU	32
9	ReLU	16
10	ReLU	8
11	Sigmoid	1

Table 2: Architecture of the convolutional neural network model for gravitational-wave detection.

The CNN constructed for the detection model is shown in Table X. This architecture was empirically determined by systematic testing. The Reshape input layer first transformed the one-dimensional input data into a two-dimensional array to enable the convolution operation. Subsequently, the data was propagated through a convolutional layer with a filter size of 3, a max-pooling

layer, and several fully-connected layers. Lastly, the output was produced at the final layer by a sigmoid activation function, which is an S-shaped curve with normalised values bounded between 0 and 1, making it particularly suitable for predicting continuous probability.

The optimiser opted was Adam and the loss function selected was binary cross-entropy, the default for binary classification. The detection CNN model was trained over 30 epochs to allow the loss functions for both the training and validation data to decrease and converge.

3.3.4 Parameter Estimation Model

	Layer Type	Output Shape
1	Reshape	(20480,1)
2	Convolution	(20478,16)
3	MaxPooling	(5119,16)
4	ReLU	(5119,16)
5	Convolution	(5117,8)
6	MaxPooling	(1279,8)
7	ReLU	(1279,8)
8	Flatten	40952
9	ReLU	256
10	ReLU	256
11	ReLU	128
12	ReLU	128
13	ReLU	64
14	ReLU	64
15	ReLU	32
16	ReLU	32
17	Linear	2

Table 3: Architecture of the convolutional neural network model for gravitational-wave parameter estimation.

The CNN built for the parameter estimation model is shown in Table X. Its architecture was derived from the detection CNN, and the final configuration was again empirically determined by systematic testing. Identically, the reshape input layer first transformed the one-dimensional input data into a two-dimensional array. Then the data was passed through a series of convolutional layers a filter size of 3, max-pooling layers, and fully-connected layers. The output was eventually created at the final layer by a linear activation function, which gives output values that are continuous and unbounded.

The optimiser chosen was Adam and the loss function implemented was mean squared error (MSE), the default for regression. The parameter estimation CNN model was trained over 30 epochs so that both the test and validation losses were reduced sufficiently and converged to a minimum value.

4 Results and Discussion

4.1 Detection

The convolutional neural network (CNN) detection model was evaluated against a validation dataset to obtain an unbiased estimate of its prediction skill. It was found that the model achieved an overall binary accuracy of 97.07%, suggesting that the model is able to successfully detect gravitational-wave signals in noisy strain data with a high degree of accuracy.

Further analysis of the prediction errors revealed the false positive rate of the model is 0.79% and the false negative rate is 5.10%. The small false positive rate indicates high detection confidence, meaning that it is unlikely that some random noise fluctuations will be mistaken as true gravitational waves. On the other hand, the moderate false negative rate implies that the probability of a true gravitational wave being wrongly classified and disregarded is non-negligible. It is inevitable that a considerable fraction of gravitational waves will be overlooked during the analysis.

However, there are the validation results of an independent single-detector experiment, which does not depend on identifying coincident signals across multiple detectors. By incorporating strain data from other detectors, the model can be modified to attain improved accuracy and reduced false rates. This self-sufficient single-detector model has addressed one of the main limitations of the current LIGO-Virgo search method, yet it can still be complemented by additional data from other detectors to provide more confident and reliable predictions.

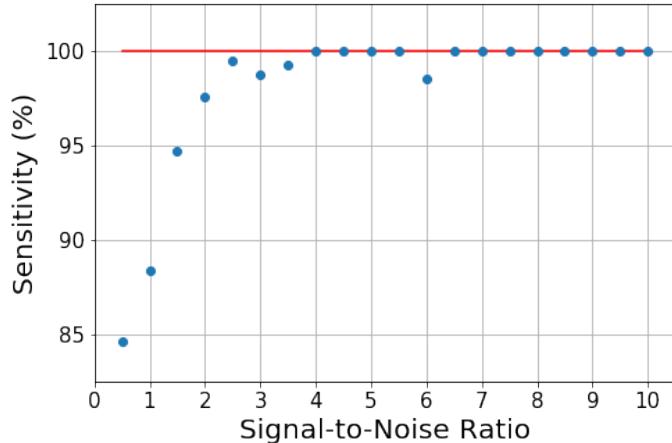


Figure 16: Relationship between the sensitivity of the detection model and signal-to-noise ratio. The sensitivity increases with SNR and converges to 100% when SNR is larger than 6.

The sensitivity, or the positive detection accuracy, is defined as the probability of successfully detecting a true gravitational-wave signal. Figure 16 shows the sensitivity of the detection model as a function of signal-to-noise ratio (SNR). It was discovered that the sensitivity increases with SNR and eventually reaches the maximum 100% when SNR is greater than 6. For weak signals with small SNR (< 0.5), the model still acquired a satisfactorily high sensitivity of 84.58%. This overall behavior is in good agreement with our knowledge of SNR, which measures the relative strength of a signal compared to noise. That is to say, for a strain data sample with large SNR, the signal can easily be distinguished from noise and detected with a high level of confidence. Nevertheless, the CNN detection model still demonstrated a strong performance in confidently identifying weak gravitational-wave signals with small amplitudes and SNR.

The sensitivity of the model was also be benchmarked against the matched-filtering algorithm

utilised by the LIGO-Virgo Collaboration and the existing CNN-based search algorithm *Deep Filtering*. 100% sensitivity was obtained by the matched-filtering approach for SNR larger than 8, and by the *Deep Filtering* method when SNR is above 10 [20]. These findings and comparisons manifest that the detection model outperforms these well-established search algorithms and succeeds in detecting gravitational waves with small amplitudes, overcoming one more major limitation of the current best practice at LIGO-Virgo.

In summary, the results evaluated using the validation dataset have substantiated that the CNN detection model is an ideal alternative for resolving the key limitations of the matched filtering search employed by LIGO and Virgo. It can achieve high overall accuracy by only analysing data from a single detector and extract weak gravitational signals hidden in noisy strain data.

Having completed the analysis with the simulated dataset, the effectiveness of the model was then verified with the LIGO real data of the 11 gravitational-wave events recorded by LIGO and Virgo. The real data was band-pass filtered and normalised so that it has the same specification as the training data. It was found that the model managed to detect all 11 events, including the one for neutron star binaries. This final verification confirms the remarkable performance of the CNN detection model and its applicability to analysing real gravitational-wave strain data.

4.2 Parameter Estimation

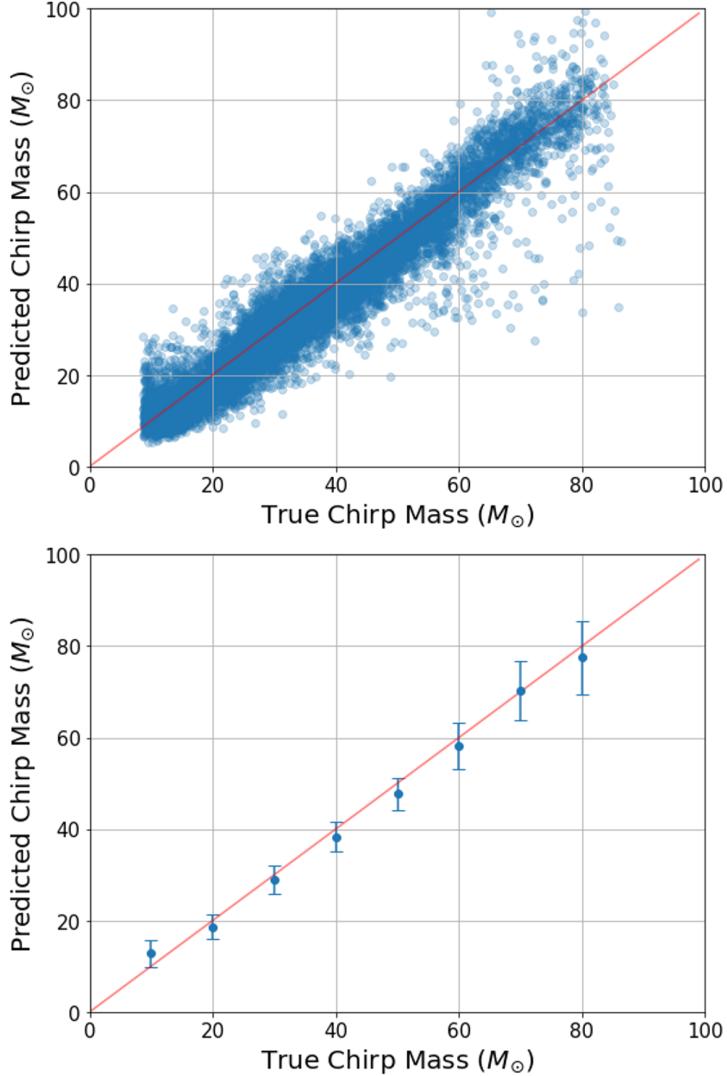


Figure 17: (Top) Predicted chirp mass values evaluated by the CNN detection model against true, labeled chirp mass values. (Bottom) Average and standard deviation of predicted chirp mass values at an interval of $10M_{\odot}$. Overall, there is a strong correlation confirming that the model is able to estimate the chirp mass accurately to some extent. The average predicted values are also in good agreement with the true values across the mass range.

The CNN parameter estimation model was evaluated against a validation dataset to assess its skill for predicting the chirp mass and luminosity distance of a compact binary system. As displayed in Figure 17 and Figure 18, there exist strong positive correlations between the predicted values and the true, labeled values for both the chirp mass and luminosity distance. The average percentage errors for the chirp mass and luminosity distance are 13.57% and 15.39% respectively. Comparing with the average percentage uncertainties of 6.16% and 37.43% provided by the LOGI-Virgo matched-filtering method [22], it was observed that the model has an acceptable accuracy for computing the chirp mass, and is especially suitable for calculating the luminosity distance. These experimental results suggest that the model has gained the ability to accurately estimate the parameters from a gravitational-wave signal to some extent. The model essentially established a

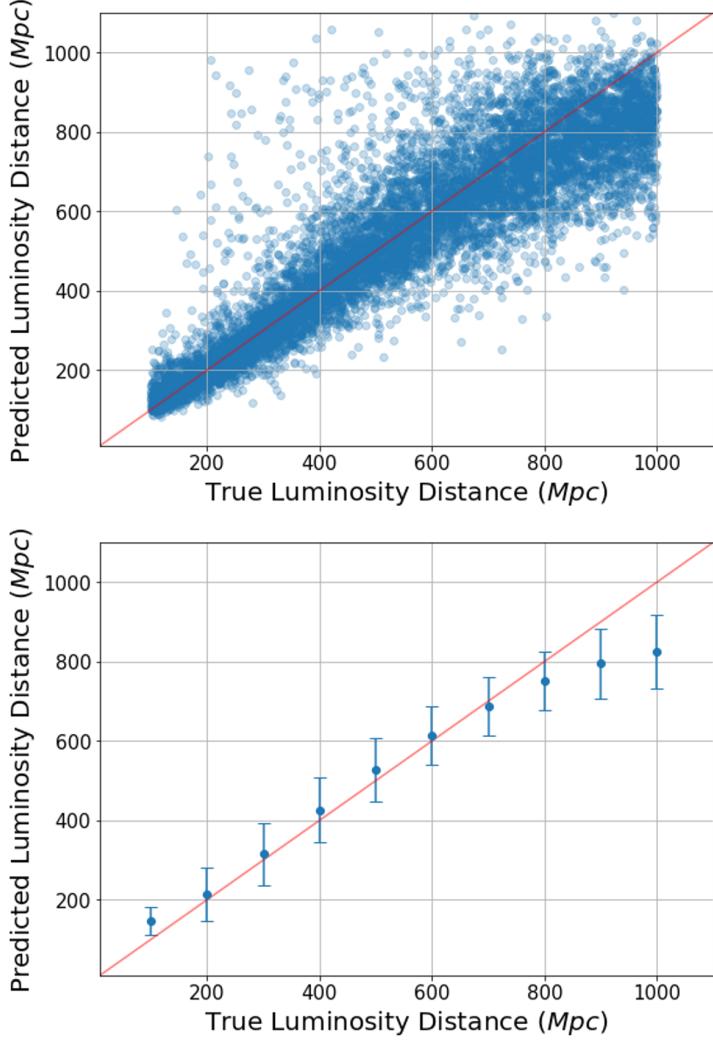


Figure 18: (Top) Predicted luminosity distance values evaluated by the CNN detection model against true, labeled luminosity distance values. (Bottom) Average and standard deviation of predicted luminosity distance values at an interval of 100Mpc . Overall, there is a strong correlation confirming that the model is able to estimate the luminosity accurately to some extent. The average predicted values are generally in good agreement with the true values, but the model is prone to underestimate in the long-distance regime.

generalised solution for predicting the parameters without modeling gravitational waveform, thus providing a more efficient and extensive parameter estimation algorithm. In other words, the model also overcomes the last main limitation of the matched-filtering method by eliminating the need for comparing a signal against thousands of modeled waveforms in a template bank.

For the chirp mass estimation, the average predicted values are in good agreement with the true values across the mass range. Whereas the average predicted values for the luminosity distance deviate slightly in the long-distance regime, showing that the model has a tendency to underestimate in this region. These findings indicate that the model is able to estimate the parameters accurately for the entire parameter space except for $d_L > 800\text{Mpc}$. One possible explanation for the departure is that long luminosity distance is often associated with small SNR, implying it is more difficult to extract signals from noisy data. As a result, the model might have overcompensated by mistakenly including some fraction of noise as signals to increase the amplitudes, resulting in giving lower

predicted values.

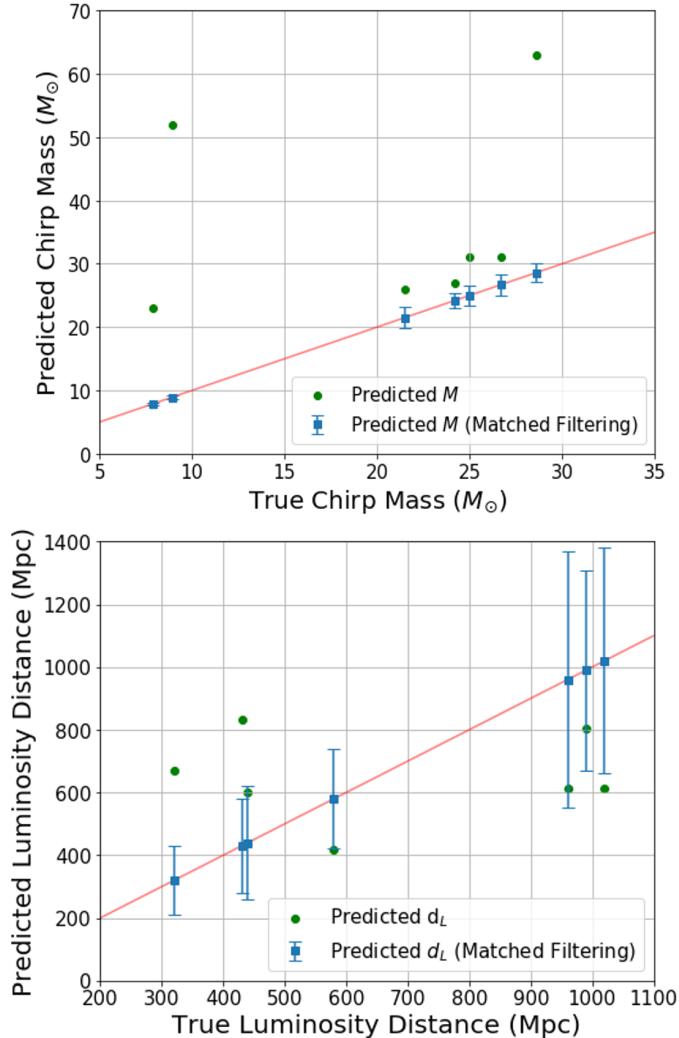


Figure 19: (Left) Chirp mass values for real gravitational-wave events predicted by the CNN parameter estimation model, compared with the values given by the matched filtering method. (Right) Luminosity distance values for real gravitational-wave events predicted by the CNN parameter estimation model, compared with the values given by the matched filtering method. The inconsistency between the two sets of estimated parameters can potentially be explained by the inaccuracies in the simplified waveform simulation.

The model was verified using the real LIGO data of the detected events with parameters compatible with the chirp mass and luminosity ranges set for the training. In Figure 19, it can be observed that the predicted values of both parameters are in general not consistent with the values computed by the LIGO-Virgo's matched filtering method, meaning that these results can not verify that the CNN parameter estimation model is effective on real gravitational-wave strain data.

Although the discrepancy can not be conclusively explained, there exist several hypothetical causes. For example, waveform simulation in this project neglected the effects of the component spins, spin alignment, cross polarisation, incident angle, etc. Also, the simulated signals were only low-passed filtered ($> 30\text{Hz}$). These inaccuracies in waveform modeling might have rendered the model less applicable to analysing real data. In principle, it is expected that the waveform simulation algorithm can be modified to generate more comprehensive, authentic gravitational-wave signals and further improve the parameter estimation model.

4.3 Limitations and Future Work

The performance of any neural network model is fundamentally determined by its training data and algorithm. In this project, the size of the training data and the complexity of the architecture are both restricted by the computational capacity and the storage space of the hardware used.

According to the previous evaluations, the parameter estimation model has a training loss of 433 after 30 epochs, while its validation loss has a considerably larger value of 6517. This implies that the model only generalised to a certain extent and was overfitting on the training data. Therefore, it is expected that a refined regression solution can be found by simply increasing the training dataset size to provide a more overarching dataset that the model can learn from. However, doing so will result in higher computational costs and greater demand for storage space. It is also difficult to estimate the amount of additional data required to significantly improve the model. As a consequence, it might be necessary to upgrade the hardware in order to allow more flexibility for the training dataset size.

Redesigning the architectures of the CNN models can potentially enhance their prediction capabilities. It is also possible to adjust other settings, such as weight initialisation, activation functions, optimiser, loss function, and number of epochs, to construct more powerful models. Unfortunately, the optimal designs can only be practically decided by systematically experimenting with different number of layers, number of nodes in each layer, arrangement of layers, and other components. This process is time-consuming and inefficient, so a more viable approach will be referencing the current CNN structures as well as the literature to guide the testing.

Moreover, the training data can be adjusted so that it becomes more compatible with the real LIGO data of the recorded events. In particular, the ranges of the component masses and luminosity distance can be extended to cover a larger parameter space. Based on the candidate events observed in O3, the majority of the compact binary systems discovered so far lies in the luminosity distance range from 100 to $5000 Mpc$. Hence it is expected that the models can become more suitable for analysing real LIGO-Virgo data by setting the luminosity distance range for the training dataset. Taking the component spins and other parameters into account will enable the model to gain more information about compact binary systems, as well as making it more robust and exhaustive against various scenarios. The incident angle can also be randomised by considering both the plus and cross polarisation and varying their proportion, so as to acquire more realistic training data.

From a statistical viewpoint, the current parameter estimation model can not calculate the uncertainty or error of each specific predicted value. To resolve this limitation, the model can be augmented with some probabilistic analysis algorithm to characterise the accuracy of each estimate. Lastly, the current models can only be applied to analyse discrete, static strain data segments. To that end, it is also desirable to develop models capable of conducting real-time analysis on continuous, dynamic strain data streams.

5 Conclusion

In this project, two convolutional neural network (CNN) models were developed for detection and parameter estimation of gravitational waves. The detection model achieved an overall accuracy of 97.07% and reached 100% sensitivity for signals with SNR greater than 6. It is able to perform single-detector analysis and identify weak signals in noisy strain data. The parameter estimation model acquired average percentage errors of 13.57% and 15.39% for predicting the chirp mass and luminosity distance respectively. It provides a general, efficient framework for accurately estimating the parameters of interest. These preliminary findings strongly suggest that the application of CNN is an appropriate alternative solution in the search of gravitational waves.

The main limitation of this CNN-based method is the lack of computational power and storage space required to generate larger, more extensive training datasets. In particular, the simplification of waveform simulation due to this constraint has compromised the parameter estimation model's applicability on real data. Further research is planned to upgrade the hardware to enable the use of larger datasets and the development of more rigorous waveform simulations.

Acknowledgements

The author gratefully acknowledge the valuable support and guidance provided by the project supervisor Professor Carlo Contaldi, as well as the substantial contributions made to all aspects of the research by the project partner.

References

- [1] Abbott, B.P. et al. (LIGO Scientific Collaboration and Virgo Collaboration). GW150914: The Advanced LIGO Detectors in the Era of First Discoveries. *Phys. Rev. Lett.* **116**, 131103 (2016).
- [2] Aasi, J. et al. (LIGO Scientific Collaboration). Advanced LIGO. *Class. Quantum Grav.* **32**, 074001 (2015).
- [3] Martynov, D.V. et al. Sensitivity of the Advanced LIGO Detectors at the Beginning of Gravitational Wave Astronomy. *Phys. Rev. D* **93**, 112004 (2016).
- [4] Belczynski, K. et al. A comprehensive study of binary compact objects as gravitational wave sources: evolutionary channels, rates, and physical properties.
- [5] (2002).
- [6] Abbott, B.P. et al. (LIGO Scientific Collaboration and Virgo Collaboration). GW150914: First results from the search for binary black hole coalescence with Advanced LIGO. *Phys. Rev. D* **93**, 122003 (2016)
- [7] Le Tiec, A. et al. The overlap of numerical relativity, perturbation theory and post-Newtonian theory in the binary black hole problem. *International Journal of Modern Physics D* (2014).
- [8] Taracchini, A. et al. Effective-one-body model for black-hole binaries with generic mass ratios and spins. *Phys. Rev. D* **89**, 061502 (2014).

- [9] Abbott, B.P. et al. (LIGO Scientific Collaboration and Virgo Collaboration). A guide to LIGO-Virgo detector noise and extraction of transient gravitational-wave signals. *Class. Quantum Grav.* **37**, 055002 (2019).
- [10] Aasi, J. et al. (LIGO Collaboration and Virgo Collaboration). Parameter Estimation for Compact Binary Coalescence Signals with the First Generation Gravitational-Wave Detector Network, *Phys. Rev. D* **88**, 062001 (2013).
- [11] Yiu, T. Understanding Neural Networks. *Towards Data Science* (2019). Available at: <https://towardsdatascience.com/understanding-neural-networks-19020b758230> [Accessed 7 Apr. 2020].
- [12] Bermudez, L. Overview of Neural Networks. *Towards Data Science* (2019). Available at: <https://medium.com/machinevision/overview-of-neural-networks-b86ce02ea3d1> [Accessed 7 Apr. 2020].
- [13] Sharma, S. Activation Functions in Neural Networks. *Towards Data Science* (2019). Available at: <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6> [Accessed 7 Apr. 2020].
- [14] Walia A.S. Activation functions and it's types-Which is better?. *Towards Data Science* (2020). Available at: <https://towardsdatascience.com/activation-functions-and-its-types-which-is-better-a9a5310cc8f> [Accessed 7 Apr. 2020].
- [15] Brown, J. A Gentle Introduction to the Rectified Linear Unit (ReLU). *Machine Learning Mastery* (2019). Available at: <https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks> [Accessed 7 Apr. 2020].
- [16] Brown, J. Gentle Introduction to the Adam Optimization Algorithm for Deep Learning. *Machine Learning Mastery* (2019). Available at: <https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning> [Accessed 7 Apr. 2020].
- [17] Brown, J. Loss and Loss Functions for Training Deep Learning Neural Networks. *Machine Learning Mastery* (2019). Available at: <https://machinelearningmastery.com/loss-and-loss-functions-for-training-deep-learning-neural-networks> [Accessed 7 Apr. 2020].
- [18] Saha, S. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. *Towards Data Science* (2018). Available at: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53> [Accessed 7 Apr. 2020].
- [19] Yamashita, R. et al. Convolutional neural networks: an overview and application in radiology. *Insights Imaging* **9**, 611–629 (2018).
- [20] George, D., Huerta, E.A. Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data. *Physics Letters B* **778** 64–70 (2017).
- [21] Purrer, M. Frequency domain reduced order model of aligned-spin effective-one-body waveforms with generic mass-ratios and spins. *Phys. Rev. D* **93**, 064041 (2016).
- [22] Abbott, B.P. et al. (LIGO Scientific Collaboration and Virgo Collaboration). GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Phys. Rev. X* **9**, 031040 (2019).