Applications of Deep Learning in Gravitational Wave Data Quality Studies: A Systematic Review

Yubin Kim[1] and Hyunggu Jung[2,3,4*]

[1]Department of Physics, Ewha Womans University, Seoul, 03760, Republic of Korea.
[2]College of Nursing, Seoul National University, Seoul, 03080, Republic of Korea.
[3]Center for World-leading Human-care Nurse Leaders for the Future by Brain Korea 21 (BK 21) four project, College of Nursing, Seoul National University, Seoul, 03080, Republic of Korea.
[4]The Research Institute of Nursing Science, Seoul National University, Seoul, 03080, Republic of Korea.

Abstract

The advent of deep learning in the field of gravitational wave (GW) data analysis has facilitated the development of effective tools for addressing the challenges encountered by traditional methods in handling the increasing volumes of real-time data. In the process of GW data analysis, data quality must be appropriately understood and improved because transient noise artifacts limit the detector sensitivity and reduce the robustness of GW detection. This study presents a systematic literature review that analyzes deep learning applications in GW data quality studies, investigating their application purposes, deep learning techniques employed, performance evaluation parameters, and training datasets. This systematic review was conducted according to PRISMA guidelines. We identified 54 studies from five databases that satisfied the eligibility criteria. The selected studies were examined to answer the research questions. The findings of the review indicate that convolutional neural networks are the most commonly used deep learning techniques, and it is expected that future research will focus on applying transformers to similar tasks. In addition, we identified that computation time was considered as one of the evaluation parameters for assessing the performance of deep learning techniques, whereas time-related parameters were not considered until 2022. Furthermore, we noted the imbalance in the use of detector-based datasets for the development of deep learning techniques. This review is expected to provide researchers with a comprehensive overview of current deep learning methods in GW data quality studies and highlight key considerations for future research directions.

Keywords: Gravitational waves, Artificial intelligence, Deep learning, Systematic literature review

1 Introduction

The first observation of a gravitational wave (GW) signal in 2015 [1] ushered in the era of GW astronomy. The detection rate of GW signals has increased with improvements in GW detectors. This has brought about a rapid growth in the field of GW data analysis, which has provided valuable information on the properties of GW sources and the nature of GWs [2]. However, ground-based detectors encounter challenges from a variety of external noise sources in addition to inherent detector noise, which constrains detector sensitivity[3]. Transient noise artifacts (i.e., glitches) deteriorate the data quality and reduce the reliability of GW detection. Some of these glitches mimic actual GW signals, potentially leading to false alarms [4]. Therefore, data quality analysis to exclude poor quality data or to assess the impact of artifacts on the actual GW signal is crucial for improving the data quality and is also an essential step for GW data analysis [5].

Several software tools such as UPV [6], hVETO [7], iDQ [8, 9], and PyChChoo [10] have been proposed to correlate glitches with numerous auxiliary channel data from the sensors monitoring the detector. However, these types of mitigation techniques reduce the analyzable data that might contain astrophysical GW signals. While it is desirable to subtract glitches through detector characterization and noise modeling, the vast volume of data and the complex coupling characteristics of noise render glitch subtraction extremely challenging [11], [12]. In addition, large amounts of data require real-time analysis in the era of multi-messenger astronomy. However, traditional methods are inefficient in handling real-time data quality assurance [13].

Artificial intelligence, particularly machine learning, has been proposed as a promising approach for GW detector data processing. Early research in the field focused on using machine learning to automate the identification and classification of transient noise events [14], [15]. The advent of deep learning has enabled more efficient analysis and management of large volumes of data related to auxiliary channels, which have been regarded as significant challenges in previous data quality analysis within the domain of GW data analysis [16]. Recently, deep learning-based techniques have been applied to every stage of GW data analysis. Many studies have applied deep learning to GW signal detection and parameter estimation, where traditional methods have been used [17], [18], [19].

The applications of deep learning in the fields of astronomy and astrophysics, including GW data analysis, will become even more crucial in the future. A comprehensive and systematic understanding of the deep learning techniques that have been applied to date is needed for future technological development. Nevertheless, as of November 2024, a systematic literature review on the application of deep learning techniques to the analysis of GW data has not yet been conducted. To date, only two review papers have been published on the use of machine learning techniques in the field of GW science. In a pioneering effort, Cuoco et al.[20] conducted a comprehensive review of the applications of machine learning techniques for ground-based GW detector data analysis up to December 2020, when their paper was published. In 2023, Benedetto et al.[21] conducted a review of artificial intelligence techniques applied to GW analysis, exploring their applications in relation to software and hardware. However, neither of these papers has focused exclusively on data quality or deep learning techniques in particular.

Thus, this study aimed to analyze the application of deep learning techniques in GW data quality studies by answering the following research questions:

- RQ1: What is the purpose of applying deep learning for GW data quality studies?
- RQ2: What deep learning techniques are used for GW data quality studies?
- RQ3: What types of deep learning techniques are employed for each application purpose?
- RQ4: Which evaluation parameters are used to assess the performance of deep learning techniques?
- RQ5: Which datasets are utilized as input for training deep learning techniques?
- RQ6: Which detectors have provided the datasets that are utilized as input data for training deep learning techniques?
- RQ7: What types of input data are utilized for the training of deep learning techniques?

The objective of this study was to address gaps in the existing literature by conducting a systematic review of studies that have employed deep learning-based artificial intelligence techniques for GW data quality studies. The contribution of this work is threefold: First, we conducted a systematic literature review that offers researchers a comprehensive understanding of the application of deep learning techniques for GW data analysis, with a particular focus on data quality studies. Second, we presented the widely used deep learning techniques, evaluation parameters, and input datasets for deep learning applications in GW data quality studies. Finally, we identified future

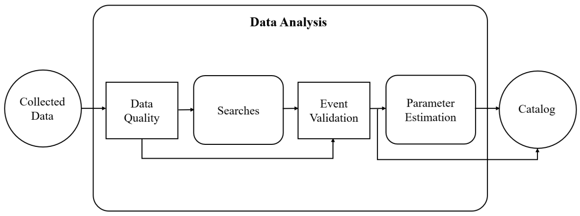

Fig. 1 Main steps in GW detector data processing.

considerations for the development of deep learning techniques to understand and improve GW data quality. To the best of our knowledge, this is the first systematic literature review on deep learning techniques applied to GW data analysis.

The remainder of this paper is organized as follows. Section 2 introduces the background of the study. Section 3 outlines the systematic review methodology used in this article. In Section 4, the review results, which include the answers to the research questions, are presented. Section 5 discusses the results of our analysis. Finally, in Section 6, the overall conclusions of the review are presented.

2 Background

We provide a concise explanation of the GW data analysis process in Section 2.1. In Section 2.2, we briefly describe the basic concepts of deep learning.

2.1 Gravitational wave data analysis

Gravitational waves (GWs) are ripples in space-time caused by the acceleration of massive objects, such as neutron stars or black holes orbiting each other, which were predicted by Einstein's general theory of relativity. Analyzing GW data from GW detectors can reveal information about the source and the nature of gravity itself [22].

Figure 1 provides a simplified overview of the GW detector data processing, including the data quality analysis addressed in this paper. This figure is based on the summary presented in the introduction of [23], with some of the detailed steps presented in an abbreviated form. The arrows indicate the primary data flows. A rectangle represents each step of the data analysis process, while a rounded rectangle indicates a multi-step procedure. The collected data from GW detectors include not only strain data representing the difference in arm length of the interferometers, but also various auxiliary channel data for monitoring the detectors. Data quality analysis, the subject of this paper, involves the analysis of the data from auxiliary channels to assess their impact on the GW data, and GW denoising to distinguish the signal from the noise and remove the identified transients [24]. Davis and Walker [25] presented an overview of the detector characterization and data quality analysis of GW detectors, focusing on the identification and mitigation of instrumental problems that impact the quality of recorded GW strain data. Based on the information from data quality analysis, researchers examine the validity of a detection candidate in the process of event validation. During the search process, GW signals are identified in the strain data by matching them with signal templates generated by waveform modeling. After signal detection, parameter estimation is conducted to extract the astrophysical source properties from the GW signal. The catalog contains information on the event validation, instrument data, and the properties of GW sources obtained through parameter estimation.

2.2 Deep learning

Deep learning, a subfield of machine learning, has rapidly emerged as the dominant research area within artificial intelligence over the last decade. Deep learning [26] is based on neural networks, which consist of input and output neurons resembling the structure of neurons and synapses in the nervous system. A layer consists of a number of horizontally aligned neurons, all of which are connected to neurons in the previous and following layers. When neural networks have multiple hidden layers consisting of hidden neurons, they can be regarded as deep neural networks (DNNs).

In this review, deep learning is defined as the use of neural networks. A distinguishing feature of deep learning is its capacity to automatically discern patterns and extract features from raw data, which renders it particularly advantageous in complex tasks such as image classification and object recognition. This is achieved by combining a computational architecture comprising interconnected layers of artificial neurons with

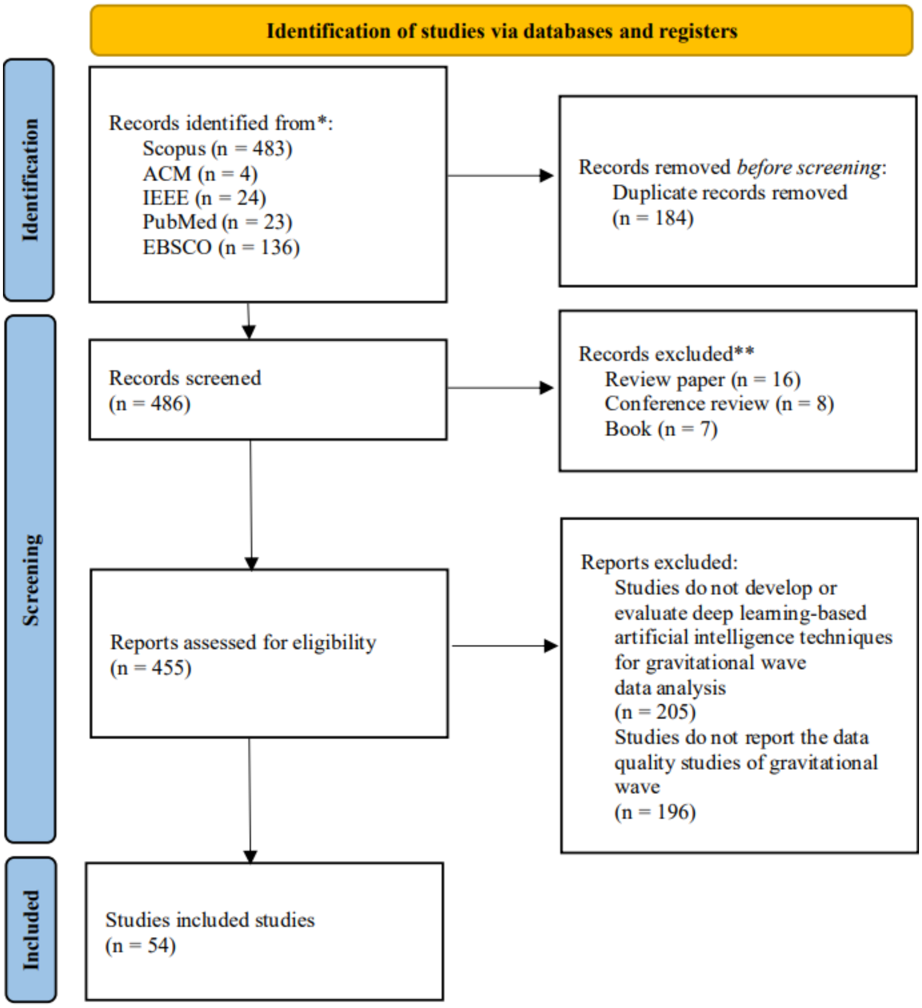

Fig. 2 Flowchart of the systematic literature review process based on PRISMA (2020) guidelines.

powerful optimization algorithms. Deep learning models can be trained in various ways to suit specific tasks and applied to a wide range of real-world domains, including healthcare and medical applications, smart agriculture, and natural language processing [27]. Various deep learning techniques employed in GW data quality studies were classified according to their architectural characteristics, including convolutional neural networks (CNNs) [28], [29], artificial neural networks (ANNs) [30], [31], [32], autoencoders [33],[34], generative adversarial networks (GANs) [35], long short-term memory (LSTM) [36], and transformers [37].

3 Method

3.1 Literature search process

To investigate the application of deep learning techniques in GW data quality studies, we systematically reviewed the relevant literature following the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines [38]. Figure 2 shows a flow diagram that reflects the systematic review strategy in adherence with the PRISMA (2020) guidelines.

In November 2024, we queried five different databases: Scopus, ACM, IEEE, PubMed, and EBSCO. The search query was applied to at least

Table 1 Search query. Abstract abbreviated as either "ABS" or "AB." Keywords abbreviated as "KEY."

Database	Search String
Scopus	TITLE-ABS-KEY(("gravitational wave") AND ("machine learning" OR "deep learning" OR "artificial neural network" OR "ann" OR "deep neural network" OR "dnn" OR "convolutional neural network" OR "cnn" OR "recurrent neural network" OR "rnn" OR "generative adversarial network" OR "gan" OR "multilayer perceptron" OR "mlp" OR "autoencoder" OR "transfer learning" OR "transformer"))
ACM Digital Library	[[Abstract: "gravitational wave"] AND [[Abstract: "machine learning"] OR [Abstract: "deep learning"] OR [Abstract: "artificial neural network"] OR [Abstract: "ann"] OR [Abstract: "deep neural network"] OR [Abstract: "dnn"] OR [Abstract: "convolutional neural network"] OR [Abstract: "cnn"] OR [Abstract: "recurrent neural network"] OR [Abstract: "rnn"] OR [Abstract: "generative adversarial network"] OR [Abstract: "gan"] OR [Abstract: "multilayer perceptron"] OR [Abstract: "mlp"] OR [Abstract: "autoencoder"] OR [Abstract: "transfer learning"] OR [Abstract: "transformer"]]
IEEE Xplore	("Abstract":"gravitational wave") AND ("Abstract":"machine learning" OR "Abstract":"deep learning" OR "Abstract":" artificial neural network" OR "Abstract":"ann" OR "Abstract":"deep neural network" OR "Abstract":"dnn" OR "Abstract":"convolutional neural network" OR "Abstract":"cnn" OR "Abstract":"recurrent neural network" OR "Abstract":"rnn" OR "Abstract":"generative adversarial network" OR "Abstract":"gan" OR "Abstract":"multilayer perceptron" OR "Abstract":"mlp" OR "Abstract":"autoencoder" OR "Abstract":"transfer learning" OR "Abstract":"transformer")
PubMed	("gravitational wave"[Title/Abstract]) AND ("machine learning"[Title/Abstract] OR "deep learning"[Title/Abstract] OR "artificial neural network"[Title/Abstract] OR "ann"[Title/Abstract] OR "deep neural network"[Title/Abstract] OR "dnn"[Title/Abstract] OR "convolutional neural network"[Title/Abstract] OR "cnn"[Title/Abstract] OR "recurrent neural network"[Title/Abstract] OR "rnn"[Title/Abstract] OR "generative adversarial network"[Title/Abstract] OR "gan"[Title/Abstract] OR "multilayer perceptron"[Title/Abstract] OR "mlp"[Title/Abstract] OR "autoencoder"[Title/Abstract] OR "transfer learning"[Title/Abstract] OR "transformer"[Title/Abstract])
EBSCO	AB("gravitational wave") AND AB("machine learning" OR "deep learning" OR "artificial neural network" OR "ann" OR "deep neural network" OR "dnn" OR "convolutional neural network" OR "cnn" OR "recurrent neural network" OR "rnn" OR "generative adversarial network" OR "gan" OR "multilayer perceptron" OR "mlp" OR "autoencoder" OR "transfer learning" OR "transformer")

one of the following items: title, abstract, or keywords (see Table 1). The initial search found 670 studies published up to November 2024 that matched the keywords. We refined our search results by removing 184 duplicates. In addition, 31 studies were excluded because they were literature reviews, book chapters or conference reviews. Finally, 401 more studies were excluded following title and abstract screening, as they did not satisfy the inclusion criteria.

3.2 Eligibility criteria

Following the retrieval of studies using the search strategy and removal of the duplicates, we applied inclusion and exclusion criteria in screening the primary studies. A publication was selected if all the exclusion criteria were false and the inclusion criteria were true [39]. The inclusion and exclusion criteria were set to ensure that the 54 results obtained were a collection of relevant studies for the systematic review (see Table 2). The study inclusion and exclusion criteria table, presented in a similar format to Table 2, has been commonly used in previous systematic literature reviews [40].

Table 2 Inclusion and exclusion criteria.

Inclusion criteria	The studies that develop and evaluate deep learning-based artificial intelligence techniques for gravitational wave data analysis
	The studies that report the data quality studies of gravitational wave
Exclusion criteria	The studies that do not meet the inclusion criteria
	The studies that are not written in English
	The studies that are not journal articles or conference journals

3.3 Data extraction

This systematic review was conducted in accordance with the PRISMA guidelines, which indicate the process for the selection of final studies included in the systematic review. From the final included studies, the following data were extracted: the study title, authors, year of publication, the purpose of applying deep learning, the deep learning technique employed, the evaluation parameters, the input dataset utilized, and the type of input data. Study selection and data extraction were performed by a single researcher. After applying the predefined inclusion and exclusion criteria, a total of 54 studies were ultimately included in the systematic review. A Google spreadsheet was used to list all the selected studies in rows with each extracted data in columns, and no automation tools were used in this process. Finally, the extracted data were synthesized to answer each research question. The results of this systematic literature review study are presented using tables and charts in the following section.

4 Result

This section presents the analysis of the studies that seek to answer the research questions. A total of 54 eligible papers were included in the systematic literature review. Figure 3 illustrates the distribution of the eligible papers according to the year of publication. Here, 28 papers (more than half of the papers published up to November 2024, when the systematic review was conducted) have been published since 2023. The yearly publication trend indicates that there has been a recent surge in research applying deep learning to GW data quality studies.

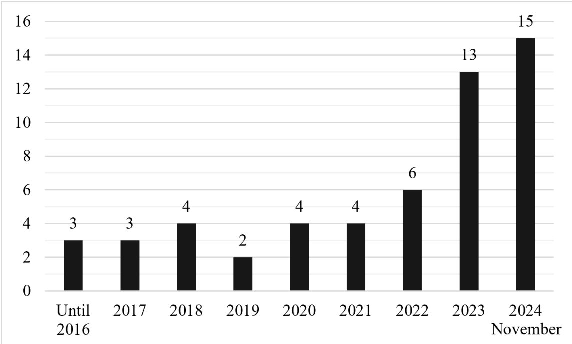

Fig. 3 Numbers of papers published by year.

4.1 What is the purpose of applying deep learning for GW data quality studies?

There are two primary purposes for the application of deep learning to GW data quality studies: glitch analysis and denoising. Figure 4 illustrates the distribution of deep learning application purposes. Out of the 54 eligible papers, 28 (51.85%) applied deep learning for the purpose of denoising, while 26 (48.15%) applied it for glitch analysis.

Glitch analysis refers to the characterization and classification of glitches to gain an understanding of their origins and characteristics. In studies aimed at glitch analysis, the main tasks performed by the applied deep learning techniques are glitch classification and glitch generation. The classification of glitches based on their morphological characteristics is an important task in data quality research. Understanding the origins and characteristics of glitches contributes to detector characterization and glitch removal from GW signals. Information about specific glitches that interfere with the detection of GW signals is used for event validation to reduce false alarms. Initial studies on this task were conducted by Biswas et al. [41] and Rampone et al. [42], demonstrating the feasibility of using artificial neural networks to characterize and classify non-Gaussian noise. This was followed by several studies using machine learning, and in 2017, Zevin et al. [43] pioneered the use of CNNs for glitch classification, spurring research on the use of deep learning for GW data analysis. Combined with the superior image classification capabilities of CNNs, the subsequent research [44–49] has led to the improvement of deep learning models for LIGO glitch classification, and a glitch classification system with high accuracy has been built. Glitch generation involves generating synthetic glitch data by emulating specific glitch classes. This task is conducted to support the aforementioned glitch analysis studies and to enhance the classification performance [50–52].

Denoising involves distinguishing and mitigating identified glitches that could be confused with true GW signals. The tasks to which deep learning is applied include detecting specific glitches in GW data [53], building a discrimination model that distinguishes GW signals from confusing glitches [54–57], and mitigating the effects of glitches [58, 59]. The purpose of this endeavor encompasses not only the removal of glitches but also the suppression of noise in GW data by employing various denoising algorithms, with the objective of improving search sensitivity. Deep learning has been successfully applied to the noise reduction task, demonstrating superior performance in removing non-linearly coupled noise from GW signals [60, 61]. Furthermore, intensive research has focused on the signal extraction task, which aims to denoise and detect astrophysical signals buried in noise [62–64].

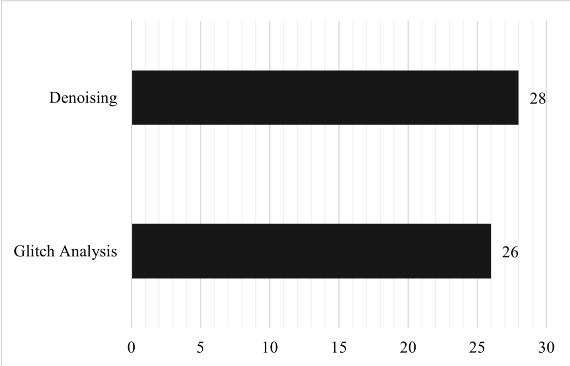

Fig. 4 Purpose of applying deep learning techniques for GW data quality studies.

4.2 What deep learning techniques are used for GW data quality studies?

Table 3 presents the summary of deep learning techniques for data quality studies. The reported deep learning techniques used for GW data quality improvement were categorized based on the deep learning architecture: CNNs, ANNs, autoencoder, GANs, Transformer, CNNs+LSTMs and LSTMs. Out of the eligible papers included in this review, 32 papers used CNNs, 7 papers used ANNs, 6 papers used autoencoders, 4 papers used GANs, 4 papers used transformers, CNNs+LSTMs were employed in 3 papers, and LSTMs were used in 1 paper.

CNNs are commonly employed for image recognition and classification. Zevin et al. [43] pioneered the introduction of citizen science into GW research and built a successful system to classify glitches in GW detectors by collaborating with CNNs and citizen scientists. Following the first application of CNNs, various models for glitch classification have been developed. Since 2020, CNNs have been used not only for glitch analysis but also for denoising purpose. Successful early studies in this field have paved the way for advancements in noise reduction techniques using CNNs. Ormiston et al. [61] applied CNNs to reduce the non-linearly and non-stationary coupled noise that limits detector sensitivity. Wei and Huerta [62] successfully demonstrated a denoising model for GW signals using WaveNet.

ANNs are one of the most basic deep learning architectures used in early research. They have been primarily implemented as a multi-layer perceptron (MLP) by Acerenese et al [65], marking the initial attempt to apply MLP in GW data quality analysis. ANN structure was used in early research on applying deep learning to glitch analysis, notably by Rampone et al. [42] and Biswas et al. [41].

Table 3: Summary of deep learning techniques used for GW data quality studies.

Architecture	Reported Paper	Application Purpose	Deep Learning Task	Deep Learning Technique (name of model employed)	Published Year
CNN	Zevin et al. [43]	Glitch analysis	Glitch classification	Deep CNN	2017
	Bahaadini et al. [45]	Glitch analysis	Glitch classification	CNN (merged-view model)	2017
	George et al. [49]	Glitch analysis	Glitch classification	CNN (InceptionV2,V3; VGG16,19; ResNet50), Deep transfer learning and Fine-tuning	2018
	Razzano and Cuoco [47]	Glitch analysis	Glitch classification	CNN	2018
	Bahaadini et al. [44]	Glitch analysis	Glitch classification	Deep CNN (single view model, merged-view model, soft fusion model, hard fusion model)	2018
	Bahaadini et al. [46]	Glitch analysis	Glitch classification	Transfer learning, CNN (pre-trained VGG16 model)	2018
	Coughlin et al. [48]	Glitch analysis	Glitch classification	Transfer learning, CNN (pre-trained VGG16 model)	2019
	Goel and Mukherjee [66]	Glitch analysis	Glitch classification	CNN	2020
	Soni et al. [67]	Glitch analysis	Glitch classification	CNN	2021
	Sakai et al. [68]	Glitch analysis	Glitch classification	Unsupervised learning using a deep CNN	2022
	Sakai et al. [69]	Glitch analysis	Glitch classification	Unsupervised learning using a deep CNN	2022
	Bişkin et al. [70]	Glitch analysis	Glitch classification	CNN (InceptionV3, ResNet50), Transfer learning and Fine-tuning	2023
	Fernandes et al. [71]	Glitch analysis	Glitch classification	CNN (ResNet18, ResNet34, ConvNeXt), Transfer learning	2023
	Razzano et al. [72]	Glitch analysis	Glitch classification	CNN	
	Koyama et al. [73]	Glitch analysis	Glitch classification	CNN	2024
	Ormiston et al. [61]	Denoising	Noise reduction	One dimensional CNN (DeepClean)	2020
	Wei and Huerta [62]	Denoising	Signal extraction	CNN (WaveNet denoising algorithm model)	2020
	Yu et al. [74]	Denoising	Noise reduction	Compound CNN	2021
	Kato et al. [75]	Denoising	Noise reduction	Denoising convolution neural network	2022
	Yu and Adhikari [60]	Denoising	Noise reduction	CNN ("slow×fast" CNN structure)	2022
	Murali and Lumley [76]	Denoising	Signal extraction	CNN	2023
	Boudart [53]	Denoising	Glitch detection	CNN	
	Choudhary et al. [54]	Denoising	Signal-glitch discrimination	Sine Gaussian projection map-convolution neural network	2023
	Sharma et al. [57]	Denoising	Signal-glitch discrimination	Inception-V3 network (THAMES)	2023
	Meijer et al. [55]	Denoising	Signal-glitch discrimination	CNN (modified WaveNet architecture)	2024
	Álvarez-López et al. [56]	Denoising	Signal-glitch discrimination	Gravity Spy Convolutional Neural Network Decision Tree	2024
	Ma et al. [77]	Denoising	Signal extraction	MSNRnet	2024
	Ma et al. [78]	Denoising	Signal extraction	MSNRnet-2	2024
	Houba et al. [79]	Denoising	Glitch detection	CNN	2024
	Saleem et al. [80]	Denoising	Noise reduction	CNN (DeepClean)	2024
	Attadio et al. [81]	Denoising	Noise reduction	CNN	2024
ANN	Rampone et al. [42]	Glitch analysis	Glitch classification	Feedforward multilayer perceptron trained by a backpropagation algorithm, Self-organizing map	2013
	Biswas et al. [41]	Glitch analysis	Glitch classification	MLP, iRPROP algorithm	2013

Table 3 continued from previous page

Architecture	Reported Paper	Application Purpose	Deep Learning Task	Deep Learning Technique (name of model employed)	Published Year
	Mukund et al. [82]	Glitch analysis	Glitch classification	Difference boosting neural network	2017
	Acernese et al. [65]	Denoising	Noise reduction	MLP	2000
	Mogushi et al. [58]	Denoising	Glitch mitigation	ANN	2021
	Macas et al. [59]	Denoising	Glitch mitigation	Dense neural network	2024
	Houba et al. [79]	Denoising	Signal-glitch discrimination	Wide and deep neural network	2024
Autoencoder	Laguarta et al. [83]	Glitch analysis	Glitch classification	Convolutional Autoencoder	2024
	Li et al. [84]	Glitch analysis	Glitch classification	Cross-Temporal Spectrogram Autoencoder	2024
	Shen et al. [63]	Denoising	Signal extraction	Enhanced deep recurrent denoising autoencoder	2019
	Hayashi et al. [85]	Denoising	Noise reduction	Denoising Autoencoder	2020
	Bini et al. [86]	Denoising	Glitch detection and mitigation	Autoencoder neural network (XGBoost+AE model)	2023
	Bacon et al. [87]	Denoising	Signal extraction	Denoising autoencoder	2023
GAN	Lopez et al. [50]	Glitch analysis	Glitch generation	CT-GAN	2022
	Yan et al. [51]	Glitch analysis	Glitch generation	ProGAN	2022
	Powell et al. [52]	Glitch analysis	Glitch generation	GAN	2023
	Dooney et al. [88]	Glitch analysis	Glitch generation	cDVGAN	2024
Transformer	Jiaxiang and Jiale [89]	Glitch analysis	Glitch classification	Vision transformer encoder+Quantum classifier	2023
	Zhao et al. [90]	Denoising	Signal extraction	Transformer based deep neural network	2023
	Ma et al. [91]	Denoising	Signal extraction	Transformer encoder+Residual blocks (time scale recursive denoising framework)	2024
	Wang et al. [92]	Denoising	Signal extraction	Deep end-to-end transformer-based pretraining model	2024
CNN+LSTM	Apostol and Truica [93]	Glitch analysis	Glitch classification	CNN+LSTM (DeepWaves Ensemble model)	2023
	Chatterjee et al. [64]	Denoising	Signal extraction	CNN-LSTM denoising autoencoder	2021
	Xu et al. [94]	Denoising	Signal extraction	CNN-LSTM denoising autoencoder	2024
LSTM	Houba et al. [79]	Denoising	Glitch mitigation	LSTM	2024

Autoencoders are neural networks comprising an encoder-decoder structure. These deep learning techniques have been employed for denoising. Shen et al. [63], Bini et al. [86] and Bacon et al. [87] reported a technique for denoising GW time series data by applying a denoising autoencoder. Autoencoders were initially used primarily for denoising, but were later introduced for glitch analysis. Laguatra et al. [83] and Li et al. [84] applied autoencoders to discover unknown glitch morphologies in the LIGO glitch population and the underlying correlations between different glitches.

GANs are a type of generative artificial intelligence technique used for generating realistic data. In the context of GW data analysis, GANs have been employed for the glitch generation task. Lopez et al. [50], Yan et al. [51] and Powell et al. [52] employed GANs for simulating glitches. The generated glitch images were used to improve the performance of glitch classification and detection, or were used in mock data challenges.

The transformer-based technique, which implements a vision transformer encoder [95], was employed to suppress noise and recover signals in signal extraction tasks [90–92]. Jiaxiang and Jiale [89] developed a model that combines a vision transformer encoder and a quantum classifier for the glitch classification task.

Chatterjee et al. [64] pioneered a denoising model that combines CNN and LSTM. They proposed the CNN-LSTM denoising autoencoder model and demonstrated its superior performance in denoising GW signals. Apostol and Truka [93] developed a deep learning ensemble model for glitch analysis that combines CNN layers and LSTMs and compared its performance with other machine learning and deep learning models. LSTM has also been employed in the glitch mitigation task. Houba et al. [79] presented a deep learning network to effectively detect and mitigate glitches that may occur in the LISA detector. This network incorporates a CNN for glitch detection, a DNN for glitch characterization, and an LSTM model for the glitch mitigation task.

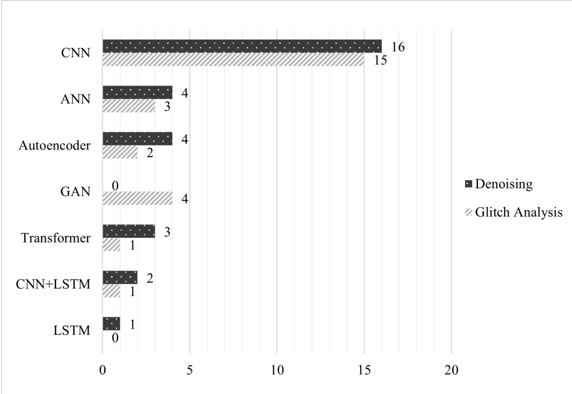

Fig. 5 Distribution of different deep learning techniques used for each application purpose.

4.3 What types of deep learning techniques are employed for each application purpose?

Figure 5 illustrates the types of deep learning techniques employed for each of the purposes identified in Section 3.1. Out of the 28 papers that used deep learning techniques for denoising, 16 papers employed CNNs, 4 employed ANNs, 4 employed autoencoders, and 3 employed transformers. Out of the 26 papers that used deep learning techniques for glitch analysis, 15 papers employed CNNs, 4 papers employed GANs, and ANNs were used in three papers.

Thus, among the deep learning techniques applied in all the 54 papers reviewed, CNNs were the most frequently used technique (31 papers). Upon analysis of the techniques employed for each application purpose, CNNs were found to be most frequently employed for both glitch analysis and denoising.

4.4 Which evaluation parameters are used to assess the performance of deep learning techniques?

The evaluation parameters used for assessing the performance of deep learning techniques in each paper are presented in Table 4, Table 5. As illustrated in both tables, a total of 40 distinct parameters were used across the studies.

Table 4: Evaluation parameters for assessing deep learning technique performance. (Part 1)

	Sharma et al.,2023[57]	Meijer et al.,2024[55]	Boudart ,2023[53]	Fernandes et al.,2023[71]	Jiaxiang and Jiale,2023[89]	Álvarez-López et al.,2024[56]	Koyama et al., 2024[73]	Houba et al.,2024[79]	Bişkin et al.,2023[70]	Apostol and Truica,2023[93]	Powell et al.,2023[52]	Razzano et al.,2023[72]	George et al.,2018[49]	Mukund et al., 2017[82]	Razzano and Cuoco,2018[47]	Bahaadini et al.,2018[44]	Bahaadini et al.,2017[45]	Soni et al.,2021[67]	Zevin et al.,2017[43]	Goel and Mukherjee,2020[66]	Attadio et al.,2024[81]	Wang et al.,2024[92]	Ma et al., 2024[77]	Ma et al., 2024[91]	Ma et al., 2024[78]	Zhao et al.,2023[90]	Murali and Lumley,2023[76]
Accuracy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	-	-	-	-	-	-
Overlap	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	✓	✓	✓	✓	✓
Detection efficiency	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-
F1 score	✓	-	-	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-
ROC curve	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-
FAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	✓	-	-
Computational cost	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	-	-	-
Detection rate	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	-	-
AUC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-
FAP	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	✓
MSE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Accuracy of unsupervised learning	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ARI	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NMI	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sensitive volume-time product	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
FPR	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Combined-F1-time	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Training time	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IFAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-
Computational time	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GAN-train score	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GAN-test score	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Weightend F1 score	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
FMG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ORL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Fractional power	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sensitivity volume	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Root-squared-sum strain amplitude	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Reconstruction error	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Misclassification error	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 4 continued from previous page

	Sharma et al.,2023[57]	Meijer et al.,2024[55]	Boudart ,2023[53]	Fernandes et al.,2023[71]	Jiaxiang and Jiale,2023[89]	Álvarez-López et al.,2024[56]	Koyama et al., 2024[73]	Houba et al.,2024[79]	Bişkin et al.,2023[70]	Apostol and Truica,2023[93]	Powell et al.,2023[52]	Razzano et al.,2023[72]	George et al.,2018[49]	Mukund et al., 2017[82]	Razzano and Cuoco,2018[47]	Bahaadini et al.,2018[44]	Bahaadini et al.,2017[45]	Soni et al.,2021[67]	Zevin et al.,2017[43]	Goel and Mukherjee,2020[66]	Attadio et al.,2024[81]	Wang et al.,2024[92]	Ma et al., 2024[77]	Ma et al., 2024[91]	Ma et al., 2024[78]	Zhao et al.,2023[90]	Murali and Lumley,2023[76]
Relative distances of each cluster center from the overall mean	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Optimal SNR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
PSNR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Wasserstein distance	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Match function	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Normalized cross-covariance	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Average value of the degree of noise reduction	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Correlations of the true signal and estimated one	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Similarity score	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 5: Evaluation parameters for assessing deep learning technique performance. (Part 2)

	Bacon et al.,2023[87]	Shen et al.,2019[63]	Wei and Huerta,2020[62]	Chatterjee et al.,2021[64]	Xu et al.,2024[94]	Choudhary et al.,2023[54]	Yu and Adhikari ,2022[60]	Saleem et al.,2024[80]	Yu et al., 2021[74]	Biswas et al.,2013[41]	Dooney et al.,2024[88]	Sakai et al.,2022[68]	Sakai et al.,2022[69]	Bahaadini et al.,2018[46]	Li et al.,2024[84]	Yan et al.,2022[51]	Moguchi et al.,2021[58]	Macas et al., 2024[59]	Bini et al.,2023[86]	Laguarta et al.,2024[83]	Rampone et al.,2013[42]	Ormiston et al.,2020[61]	Kato et al.,2022[75]	Lopez et al.,2022[50]	Hayashi et al.,2020[85]	Acernese et al.,2000[65]	Coughlin et al.,2019[48]
Accuracy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Overlap	✓	✓	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Detection efficiency	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
F1 score	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASD	-	-	-	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ROC curve	-	-	-	-	-	✓	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
FAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computa-tional cost	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Detection rate	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AUC	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
FAP	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 5 continued from previous page

	Bacon et al.,2023[87]	Shen et al.,2019[63]	Wei and Huerta,2020[62]	Chatterjee et al.,2021[64]	Xu et al.,2024[94]	Choudhary et al.,2023[54]	Yu and Adhikari ,2022[60]	Saleem et al.,2024[80]	Yu et al., 2021[74]	Biswas et al.,2013[41]	Dooney et al.,2024[88]	Sakai et al.,2022[68]	Sakai et al.,2022[69]	Bahaadini et al.,2018[46]	Li et al.,2024[84]	Yan et al.,2022[51]	Mogushi et al.,2021[58]	Macas et al., 2024[59]	Bini et al.,2023[86]	Laguarta et al.,2024[83]	Rampone et al.,2013[42]	Ormiston et al.,2020[61]	Kato et al.,2022[75]	Lopez et al.,2022[50]	Hayashi et al.,2020[85]	Acernese et al.,2000[65]	Coughlin et al.,2019[48]
MSE	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Accuracy of unsupervised learning	-	-	-	-	-	-	-	-	-	-	-	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ARI	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-
NMI	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	✓	-	-	-	-	-	-	-	-	-	-	-	-
Sensitive volume-time product	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
FPR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Combined-F1-time	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Training time	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IFAR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computational time	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GAN-train score	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-
GAN-test score	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-
Weightend F1 score	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-
FMG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-
ORL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-
Fractional power	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-
Sensitivity volume	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-
Root-squared-sum strain amplitude	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-
Reconstruction error	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-
Misclassification error	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-
Relative distances of each cluster center from the overall mean	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-
Optimal SNR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-
PSNR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-
Wasserstein distance	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-
Match function	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-
Normalized cross-covariance	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-
Average value of the degree of noise reduction	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-
Correlations of the true signal and estimated one	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-
Similarity score	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓

In both table, "✓" indicates that the parameter was used by the researcher to evaluate the performance while "-" indicates that the parameter was not used by the researcher. The most frequently used evaluation parameter was accuracy, which was used in 20 papers. Unlike other CNN techniques applied to glitch classification, Sakai et al. [69] presented an unsupervised learning-based approach that reduced the annotation work on the training data and ensured the objectivity of the classification. They defined the accuracy of unsupervised learning to evaluate the performance of their model using the CNN and classification architecture proposed in the paper. Among the papers that applied deep learning techniques for denoising, 12 studies evaluated the performance of the proposed models by comparing the denoised GW signal and the template waveform generated by a simulation program. The F1 score, which is the harmonic mean of precision and recall, was used in 4 papers. Yan et al.[51] used weighted F1 score, which is a variant of macro-averaged F1-score, to evaluate the classification performance in their work. Fernandes et al. [71] employed a newly defined metric called combined-F1-time, which combines F1 score and training time to evaluate glitch classification performance.

In addition, various other evaluation parameters were used across different papers: detection efficiency, comparison of amplitude spectral density (ASD), receiver operating characteristic (ROC) curves, area under the curve (AUC) which refers to the area under the ROC curve, false alarm rate (FAR) that measures significance of each coincident trigger in GW astronomy, inverse false alarm rate (IFAR), detection rate of GW signal, mean square error (MSE), false alarm probability (FAP), normalized mutual information score (NMI), adjusted rand score or adjusted rand index (ARI), false positive rate (FPR), fractional match gain (FMG) and overlap log ratio (ORL) defined in [60], fractional power, sensitive volume-time product, sensitivity volume, root-squared-sum strain amplitude, reconstruction error, misclassification error (false alarm and false dismissal), relative distances between the centers and from the overall mean, optimal signal-to-noise ratio (SNR) of the injected signal, peak SNR (PSNR), Wasserstein distance, match function, normalized cross-covariance, GAN-train and GAN-test score proposed by [96], the average value of the degree of noise reduction defined in [85], correlations of the true signal and estimated one, and similarity score. Notably, several papers published after 2023 included "time" in their evaluation parameters: combined-F1-time, training time, computational time, and computational cost which measured the time consumption of a single CPU or GPU as an indicator. [54], [70],[71],[77], [91]

4.5 Which datasets are utilized as input for training deep learning techniques?

We grouped the datasets utilized for training deep learning techniques into four groups: the GW detector, glitch spectrogram, simulated, and detector sensitivity datasets. The GW detector dataset involves a time series of strain data from the current GW detectors: advanced LIGO (aLIGO) [97], Virgo [98], and KAGRA [99]. As of November 2024, none of the studies included in this review used data from KAGRA, which was the last of the ground-based detectors to join the observations. The datasets from each detector have been uploaded to the GW open science center (GWOSC) following the observation run.

The glitch spectrogram dataset involves information regarding various glitches that were identified by the Omicron transient search algorithm [100] during the observation run of aLIGO. They were classified and labeled by Gravity Spy, a citizen science project on the Zooniverse platform. The classification dataset of the Gravity Spy project, which comprises identified glitches in the LIGO detector, was released in October 2018 (Gravity Spy v1.0.0) [44]. The data pertaining to the labeled glitches, results of machine learning classifications and the spectrogram images have been made available in Zenodo.

The detector sensitivity dataset consists of the sensitivity curves of GW detectors that are currently operational or under development. These datasets have been primarily utilized for the simulation of noise data, which serves as the input training data. For example, Meijer et al. utilized the design sensitivity dataset of the future Einstein Telescope (ET) [101] to develop a technique that distinguishes GW signals from simulated blip glitches in the ET's design sensitivity data [55].

Table 6 Summary of input datasets for training deep learning techniques.

	Dataset Grouping	Dataset	Source of Data Collection	Description	No. of Papers
Detector -based	GW detector dataset	aLIGO O1 data	GWOSC	Data collected by the aLIGO detector during their first observation runs (September 12, 2015 through January 19, 2016)	9
		aLIGO O2 data	GWOSC	Data collected by the aLIGO detector during their second observation runs (November 30, 2016 through August 25, 2017)	8
		aLIGO O3a data	GWOSC	Data collected by the aLIGO detector during the first half of their third observation runs (April 1, 2019 through October 1, 2019)	8
		aLIGO O3b data	GWOSC	Data collected by the aLIGO detector during the last half of their third observation runs (November 1, 2019 through March 27, 2020)	6
		LIGO S4 data	GWOSC	Data collected by the LIGO detector during their fourth science runs (from February 22 to March 24, 2005)	1
		LIGO S6 data	GWOSC	Data collected by the LIGO detector during their sixth science runs (July 7, 2009 through October 20, 2010)	2
		Saulson catalog	LIGO DCC	Representative high SNR subset of glitches from the LIGO detector during their fifth science runs (November 4, 2005 through October 1, 2007)	1
		LIGO auxiliary channels data	GWOSC	Data collected by a large number of sensors that are used to record the state of the LIGO instruments and their environment	4
		Virgo O3 data	GWOSC	Data collected by the Virgo detector during their third observation runs (April 1, 2019 through March 27, 2020)	1
	Glitch spectrogram dataset	Gravity Spy	Zenodo	LIGO glitch spectrogram labeled by Gravity Spy project	11
		Gravity Spy v1.0	Zenodo	LIGO glitch spectrogram during O1, O2 labeled by Gravity Spy project	5
	Detector sensitivity dataset	Sensitivity of LIGO	GWOSC	LIGO detector sensitivity used to simulate noise	7
		Sensitivity of Virgo	LIGO DCC	Virgo detector sensitivity used to simulate noise	1
		Sensitivity of ET	LIGO DCC	ET detector sensitivity used to simulate noise	1
		Sensitivity of LISA	[102]	LISA detector sensitivity used to simulate noise	1
Simulated	Simulation Dataset	Simulated signals	-	Simulated GW signals	25
		Simulated glitches	-	Simulated glitches	4
		Simulated noise	-	Simulated various kinds of noise	18

The design sensitivity of current and future GW detectors, including ET, has been uploaded to the LIGO document control center (LIGO DCC) portal[1] and made available on each detector website. To develop a signal extraction method for space-based GW sources, Zhao et al. [90] used the sensitivity curve of LISA in a noise simulation. The sensitivity curve of LISA, the next-generation space-based gravitational wave detector, has been proposed in [102]. The simulation dataset comprises GW signals and noise data generated by a computer program, instead of real GW data from observation.

A summary of the input datasets used for training deep learning techniques is presented in Table 6. The datasets under consideration were classified as either detector-based or simulated according to their origin of resources. Detector-based datasets are derived from GW detectors: LIGO, Virgo, ET and LISA, reflecting the characteristics inherent to each detector. These datasets are publicly accessible via the internet. The GW detector, glitch spectrogram, and detector sensitivity datasets are classified as the detector-based dataset. On the other hand, the simulation

[1] https://dcc.ligo.org/LIGO-T2000012/public

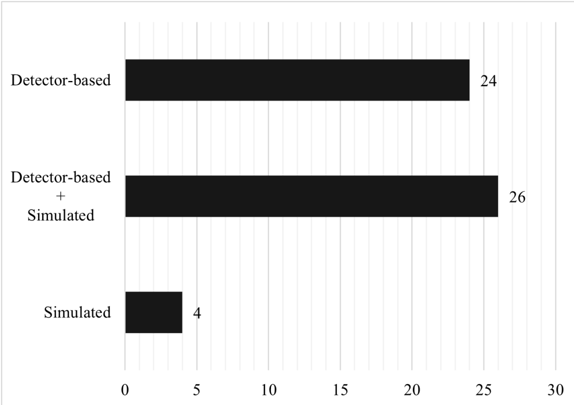

Fig. 6 Distribution of sources of input datasets reported in papers.

dataset, which is created by researchers using computer programs, is classified as the simulated dataset. Upon examination of the distribution of input datasets in Table 6, we found that 24 (44.44%) of the papers utilized detector-based datasets, whereas 26 (48.15%) employed both detector based and simulated datasets to create their training dataset. Four (7.41%) papers employed only simulated datasets as inputs for training their deep learning techniques (see Figure 6).

4.6 Which detectors have provided the datasets that are utilized as input data for training deep learning techniques?

The upper panel of Figure 7 indicates the detectors that provided detector-based datasets utilized in the glitch analysis purpose papers. Among the detector-based datasets employed in the 26 glitch analysis papers, 25 papers utilized datasets from LIGO, including aLIGO O1, O2, O3a, and O3b data; LIGO S4 and S6 data; the Saulson catalog; LIGO auxiliary channels data; Gravity Spy and Gravity Spy v1.0; and LIGO sensitivity. One paper used datasets from Virgo, including Virgo O3 data and Virgo sensitivity. The use of detector-based datasets for training deep learning techniques for glitch analysis is predominantly derived from the LIGO detector. Examining the distribution of detector-based datasets sources across all papers included in the review (see lower panel of Figure 7), we found that 46 papers utilized

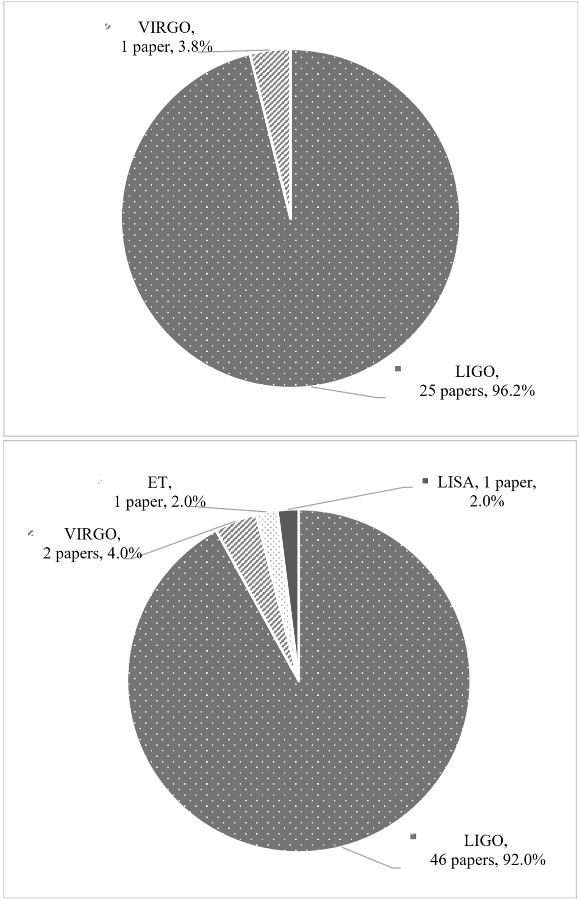

Fig. 7 (upper) Distribution of detectors that provided detector-based datasets used in glitch analysis papers. (lower) Distribution of detectors that provided detector-based datasets used in total papers.

datasets from LIGO, two papers utilized datasets from Virgo, and one paper each utilized a dataset from ET and LISA: ET sensitivity and LISA sensitivity. It is evident that the entire set of papers, not just the glitch analysis papers, relies heavily on the LIGO dataset for deep learning model development.

4.7 What types of input data are utilized for the training of deep learning techniques?

Figure 8 illustrates the distribution of types of input data reported in the eligible papers. An examination of the type of input data employed in all 54 papers indicated that 28 (51.85%) papers utilized image data, 19 (35.19%) papers used time

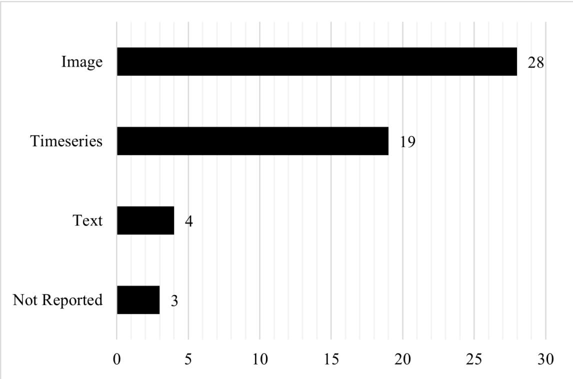

Fig. 8 Distribution of types of input data reported in papers.

series data, 4 (7.41%) papers used text data, and 3 (5.56%) papers did not explicitly present the type of input data.

The image data mainly consist of two-dimensional spectrogram images, represented as time-frequency maps (TF maps), which are generated by applying the Q transform [103] to the time-series signals. Only one paper employed a sine-Gaussian projection map (SGP map) instead of a TF map spectrogram. The SGP map represents the projection of the GW signals onto a sine-Gaussian parameter space. The time series data consist of numerical values arranged at constant time intervals. This includes the original time series data observed by the GW detector or the numerically generated data. The text data consist of a set of parameters, which are feature vectors of GW signals or glitches.

5 Discussion

This review examined various studies that have reported the use of deep learning techniques for GW data quality studies by conducting a systematic review. We attempted to understand the current state of deep learning techniques applied to data quality studies in the field of GW data analysis. In this section, we reflect on our findings and discuss the limitations of our study and directions for future work.

5.1 Findings and discussion

Based on the analysis in Section 4, we revealed three key findings from the answers to the research questions.

5.1.1 CNNs to transformers

Firstly, an analysis of Section 4.2 and 4.3 revealed that CNNs were the most frequently utilized techniques, for both glitch analysis and denoising (see Figure 5). The deep learning techniques applied across all 54 papers in the review, categorized by the architecture used, are summarized in Table 3. Among the deep learning techniques shown in this table, CNNs were used in 31 papers.

As demonstrated by Razzano and Cuoco [47], CNNs typically exhibit superior accuracy when distinguishing glitches with similar morphology compared to other machine learning techniques. CNNs have been employed primarily for the classification of glitches and unsupervised clustering of transient anomalies. However, Wei and Huerta [62] demonstrated that their CNN-based model was more suitable for denoising GW signals than the existing autoencoder models. While recurrent autoencoders outperformed traditional machine learning methods [104], they struggled in the case of denoising second-long GW signals. In contrast, CNNs are effective in denoising longer and weaker GW signals. Consequently, the widespread utilization of CNNs for both glitch analysis and denoising is attributed to their versatility. Their superior performance as a classifier compared to other machine learning techniques, combined with their less computationally intensive nature and ability to make rapid inferences, makes them an effective tool for GW data analysis [105]. Research included in the review has shown that CNNs achieve excellent performance for both image and non-image data inputs, and are more suitable than existing autoencoder models for denoising GW signals.

It is noteworthy that since 2023, transformers have been incorporated into GW data quality studies for both glitch analysis and denoising, as shown in Table 3. The vision transformer encoder is a model that adapts the transformer architecture, originally used in natural language processing, to computer vision tasks. It has been demonstrated to exhibit superior performance in

Table 7 Comparative analysis of deep learning model's glitch classification performance on GravitySpy dataset.

Reported paper	Published year	Model	Accuracy (%)
Bahaadini et al. [44]	2018	CNN	98.21
George et al. [49]	2018	CNN	98.8
Bişkin et al. [70]	2023	CNN	93.98
Fernandes et al. [71]	2023	CNN	96.5
Jiaxiang and Jiale [89]	2023	Transformer	94.27
Koyama et al. [73]	2024	CNN	98.1

Table 8 Comparative analysis of deep learning model's denoising performance on GW150914. "H" indicates LIGO Hanford data, while "L" refers to LIGO Livingston data.

Reported paper	Published year	Model	Overlap
Wei and Huerta [62]	2020	CNN	0.99
Chatterjee et al. [64]	2021	CNN+LSTM	H 0.983 / L 0.995
Murali and Lumley [76]	2023	CNN	H 0.861 / L 0.80
Ma et al. [77]	2024	CNN	H 0.97
Wang et al. [92]	2024	Transformer	H 0.9919 / L 0.9917

computer vision tasks, such as image classification and object detection, often complementing or outperforming traditional CNN techniques [106]. From the results of the review, we found that there is growing research in GW data quality studies that applies transformers to tasks that are typically performed by CNNs. Li et al. [84] and Jiaxiang and Jiale [89] incorporated a vision transformer encoder as a part of their deep learning model for analysis of glitch images. Zhao's integration of the transformer into signal denoising [90] demonstrates the transformer's capacity to effectively process complex GW data.

Table 7 provides a comparative analysis of the results of the studies that applied deep learning techniques to glitch analysis, specifically focusing on studies that tested model performance on glitch classification tasks using the GravitySpy dataset. The results of classifying over 20 glitch image classes in the GravitySpy dataset are presented as the average classification accuracy. Table 8 presents the results of the studies that applied deep learning techniques for denoising, specifically those that evaluated denoising performance on real GW event GW150914 data. The results of the signal extraction tasks performed by these models are presented using the "overlap" parameter. The overlap is calculated between the denoised GW signal extracted from the LIGO data during the period containing the GW event and the simulated GW waveform template. The papers presented in Table 7 and Table 8 are part of the papers included in our review, and details of the deep learning techniques used are presented in Table 3. This shows that since 2023, transformers have been introduced to tasks where CNNs were used for both glitch analysis and denoising, and have performed comparably to CNNs. Thus, the role of CNNs, which have been widely used in GW data analysis, is notable when considering their potential transition to vision transformers in the future development of deep learning techniques.

5.1.2 Time consideration

Secondly, our review found that three recent papers published since 2023 considered the time-related parameters as evaluation parameters to assess the performance of deep learning techniques: training time, computation time, computational cost and combined F1 time (F1 score$-$training time$/3\times10^4$). In Section 4.4, we have compiled the evaluation parameters used to assess the performance of the papers included in our review (see Table 4 and Table 5). The result of this table shows that papers published up to

2022 did not use time-related parameters to evaluate the performance of deep learning, but such parameters have begun appearing in more recent studies.

As highlighted by Bişkin et al. [70], the time required for model training must be reduced without compromising the classification accuracy. This is because, following each observation, the measurement system typically undergoes software and hardware upgrades that may result in the emergence of the previously unidentified types of glitches [67]. This necessitates the retraining of the model with new data to accommodate these new forms of noise. Thus, the consideration of time-related parameters in the evaluation of deep learning techniques may be attributed to the increased quantity and diversity of noise data that must be analyzed and removed.

Another reason for the increasing importance of "time" is the growing necessity for real-time data quality assurance, which has been driven by the advent of multi-messenger astronomy. This has facilitated the development of "fast" and efficient pipelines for real-time data analysis across the entire field of GW data analysis. Ma et al. [77, 91] incorporated the computation cost, measured by the time consumption of each stage, as an evaluation parameter in the development of their GW signal denoising and detection framework. As stated in the foundational research by George and Huerta [19], a primary advantage of employing deep learning applications in GW data analysis is the reduction of computational cost. This enables the development of deeper networks and faster training and evaluation speeds. The traditional GW detection methods are not sufficient for real-time analysis in GW events, where a rapid follow-up is essential for the successful observation of their electromagnetic counterparts. This highlights the necessity for the development of rapid and precise GW detection methods. Considering the crucial role of data quality checks in the GW detection pipeline, it is suggested that future deep learning techniques should consider time as a key evaluation parameter in the development of future deep learning techniques.

5.1.3 LIGO dataset usage

Finally, our analysis in Section 4.6 demonstrated that in all but one of the 26 papers

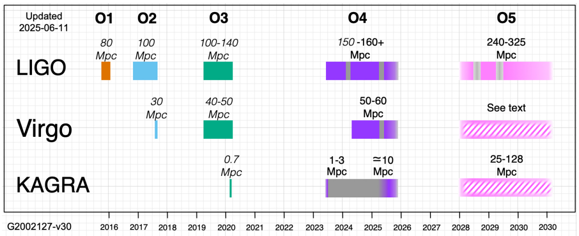

Fig. 9 Timeline of GW observing schedule. Figure reproduced from [107].

focused on glitch analysis, LIGO was the primary source of detector-based datasets utilized as input for training deep learning techniques (see Figure 7 upper panel). Moreover, out of 50 papers that used detector-based dataset, the majority—46 papers—utilized datasets from LIGO (see Figure 7 lower panel). This imbalance in the usage of detector-based datasets suggests that the development of deep learning techniques in the field of GW data quality tends to rely on LIGO data.

One possible explanation for the high usage of LIGO-based data is the Gravity Spy dataset. As illustrated in Table 6, 16 papers utilized the Gravity Spy dataset as training data. Thus, the Gravity Spy dataset is a contributing factor in the increased usage of LIGO-based data. A similar project for Virgo-based data, GWitchHunters [72, 108] released its classification dataset in March 2023, but we found no examples of it being used as training data for deep learning techniques in GW data quality analysis.

Figure 9 shows the timeline of the current ground-based detectors since they began their observations, along with their future observing plans. This information has been uploaded to LIGO DCC[2]. As indicated by the sensitivity distance presented in Figure 9, LIGO's superior sensitivity compared to Virgo and KAGRA may have led to its more frequent data usage. Additionally, the data from LIGO were primarily utilized during the early observation runs because Virgo and KAGRA joined the observation later, resulting in the exclusive availability of the LIGO dataset prior to the O2 data release in 2019 [109]. However, even after the release of Virgo data in 2019, there are few papers that have

[2] https://dcc.ligo.org/LIGO-G2002127/public

used Virgo observational data for training deep learning techniques.

Data quality analysis revealed that the distribution of glitches differs even between two different detectors of LIGO, namely, LIGO Hanford and LIGO Livingston [110]. The causes of glitches are related to detector characteristics. It is evident that the exclusive reliance on LIGO for training data in the development of deep learning techniques represents a limitation, as it does not reflect the characteristics of other detectors. Therefore, it is imperative to develop deep learning tools for glitch analysis using data from various detectors.

5.2 Limitations and future work

This study has several limitations. First, this study focused solely on the aspect of data quality within the process of analyzing GW data. Studies that have applied deep learning to other data analysis processes such as signal detection, parameter estimation, or waveform modeling were excluded. A more comprehensive analysis of deep learning techniques across the entire process of GW data analysis could potentially yield further meaningful insights. Second, this study focused on deep learning-based techniques; Thus, traditional machine learning techniques currently employed in GW data analysis were excluded from this review.

Future work should extend the scope of analysis to include the deep learning techniques employed throughout the entire process of the GW data analysis, beyond data quality part. Moreover, the applicability of our findings to other parts of GW data analysis must be ascertained. Another future study prospect is the examination of other artificial intelligence techniques (in addition to deep learning), that are currently utilized or have the potential to be utilized for the analysis of GW data. In addition, it would be beneficial to examine the potential for generative models, such as large language models [111], to support the GW data analysis process.

6 Conclusion

This study aimed to examine the application of deep learning techniques to understand and improve GW data quality. The key contributions of this study include providing a comprehensive understanding of the application of deep learning techniques to GW data quality studies. The findings of this study are expected to provide researchers with insights into the development of deep learning techniques to enhance GW data quality. A systematic literature review was conducted to examine 486 studies published in the field of GW data analysis that used deep learning. Consequently, a total of 54 studies were identified based on the eligibility criteria. The 54 studies that developed and evaluated deep learning techniques for GW data quality studies were then examined and categorized. The results of this review reveal that CNNs were the most prevalent deep learning technique and future research will concentrate on the implementation of transformers for data quality analysis. In addition, the review identified that recently published papers have considered time-related parameters as the evaluation parameters to assess the performance of deep learning techniques. Furthermore, an analysis of the source of input datasets used for training deep learning techniques revealed that the examined studies primarily used datasets derived from the LIGO detector. This indicates an imbalance in detector-based datasets used for the development of deep learning techniques.

This study presents the first systematic literature review on the application of deep learning techniques in GW data quality analysis. However, there are some limitations to the review, which has a restricted scope. Future work should be directed towards expanding the scope of the systematic review beyond data quality and deep learning to encompass a wider range of artificial intelligence techniques applied to the entire pipeline of the GW data analysis.

Declarations

- **Funding** This work was supported by the New Faculty Startup Fund from Seoul National University. This work was also supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2024-00407105).

- **Competing interests** The authors declare no competing interests.
- **Ethics approval and consent to participate** Not applicable.

- **Data availability** No datasets were generated during the current study.
- **Author contribution** Yubin Kim: conceptualization, data curation, formal analysis, investigation, methodology, project administration, visualization, writing the main manuscript text, and preparing all the figures and tables. Hyunggu Jung: conceptualization, project administration, funding acquisition, and supervision All authors reviewed the manuscript.

Acknowledgment

We appreciate HCAIL members for their constructive feedback on our initial manuscript.

References

[1] Abbott, B.P., *et al.*: Observation of gravitational waves from a binary black hole merger. Phys. Rev. Lett. **116**, 061102 (2016) https://doi.org/10.1103/PhysRevLett.116.061102

[2] Abbott, R., *et al.*: Gwtc-3: Compact binary coalescences observed by ligo and virgo during the second part of the third observing run. Phys. Rev. X **13**, 041039 (2023) https://doi.org/10.1103/PhysRevX.13.041039

[3] Covas, P.B., *et al.*: Identification and mitigation of narrow spectral artifacts that degrade searches for persistent gravitational waves in the first two observing runs of advanced ligo. Phys. Rev. D **97**, 082002 (2018) https://doi.org/10.1103/PhysRevD.97.082002

[4] Cabero, M., Lundgren, A., Nitz, A.H., Dent, T., Barker, D., Goetz, E., Kissel, J.S., Nuttall, L.K., Schale, P., Schofield, R., Davis, D.: Blip glitches in advanced ligo data. Classical and Quantum Gravity **36**(15), 155010 (2019) https://doi.org/10.1088/1361-6382/ab2e14

[5] Abbott, B.P., *et al.*: Characterization of transient noise in advanced ligo relevant to gravitational wave signal gw150914. Class Quantum Gravity **33**(13), 134001 (2016) https://doi.org/10.1088/0264-9381/33/13/134001

[6] Isogai, T., Ligo Scientific Collaboration, Virgo Collaboration): Used percentage veto for ligo and virgo binary inspiral searches. Journal of Physics: Conference Series **243**(1), 012005 (2010) https://doi.org/10.1088/1742-6596/243/1/012005

[7] Smith, J., et al.: A hierarchical method for vetoing noise transients in gravitational-wave detectors. Class Quantum Gravity **28** (2011) https://doi.org/10.1088/0264-9381/28/23/235005

[8] Essick, R., Blackburn, L., Katsavounidis, E.: Optimizing vetoes for gravitational-wave transient searches. Classical and Quantum Gravity **30**(15), 155010 (2013) https://doi.org/10.1088/0264-9381/30/15/155010

[9] Essick, R., Godwin, P., Hanna, C., Blackburn, L., Katsavounidis, E.: idq: Statistical inference of non-gaussian noise with auxiliary degrees of freedom in gravitational-wave detectors. Mach. Learn.: Sci. Technol. **2**(1), 015004 (2020) https://doi.org/10.1088/2632-2153/abab5f

[10] Mogushi, K.: Application of a new transient-noise analysis tool for an unmodeled gravitational-wave search pipeline. Class Quantum Gravity **38**(15), 155004 (2021) https://doi.org/10.1088/1361-6382/ac08a7

[11] Davis, D., *et al.*: Subtracting glitches from gravitational-wave detector data during the third LIGO-Virgo observing run. Classical and Quantum Gravity **39**(24), 245013 (2022) https://doi.org/10.1088/1361-6382/aca238 arXiv:2207.03429 [astro-ph.IM]

[12] Mogushi, K.: Reduction of transient noise artifacts in gravitational-wave data using deep learning. arXiv e-prints, 2105–10522 (2021) https://doi.org/10.48550/arXiv.2105.10522 arXiv:2105.10522 [gr-qc]

[13] Godwin, P.: Low-latency Statistical Data Quality in the Era of

Multi-Messenger Astronomy. Pennsylvania State University, ??? (2020). https://books.google.co.kr/books?id=vrv0zQEACAAJ

[14] Powell, J., Trifirò, D., Cuoco, E., Heng, I.S., Cavaglià, M.: Classification methods for noise transients in advanced gravitational-wave detectors. Class Quantum Gravity **32** (2015)

[15] Powell, J., Torres-Forné, A., Lynch, R., Trifirò, D., Cuoco, E., Cavaglià, M., Heng, I.S., Font, J.A.: Classification methods for noise transients in advanced gravitational-wave detectors ii: performance tests on advanced ligo data. Class Quantum Gravity **34**(3), 034002 (2017) https://doi.org/10.1088/1361-6382/34/3/034002

[16] Negri, L.: Deep learning methods to study the auxiliary channels in the advanced virgo detector. Master's thesis, Universita di Padova (2022)

[17] Xia, H., Shao, L., Zhao, J., Cao, Z.: Improved deep learning techniques in gravitational-wave data analysis. Phys. Rev. D **103**, 024040 (2021) https://doi.org/10.1103/PhysRevD.103.024040

[18] 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 (2018) https://doi.org/10.1016/j.physletb.2017.12.053

[19] George, D., Huerta, E.A.: Deep neural networks to enable real-time multimessenger astrophysics. Phys. Rev. D **97**, 044039 (2018) https://doi.org/10.1103/PhysRevD.97.044039

[20] Cuoco, E., *et al.*: Enhancing gravitational-wave science with machine learning. Mach. Learn.: Sci. Technol. **2**(1), 011002 (2020) https://doi.org/10.1088/2632-2153/abb93a

[21] Benedetto, V., Gissi, F., Ciaparrone, G., Troiano, L.: Ai in gravitational wave analysis, an overview. Applied Sciences **13**(17) (2023) https://doi.org/10.3390/app13179886

[22] Cai, R.-G., Cao, Z., Guo, Z.-K., Wang, S.-J., Yang, T.: The gravitational-wave physics. National Science Review **4**(5), 687–706 (2017) https://doi.org/10.1093/nsr/nwx029 https://academic.oup.com/nsr/article-pdf/4/5/687/31566614/nwx029.pdf

[23] Abbott, B.P., *et al.*: A guide to ligo-virgo detector noise and extraction of transient gravitational-wave signals. Class Quantum Gravity **37**(5), 055002 (2020) https://doi.org/10.1088/1361-6382/ab685e

[24] McIver, J.: Data quality studies of enhanced interferometric gravitational wave detectors. Classical and Quantum Gravity **29**(12), 124010 (2012) https://doi.org/10.1088/0264-9381/29/12/124010 arXiv:1204.2497 [gr-qc]

[25] Davis, D., Walker, M.: Detector characterization and mitigation of noise in ground-based gravitational-wave interferometers. Galaxies **10**(1) (2022) https://doi.org/10.3390/galaxies10010012

[26] LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521** (2015) https://doi.org/10.1038/nature14539

[27] Sarker, I.H.: Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. SN Comput. Sci. **2**(6) (2021) https://doi.org/10.1007/s42979-021-00815-1

[28] Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998) https://doi.org/10.1109/5.726791

[29] Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J.: A survey of convolutional neural networks: Analysis, applications, and prospects. IEEE Transactions on Neural Networks and Learning Systems **33**(12),

6999–7019 (2022) https://doi.org/10.1109/TNNLS.2021.3084827

[30] Rosenblatt, F.: The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review **65 6**, 386–408 (1958)

[31] Hecht-Nielsen, R.: Theory of the backpropagation neural network. International 1989 Joint Conference on Neural Networks, 593–6051 (1989)

[32] Hornik, K., Stinchcombe, M., White, H.: Multilayer feedforward networks are universal approximators. Neural Netw **2**(5), 359–366 (1989) https://doi.org/10.1016/0893-6080(89)90020-8

[33] Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. Science **313**(5786), 504–507 (2006) https://doi.org/10.1126/science.1127647 https://www.science.org/doi/pdf/10.1126/science.1127647

[34] Kingma, D.P., Welling, M.: Auto-Encoding Variational Bayes (2022)

[35] Goodfellow, I., *et al.*: Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems, vol. 27. Curran Associates, Inc., ??? (2014). https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf

[36] Hochreiter, S., Schmidhuber, J.: Long Short-Term Memory. Neural Comput **9**(8), 1735–1780 (1997) https://doi.org/10.1162/neco.1997.9.8.1735 https://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf

[37] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L.u., Polosukhin, I.: Attention is all you need. In: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (eds.) Advances in Neural Information Processing Systems, vol. 30. Curran Associates, Inc., ??? (2017). https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[38] Page, M.J., et al.: The prisma 2020 statement: an updated guideline for reporting systematic reviews. BMJ **372** (2021) https://doi.org/10.1136/bmj.n71 https://www.bmj.com/content/372/bmj.n71.full.pdf

[39] Kitchenham, B.A., Charters, S.: Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE 2007-001, Keele University and Durham University Joint Report (July 2007). https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf

[40] Sharma, V., Gupta, M., Kumar, A., Mishra, D.: Video Processing Using Deep Learning Techniques: A Systematic Literature Review. IEEE Access **9**, 139489–139507 (2021) https://doi.org/10.1109/ACCESS.2021.3118541

[41] Biswas, R., *et al.*: Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data. Phys. Rev. D **88**, 062003 (2013) https://doi.org/10.1103/PhysRevD.88.062003

[42] Rampone, S., Pierro, V., Troiano, L., Pinto, I.M.: Neural network aided glitch-burst discrimination and glitch classification. International Journal of Modern Physics C **24**(11), 1350084 (2013) https://doi.org/10.1142/S0129183113500848 https://doi.org/10.1142/S0129183113500848

[43] Zevin, M., *et al.*: Gravity spy: integrating advanced ligo detector characterization, machine learning, and citizen science. Class Quantum Gravity **34**(6), 064003 (2017) https://doi.org/10.1088/1361-6382/aa5cea

[44] Bahaadini, S., *et al.*: Machine learning for gravity spy: Glitch classification and dataset. Inf. Sci. **444**, 172–186 (2018) https://doi.org/10.1016/j.ins.2018.02.068

[45] Bahaadini, S., Rohani, N., Coughlin, S., Zevin, M., Kalogera, V., Katsaggelos, A.K.: Deep multi-view models for glitch classification. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2931–2935 (2017). https://doi.org/10.1109/ICASSP.2017.7952693

[46] Bahaadini, S., Rohani, N., Katsaggelos, A.K., Noroozi, V., Coughlin, S., Zevin, M.: Direct: Deep discriminative embedding for clustering of ligo data. In: 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 748–752 (2018). https://doi.org/10.1109/ICIP.2018.8451708

[47] Razzano, M., Cuoco, E.: Image-based deep learning for classification of noise transients in gravitational wave detectors. Class Quantum Gravity **35**(9), 095016 (2018) https://doi.org/10.1088/1361-6382/aab793

[48] Coughlin, S., *et al.*: Classifying the unknown: Discovering novel gravitational-wave detector glitches using similarity learning. Phys. Rev. D **99**, 082002 (2019) https://doi.org/10.1103/PhysRevD.99.082002

[49] George, D., Shen, H., Huerta, E.A.: Classification and unsupervised clustering of ligo data with deep transfer learning. Phys. Rev. D **97**, 101501 (2018) https://doi.org/10.1103/PhysRevD.97.101501

[50] Lopez, M., Boudart, V., Buijsman, K., Reza, A., Caudill, S.: Simulating transient noise bursts in ligo with generative adversarial networks. Phys. Rev. D **106**, 023027 (2022) https://doi.org/10.1103/PhysRevD.106.023027

[51] Yan, J., Leung, A.P., Hui, C.Y.: On improving the performance of glitch classification for gravitational wave detection by using Generative Adversarial Networks. Monthly Notices of the Royal Astronomical Society **515**(3), 4606–4621 (2022) https://doi.org/10.1093/mnras/stac1996 https://academic.oup.com/mnras/article-pdf/515/3/4606/45475475/stac1996.pdf

[52] Powell, J., Sun, L., Gereb, K., Lasky, P.D., Dollmann, M.: Generating transient noise artefacts in gravitational-wave detector data with generative adversarial networks. Class Quantum Gravity **40**(3), 035006 (2023) https://doi.org/10.1088/1361-6382/acb038

[53] Boudart, V.: Convolutional neural network to distinguish glitches from minute-long gravitational wave transients. Phys. Rev. D **107**, 024007 (2023) https://doi.org/10.1103/PhysRevD.107.024007

[54] Choudhary, S., More, A., Suyamprakasam, S., Bose, S.: Deep learning network to distinguish binary black hole signals from short-duration noise transients. Phys. Rev. D **107**, 024030 (2023) https://doi.org/10.1103/PhysRevD.107.024030

[55] Meijer, Q., Lopez, M., Tsuna, D., Caudill, S.: Gravitational-wave searches for cosmic string cusps in einstein telescope data using deep learning. Phys. Rev. D **109**, 022006 (2024) https://doi.org/10.1103/PhysRevD.109.022006

[56] Álvarez-López, S., Liyanage, A., Ding, J., Ng, R., McIver, J.: Gspynettree: a signal-vs-glitch classifier for gravitational-wave event candidates. Classical and Quantum Gravity **41**(8), 085007 (2024) https://doi.org/10.1088/1361-6382/ad2194

[57] Sharma, K., Chandra, K., Pai, A.: Searching for massive black hole binaries with a transfer learning algorithm. Phys. Rev. D **108**, 124061 (2023) https://doi.org/10.1103/PhysRevD.108.124061

[58] Mogushi, K., Quitzow-James, R., Cavaglia, M., Kulkarni, S., Hayes, F.: Nnetfix: an artificial neural network-based denoising engine for gravitational-wave signals. Mach. Learn.: Sci. Technol. **2**(3), 035018 (2021) https://doi.org/10.1088/2632-2153/abea69

[59] Macas, R., Lundgren, A., Ashton, G.: Revisiting the evidence for precession in gw200129 with machine learning noise mitigation. Phys. Rev. D **109**, 062006 (2024) https://doi.org/10.1103/PhysRevD.109.062006

[60] Yu, H., Adhikari, R.X.: Nonlinear noise cleaning in gravitational-wave detectors with convolutional neural networks. Front. Artif. Intell. Appl. **5** (2022) https://doi.org/10.3389/frai.2022.811563

[61] Ormiston, R., Nguyen, T., Coughlin, M., Adhikari, R.X., Katsavounidis, E.: Noise reduction in gravitational-wave data via deep learning. Phys. Rev. Res. **2**, 033066 (2020) https://doi.org/10.1103/PhysRevResearch.2.033066

[62] Wei, W., Huerta, E.A.: Gravitational wave denoising of binary black hole mergers with deep learning. Physics Letters B **800**, 135081 (2020) https://doi.org/10.1016/j.physletb.2019.135081

[63] Shen, H., George, D., Huerta, E.A., Zhao, Z.: Denoising gravitational waves with enhanced deep recurrent denoising auto-encoders. In: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3237–3241 (2019). https://doi.org/10.1109/ICASSP.2019.8683061

[64] Chatterjee, C., Wen, L., Diakogiannis, F.I., Vinsen, K.: Extraction of binary black hole gravitational wave signals from detector data using deep learning. Physical Review D (2021)

[65] Acernese, F., Barone, F., Eleuteri, A., Garufi, F., Milano, L., Tagliaferri, R.: A hierarchical neural network-based approach to virgo noise identification. In: Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, vol. 6, pp. 253–2586 (2000). https://doi.org/10.1109/IJCNN.2000.859405

[66] Goel, L., Mukherjee, J.: Characterizing high level ligo gravitational wave data using deep learning. In: Bi, Y., Bhatia, R., Kapoor, S. (eds.) Intelligent Systems and Applications, pp. 848–860. Springer, Cham (2020)

[67] Soni, S., *et al.*: Discovering features in gravitational-wave data through detector characterization, citizen science and machine learning. Class Quantum Gravity **38**(19), 195016 (2021) https://doi.org/10.1088/1361-6382/ac1ccb

[68] Sakai, Y., *et al.*: Training process of unsupervised learning architecture for gravity spy dataset. Annalen der Physik **536**(2), 2200140 (2022) https://doi.org/10.1002/andp.202200140 https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.20220

[69] Sakai, Y., *et al.*: Unsupervised learning architecture for classifying the transient noise of interferometric gravitational-wave detectors. Scientific Reports volume **12**, 9935 (2022) https://doi.org/10.1038/s41598-022-13329-4

[70] Bişkin, O.T., Kirbaş, .I., Çelik, A.: A fast and time-efficient glitch classification method: A deep learning-based visual feature extractor for machine learning algorithms. Astronomy and Computing **42**, 100683 (2023) https://doi.org/10.1016/j.ascom.2022.100683

[71] Fernandes, T., Vieira, S., Onofre, A., Bustillo, J.C., Torres-Forne, A., Font, J.A.: Convolutional neural networks for the classification of glitches in gravitational-wave data streams. Class Quantum Gravity **40**(19), 195018 (2023) https://doi.org/10.1088/1361-6382/acf26c

[72] Razzano, M., Di Renzo, F., Fidecaro, F., Hemming, G., Katsanevas, S.: Gwitchhunters: Machine learning and citizen science to improve the performance of gravitational wave detector. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1048**,

167959 (2023) https://doi.org/10.1016/j.nima.2022.167959

[73] Koyama, N., Sakai, Y., Sasaoka, S., Dominguez, D., Somiya, K., Omae, Y., Terada, Y., Meyer-Conde, M., Takahashi, H.: Enhancing the rationale of convolutional neural networks for glitch classification in gravitational wave detectors: a visual explanation. Machine Learning: Science and Technology **5**(3), 035028 (2024) https://doi.org/10.1088/2632-2153/ad6391

[74] Yu, H., Adhikari, R., Magee, R., Sachdev, S., Chen, Y.: Early warning of coalescing neutron-star and neutron-star-black-hole binaries from the nonstationary noise background using neural networks. Physical Review D **104** (2021) https://doi.org/10.1103/PhysRevD.104.062004

[75] Kato, T., Hasegawa, M., Hirobayashi, S.: Validation of denoising system using non-harmonic analysis and denoising convolutional neural network for removal of gaussian noise from gravitational waves observed by ligo. Astronomy and Computing **40**, 100607 (2022) https://doi.org/10.1016/j.ascom.2022.100607

[76] Murali, C., Lumley, D.: Detecting and denoising gravitational wave signals from binary black holes using deep learning. Phys. Rev. D **108**, 043024 (2023) https://doi.org/10.1103/PhysRevD.108.043024

[77] Ma, C., Wang, S., Wang, W., Cao, Z.: Using deep learning to predict matched signal-to-noise ratio of gravitational waves. Phys. Rev. D **109**, 043009 (2024) https://doi.org/10.1103/PhysRevD.109.043009

[78] Ma, C., Li, S., Cao, Z., Jia, M.: Using deep learning to denoise and detect gravitational waves. Phys. Rev. D **110**, 063010 (2024) https://doi.org/10.1103/PhysRevD.110.063010

[79] Houba, N., Ferraioli, L., Giardini, D.: Detection and mitigation of glitches in lisa data: A machine learning approach. Physical Review D **109** (2024) https://doi.org/10.1103/PhysRevD.109.083027

[80] Saleem, M., Gunny, A., Chou, C.-J., Yang, L.-C., Yeh, S.-W., Chen, A.H.Y., Magee, R., Benoit, W., Nguyen, T., Fan, P., Chatterjee, D., Marx, E., Moreno, E., Omer, R., Raikman, R., Rankin, D., Sharma, R., Coughlin, M., Harris, P., Katsavounidis, E.: Demonstration of machine learning-assisted low-latency noise regression in gravitational wave detectors. Classical and Quantum Gravity **41**(19), 195024 (2024) https://doi.org/10.1088/1361-6382/ad708a

[81] Attadio, F., *et al.*: Neural network method to search for long transient gravitational waves. Phys. Rev. D **110**(10), 103047 (2024) https://doi.org/10.1103/PhysRevD.110.103047 arXiv:2407.02391 [astro-ph.IM]

[82] Mukund, N., Abraham, S., Kandhasamy, S., Mitra, S., Philip, N.S.: Transient classification in ligo data using difference boosting neural network. Phys. Rev. D **95**, 104059 (2017) https://doi.org/10.1103/PhysRevD.95.104059

[83] Laguarta, P., Laag, R., Lopez, M., Dooney, T., Miller, A.L., Schmidt, S., Cavaglia, M., Caudill, S., Driessens, K., Karel, J., Lenders, R., Van Den Broeck, C.: Detection of anomalies amongst ligo's glitch populations with autoencoders. Classical and Quantum Gravity **41**(5), 055004 (2024) https://doi.org/10.1088/1361-6382/ad1f26

[84] Li, Y., Wu, Y., Katsaggelos, A.K.: Cross-temporal spectrogram autoencoder (ctsae): Unsupervised dimensionality reduction for clustering gravitational wave glitches. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 6837–6846 (2024)

[85] Hayashi, H., Sakai, K., Hamazumi, H., Takahashi, H., Omae, Y.: A basic study of noise reduction on the analysis of burst gravitational waves by direct and parallel denoising autoencoder. ICIC Express Letters **14**(4), 337–345 (2020) https://doi.org/10.24507/icicel.14.04.337 . Publisher Copyright: © 2020 ISSN 1881-803X.

[86] Bini, S., Vedovato, G., Drago, M., Salemi, F., Prodi, G.A.: An autoencoder neural network integrated into gravitational-wave burst searches to improve the rejection of noise transients. Class Quantum Gravity **40**(13), 135008 (2023) https://doi.org/10.1088/1361-6382/acd981

[87] Bacon, P., Trovato, A., Bejger, M.: Denoising gravitational-wave signals from binary black holes with a dilated convolutional autoencoder. Mach. Learn.: Sci. Technol. **4**(3), 035024 (2023) https://doi.org/10.1088/2632-2153/acd90f

[88] Dooney, T., Curier, R.L., Tan, D.S., Lopez, M., Van Den Broeck, C., Bromuri, S.: One flexible model for multiclass gravitational wave signal and glitch generation. Phys. Rev. D **110**, 022004 (2024) https://doi.org/10.1103/PhysRevD.110.022004

[89] Jiaxiang, H., Jiale, L.: Qgformer: Quantum-classical hybrid transformer architecture for gravitational wave detection. In: 2023 20th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), pp. 1–5 (2023). https://doi.org/10.1109/ICCWAMTIP60502.2023.10387052

[90] Zhao, T., Lyu, R., Wang, H., Cao, Z., Ren, Z.: Space-based gravitational wave signal detection and extraction with deep neural network. Commun. Phys. **6**(1), 212 (2023) https://doi.org/10.1038/s42005-023-01334-6 arXiv:2207.07414 [gr-qc]

[91] Ma, C., Ma, C., Cao, Z., Jia, M.: Gravitational wave search by time-scale-recursive denoising and matched filtering. Sci. China Phys. Mech. Astron. **67**(12), 129511 (2024) https://doi.org/10.1007/s11433-024-2469-4

[92] Wang, H., Zhou, Y., Cao, Z., Guo, Z., Ren, Z.: Waveformer: transformer-based denoising method for gravitational-wave data. Machine Learning: Science and Technology **5**(1), 015046 (2024) https://doi.org/10.1088/2632-2153/ad2f54

[93] Apostol, E.S., Truică, C.-O.: Efficient machine learning ensemble methods for detecting gravitational wave glitches in ligo time series. 2023 IEEE 19th International Conference on Intelligent Computer Communication and Processing (ICCP), 79–86 (2023)

[94] Xu, Y., Du, M., Xu, P., Liang, B., Wang, H.: Gravitational wave signal extraction against non-stationary instrumental noises with deep neural network. Physics Letters B **858**, 139016 (2024) https://doi.org/10.1016/j.physletb.2024.139016

[95] Dosovitskiy, A., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. ArXiv **abs/2010.11929** (2020)

[96] Shmelkov, K., Schmid, C., Alahari, K.: How good is my gan? CoRR **abs/1807.09499** (2018) 1807.09499

[97] Aasi, J., *et al.*: Advanced ligo. Class Quantum Gravity **32**(7), 074001 (2015) https://doi.org/10.1088/0264-9381/32/7/074001

[98] Acernese, F., *et al.*: Advanced virgo: a second-generation interferometric gravitational wave detector. Class Quantum Gravity **32**(2), 024001 (2014) https://doi.org/10.1088/0264-9381/32/2/024001

[99] Akutsu, T., *et al.*: KAGRA: 2.5 generation interferometric gravitational wave detector. Nat. Astron. **3**, 35–40 (2019) https://doi.org/10.1038/s41550-018-0658-y arXiv:1811.08079 [gr-qc]

[100] Lynch, R., Vitale, S., Essick, R., Katsavounidis, E., Robinet, F.: Information-theoretic approach to the gravitational-wave burst detection problem. Phys. Rev. D **95**, 104046 (2017) https://doi.org/10.1103/PhysRevD.95.104046

[101] Punturo, M., *et al.*: The einstein telescope: a third-generation gravitational wave observatory. Class Quantum Gravity **27**(19), 194002 (2010) https://doi.org/10.1088/0264-9381/27/19/194002

[102] Robson, T., Cornish, N.J., Liu, C.: The construction and use of lisa sensitivity curves. Classical and Quantum Gravity **36**(10), 105011 (2019) https://doi.org/10.1088/1361-6382/ab1101

[103] Chatterji, S., Blackburn, L., Martin, G., Katsavounidis, E.: Multiresolution techniques for the detection of gravitational-wave bursts. Class Quantum Gravity **21**(20), 1809 (2004) https://doi.org/10.1088/0264-9381/21/20/024

[104] Shen, H., George, D., Huerta, E.A., Zhao, Z.: Denoising Gravitational Waves using Deep Learning with Recurrent Denoising Autoencoders. arXiv e-prints, 1711–09919 (2017) https://doi.org/10.48550/arXiv.1711.09919 arXiv:1711.09919 [gr-qc]

[105] Gebhard, T.D., Kilbertus, N., Harry, I., Schölkopf, B.: Convolutional neural networks: A magic bullet for gravitational-wave detection? Phys. Rev. D **100**, 063015 (2019) https://doi.org/10.1103/PhysRevD.100.063015

[106] Tuli, S., Dasgupta, I., Grant, E., Griffiths, T.L.: Are Convolutional Neural Networks or Transformers more like human vision? (2021). https://arxiv.org/abs/2105.07197

[107] Shoemaker, D., Rocchi, A., Miyoki, S.: LIGO Document G2002127-v25. [Online]. Accessed: 9 July 2024 (2024). https://dcc.ligo.org/LIGO-G2002127/public

[108] Di Renzo, F., Fidecaro, F., Hemming, G., Katsanevas, S., Razzano, M.: GWitch-Hunters – A citizen science project for the improvement of gravitational wave detectors. PoS **ICHEP2022**, 1152 (2022) https://doi.org/10.22323/1.414.1152

[109] Abbott, B.P., *et al.*: 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) https://doi.org/10.1103/PhysRevX.9.031040

[110] Glanzer, J., *et al.*: Data quality up to the third observing run of advanced ligo: Gravity spy glitch classifications. Class Quantum Gravity **40**(6), 065004 (2023) https://doi.org/10.1088/1361-6382/acb633

[111] OpenAI, et al.: GPT-4 Technical Report (2024)