

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,500

Open access books available

176,000

International authors and editors

190M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Anomaly Detection in Medical Time Series with Generative Adversarial Networks: A Selective Review

Miloš Cekić

Abstract

Anomaly detection in medical data is often of critical importance, from diagnosing and potentially localizing disease processes such as epilepsy to detecting and preventing fatal events such as cardiac arrhythmias. Generative adversarial networks (GANs) have since their inception shown promise in various applications and have been shown to be effective in cybersecurity, data denoising, and data augmentation, and have more recently found a potentially important place in the detection of anomalies in medical time series. This chapter provides a selective review of this novel use of GANs, in the process highlighting the nature of anomalies in time series, special challenges related to medical time series, and some general issues in approaching time series anomaly detection with deep learning. We cover the most frequently applied GAN models and briefly detail the current landscape of applying GANs to anomaly detection in two commonly used medical time series, electrocardiography (ECG) and electroencephalography (EEG).

Keywords: anomaly detection, medical time series, generative adversarial network (GAN), electrocardiogram (ECG), electroencephalogram (EEG)

1. Introduction

The increasingly widespread deployment of advanced technology in modern healthcare has led to exponential growth in the generation and collection of medical time series data, which offer unprecedented opportunities for the detection and diagnosis of disease processes. Being able to use these “big” data to detect and localize anomalies can lead to early and precise disease recognition, timely and proactive intervention, development of personalized treatment plans, improved patient outcomes, and better risk management while at the same time offering opportunities for better understanding and classification of disease pathology [1–5].

The extraction of meaningful insights from medical time series data, however, poses a number of significant challenges. For one, the data are highly complex and

often multidimensional/multimodal and nonstationary. They also usually suffer from noise, missing values, and artifacts as well as from the inherent variability of human physiology and broad range of normality across individuals, which makes the identification of “abnormal” a nontrivial task [6]. Anomalies within the data can be highly heterogeneous and can manifest as subtle deviations or sudden, drastic changes. Additionally, specific disease data that could be used to train analytical models are at best highly heterogeneous and more likely simply not available; even when data are available, properly labeled data are generally lacking. Finally, at the present time, there is no expert knowledge related to large medical datasets that can parallel classical medical knowledge—it is sometimes not even clear what exactly to look for since traditional medical semiology operates in a different conceptual space (defined signs/symptoms vs. patterns in massive data) [1–3, 6].

Traditionally, statistical and rule-based methods have been employed for anomaly detection in medical time series data [4, 7, 8]. These methods rely on predefined thresholds, statistical models, or expert knowledge to identify deviations from normal. These approaches, however, often struggle to capture complex and nonlinear patterns that may be present in the data. This has recently led to a growing interest in leveraging machine learning and deep learning techniques for anomaly detection in medical imaging [1–3, 6, 7, 9–11]. A widely used type of model is the generative adversarial network (GAN), which has demonstrated superiority in a variety of tasks in medical imaging due to its powerful ability to learn the distribution of the training data and to generate novel but realistic samples that reflect the underlying data characteristics [12–26]. A GAN trained on normal instances only, for example, can capture the complex patterns and dependencies inherent in the data, enabling the generation of synthetic samples that closely adhere to the learned distribution. Anomalies, being significantly different from normal, can then be identified by how far they deviate from their reconstruction [27–32]. The application of GANs in anomaly detection for medical time series has demonstrated promising results, and GAN-based approaches have been shown to be able to effectively capture temporal dependencies, handle complex patterns, and adapt to individual patient variations. Moreover, they can detect subtle anomalies that may go unnoticed by traditional methods [29, 31, 32].

This chapter endeavors to provide a brief review of the current landscape of the use of GANs in anomaly detection in medical time series data. We first present a brief overview of properties of anomalies, time series data, and specific challenges related to medical time series. We then discuss general ways of approaching time series with deep learning methods before discussing GANs and GAN applications to anomaly detection in general and to time series in particular. We then review the current state of the use of GANs in medical imaging and anomaly detection in specific fields of electrocardiography (ECG) and electroencephalography (EEG). Finally, we briefly discuss some challenges and future directions.

2. Anomalies in time series: problem complexities and challenges

2.1 Properties of time series data

While time is a fundamental concept in nearly all data, time series explicitly involve the temporal dimension. The following is a brief summary of the specific properties of time series and how they affect anomaly detection.

2.1.1 Temporality

A time series is an ordered sequence of data points indexed by time (usually but not always across equal temporal intervals) [33–36]. We can define a time series as a vector X such that: $X = \{x_1, x_2, \dots, x_t\}$, where x_i represents the datum at time $i \in T$ and $T = \{1, 2, \dots, t\}$. The (necessary) assumption of continuity of the underlying generative process implies that each point is in some way conditioned on previous values (past states of the process), with this dependency captured as a joint distribution of a set of observations: $p(x_1, x_2, \dots, x_t) = p(x_1) \prod_2^t p(x_t | x_1, x_2, \dots, x_{t-1})$. The influence of the past is generally assumed to decrease with time, though this may not necessarily be the case [33].

2.1.2 Dimensionality

Time series data may be univariate or multivariate, with the dimension representing the number of individual data attributes captured at each time point. The above specification of a time series vector is univariate. Multivariate series can be defined as a time-ordered set of multidimensional vectors X_t (rather than points), with $X_t = (x_t^1, x_t^2, \dots, x_t^d)$, where d is the number of dimensions; the multivariate time series is then a rank $d + 1$ tensor X_i^j , where $j \in D$ and $D = \{1, 2, \dots, d\}$ is the number of dimensions and $i \in T$ and $T = \{1, 2, \dots, t\}$ denotes time [9, 33]. Alternatively, multivariate series can also be conceptualized as a collection of univariate time series. While analysis of univariate time series needs to consider only the relationship between the current state and previous states (temporal dependency), multivariate series entail dependencies and correlations (semantics) across both previous states (temporal) within a series and other dimensions (spatial) at any given time point, keeping in mind that any given datum may also depend on a mixture within and across different time series (spatiotemporal dependencies). These dependencies may be multiscale (short-, medium-, or long-range) and in some cases nonstationary or dynamic, meaning that the scale and structure of dependencies itself may vary in time [37–40].

2.1.3 Nonstationarity

A time series is assumed to be stationary if its statistical properties do not change over time. Most real-world time series are not stationary, however, meaning the mean and variance (and other moments of the distribution) vary. Common sources of nonstationarity include trends (baseline drift that may be local or global and linear or nonlinear), seasonal cycles (with a stable period), nonseasonal cycles (with a variable period), pulses and steps (including concept drift and change points, instances where the relationship between input and output changes), and random/irregular movement. Because nonstationarity implies that the data distribution itself changes in some way, an appropriate model will need to somehow capture the underlying generative process rather than the statistics of the apparent data [33–36].

2.1.4 Noise

Real-world datasets typically contain a significant amount of noise or unwanted signal, which represents the semantic boundary between normal data and true

anomalies. The classical definition of an anomaly or outlier is “an observation which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism” [41]. In this context, it is helpful to differentiate between outlier and anomaly: *outlier* refers to “unusual data objects that are statistically rare but may or may not result from the same generative process as other data objects,” while, in contrast, *anomalies* are defined as “data objects that have resulted from a distinct generative process that may or may not appear in the data as outliers” [42]. This distinction is necessarily contextual and application-specific, although typically anomalies will have a much higher outlier score than noise [43] since they are presumably generated by a different underlying process.

2.2 Properties of anomalies

The detection of anomalies is a specific problem in pattern recognition that is distinct from other analytical and learning tasks. Key complexities are discussed below. These issues apply to any anomaly detection and not just time series, but with time series and medical time series in particular additional specific challenges arise.

2.2.1 Unknown nature of anomalies

Anomalies are by definition unknown and may involve unknown abrupt behaviors, patterns, or distributions, which remain unknown until they occur. Even if a particular type of anomaly is known and categorized (there are two distinct issues here, since even if an anomaly type exists, it may not be immediately classifiable due to the heterogeneity of its manifestations), recognizing it may still be difficult. With machine learning, there is the added complexity that, even if various diseases are categorized in terms of their specific symptoms, the disease process may not be clearly defined within the particular modality performed for its detection [10, 11, 33, 37–39].

2.2.2 Anomaly class heterogeneity

Since anomalies are irregular and heterogeneous, one class of anomalies may have very different abnormal characteristics when compared to another class of anomalies; in other words, not only are anomaly classes themselves heterogeneous, the heterogeneity *within* anomaly classes or types is itself heterogeneous [11, 36, 39].

2.2.3 Dataset/class imbalance

Anomalies are typically rare events, occurring much less often than normal instances, which account for the overwhelming majority of the data. It is therefore extremely difficult and labor-intensive (if not impossible) to collect a sufficient amount of labeled and/or clearly defined abnormal instances that could be used for anomaly definition and model training. The result is severe class imbalance in any potential training set [29, 32, 40].

2.2.4 Types of anomalies

Anomalies can generally be classified into three types, point anomalies, contextual anomalies, and collective anomalies, which in time series correspond to

abnormal time points, abnormal time intervals/subsequences, and abnormal time series [10, 29, 33, 37, 39].

2.2.4.1 Point anomalies

A point anomaly is a single datum at which the value of the observed variable is significantly different either from the entire time series (global) or from neighboring points in a time series (local or contextual). Point anomalies may be univariate or multivariate and usually entail extreme values. An example may include an abnormal blood pressure reading, which would then need to be defined as noise or something potentially indicating a deeper problem.

2.2.4.2 Time interval/subsequence anomalies

This type involves a subsequence of points that does not reflect the normal behavior of the system and in which each individual observation may be within normal range but the subsequence as a whole is anomalous. The subsequence may affect a single (univariate) or multiple (multivariate) time-dependent variables. An example would be an epileptic seizure on electroencephalogram (EEG), which may not be out of normal range in any individual point, but in which the multivariate pattern across multiple electrodes over a finite time period reflects the abnormality.

2.2.4.3 Collective or time series anomalies

This class includes cases where the entirety of a (or several) time series of a multivariate dataset is anomalous with respect to the dataset as a whole. This type is distinct from subsequence anomalies due to the length of the anomaly, which extends to the same length as the sequence. The idea here is that what is represented is not a temporary anomaly in the functioning of a part of a multidimensional system that returns to normal at some point, but rather that there is a persistent underlying abnormality in some portion of the system that can only be detected in the context of the entire dataset. An example would be abnormal signal related to a voxel (volume element) or group of voxels in a resting-state functional magnetic resonance imaging (MRI) (rs-fMRI) study representing a brain region that might reflect an underlying disease process such as an epileptogenic focus or tumor. The only way to detect this is by looking at the relationship of the various time series to each other both within individual patients and across individuals.

2.3 Challenges specific to medical time series

Anomaly detection in medical time series data comes with unique challenges, stemming from the nature of the data, the inherent variability in human physiology, and the requirement for the results to be interpretable by healthcare professionals. The following are the major domain-specific challenges.

2.3.1 Noise and artifacts

Medical time series data can contain a significant amount of noise due to sensor inaccuracies, patient movement, measurement errors, and physiological

artifacts [1, 2]. For example, an ECG signal may contain noise from muscle contractions, or a glucose monitor may have inaccuracies due to calibration errors. These artifacts can distort the underlying physiological signal and lead to false detections of anomalies.

2.3.2 Missing and irregularly sampled data

Medical time series data often suffer from missing values and irregular sampling intervals. For example, a patient might remove a wearable device for a period, leading to missing data, or a sensor might malfunction. Irregularly sampled data can arise in outpatient settings where measurements are taken at each visit, but the visits occur at irregular intervals. These irregularities pose challenges for conventional time series analysis methods, which typically assume regular sampling, and require specialized techniques to handle missing values, synchronize timestamps, and ensure consistent analysis across different time series [1, 2, 6].

2.3.3 Nonstationarity

Medical time series data often exhibit nonstationarity, meaning that their statistical properties change over time. This could be due to a patient's changing health status, the effect of medications or interventions, or changes in external conditions such as time of day or physical activity. Traditional time series analysis techniques often assume stationarity, so nonstationarity poses a significant challenge [35, 43].

2.3.4 High dimensionality

Medical time series data can involve a high number of variables or channels, possibly with different types of data that are collected differently (e.g., blood pressure, heart rate, electrocardiogram, oxygen (O₂) saturation, and respiration may all be monitored simultaneously in the intensive care unit (ICU)) or may consist of massively multidimensional imaging data (e.g., for the open source dataset in the Human Connectome Project, each raw rs-fMRI time point contains 673,920 voxels or dimensions, which over the span of an approximately 15 min scan generates 8×10^8 data points per run per subject). This high dimensionality presents challenges in data storage, computation, visualization, and analysis. It also increases the need for complexity of anomaly detection algorithms, as they need to be able to handle and interpret data across multiple dimensions [1, 4, 5, 38].

2.3.5 Intra- and interpatient variability

There is a high degree of variability in physiological parameters both within the same individual over time (inpatient variability) and between different individuals (interpatient variability). This variability complicates the task of defining what constitutes an "anomaly." An anomalous value for one individual might be normal for another, and a reading that is normal for a person at rest might be abnormal during exercise. Anomaly detection methods need to account for these interpatient variabilities and develop patient-specific or subgroup-specific models to accurately capture normal and abnormal patterns [1].

2.3.6 Lack of labeled data

Supervised learning methods, which can be very effective for anomaly detection, require labeled data for training. Obtaining such labeled data in healthcare settings is challenging, however, as it typically requires expert clinicians to manually review and label the data, which is time-consuming and expensive [2]. This essentially forces most anomaly detection (AD) models to be semi-supervised or unsupervised (see below).

2.3.7 Privacy and security

Medical data are highly sensitive, and strong safeguards are required to protect patient privacy. This can make it challenging to gather sufficient data for building robust anomaly detection models, and it also requires careful consideration when deploying these models to ensure that patient data are handled securely [1, 27, 28].

2.3.8 Interpretability and explainability

Healthcare professionals need to understand the detected anomalies and trust the reasoning behind them to make informed decisions. Anomaly detection methods should therefore provide interpretable results, visualizations, or explanations that can lead to improved or optimized treatment decisions. Many advanced machine learning models, however, and particularly deep learning models, often act as “black boxes,” making their predictions difficult to interpret. This lack of interpretability can hinder the adoption of these models in clinical practice [1, 28].

2.4 Summary

Anomaly detection in multidimensional time series data is a significant challenge both from conceptual and technical points of view. Anomalies are by nature rare and heterogeneous, and therefore not easily amenable to classification, which means that they must generally be classified not as what they are but rather as what they are not. This makes not just detection but even simply definition of anomalies challenging: what is normal? How “different” must something be from some generalization (e.g., mean or average) of normal to constitute an anomaly? How do we define this difference? If the normal data change over time, how do we learn and describe this aspect of normality? What is the normal evolution of the behavior of our system? For multivariate time series, the issues multiply due to dependencies in the data that may extend over time and across multiple variables and may involve different spatiotemporal scales simultaneously. Determining normal behavior then is quite challenging, and the detection of potentially multiscale spatiotemporal anomalies even more so. As an extended example, consider rs-fMRI once again: say a brain region comprising a set of voxels is abnormal and therefore generates an anomalous signal. Detecting this anomaly requires that we are somehow able to learn all the multivariate dependencies at multiple scales across multiple subjects (the normal population): spatially between not just voxels but groups of voxels comprising brain regions, which may vary in size; temporally in that the length of the anomalous series may vary from short unusual behaviors to longer term abnormalities; and spatiotemporally in that different length signals from regions of different sizes may affect other regions at different spatial scales. Since anomalies are heterogeneous and we cannot know *a priori* exactly what

we are looking for, detecting an “anomalous brain region” in the highly multidimensional rs-fMRI signal requires significant model complexity and computational power.

Finally, there are additional practical concerns: generalized unavailability of labeled training data, unreliable or noisy data, heterogeneously organized or collected data, data privacy (extremely important in medical datasets), interpretability, throughput and automatization, and scalability from specific research purposes to clinical applicability.

3. Anomaly detection in time series: concepts and models

3.1 Basic paradigms

Traditional anomaly detection methods in healthcare include statistical approaches, rule-based methods, and machine learning techniques. Statistical methods often rely on distribution-based models, such as Gaussian or Hidden Markov Models, to capture the normal patterns and detect deviations from them. More sophisticated classical models include autoregressive integrated moving average (ARIMA) models, exponential smoothing state space model (ETS), and Seasonal-Trend decomposition using LOESS (STL). These approaches are useful for detecting point or global anomalies, but are not helpful in the detection of contextual or collective anomalies [42–45] and are not useful for multivariate or multidimensional data [37–39]. Rule-based methods utilize expert-defined rules or thresholds to identify anomalies based on predefined criteria. Machine learning models, especially unsupervised machine learning algorithms, like clustering (K-means, density-based spatial clustering of applications with noise (DBSCAN)) or K-nearest neighbors (K-NN), are often used to detect outliers in time series data. Supervised learning methods, like support vector machine (SVM) or Random Forests, can also be used when labeled data are available. Deep learning methods, however, are best for capturing complex nonlinear relationships and dependencies in time series data [33, 35, 44–47]. Long Short-Term Memory (LSTM) networks and autoencoders (AEs) have been common choices due to their ability to model temporal dependencies, with GANs being utilized more recently for anomaly detection in time series [29, 33, 39].

Generally speaking, machine (and deep) learning methods can be grossly categorized along two axes: learning scheme vs. anomaly determination. Note that additional axes are possible, e.g., univariate vs. multivariate, anomaly type (point, subsequence, or time series); however, these will not be specifically considered here since most medical time series are multivariate and most deep learning approaches attempt to find anomalies in an anomaly-type agnostic manner [33].

3.1.1 Learning scheme

The three major learning schemes in machine learning are supervised, unsupervised, and semi-supervised. Supervised models aim to learn a mapping from data to their corresponding annotations and then use this mapping to perform classification on test data. Important examples of a supervised approach in the medical field are applications to automated brain tumor and ischemic stroke recognition and segmentation (using the extensively labeled Multimodal Brain Tumor Image

Segmentation, BraTS [48], and Ischemic Stroke Lesion Segmentation, ISLES, datasets [49]). Although these networks can successfully learn to recognize visual anomalies in medical images [2, 4], they require accurate labels and annotations and can only learn to recognize specific predefined abnormalities that are already present in the training set. Given the difficulty and time-consuming nature of labeling data manually, this approach is usually only appropriate for very specific use-cases.

Most approaches to anomaly detection time series therefore take the form of unsupervised or semi-supervised learning [32, 38, 45]. In unsupervised anomaly detection, the algorithms separate anomalies without prior knowledge or any explicit distinction between normal and abnormal, whereas in semi-supervised approaches all training data are assumed to be normal. Unsupervised methods are the most flexible since they depend entirely on the internal features of the data; however, this type of approach generates its own set of problems including potentially nonconvergent training, unclear recognition of anomalies, and difficulty in interpretation. For most medical data, including time series, a semi-supervised approach is generally considered to be most appropriate [6, 10, 11, 45–47], given the fact that the overwhelming amount of data collected is normal and that the anomalies themselves are highly heterogeneous.

3.1.2 Anomaly determination

Most anomaly detection involves an “anomaly score”—a number that is calculated based on the analytical technique that can then be compared to what is expected from a normal dataset. Once the data are learned (the model or distribution is assumed and fitted), a measure of the difference of each particular datum (whatever form this takes, point, subsequence, etc.) from the learned distribution must be determined, and this is the anomaly score. The following approach is nearly universal in machine (and by extension deep) learning for anomaly detection in time series: pick or create a model/architecture that we think will appropriately model the dataset, train the model on normal data in order to learn the data distribution (presumably with all dependencies), and then test new data with reference to the learned distribution using an anomaly score, which then determines if the datum in question was generated by the same underlying process the model was trained on or something different [11, 29].

Three basic approaches to determining anomaly scores are used, regardless of the time-modeling approach taken (see below): forecasting/prediction, reconstruction, and distance/dissimilarity [45].

3.1.2.1 Prediction

Prediction or forecasting-based models learn to predict expected future values based on the learned data, with anomalies determined based on the residual between the predicted value and the observed quantity. Most forecasting models use a sliding window to forecast one point at a time, although short sequences can also be generated. There is no robust forecasting-based model for rapidly and continuously changing time series, however, since such time series can only be predicted in the very short term if at all, and forecasting models are known to demonstrate significantly increased prediction errors as the number of time points increases [39]. This also makes them generally unsuitable for subsequence anomalies. Even in the deep learning context, forecasting-based models can only make short-term predictions with acceptable accuracy.

There is certainly a place for forecasting-based models in medical time series analysis, however. For any real-time applications where early or real-time anomaly detection is important, forecasting models are crucial. An example would be patients on telemetry or in the ICU, where blood pressure, respiration, and ECG may, individually or in combination, signal an impending collapse. This may also be the case in predicting the onset of seizures with EEG or cardiac arrest with ECG, where any advance notice of a critical event may significantly alter outcomes.

3.1.2.2 Reconstruction

Reconstruction-based models are not subject to the constraints of prediction models. With this type of model, normal behavior is modeled by encoding subsequences of normal training data (usually input as a sliding window that provides the temporal context for each datum) into a lower-dimensional latent space. In a semi-supervised context where the model is trained on only normal data, the model should be incapable of generating an anomalous output sequence. Anomalies are therefore detected by reconstructing a subsequence/sliding window from the test data and comparing it to the actual values, which generates a “reconstruction” error. Anomalies are generally flagged when the reconstruction probability falls below a specified threshold.

Most deep learning methods including generative models such as autoencoders (AEs), variational autoencoders (VAEs), generative adversarial networks (GANs), and transformers use reconstruction error as the anomaly score [29, 33]. Although these models are different in their architectures, training, and objective functions, most approaches using one of these models calculate anomaly scores as reconstruction errors. Note that in multivariable time series with multiscale spatiotemporal dependencies deciding what exactly constitutes “similarity” may be difficult. Fortunately, analytically defining similarity is usually not necessary and the anomaly score is often related to the loss function of the model. While reconstruction-based AD methods are fairly intuitive and quite widely used, they can be plagued with difficulties such as computational cost for data reconstruction, mode collapse, nonconvergence, and instability [33, 50].

3.1.2.3 Distance/dissimilarity

Distance-based models are based on a similarity metric that flags anomalies if their distance from normal is past a certain threshold. Clustering is an unsupervised machine learning model that is effective for grouping data and detecting anomalies; it involves mapping the time series data into a multidimensional space where the data are grouped near centroids based on feature similarity. Anomalies are then detected if they are far from existing clusters or have low probability of belonging to a cluster. Examples of clustering methods include K-means algorithm, one-class support vector machine (OCSVM), and Gaussian mixture model (GMM) [33]. More sophisticated machine learning methods such as Dynamic Time Warping provide a more complex comparison of temporal sequences (or subsequences, usually determined with a sliding window of fixed length) by allowing nonlinear alignment between sequences that are locally out of phase [11, 45]. Clustering methods are currently the benchmarks for anomaly detection in time series [39] and some recent studies demonstrate that many advanced algorithms do not deliver improved performance on basic univariate time series in comparison to more traditional methods and may in fact result in inferior

performance [47]. The performance of clustering methods is generally degraded on complex high-dimensional datasets [33, 43], however. In such cases, methods to reduce data dimensionality such as expert opinion or various feature selection and extraction techniques such as deep autoencoders, principal component analysis (PCA) [51], and multidimensional scaling (MDS) [52, 53] can be used, in effect utilizing a hybrid approach using deep learning for dimensionality reduction and clustering methods for anomaly detection.

3.2 Capturing temporal context

The history of a sequence contains significant information regarding its potential future behavior and most deep learning approaches in some way depend on modeling the temporal dimension in order to explicitly capture the past during reconstruction or prediction.

3.2.1 Input

Input shape is essential to capturing temporal context and can take the form of individual (multidimensional) points or windows, which consist of a subsequence that contains some portion of the historical information. The width of the window is usually predetermined and can be based on the known or estimated/expected characteristics of the dataset and the presumed underlying process. Windows can be advanced by some number of data points (window step) and used in order (sliding windows) or they can be shuffled and entered out of order depending on the application and dataset. To specifically address the challenge of comparing subsequences rather than points, many models use representations of subsequences instead of the raw data. A sliding window decomposition/extraction is usually performed in the preprocessing stage after other operations such as missing value imputation, changing the sampling rate, or data normalization, have been completed [38, 39].

3.2.2 Temporal modeling

Several approaches to modeling temporal context and dependencies are commonly used in deep learning models. These essentially provide ways of organizing “memory,” which amounts to in some way utilizing appropriately weighted previously encountered data in order to generate current or future output. These approaches may constitute the model architecture itself or they may be utilized at the midlevel of network dynamics and can be combined with various more basic activation functions as well as higher-level deep learning architectures.

3.2.2.1 Recurrent neural networks

Recurrent neural networks (RNNs) are a class of neural networks specifically designed for modeling sequential data and are well suited for capturing temporal and long-term dependencies. Unlike traditional feedforward neural networks, RNNs have a recurrent connection that allows information to be looped back and processed at each time step. This enables RNNs to retain memory of previous time steps that they can use to inform predictions at subsequent time steps. The hidden state of the RNN serves as an internal representation that evolves over time, capturing the context and

history of the time series [52, 53]. A major shortcoming of RNNs is their instability due to the vanishing or exploding gradient problem, where the learning gradient becomes extreme as the network becomes deeper [54]. The earliest and most widely used RNN modification designed to address this problem is the long short-term memory (LSTM) unit [55], which avoids the problem by controlling retained information through a memory cell and input, output, and forget gates. If the LSTM unit detects an important feature from an earlier input sequence, it can carry this information over an extended distance, sometimes up to thousands of steps [39, 55]. A simpler and more computationally efficient but similarly effective modification is the gated recurrent unit (GRU) [56], which modulates the flow of information inside the unit but without a separate memory cell [52, 53]. Both LSTM and GRU cells are able to learn long-term dependencies by determining the number of weighted previous states to keep or forget at each time step. Both types of cells have been used with success in time series anomaly detection [39].

3.2.2.2 Convolutional neural networks

Convolutional neural networks (CNNs) have traditionally been used for image analysis but can be adapted to time series data and in some applications demonstrate better performance than RNNs [57–59], which still remain the most commonly used approach to temporal modeling. CNNs treat time series data as a one-dimensional (1D) array rather than a sequence and employ convolutional operations to capture local patterns and dependencies within the data. By applying one-dimensional convolutions, CNNs can learn hierarchical representation of time series and automatically extract relevant features at different timescales. Pooling layers can be added to downsample the output and reduce dimensionality. The learned features can then be fed into fully connected layers for classification or regression tasks [29, 33]. Anomaly detection can be performed by training the model on a normal dataset and then computing prediction or reconstruction error during inference. CNNs can be used for multidimensional time series as well as for real-time detection of anomalies; however, the extensive computational demands of CNNs make them less efficient for real-time monitoring. A specific approach for time series, the temporal convolutional network (TCN) [60], uses one-way convolutions in order to maintain temporally ordered/causal relations in the data. The TCN can generate sequences of any length and employs dilated convolutions, where the receptive field of the convolutional filters expands exponentially, which allow the model to capture long-range dependencies in time series.

Convolutional neural networks and RNNs can be combined in the same network architecture in order to capture spatial and temporal dependencies distinctly but simultaneously. In such models, CNNs are usually used to capture local patterns and features, with the output then fed into a RNN (either LSTM or GRU) that then processes the features extracted from the CNN in a sequential manner. The RNN performs the task of modeling the temporal context and relationships between the features extracted by the CNN [34, 61].

3.2.2.3 Attention

The attention mechanism, initially popularized in the context of natural language processing [62], has been extended to handle general time series data and provides a

recent alternative to both RNN and CNN models. In the context of time series analysis, the attention mechanism allows the model to focus on specific temporal segments or patterns within the input sequence that are most relevant for making predictions. Using attention, the model can dynamically weigh the contribution of each time step based on its relevance, rather than relying solely on a fixed window or fixed-size context. Attention mechanisms typically involve a scoring mechanism that calculates the relevance or attention weight for each time step, followed by a weighted combination of the time step representations to produce a context vector that is then used for making predictions or further processing. Attention mechanisms in time series analysis have shown effectiveness in tasks such as sequence classification, forecasting, and anomaly detection, as they enable models to selectively attend to relevant temporal information while disregarding irrelevant or noisy segments of the time series [63–65].

3.2.2.4 Graph neural networks

Graph neural networks (GNNs) for time series data combine the power of graph structures with the ability to model temporal dependencies. GNNs enable the representation and analysis of time series data as graphs, where each data point is a node and the temporal relationships between them are represented as edges [66]. By incorporating recurrent or convolutional mechanisms, GNNs can capture the dynamics and patterns of time series data within the graph framework. These models leverage information from neighboring nodes and the temporal context to make predictions or perform tasks such as forecasting, classification, or anomaly detection. GNNs for time series offer a flexible and effective approach for handling complex temporal dependencies while leveraging the benefits of graph representations, enabling improved understanding and analysis of time-varying data [66]. GNN ideas have been implemented in graph convolutional networks [67] and graph attention networks [68] and are a promising future direction.

3.3 Summary

Deep learning models can automatically learn representations from raw time series data and capture both local and global dependencies, making them well suited for anomaly detection tasks. One of the key advantages of using deep learning for anomaly detection in time series is the ability to handle high-dimensional and complex data. Deep learning models such as RNNs, CNNs, and their variants, have been successfully applied to capture temporal dependencies, spatial patterns, and multiscale features in time series data and can effectively extract meaningful representations from the input data and detect anomalies based on deviations from learned normal patterns. These medium-scale models can be combined and integrated into larger scale unsupervised architectures such as autoencoders and GANs, which have been widely employed to learn compact representations of normal patterns in data. By reconstructing the input data or generating synthetic data samples, these models can detect anomalies by measuring the reconstruction error or the divergence between the real and generated data. More recent ways of capturing the temporal context include attention and GNNs, and these have shown promise in anomaly detection in time series.

4. Generative adversarial models for anomaly detection

4.1 Motivation for the use of GANs in medical data

The key motivation for utilizing GANs for anomaly detection in medical time series is that they almost directly address many of the challenges specific to medical time series discussed above [1, 14, 16–21]. Since GANs can learn to recreate normal data patterns and detect anomalies as deviations from these patterns, they are optimized to operate in an unsupervised setting; they can easily be combined with a temporal model (e.g., RNN or Attention) to model complex temporal and multivariate dependencies due to their ability to capture intricate data structures and patterns, nonlinearities, and high-dimensional relationships and they can do so in high-dimensional datasets involving multiple variables recorded over time; they can generate synthetic realistic data that can supplement original medical datasets, which often suffer from lack of volume and diversity, thus aiding in the training of more robust and generalizable models; and they have shown impressive generalization capabilities, allowing them to learn representations that generalize well to unseen data, which are particularly valuable in anomaly detection in medical time series, where the ability to accurately detect anomalies in previously unseen or rare cases is crucial.

4.2 GAN overview

Generative adversarial networks are a class of deep learning models introduced by Goodfellow et al. in 2014 [69] designed to generate synthetic data that resemble the training dataset by learning and reproducing the distribution inherent in the data. A GAN consists of two parts, a generator (\mathcal{G}) and a discriminator (\mathcal{D}), which are usually implemented as neural networks. Both networks play a two-player minimax game, in which \mathcal{G} tries to generate data that \mathcal{D} cannot distinguish from the real training data, while \mathcal{D} tries to correctly classify data as real (from the training data) or fake (from \mathcal{G}).

The generator's role is to produce synthetic data samples, such as images, texts, or audio, which mimic the distribution of the training data. The generator is a mapping function that projects random noise or an input vector sampled from a prior distribution (which may be uniform or Gaussian) to the data space. The network gradually learns to transform the noise into output that is generated from the dataset distribution and is ideally indistinguishable from the training data. The discriminator, on the other hand, acts as a binary classifier. It is trained to distinguish between real data samples from the training set and synthetic samples generated by the generator. The discriminator takes both real and generated samples as inputs and outputs a probability score indicating the likelihood that the input is real. The objective of the discriminator is to learn to accurately classify the data samples.

The generator and the discriminator are trained simultaneously. Initially, the generator produces poor quality outputs and the discriminator can easily tell the difference between real and fake. As training progresses, the generator becomes better at generating fake outputs that appear real, and the discriminator becomes better at distinguishing between the real and the fake. The training continues until a point at which the discriminator can no longer distinguish fake data from real data (Nash equilibrium). The optimization process uses backpropagation and an optimization algorithm (like stochastic gradient descent) to adjust the parameters of \mathcal{G} and \mathcal{D} . The loss function used for the training is typically binary cross-entropy.

In the context of anomaly detection, after the GAN is trained, new data instances can be fed to the trained generator, which then attempts to reconstruct or regenerate these instances. The regenerated samples are then compared to the original data. If the difference between the original and the regenerated sample surpasses a defined threshold, the data instance is flagged as an anomaly. The underlying assumption is that the GAN will be less successful in accurately reproducing instances that are significantly different from the distribution it was trained on, i.e., anomalous instances. This procedure allows for effective anomaly detection in an unsupervised setting [27, 29–33].

The objective function summarizing this process is:

$$\text{Min}_G \text{Max}_D f(\mathcal{D}, \mathcal{G}) = \mathbb{E}_{x \sim p_{data}(x)} [\log(\mathcal{D}(x))] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - \mathcal{D}(\mathcal{G}(z)))] \quad (1)$$

Where \mathbb{E}_x and \mathbb{E}_z are the expectation values over real and random data samples, respectively, $\mathcal{D}(x)$ is the probability estimate of the discriminator \mathcal{D} if x is real, $\mathcal{G}(z)$ is the output of the generator \mathcal{G} for a given vector z as input, and $\mathcal{D}(\mathcal{G}(z))$ is the probability estimate of the discriminator \mathcal{D} that the fake generated sample $\mathcal{G}(z)$ is real.

Generative adversarial networks have achieved remarkable success in various domains, including image synthesis, video generation, text generation, and more and have been applied to additional tasks such as image-to-image translation, style transfer, super-resolution, and data augmentation. The model structure described above is considered to be “vanilla” GAN. Since it is the most basic, the GAN model has been modified and extended to address specific challenges. For example, Wasserstein GAN [70] was developed to address unstable optimization and mode collapse; it shares the same minmax training procedure with the original GAN model but adjusts the loss function to minimize the Wasserstein distance between the real and fake data distributions. Conditional GAN (C-GAN) [71] adds conditioning information to both the generator and discriminator, which allows the model to generate data based on the chosen set of conditional parameters. Deep Convolutional GAN (DC-GAN) [72] uses convolutional layers in both generator and discriminator in order to generate more realistic images. Self-attention GAN (SA-GAN) [73] adds the self-attention mechanism to the convolutional GAN in order to model long-range, multilevel dependencies across image regions in order to avoid using only spatially local properties for generating high-resolution images. In Bidirectional GAN (Bi-GAN) [74], an encoder is added to the generator and discriminator in order to map the data into a latent space, enabling learning of a bidirectional mapping between the data space and the latent space. Finally, Cycle-consistent GAN (CycleGAN) involves training two GANs simultaneously using cycle consistency loss in order to encourage learned mappings in both directions and optimize image-to-image translation [75].

4.3 GANs for anomaly detection

Generative adversarial networks have proven to be effective as an unsupervised anomaly detection technique and have overcome significant challenges that are common to medical datasets such as a lack of adequately labeled datasets, dearth of anomalous data, and unbalanced datasets [28–31]. The following is a brief summary of the major models, which constitute the core models that have been adapted and modified for different applications.

4.3.1 AnoGAN

The AnoGAN model [76] was one of the earliest uses of GANs for anomaly detection. It uses a deep convolutional GAN (DC-GAN) architecture (see above) trained on normal data. Once the model is trained, anomaly scoring for a new instance is performed by calculating the anomaly score as a discrepancy between the instance and its reconstructed version obtained from the latent (random) space of the GAN. To reconstruct an instance, AnoGAN employs an optimization process that finds the closest point to it in the latent space by minimizing the difference between the reconstructed output and the original input using gradient descent and updating the latent code iteratively until convergence. The anomaly score is then calculated based on the difference between the original instance and the reconstructed output. A problem with this approach is that the GAN only implicitly models the data distribution and the optimization procedure for recovering the latent representation of a given sample is computationally costly and not practical for large datasets [29, 31]. The same authors followed up [77] with a modified model based on Wasserstein GAN, fast unsupervised anomaly detection with generative adversarial networks (f-AnoGAN), which substantially sped up the process of mapping to the latent space by moving from an iterative gradient descent approach to a learned mapping, which made the model more suitable for real-time anomaly detection.

4.3.2 Efficient GAN (E-GAN)

Efficient GAN [78] is based on AnoGAN, but it uses the bidirectional GAN (Bi-GAN) rather than a DC-GAN and incorporates an encoder into the architecture in order to alleviate the computational complexity associated with inference. Here, the discriminator separates two joint distributions: the given sample and the corresponding latent space (the output of the encoder) versus the original latent space and its generated synthetic sample (the output of the generator). The encoder acts as a regularization mechanism, helping to mitigate mode collapse and stabilizing the training process as well as resulting in significantly improved efficiency in detecting anomalies.

4.3.3 GANomaly

GANomaly [79] represents an addition of an autoencoder to AnoGAN in order to learn both the image and latent representations jointly. Here, the generator is constructed of an encoder and decoder, with an additional encoder that takes the output from the generated sample space and maps it back to a latent space. The discriminator as usual compares the generated samples to the original data. The total training loss then consists of the reconstructed loss in the latent space, the reconstructed loss in the sample space, and the adversarial loss in the sample space. The anomaly score is based on the encoder loss [31]. An important aspect of this model is that the space used for comparison is from the original sample space rather than depending on random sampling from a latent space as in other GAN-based models. This results in a model with high detection accuracy that has been adapted extensively to specific applications. Skip-GANomaly, for example, adds skip connections between the encoder and generator, leading to improved reconstruction accuracy and thus more accurate anomaly detection [29].

4.4 GANs for anomaly detection in time series

Generative adversarial networks offer a powerful framework for modeling and generating complex data distributions, making them well suited for capturing the intricate patterns and dynamics present in time series data. Several types of GAN architectures have been adapted for anomaly detection in time series. The following is a selected review of several important models.

4.4.1 *TAnoGAN*

Time Series Anomaly Detection with Generative Adversarial Networks (TAnoGAN) [80] is the simplest GAN-based model adapted specifically for time series. It is similar to the AnoGAN model, except that rather than using deep CNNs, both the generator and the discriminator are composed of LSTM layers in order to model temporality. A sliding window inputs real subsequences and the generator outputs simulated sequences of the same length, which are then compared by the discriminator using pointwise distance. A shortcoming of this model is that it is incapable of dealing with multivariate time series.

4.4.2 *MAD-GAN*

Multivariate Anomaly Detection GAN (MAD-GAN) [81] was designed specifically to address anomaly detection in multivariate time series. MAD-GAN also consists of LSTM layers in the generator and discriminator, and its training is similar to that of TAnoGAN. The model generates a residual loss based on the idea that the generator implicitly models the data distribution by learning to map it back into its latent space. During inference, this residual loss is combined with the standard discrimination loss to determine whether the sample is abnormal or not.

4.4.3 *TadGAN*

Time series anomaly detection using generative adversarial network (TadGAN) [82] also uses the LSTM structure in its generator and discriminator, but introduces an encoder in order to generate the latent space from which the generator produces synthetic data samples (rather than from a random latent space as usual). There are two discriminators, one for the samples and one for the latent space. The model uses Wasserstein distance for the discriminator loss and cycle consistency loss for the reconstruction error, which are computed through a combination of pointwise difference, area difference, and dynamic time warping. Anomalies are detected through a combination of reconstruction and discriminator errors.

4.4.4 *Beat-GAN*

Beat-GAN [59] was specifically developed for ECG anomaly detection. It utilizes an encoder-decoder architecture as the generator, and both the encoder and decoder are CNNs rather than LSTMs. They deal with temporality by utilizing a one-dimensional filter sliding along the temporal dimension. An interesting contribution that this model uses in order to deal with the inherent nonstationarity of their data (heart rate variability) is a modified form of time warping, where data are imputed during decelerations and removed during accelerations in order to generate a steady

“beat.” The model functions essentially like an autoencoder, but uses the discriminator for regularization to improve stability. During inference, anomalies are detected by a combination of pairwise reconstruction error and discriminator error.

5. GANs in medicine

5.1 GANs in medical imaging

Due to their capacity to generate realistic images, it is not surprising that the most common use of GANs in medical imaging has been primarily in data synthesis [83–85]. While some of these uses do not directly involve anomaly detection, given that they are in some way involved in detecting and diagnosing specific disease or behavioral states, they are at least obliquely related [21, 28]. The following are some of the most common current applications.

5.1.1 Data augmentation

As previously discussed, scarcity of labeled data represents one of the main limitations to the application of deep learning in medicine [86, 87]. Medical datasets are often imbalanced and lack diversity, which can lead to biased or poor-performing models. GANs have been shown to be able to generate synthetic medical data, helping to augment existing datasets, rectify imbalance, increase diversity, and improve the performance of machine learning models. A key advantage of being able to generate synthetic data with the same statistical characteristics as the original data but without personal health information is the ability to widely share and analyze data without the risk of violating patient privacy, which is often a barrier to producing large public datasets.

From an anomaly detection perspective, data synthesis can be used to turn an unsupervised or semi-supervised anomaly problem into a supervised binary (or multiclass) classification problem: GANs are used to generate synthetic data that statistically resemble anomalies, which can then be used for balanced classification training in a supervised manner. This is, in fact, the most common current application of GANs in anomaly detection [21, 29, 30] and has been applied to images [83, 84] as well as time series such as ECG [88] and EEG [89]. While this approach to anomaly recognition is more straightforward than classical anomaly detection, it assumes that the known anomalous data cover the entire distribution of possible anomalies, which may not be the case and could result in excellent classification of known anomalies but possible misclassification of unknown ones [39, 85].

5.1.2 Image-to-image translation

A powerful application specific to GANs is their ability for image-to-image translation such as converting MRI images to computed tomography (CT) images or vice versa, enhancing the quality of medical images, or generating angiography images from MRI images. There are multiple potential benefits to this, including decreased need for multimodal studies, reduced acquisition time and radiation exposure, and increased availability of appropriate imaging in cases where there is limited access to multimodal imaging, e.g., in a clinical scenario where a certain imaging modality (such as MRI) might be optimal for diagnosis or treatment planning but only another

modality (such as CT) is available to the patient. The predominant type of GAN used for this application is CycleGAN [21, 90–93].

5.1.3 Super-resolution/denoising

Generative adversarial networks can also be used to increase the resolution of medical images, known as super-resolution, and to reduce noise and artifacts. Since the quality of medical images is often degraded by various factors such as hardware limitation or patient movement, this can be extremely helpful for discernment of fine details that can be critical for correct diagnosis. Super-resolution and noise reduction are types of image-to-image translation that involve converting low-resolution images into high-resolution images by imputing data. This is especially interesting in cases where the imaging modality is intrinsically low-resolution, such as positron emission tomography (PET) [94–99].

5.1.4 Image segmentation

Image segmentation is important for measuring and visualizing anatomical structures, delineating pathological regions, and for surgical planning and image-guided interventions. The process of applying GANs to image segmentation is slightly different than that of applying vanilla GAN: the generator now aims to create an image where each pixel corresponds to a particular class label and the discriminator attempts to differentiate between the ground-truth segmentation (real) and the generator's segmentation (synthetic). This is again a type of image-to-image translation for which GANs have been used successfully [100–102].

5.1.5 Disease progression modeling

Disease progression modeling involves predicting how a disease will develop in a patient over time, which can allow for early intervention, optimized treatment plans, and better patient outcomes. In the context of disease progression modeling, the generator could be conditioned on a particular disease stage or on past medical history in order to generate synthetic data predicting what could potentially happen at the next disease stage or time point. Different models or approaches might be used to handle different kinds of data (continuous, discrete, etc.) and different diseases. GANs have been successfully applied to tumor growth [103] and Alzheimer's disease prediction [104].

5.1.6 Brain decoding

Brain decoding involves using machine learning algorithms to map patterns of brain activity (measured via EEG or fMRI) to mental states or processes. For example, using visual image reconstruction decoding researchers have been able to reproduce images a person is viewing directly from brain activity [105–107]. This process is often referred to as “mind reading” or “brain-to-image reconstruction”. The most direct clinical applicability of this technique at this time is in brain-computer interfaces, which would potentially allow disabled and paralyzed individuals to communicate and control external devices more easily. The technique also offers the potential for significant advancements in the understanding of the biology of consciousness and mind-body medicine.

5.2 GANs in medical time series: ECG

5.2.1 ECG overview

Electrocardiography (ECG or EKG) is a diagnostic tool that records the electrical activity of the heart over a period of time through electrodes placed on the skin, typically in a standard 12-lead setup for a clinical ECG. The ECG waveform represents the electrical depolarization and repolarization of the cardiac muscle during each heartbeat and can provide a large amount of valuable information such as the heart rate, rhythm, and the size and position of the chambers. It can also show evidence of damage to the cardiac muscle (ischemia or infarction), effects of drugs or devices (such as a pacemaker), and other types of heart disease or conditions (e.g., pericarditis, electrolyte imbalances). A significant advantage of ECG is that it is noninvasive, inexpensive, and relatively quick to perform. However, it requires expert interpretation, and while it is highly valuable, it may not provide a definitive diagnosis on its own and may need to be combined with other studies [108, 109].

5.2.2 GANs in ECG anomaly detection

As discussed above, Beat-GAN was specifically designed to detect anomalies in ECG. It outperformed other anomaly detection methods (including OCSVM) and achieved high accuracy and fast inference. It was also to some extent interpretable since it was able to pinpoint anomalies in sample space. The model has also been applied successfully to time series in other domains [59]. Shin et al. [110] deployed the AnoGAN architecture, but modified it with extensive data preprocessing as well as dimensionality reduction with t-distributed stochastic neighbor embedding (t-SNE), which they utilized to generate an experimentally determined objective decision boundary that could effectively differentiate between normal signal and arrhythmia based on anomaly score.

Li et al. [111] proposed Single-Lead convolutional generative adversarial network (SLC-GAN) for automated myocardial infarction (MI) in single-lead ECG. The model involves a GAN with multiple convolutional layers (DC-GAN) in the generator and discriminator with an added CNN classifier for MI detection. The GAN portion learns to generate synthetic ECG data, which are then used to augment the volume of the training data for the classifier. The model achieved excellent classification accuracy and provides a good example of using synthetic data in order to change the problem parameters from unsupervised to supervised and improve performance. Xia et al. [112] extended this idea by applying a transformer model in the generator with a CNN discriminator. This model was then used to generate synthetic data, which were used to augment training of a classifier that combined a CNN-based feature extraction block followed by a bidirectional LSTM (Bi-LSTM) architecture. The overall model achieved superior performance and demonstrated improved classification than models that did not use added synthetic data. A similar conclusion was reached by Rath et al. [113], who tested several machine and deep learning methods and found that the best performance on ECG classification was achieved by a GAN-LSTM ensemble model.

Qin et al. [114] proposed ECG-ADGAN, a semi-supervised model that incorporates a Bi-LSTM network in addition to multiple 1D convolutional layers into the GAN generator in order to preserve temporal patterns of the ECG signal. Training takes place in two stages, with stage I following normal GAN training to Nash equilibrium

between the generator and discriminator and stage II freezing the generator and training the discriminator/classifier separately specifically for anomaly detection. The authors also utilized mini-batch training during stage I to improve convergence. The model demonstrated superior detection of unknown anomalies when compared to supervised learning methods.

Wang and colleagues [115] further extended this approach beyond binary classification (normal vs. abnormal) by incorporating GAN into a two-level hierarchical framework in order to not only detect but also classify different types of arrhythmias. The first level consists of a memory-augmented deep autoencoder with GAN (MadeGAN) designed to perform anomaly detection; the second level consists of a multibranching deep CNN architecture utilizing transfer learning to allow classification of different types of heart disease given the fundamental imbalance in the training dataset. This framework was able to effectively capture disease-altered features of ECG signals and accurately predict and classify heart disease with better performance compared to existing methods.

5.3 GANs in medical time series: EEG

5.3.1 EEG overview

An electroencephalogram (EEG) is a neuroimaging technique used to record the electrical activity of the brain. It is carried out using multiple electrodes placed on the scalp according to a standardized placement system, usually the 10–20 system. These electrodes measure voltage fluctuations resulting from ionic current flows within neural populations in the brain. The resulting traces, EEG waves, represent the summation of postsynaptic potentials (PSPs) from a large number of neurons, specifically from cortical pyramidal neurons, detected as fluctuations in voltage over time [116].

EEG waves are characterized by their frequency (measured in Hertz), amplitude (measured in microvolts), and waveform morphology. They are typically categorized into bandwidths: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (30–100 Hz), each of which may correspond to different states of brain activity or consciousness. For instance, alpha waves are typically associated with relaxed, closed-eye states, while beta waves are associated with active thinking, attentional focus, or rapid eye movement (REM) sleep [116].

EEG can be used for the detection and study of various neurological and psychiatric conditions such as epilepsy, sleep disorders, encephalopathies, and even cognitive processes. The main advantage of EEG in cognitive imaging is its high temporal resolution, which allows for the study of fast-dynamic processes within the brain. Its spatial resolution is limited, however, and it is less effective for capturing activity occurring deep within the brain. Anomaly detection in EEG traces involves all the challenges discussed above: multidimensional time series with complex dependencies, highly complex and nonstationary normal behavior, and rarity and extreme heterogeneity of disease patterns, from normal brain aging to active seizures.

5.3.2 GANs in EEG epilepsy detection

Epilepsy is a chronic neurological disorder marked by recurrent seizures, which are symptoms of abnormal excessive or synchronous neuronal activity in the brain. Seizures can vary greatly in their presentation, from minor sensory disturbances or momentary lapses in consciousness (“absence seizures”) to full-body convulsions

(“grand mal seizure”). EEG is crucial in diagnosing and localizing the disorder, as well as in detecting seizures when they occur. Seizures are characterized by a variety of EEG patterns and frequencies such as spike-and-wave discharges (70 ms waveforms often followed by a slow wave, with specificity based on the frequency band), sharp waves (typically seen in focal seizures), polyspikes (typically seen in generalized seizures), focal slowing (which helps with localization and can be seen before, during, or after focal seizures), or generalized paroxysmal fast activity (rapid continuous spiking typically seen in severe diseases such as Lennox-Gastaut syndrome). Seizure detection and monitoring has an important role in diagnosis, improving quality of life, and general understanding of the disease. For example, alerting a patient about an impending seizure might allow them to take appropriate safety precautions or breakthrough medications for seizure control [117–121]. In this context, automatic seizure detection or prediction essentially consists of a binary classification task between the ictal (seizure) or pre-ictal and nonictal (nonseizure) EEG patterns.

A number of machine and deep learning approaches have been applied to epileptic seizure detection and prediction, but these predominantly apply feature extraction such as wavelets or independent component analysis (ICA) followed by a classifier such as random forest or support vector machine (SVM) [122]. One of the only models to directly apply a semi-supervised GAN model to seizure detection was introduced by You et al. [119], who modified the AnoGAN model (DC-GAN architecture) for seizure detection in a behind-the-ear EEG two-channel signal. The EEG signal channels were filtered and transformed into spectrogram images, which were combined to form a virtual channel image that was fed to the network for training. The GAN was trained on normal data and anomalies were detected using a combination of residual and discriminative loss. The authors noted that the addition of a Gram matrix of the feature maps from each convolutional layer was shown to improve performance. The model demonstrated a 96.3% sensitivity for automated seizure detection.

Zhu and Shoaran [123] utilized an unsupervised adversarial model to map power spectrum features from intracranial EEG recordings into a subject-invariant feature space via domain transfer learning. The model consisted of an encoder and decoder functioning as a generator for both the source (labeled data) and target (unlabeled or new data) domains, with a discriminator designed to try to differentiate between the domains. A discriminative model was then trained on the resultant subject-invariant features to generate predictions about the target patients. The model demonstrated improved performance compared to the more conventional subject-specific approach, allowing for better generalization.

In contrast, Truong et al. [124] used a GAN to extract features from the EEG data that could then undergo binary classification. The generator was trained to synthesize realistic short-time Fourier transform images from a noise vector that were then passed through the discriminator. After training, the discriminator was able to collect and flatten features that could be used with any generic classifier, in this case two fully connected layers. The model again consisted of a modified DC-GAN architecture that was trained in an unsupervised manner where information regarding seizure onset was disregarded.

While there are very few applications of GANs directly to seizure detection, most applications use GANs in order to produce balanced training datasets with generated synthetic data, again effectively changing the problem from an unsupervised to a supervised one. One of the first applications of GANs to EEG data generation was proposed by Pascual et al. [125] who used their conditional GAN model to generate

synthetic ictal signals conditioned on interictal data from individual patients. The model consisted of a convolutional autoencoder as the generator, in this case translating the latent code into an ictal sample rather than restoring the original sample. The discriminator had the same architecture as the encoder of the generator and was trained to distinguish between real and fake ictal signals. The inclusion of synthetic samples resulted in improved classification performance compared to training on only real samples. The additional benefit of the synthetic procedure, as the authors point out, was deidentification of the original data and significant improvement in data privacy. Multiple additional groups have applied similar data augmentation approaches with various modifications, including different feature extraction methods, different generator and discriminator architectures, different loss functions, utilization of LSTM/GRU cells or attention instead of CNNs, and the application of different classifiers [126–143].

6. Conclusion: challenges and future directions

Although GANs have shown significant promise in anomaly detection in medical time series, serious challenges remain, such as (notorious) training instability, lack of clear evaluation metrics for generated data, limited interpretability, inability to explicitly model causal relationships, question of preservation of temporal dynamics, privacy concerns, and ethical considerations in the use of generated data. For example, evaluating the performance of GAN-based anomaly detection in time series is nontrivial since traditional evaluation metrics such as precision, recall, or F1-score or even more advanced techniques such as maximum mean discrepancy or Fréchet inception distance [144, 145] may be unable to fully capture the highly complex characteristics of the data. Additional issues such as fairness in AI models also need to be considered to ensure that algorithmic decisions do not discriminate against certain demographic groups.

Future research efforts should focus on refining training techniques, incorporating domain knowledge, and developing hybrid approaches to enhance the performance and applicability of GANs in anomaly detection in time series data. Since GANs were originally designed to generate realistic images, their use is still predominantly focused on the generation of synthetic data. This is, of course, extremely useful as it can help resolve issues of unbalanced datasets and lack of anomalous and labeled data and transform unsupervised or semi-supervised approaches to anomaly detection into more manageable supervised problems. Direct anomaly detection with GANs is becoming more common, however, as different research groups realize that applying generative deep learning to data directly is feasible and provides significant advantages (and since generation of synthetic data inherently biases the network away from correct classification or detection of unknown or unusual anomalies). While GANs have been applied successfully, at least initially, in anomaly detection in time series such as ECG and EEG, there are many other areas where they could be extremely useful, such as rs-fMRI, ICU physiological data, or even monitoring of medical records for the possibility of subtle signs of chronic disease. In these areas, the application would leverage the model's ability to learn the underlying distribution, and modifications such as adding or embedding different modules (such as autoencoders) into the adversarial structure and/or including recurrence, attention, or graph structures could serve to better model the long-range spatiotemporal dynamics and dependencies in the data. Additional potential areas of future research include building patient-specific

models, improving generality through transfer learning, and real-time applications in the monitoring of healthcare data. Finally, there is a need for more research on the practical deployment of GANs in clinical settings, which involves not only technical considerations but also evaluation in terms of clinical outcomes, cost-effectiveness, user experience, and ethical, legal, and societal implications.

IntechOpen

IntechOpen


Author details

Miloš Cekić

University of California, Los Angeles, Los Angeles, California, USA

*Address all correspondence to: mcekic@mednet.ucla.edu

IntechOpen

© 2023 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, et al. Opportunities and obstacles for deep learning in biology and medicine. *Journal of the Royal Society Interface*. 2018;**15**:2017038. DOI: 10.1098/rsif.2017.0387
- [2] Roy S, Meena T, Lim SJ. Demystifying supervised learning in healthcare 4.0: A new reality of transforming diagnostic medicine. *Diagnostics*. 2022;**12**:2549. DOI: 10.3390/diagnostics12102549
- [3] Kaushik S, Choudhury A, Sheron PK, Dasgupta N, Natarajan S, Pickett LA, et al. AI in healthcare: Time-series forecasting using statistical, neural, and ensemble architectures. *Frontiers in Big Data*. 2020;**3**:4. DOI: 10.3389/fdata.2020.00004
- [4] Wang WK, Chen I, Hershkovich L, Yang J, Shetty A, Singh G, et al. A systematic review of time series classification techniques used in biomedical applications. *Sensors*. 2022;**22**(20):8016. DOI: 10.3390/s22208016
- [5] Kline A, Wang H, Li Y, Dennis S, Hutch M, Xu Z, et al. Multimodal machine learning in precision health: A scoping review. *NPJ Digital Medicine*. 2022;**5**:171. DOI: 10.1038/s41746-022-00712-8
- [6] Fernando T, Gammulle H, Denman S, Sridharan S, Fookes C. Deep learning for medical anomaly detection—A survey. *ACM Computing Surveys*. 2021;**54**(7):141. DOI: 10.1145/3464423
- [7] Tschuchnig ME, Gadermayr M. Anomaly detection in medical imaging—A mini review. *arXiv. arXiv preprint arXiv:2108.11986*. 2021. DOI: 10.48550/arXiv.2108.11986
- [8] Samariya D, Ma J. Anomaly detection on health data. In: Traina A, Wang H, Zhang Y, Siuly S, Zhou R, Chen L, editors. *Health Information Science (HIS 2022)*. Cham, Switzerland: Springer Nature; 2022; LNCS, (13705):34-41. DOI: 10.1007/978-3-031-20627-6_4
- [9] Chalapathy R, Chawla S. Deep learning for anomaly detection: A survey. *arXiv. arXiv preprint arXiv:1901.03407*. 2019. DOI: 10.48550/arXiv.1901.03407
- [10] Pang G, Shen C, Cao L, Van Den Hengel A. Deep learning for anomaly detection: A review. *ACM Computing Surveys*. 2022;**54**(2):1-38. DOI: 10.1145/3439950
- [11] Li G, Jung JJ. Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges. *Information Fusion*. 2023;**91**:93-102. DOI: 10.1016/j.inffus.2022.10.008
- [12] Gui J, Sun Z, Wen Y, Tao D, Ye J. A review on generative adversarial networks: Algorithms, theory, and applications. *IEEE Transactions on Knowledge and Data Engineering*. 2023;**35**(4):3313-3332. DOI: 10.1109/TKDE.2021.3130191
- [13] Aggarwal A, Mittal M, Battineni G. Generative adversarial network: An overview of theory and applications. *International Journal of Information Management, Data Insights*. 2021;**1**:100004. DOI: 10.1016/j.jjime.2020.100004
- [14] Dash A, Ye J, Wang G. A review of generative adversarial networks (GANs) and its applications in a wide variety of disciplines—from medical to remote sensing. *International Journal of Applied Earth Observation and Geoinformation*.

2021;**108**:102734. DOI: 10.48550/arXiv.2110.01442

[15] Jabbar A, Li X, Omar B. A survey on generative adversarial networks: Variants, applications, and training. *ACM Computing Surveys (CSUR)*. 2021; **54**(8):1-49. DOI: 10.1145/3463475

[16] Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: A review. *Medical Image Analysis*. 2019; **2019**(58):101552. DOI: 10.1016/j.media.2019.101552

[17] Koshino K, Werner RA, Pomper MG, Bundschuh RA, Toriumi F, Higuchi T, et al. Narrative review of generative adversarial networks in medical and molecular imaging. *The Annals of Translational Medicine*. 2021;**9**(9):821. DOI: 10.21037/atm-20-6325

[18] Lan L, You L, Zhang Z, Fan Z, Zhao W, Zeng N, et al. Generative adversarial networks and its applications in biomedical informatics. *Frontiers in Public Health*. 2020;**8**:164. DOI: 10.3389/fpubh.2020.00164

[19] Kazemina S, Baur C, Kuijper A, van Ginneken B, Navab N, Albarqouni S, et al. GANs for medical image analysis. *Artificial Intelligence in Medicine*. 2020; **109**:101938. DOI: 10.1016/j.artmed.2020.101938

[20] Iqbal A, Sharif M, Yasmin M, Raza M, Aftab S. Generative adversarial networks and its applications in the biomedical image segmentation: A comprehensive survey. *The International Journal of Multimedia Information Retrieval*. 2022;**11**:333-368. DOI: 10.1007/s13735-022-00240-x

[21] Laino ME, Cancian P, Politi LS, Della Porta MG, Saba L, Savevski V. Generative adversarial networks in brain

imaging: A narrative review. *Journal of Imaging*. 2022;**8**:83. DOI: 10.3390/jimaging8040083

[22] Soomro TA, Zheng L, Afifi AJ, Ali A, Soomro S, Yin M, et al. Image segmentation for MR brain tumor detection using machine learning: A review. *IEEE Reviews in Biomedical Engineering*. 2023;**16**:70-90. DOI: 10.1109/RBME.2022.3185292

[23] Krithika M, Suganthi K. Review of medical image synthesis using GAN techniques. *ITM Web of Conferences*. 2021;**37**:01005. DOI: 10.1051/itmconf/20213701005

[24] Ali H, Biswas R, Mohsen F, Shah U, Alamgir A, Mousa O, et al. The role of generative adversarial networks in brain MRI: A scoping review. *Insights Into Imaging*. 2022;**13**:98. DOI: 10.1186/s13244-022-01237-0

[25] Jeong JJ, Tariq A, Adejumo T, Trivedi H, Gichoya JW, Banerjee I. Systematic review of generative adversarial networks (GANs) for medical image classification and segmentation. *Journal of Digital Imaging*. 2022;**35**:137-152. DOI: 10.1007/s10278-021-00556-w

[26] Yahaya MSM, Teo J. Data augmentation using generative adversarial networks for images and biomarkers in medicine and neuroscience. *Frontiers in Applied Mathematics and Statistics*. 2023;**9**:1162760. DOI: 10.3389/fams.2023.1162760

[27] Sabuhi M, Zhou M, Bezemer CP, Musilek P. Applications of generative adversarial models in anomaly detection: A systematic literature review. *IEEE Access*. 2021;**9**:161003-161029. DOI: 10.1109/ACCESS.2021.3131949

- [28] Wang R, Bashyam V, Yang Z, Yu F, Tassopoulou V, Chitapalli SS, et al. Applications of generative adversarial networks in neuroimaging and clinical neuroscience. *NeuroImage*. 2023;**269**: 119898. DOI: 10.1016/j.neuroimage.2023.119898
- [29] Li H, Li Y. Anomaly detection based on GAN: A survey. *Applied Intelligence*. 2023;**53**:8209-8231. DOI: 10.1007/s10489-022-03905-6
- [30] Di Mattia F, Galeone P, De Simoni M, Ghelfi E. A survey on GANs for anomaly detection. *arXiv preprint arXiv:1906.11632*. 2019. DOI: 10.48550/arXiv.1906.11632
- [31] Esmaeili M, Toosi A, Roshanpoor A, Changizi V, Ghazisaeedi M, Rahmim A, et al. Generative adversarial networks for anomaly detection in biomedical imaging: A study on seven medical image datasets. *IEEE Access*. 2023;**11**: 17906. DOI: 10.1109/ACCESS.2023.3244741
- [32] Chen X, Konukoglu E. Unsupervised abnormality detection in medical images with deep generative methods. In: *Biomedical Image Synthesis and Simulation: Methods and Applications*. London, UK; Academic Press; 2022. DOI: 10.1016/B978-0-12-824349-7.00022-0
- [33] Choi K, Yi J, Park C, Yoon S. Deep learning for anomaly detection in time-series data: Review, analysis, and guidelines. *IEEE Access*. 2021;**9**:120043. DOI: 10.1109/ACCESS.2021.3107975
- [34] Hamilton JD. *Time Series Analysis*. Princeton, NJ, USA: Princeton University Press; 2020. DOI: 10.1515/9780691218632
- [35] Mills T. *Applied Time Series Analysis: A Practical Guide to Modeling and Forecasting*. London, UK: Academic Press; 2019. DOI: 10.1016/B978-0-12-813117-6.00001-6
- [36] Shumway RH, Stoffer DS. *Time Series Analysis and its Applications*. 4th ed. New York, NY, USA: Springer; 2017. DOI: 10.1007/978-3-319-52452-8
- [37] Chandola V, Banerjee A, Kumar V. Anomaly detection for discrete sequences: A survey. *IEEE Transactions on Knowledge and Data Engineering*. 2012;**24**(5):823-839. DOI: 10.1109/TKDE.2010.235
- [38] Blazquez-Garcia A, Conde A, Mori U, Lozano JA. A review on outlier/anomaly detection in time series data. *ACM Computing Surveys*. 2021;**54**(3): 1-33. DOI: 10.1145/34444690
- [39] Darban AA, Webb GI, Pan S, Aggarwal CC, Salehi M. Deep learning for time series anomaly detection: A survey. *arXiv preprint arXiv:2211.05244*. 2022. DOI: 10.48550/arXiv.2211.05244
- [40] Brophy E, Wang Z, She Q, Ward T. Generative adversarial networks in time series: A systematic literature review. *ACM Computing Surveys*. 2023;**55**(10): 199. DOI: 10.1145.3559540
- [41] Hawkins DM. *Identification of Outliers*. London, UK: Springer Netherlands; 1980. DOI: 10.1007/978-94-015-3994-4
- [42] Ranga, Suri NNR, Murty N, Athithan MG. *Outlier Detection: Techniques and Applications*. New York, NY, USA: Springer; 2019. DOI: 10.1007/978-3-030-05127-3
- [43] Aggarwal CC. *Outlier Analysis*. 2nd ed. New York, NY, USA: Springer; 2017. DOI: 10.1007/978-3-319-47578-3

- [44] Munir M, Chattha MA, Dengel A, Ahmed S. A comparative analysis of traditional and deep learning-based anomaly detection methods for streaming data. In: 18th IEEE International Conference on Machine Learning and Applications (ICMLA). Los Alamitos, CA, USA: IEEE Computer Society Conference Publishing Services; 2019. pp. 561-566. DOI: 10.1109/ICMLA.2019.00105
- [45] Ruff L, Kauffmann JR, Vandermeulen RA, Montavon G, Samek W, Kloft M, et al. A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*. 2021; **109**(5):756-795. DOI: 10.1109/JPROC.2021.3052449
- [46] Rewicki F, Denzler J, Niebling J. Is it worth it? Comparing six deep and classical methods for unsupervised anomaly detection in time series. *Applied Sciences*. 2023;**13**(3):1778. DOI: 10.3390/app13031778
- [47] Audibert J, Michiardi P, Guyard F, Marti S, Zuluaga MA. Do deep neural networks contribute to multivariate time series anomaly detection? *Pattern Recognition*. 2022;**132**:108945. DOI: 10.1016/j.patcog.2022.108945
- [48] Baid U et al. The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification. *arXiv preprint arXiv:2107.02314*. 2020. DOI: 10.48550/arXiv.2107.02314
- [49] Petzsche MRH et al. ISLES 2022: a multi-center magnetic resonance imaging stroke lesion segmentation dataset. *Scientific Data*. 2022;**9**:762. DOI: 10.1038/s41597-022-01875-5
- [50] Hsu CY, Liu WC. Multiple time-series convolutional neural network for fault detection and diagnosis and empirical study in semiconductor manufacturing. *Journal of Intelligent Manufacturing*. 2020;**32**:1-14. DOI: 10.1007/s10845-020-01591-0
- [51] Bao Y, Tang Z, Li H, Zhang Y. Computer vision and deep learning-based data anomaly detection method for structural health monitoring. *Structural Health Monitoring*. 2019; **18**(2):401-421. DOI: 10.1177/1475921718757405
- [52] Tang Z, Chen Z, Bao Y, Li H. Convolutional neural network-based data anomaly detection method using multiple information for structural health monitoring. *Structural Control and Health Monitoring*. 2019;**26**(1): e2296. DOI: 10.1002/stc.2296
- [53] Chung J, Gulcehre C, Cho KH, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*. 2014; *arXiv*. DOI: 10.48550/arXiv.1412.3555
- [54] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*. 1994; **5**(2):157-166. DOI: 10.1109/72.279181
- [55] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*. 1997;**9**(8):1735-1780. DOI: 10.1162/neco.1997.9.8
- [56] Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*. 2014. DOI: 10.48550/arXiv.1406.1078
- [57] Choi Y, Lim H, Choi H, Kim IJ. GAN-based anomaly detection and

- localization of multivariate time series data for power plant. In: Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp). Los Alamitos, CA, USA: IEEE Computer Society Conference Publishing Services; 2020. pp. 71-74. DOI: 10.1109/BigComp48618.2020.00-97
- [58] Wen T, Keyes R. Time series anomaly detection using convolutional neural networks and transfer learning. arXiv. arXiv preprint arXiv:1905.13628. 2019. DOI: 10.48550/arXiv.1905.13628
- [59] Zhou B, Liu S, Hooi B, Cheng X, Ye J. BeatGAN: Anomalous rhythm detection using adversarially generated time series. In: Proc. 28th Int. Joint Conf. Artif. Intell. Menlo Park, CA, USA: AAAI Press; 2019. pp. 4433-4439. DOI: 10.24963/ijcai.2019/616
- [60] Lea C, Vidal R, Reiter A, Hager GD. Temporal convolutional networks: A unified approach to action segmentation. arXiv. arXiv preprint arXiv:1611.05267. 2016. DOI: 10.48550/arXiv.1611.05267
- [61] Mamandipoor B, Majd M, Sheikhalishahi S, Modena C, Osmani V. Monitoring and detecting faults in wastewater treatment plants using deep learning. Environmental Monitoring and Assessment. 2020;**192**(2):1-12. DOI: 10.1007/s10661-020-8064-1
- [62] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. arXiv. arXiv preprint arXiv:1706.03762. 2017. DOI: 10.48550/arXiv.1706.03762
- [63] Guo Y, Liao W, Wang Q, Yu L, Ji T, Li P. Multidimensional time series anomaly detection: A GRU-based Gaussian mixture variational autoencoder approach. In: Proceedings of the 10th Asian Conference on Machine Learning. Brookline, MA, USA: MLR Press/Microtome Publishing; 2018. pp. 97-112
- [64] Lee TJ, Gottschlich J, Tatbul N, Metcalf E, Zdonik S. Greenhouse: A zero-positive machine learning system for time-series anomaly detection. arXiv. arXiv preprint arXiv:1801.03168. 2018. DOI: 10.48550/arXiv.1801.03168
- [65] Lu Z, Lv W, Xie Z, Du B, Xiong G, Sun L, et al. Graph sequence neural network with an attention mechanism for traffic speed prediction. ACM Transactions on Intelligent Systems and Technology (TIST). 2022, 2022;**13**(2): 1-24. DOI: 10.1145/3470889
- [66] Wu L, Cui P, Pei J, Zhao L, Song L. Graph neural networks. In: Wu L, Cui P, Pei J, Zhao L, editors. Graph Neural Networks: Foundations, Frontiers, and Applications. Singapore: Springer; 2022. DOI: 10.1007/978-981-16-6054-2_3
- [67] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv. arXiv preprint arXiv:1609.02907. 2016. DOI: 10.48550/arXiv.1609.02907
- [68] Veličković P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y. Graph attention networks. arXiv. arXiv preprint arXiv:1701.10903. 2017. DOI: 10.48550/arXiv.1710.10903
- [69] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. Neural Information Processing Systems. 2014;**2**:2672-2680. DOI: 10.3156/jsoft.29.5_177_2
- [70] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. Proceedings of the 34th international conference on machine learning. PMLR. 2017;**70**:214-223

- [71] Mirza M, Osindero S. Conditional generative adversarial nets. arXiv. arXiv preprint arXiv:1411.1784. 2014. DOI: 10.48550/arXiv.1411.1784
- [72] Radford A, Metz L, Chitala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv. arXiv preprint arXiv:1511.06434. 2015. DOI: 10.48550/arXiv.1511.06434
- [73] Zhang H, Goodfellow I, Metaxas D, Odena A. Self-attention generative adversarial networks. arXiv. arXiv preprint arXiv:1805.08318. 2018. DOI: 10.48550/arXiv.1805.08318
- [74] Donahue J, Krahenbuhl P, Darrell T. Adversarial feature learning. arXiv. arXiv preprint arXiv:1605.09782. 2016. DOI: 10.48550/arXiv.1605.09782
- [75] Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv. arXiv preprint arXiv:1703.10593. 2017. DOI: 10.48550/arXiv.1703.10593
- [76] Schlegl T, Seebock P, Waldstein SM, Schmidt-Erfurth U, Langs G. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: Information Processing in Medical Imaging, Lecture Notes in Computer Science. Cham: Springer; 2017. pp. 146-147. DOI: 10.1007/978-3-319-59050-9_12
- [77] Schlegl T, Seebock P, Waldstein SM, Langs G, Schmidt-Erfurth U. f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. Medical Image Analysis. 2019;54:30-44. DOI: 10.1016/j.media.2019.01.010
- [78] Zenati H, Foo CS, Lecouat B, Manek G, Chandrasekhar VR. Efficient GAN-based anomaly detection. arXiv. arXiv preprint arXiv:1902.03984. 2018, 2019. DOI: 10.48550/arXiv.1902.03984
- [79] Akcay S, Atapour-Abarghouei A, Breckon TP. GANomaly: Semi-supervised anomaly detection via adversarial training. In: Lecture Notes in Computer Science. Berlin, Germany: Springer; 2019. pp. 622-637. DOI: 10.1007/978-3-030-20893-6_39
- [80] Bashar MA, Nayak R. TANOAN: Time series anomaly detection with generative adversarial networks. In: 2020 IEEE Symposium Series on Computational Intelligence (SSCI). Piscataway, NJ, USA: IEEE Publishing; 2020. pp. 1778-1785. DOI: 10.1109/SSCI47803.2020.9308512
- [81] Li D, Chen D, Jin B, Shi L, Goh J, Ng SK. MADGAN: Multivariate anomaly detection for time series data with generative adversarial networks. In: Lecture Notes in Computer Science. Berlin, Germany: Springer; 2019. pp. 703-716. DOI: 10.1007/978-3-030-30490-4_56
- [82] Geiger A, Liu D, Alnegheimish S, Cuesta-Infante A, Veeramachaneni K. TadGAN: Time series anomaly detection using generative adversarial networks. In: Proceedings - 2020 IEEE International Conference on Big Data. Piscataway, NJ, USA: IEEE Publishing; 2020. pp. 33-43. DOI: 10.1109/bigdata50022.2020.9378139
- [83] Sorin V, Barash Y, Konen E, Klang E. Creating artificial images for radiology applications using generative adversarial networks (GANs) – A systematic review. Acta Radiologica. 2020;27:1175-1185. DOI: 10.1016/j.acra.2019.12.024
- [84] Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nature

Reviews. *Cancer*. 2018;**18**:500-510.
DOI: 10.1038/s41568-018-0016-5

Analysis. 2021;**70**:101944. DOI: 10.1016/j.media.2020.101944

[85] Festag S, Denzler J, Spreckelsen C. Generative adversarial networks for biomedical time series forecasting and imputation. *Journal of Biomedical Informatics*. 2022;**129**:104058. DOI: 10.1016/j.jbi.2022.104058

[92] Jin CB, Kim H, Liu M, Jung W, Joo S, Park E, et al. Deep CT to MR synthesis using paired and unpaired data. *Sensors*. 2019;**22**(19):2361. DOI: 10.3390/s19102361

[86] Islam J, Zhang Y. GAN-based synthetic brain PET image generation. *Brain Informatics*. 2020;**7**(1):3. DOI: 10.1186/s40708-020-00104-2

[93] Lan H, Alzheimer Disease Neuroimaging Initiative, Toga AW, Sepelband F. Three-dimensional self-attention conditional GAN with spectral normalization for multimodal neuroimaging synthesis. *Magnetic Resonance in Medicine*. 2021;**86**:1718-1733. DOI: 10.1002/mrm.28819

[87] Hirte AU, Platscher M, Joyce T, Heit JJ, Tranvinh E, Federau C. Realistic generation of diffusion-weighted magnetic resonance brain images with deep generative models. *Magnetic Resonance Imaging*. 2021;**81**:60-66. DOI: 10.1016/j.mri.2021.06.001

[94] Zhao K, Zhou L, Gao S, Wang X, Wang Y, Zhao X, et al. Study of low-dose PET image recovery using supervised learning with CycleGAN. *PLoS One*. 2020;**15**:e0238455. DOI: 10.1371/journal.pone.0238455

[88] Thambawita V et al. DeepFake electrocardiograms using generative adversarial networks are the beginning of the end for privacy issues in medicine. *Scientific Reports*. 2021;**11**:21896. DOI: 10.1038/s41598-021-01295-2

[95] Sundar LKS, Iommi D, Muzik O, Chalampalakis Z, Klebermass EV, Hienert M, et al. Conditional generative adversarial networks aided motion correction of dynamic 18F-FDG PET brain studies. *Journal of Nuclear Medicine*. 2021;**62**:871-880. DOI: 10.2967/jnumed.120.248856

[89] Lashgari E, Liang D, Maoz U. Data augmentation for deep-learning-based electroencephalography. *Journal of Neuroscience Methods*. 2020;**346**:108885. DOI: 10.1016/j.jneumeth.2020.108885

[96] Delannoy Q, Pham CH, Cazorla C, Tor-Díez C, Dollé G, Meunier H, et al. SegSRGAN: Super-resolution and segmentation using generative adversarial networks—Application to neonatal brain MRI. *Computers in Biology and Medicine*. 2020;**120**:103755. DOI: 10.1016/j.compbmed.2020.103755

[90] Cheng D, Qiu N, Zhao F, Mao Y, Li C. Research on the modality transfer method of brain imaging based on generative adversarial network. *Frontiers in Neuroscience*. 2021;**15**:655019. DOI: 10.3389/fnins.2021.655019

[97] Shaul R, David I, Shitrit O, Raviv TR. Subsampled brain MRI reconstruction by generative adversarial neural networks. *Medical Image Analysis*. 2020;**65**:101747. DOI: 10.1016/j.media.2020.101747

[91] Yurt M, Dar SU, Erdem A, Erdem E, Oguz KK, Çukur T. mustGAN: Multi-stream generative adversarial networks for MR image synthesis. *Medical Image*

- [98] An Y, Lam HK, Ling SH. Auto-denoising for EEG signals using generative adversarial network. *Sensors*. 2022;**22**:1750. DOI: 10.3390/s22051750
- [99] Wolterink JM, Leiner T, Viergever MA, Isgum I. Generative adversarial networks for noise reduction in low-dose CT. *IEEE Transactions on Medical Imaging*. 2017;**36**(12):2536-2545. DOI: 10.1109/TMI.2017.2708987
- [100] Sille R, Choudhury T, Sharma A, Chauhan P, Tomar R, Sharma D. A novel generative adversarial network-based approach for automated brain tumour segmentation. *Medicina*. 2023;**59**(1):119. DOI: 10.3390/medicina59010119
- [101] Yuan W, Wei J, Wang J, Ma Q, Tasdizen T. Unified generative adversarial networks for multimodal segmentation from unpaired 3D medical images. *Medical Image Analysis*. 2020; **64**:101731. DOI: 10.1016/j.media.2020.101731
- [102] Oh KT, Lee S, Lee H, Yun M, Yoo SK. Semantic segmentation of white matter in FDG-PET using generative adversarial network. *Journal of Digital Imaging*. 2020;**33**:816-825. DOI: 10.1007/s10278-020-00321-5
- [103] Elazab A, Wang C, Gardezi SJS, Bai H, Hu Q, Wang T, et al. GP-GAN: Brain tumor growth prediction using stacked 3D generative adversarial networks from longitudinal MR images. *Neural Networks*. 2020;**132**:321-332. DOI: 10.1016/j.neunet.2020.09.004
- [104] Han C, Rundo L, Murao K, Noguchi T, Shimahara Y, Milacski ZÁ, et al. MADGAN: Unsupervised medical anomaly detection GAN using multiple adjacent brain MRI slice reconstruction. *BMC Bioinformatics*. 2021;**22**(Suppl 2):31. DOI: 10.1186/s12859-020-03936-1
- [105] Ren Z, Li J, Xue X, Li X, Yang F, Jiao Z, et al. Reconstructing seen image from brain activity by visually-guided cognitive representation and adversarial learning. *NeuroImage*. 2021; **228**:117602. DOI: 10.1016/j.neuroimage.2020.117602
- [106] Huang W, Yan H, Wang C, Yang X, Li J, Zuo Z, et al. Deep natural image reconstruction from human brain activity based on conditional progressively growing generative adversarial networks. *Neuroscience Bulletin*. 2021;**37**:369-379. DOI: 10.1007/s12264-020-00613-4
- [107] Al-Tahan H, Mohsenzadeh Y. Reconstructing feedback representations in the ventral visual pathway with a generative adversarial autoencoder. *PLoS Computational Biology*. 2021;**17**: 1-19. DOI: 10.1371/journal.pcbi.1008775
- [108] Goldberger AL, Goldberger ZD, Shvilkin A. *Golberger's Clinical Electrocardiography: A Simplified Approach*. 9th ed. Philadelphia, PA, USA: Elsevier; 2017. DOI: 10.1016/C2014-0-03319-9
- [109] Skandarani Y, Lalande A, Afilalo J, Jodoin PM. Generative adversarial networks in cardiology. *The Canadian Journal of Cardiology*. 2022;**38**:196-203. DOI: 10.1016/j.cjca.2021.11.003
- [110] Shin DH, Park RC, Chung K. Decision boundary-based anomaly detection model using improved AnoGAN from ECG data. *IEEE Access*. 2020;**8**:108664-108674. DOI: 10.1109/ACCESS.2020.3000638
- [111] Li W, Tang YM, Yu KM, To S. SLC-GAN: An automated myocardial infarction detection model based on generative adversarial networks and convolutional neural networks with single-lead electrocardiogram synthesis.

Information Sciences. 2022;**589**:738-750.
DOI: 10.1016/j.ins.2021.12.083

[112] Xia Y, Xu Y, Chen P, Zhang J, Zhang Y. Generative adversarial network with transformer generator for boosting ECG classification. *Biomedical Signal Processing and Control*. 2023;**80**: 104276. DOI: 10.1016/j.bspc.2022.104276

[113] Rath A, Mishra D, Panda G, Satapathy SC. Heart disease detection using deep learning methods from imbalanced ECG samples. *Biomedical Signal Processing and Control*. 2021;**68**: 102820. DOI: 10.1016/j.bspc.2021.102820

[114] Qin J, Gao F, Wang Z, Wong DC, Zhao Z, Relton SD, et al. A novel temporal generative adversarial network for electrocardiography anomaly detection. *Artificial Intelligence in Medicine*. 2023;**136**:102489. DOI: 10.1016/j.artmed.2023.102489

[115] Wang Z, Stavrakis S, Yao B. Hierarchical deep learning with generative adversarial network for automatic cardiac diagnosis from ECG signals. *Computers in Biology and Medicine*. 2023;**155**:106641. DOI: 10.1016/j.combiomed.2023.106641

[116] Nunez PL, Srinivasan R. *Electric Fields of the Brain: The Neurophysics of EEG*. 2nd ed. Oxford, UK: Oxford University Press; 2006. DOI: 10.1093/acprof:oso/9780195050387.001.0001

[117] Habashi AG, Azab AM, Eldawlatly S, Aly GM. Generative adversarial networks in EEG analysis: An overview. *Journal of Neuro-engineering and Rehabilitation*. 2023;**20**:40. DOI: 10.1186/s12984-023-01169-w

[118] Wei Z, Zou J, Zhang J, Xu J. Automatic epileptic EEG detection using

convolutional neural network with improvements in time-domain. *Biomedical Signal Processing and Control*. 2019;**53**:101551. DOI: 10.1016/j.bspc.2019.04.028

[119] You S, Cho BH, Yook S, Kim JY, Shon YM, Seo DW, et al. Unsupervised automatic seizure detection for focal-onset seizures recorded with behind-the-ear EEG using an anomaly-detecting generative adversarial network. *Computer Methods and Programs in Biomedicine*. 2020;**193**:105472. DOI: 10.1016/j.cmpb.2020.105472

[120] Tomson T, Nashef L, Ryvlin P. Sudden unexpected death in epilepsy: Current knowledge and future directions. *Lancet Neurology*. 2008;**7**: 1021-1031. DOI: 10.1016/S1474-4422(08)70202-3

[121] Usman SM, Khalid S, Bashir Z. Epileptic seizure prediction using scalp electroencephalogram signals. *Biocybernetics and Biomedical Engineering*. 2021;**41**:211-220. DOI: 10.1016/j.bbe.2021.01.001

[122] Natsu M, Bachute M, Gite S, Kotecha K, Vidyarthi A. Review on epileptic seizure prediction: Machine learning and deep learning approaches. *Computational and Mathematical Methods in Medicine*. 2022;**2022**: 7751263. DOI: 10.1155/2022/7751263

[123] Zhu B, Shoaran M. Unsupervised domain adaptation for cross-subject few-shot neurological symptom detection. In: *International IEEE/EMBS Conference on Neural Engineering*. Piscataway, NJ, USA: IEEE Publishing; 2021. DOI: 10.1109/NER49283.2021.9441235

[124] Truong ND, Kuhlmann L, Bonyadi MR, Querlioz D, Zhou L, Kavehei O. Epileptic seizure forecasting

with generative adversarial networks. IEEE Access. 2019;7:143999-144009. DOI: 10.1109/ACCESS.2019.2944691

[125] Pascual D, Amirshahi A, Aminifar A, Atienza D, Ryvlin P, Wattenhofer R. EpilepsyGAN: Synthetic epileptic brain activities with privacy preservation. IEEE Transactions on Biomedical Engineering. 2021;68(8):2435-2446. DOI: 10.1109/TBME.2020.3042574

[126] Yin X, Han Y, Sun H, Xu Z, Yu H, Duan X. Multi-attention generative adversarial network for multivariate time series prediction. IEEE Access. 2021;9:57351-57363

[127] Usman SM, Khalid S, Bashir S. A deep learning based ensemble learning method for epileptic seizure prediction. Computers in Biology and Medicine. 2021;136:1104710. DOI: 10.1016/j.compbiomed.2021.104710

[128] Salazar A, Vergara L, Safont G. Generative adversarial networks and Markov random fields for oversampling very small training sets. Expert Systems with Applications. 2021;163:113819. DOI: 10.1016/j.eswa.2020.113819

[129] Yin X, Han Y, Xu Z, Liu J. VAECGAN: A generating framework for longterm prediction in multivariate time series. Cybersecurity. 2021;4:22. DOI: 10.1186/s42400-021-00090-w

[130] Rasheed K, Qadir J, O'Brien TJ, Kuhlmann L, Razi A. A generative model to synthesize EEG data for epileptic seizure prediction. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2020;29:2322-2332. DOI: 10.1109/TNSRE.2021.3125023

[131] Gang D, Alkhachroum A, Bicchi MAM, Jagged JR, Cajigas I,

Chen ZS. Deep learning for robust detection of interictal epileptiform discharges. Journal of Neural Engineering. 2021;18:056015. DOI: 10.1088/1741-2552/abf28e

[132] Luo TJ, Fan Y, Chen L, Guo G, Zhou C. EEG signal reconstruction using a generative adversarial network with Wasserstein distance and temporal-spatial-frequency loss. Frontiers in Neuroinformatics. 2020;14:15. DOI: 10.3389/fninf.2020.00015

[133] Wang J, Mu W, Wang A, Wang L, Han J, Wang P, et al. Generative adversarial networks for electroencephalogram signal analysis: A mini review. In: International Winter Conference on Brain Computer Interface (BCI). Piscataway, NJ, USA: IEEE Publishing; 2023. DOI: 10.1109/BCI57258.2023.10078666

[134] Handa P, Gupta E, Muskan S, Goel N. A review on software and hardware developments in automatic epilepsy diagnosis using EEG datasets. Expert Systems. 2023:e13374. DOI: 10.1111/exsy.13374

[135] Daoud H, Bayoumi M. Generative adversarial network based semi-supervised learning for epileptic focus localization. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Piscataway, NJ, USA: IEEE Publishing; 2021. DOI: 10.1109/BIBM52615.2021.9669695

[136] Dong Z, Zhou S. EEG-based seizure detection using generative model and deep learning. In: IEEE International Conference on E-Health and Bioengineering (EHB). Piscataway, NJ, USA: IEEE Publishing; 2022. DOI: 10.1109/EHB55594.2022.9991438

[137] Ganti B, Chaitanya G, Balamurugan S, Nagaraj N,

Balasubramanian K, Pati S. Time-series generative adversarial network approach of deep learning improves seizure detection from the human thalamic SEEG. *Frontiers in Neurology*. 2022;**13**:755094. DOI: 10.3389/fneur.2022.755094

[138] Xu M, Jie J, Zhou W, Zhou H, Jin S. Synthetic epileptic brain activities with TripleGAN. *Computational and Mathematical Methods in Medicine*. 2022;**2022**:2841228. DOI: 10.1155/2022/2841228

[139] Zhang X, Yao L, Dong M, Liu Z, Zhang Y, Li Y. Adversarial representation learning for robust patient-independent epileptic seizure detection. *IEEE Journal of Biomedical and Health Informatics*. 2020;**24**(10):2852-2859. DOI: 10.1109/JBHI.2020.2971610

[140] Boonyakitanont P, Lek-uthai A, Chomtho K, Songsiri J. A review of feature extraction and performance evaluation in epileptic seizure detection using EEG. *Biomedical Signal Processing and Control*. 2020;**57**:101702. DOI: 10.1016/j.bspc.2019.101702

[141] Cherian R, Kanaga EG. Theoretical and methodological analysis of EEG based seizure detection and prediction: An exhaustive review. *Journal of Neuroscience Methods*. 2022;**369**:109483. DOI: 10.1016/j.jneumeth.2022.109483

[142] Nafea MS, Ismail ZH. Supervised machine learning and deep learning techniques for epileptic seizure recognition using EEG signals—A systematic literature review. *Bioengineering*. 2022;**9**:781. DOI: 10.3390/bioengineering9120781

[143] Yuan J, Ran X, Liu K, Yao C, Yao Y, Wu H, et al. Machine learning applications on neuroimaging for

diagnosis and prognosis of epilepsy: A review. *Journal of Neuroscience Methods*. 2022;**368**:109441.

DOI: 10.1016/j.jneumeth.2021.109441

[144] Xu Q, Huang G, Yuan Y, Guo C, Sun Y, Wu F, et al. An empirical study on evaluation metrics of generative adversarial networks. *arXiv preprint arXiv:1806.07755*. 2018. DOI: 10.48550/arXiv.1806.07755

[145] Borji A. Pros and cons of GAN evaluation measures: New developments. *Computer Vision and Image Understanding*. 2022;**215**:103329. DOI: 10.1016/j.cviu.2021.103329