Deep Generative Models: The winning key for large and easily accessible ECG datasets?

Giuliana Monachino [a,b,*], Beatrice Zanchi [a,c], Luigi Fiorillo [a], Giulio Conte [d,e], Angelo Auricchio [d,e], Athina Tzovara [b,f], Francesca Dalia Faraci [a]

[a] Institute of Digital Technologies for Personalized Healthcare - MeDiTech, Department of Innovative Technologies, University of Applied Sciences and Arts of Southern Switzerland, Via la Santa 1, Lugano 6900, Switzerland

[b] Institute of Informatics, University of Bern, Neubrückstrasse 10, Bern 3012, Switzerland

[c] Department of Quantitative Biomedicine, University of Zurich, Schmelzbergstrasse 26, Zurich 8091, Switzerland

[d] Division of Cardiology, Fondazione Cardiocentro Ticino, Via Tesserete 48, Lugano 6900, Switzerland

[e] Centre for Computational Medicine in Cardiology, Faculty of Informatics, Università della Svizzera Italiana, Via la Santa 1, Lugano 6900, Switzerland

[f] Sleep Wake Epilepsy Center | NeuroTec, Department of Neurology, Inselspital, Bern University Hospital, University of Bern, Freiburgstrasse 16, Bern 3010, Switzerland

A R T I C L E   I N F O

A B S T R A C T

Large high-quality datasets are essential for building powerful artificial intelligence (AI) algorithms capable of supporting advancement in cardiac clinical research. However, researchers working with electrocardiogram (ECG) signals struggle to get access and/or to build one. The aim of the present work is to shed light on a potential solution to address the lack of large and easily accessible ECG datasets. Firstly, the main causes of such a lack are identified and examined. Afterward, the potentials and limitations of cardiac data generation via deep generative models (DGMs) are deeply analyzed. These very promising algorithms have been found capable not only of generating large quantities of ECG signals but also of supporting data anonymization processes, to simplify data sharing while respecting patients' privacy. Their application could help research progress and cooperation in the name of open science. However several aspects, such as a standardized synthetic data quality evaluation and algorithm stability, need to be further explored.

1. Introduction

The electrocardiogram (ECG) is a powerful, non-invasive, and widely used diagnostic tool to investigate the presence of electrophysiological alterations in cardiac electrical conduction.

In the last decades, AI has made significant advancements in the medical domain, particularly in the area of machine learning (ML) and, more specifically, deep learning (DL). The quality of the dataset exploited to train the models does have a significant influence on the final performance. A large and heterogeneous dataset is more likely to be representative of the underlying data distribution of the whole population of interest, enhancing model performance and generalizability. Moreover, a balanced dataset is desirable in model training as it prevents biases toward one class.

Despite the extensive recording of ECG worldwide, it is quite challenging to find large and good quality ECG databases openly shared among the scientific community. The reasons are related to the huge challenges that have to be faced for both dataset building and sharing.

Building a sufficiently large database requires time, resources, and often the coordination and collaboration of several specialized centers located in different countries with different and complex regulations. Moreover, the low incidence of certain pathologies and the clear difficulty in recording data related to paroxysmal events lead to the scarcity of certain categories of pathological data. The secondary usage of ECGs recorded and stored in clinical centers involves further complications related to data sharing regulations to ensure patients' privacy. Furthermore, an ECG signal can be considered as a human fingerprint, making its anonymization and sharing even more complex.

Generating synthetic ECG signals can be a potential solution to overcome the data scarcity problem. The very first attempts to generate synthetic ECG signals have been made already in the 90 s. Mathematical equations were exploited to model the cardiac signal. Some focused on the underlying physiological mechanism aiming at reproducing cardiac dynamics to simulate healthy or arrhythmic conditions [1–8].

---

* Corresponding author at: Institute of Digital Technologies for Personalized Healthcare - MeDiTech, Department of Innovative Technologies, University of Applied Sciences and Arts of Southern Switzerland, Via la Santa 1, Lugano 6900, Switzerland.

E-mail address: giuliana.monachino@supsi.ch (G. Monachino).

Other authors, instead, focused directly on reproducing the morphology and rhythmicity of ECG signals [9–18]. However, these approaches have been designed mainly for studying cardiac dynamics. Therefore, they are not the optimal solution to generate a large quantity of heterogeneous ECG signals.

An alternative approach, inherited from the computer vision field, has been proposed and applied to ECG signals [19–23] with data augmentation purposes. It is based on direct manipulation and slight distortion of the original data, through geometric transformations and filtering techniques. These methods have proven to be capable to enhance ML model performances by increasing the training set size. On the other hand, the variability they are able to produce with respect to the original data is quite limited.

Recently, the so-called deep generative models (DGMs) that were mainly exploited for image generation [24], have been applied also for time-series and ECG signals generation. These ML-based models are able to generate a large quantity of synthetic data, with greater variability, starting from a set of real data. The generated data share general characteristics with the source dataset but they are not directly linked to an individual patient recording.

The present work sheds light on the exploitation of DGMs as a tool for ECG data augmentation and anonymization, by investigating their potential and limitations, to overcome data scarcity issues and data sharing regulation complexity.

As shown in the graphical abstract in Fig. 1, we firstly analyze the reasons leading to the poor availability of ECG data, which are related to several challenges in dataset creation and data sharing procedures. Then, we introduce DGMs and the main categories of models that have been applied to ECG signals in the literature. Lastly, we analyze the application of this methodology in solving both the problems through data augmentation and data anonymization. The characteristics of the synthetic ECG generated and the criteria adopted for their evaluation are then presented. Conclusive remarks are then provided, emphasizing the strengths and limitations of DGMs application and suggesting future research.

The main contributions of this work can be summarized as follows:

- identifying and analyzing the typical challenges of ECG dataset building and sharing
- providing a comprehensive overview of DGMs for ECG signals generation. This analysis includes the type of DGMs (variational autoencoders, generative adversarial networks, and diffusion models), the practical employment of these models, the characteristics of the synthetic ECGs generated and the evaluation criteria proposed in the literature
- identifying the key potentialities and weaknesses of DGMs and suggesting potential usage of DGMs and other recent techniques and directions for future research.

2. Limited ECG data availability

AI researchers struggle in finding, accessing and/or building high quality ECG datasets that can be exploited to train their algorithms. An efficient training of a ML model is strictly related to the quality of the training set. An ideal dataset should be heterogeneous and large enough to be representative of the population of interest. Moreover, data balancing is needed to avoid class biases [25].

The open access ECG datasets that can be exploited to better characterize certain specific ECG pathologies are usually limited in numbers. A proper accurate labeling is often missing, and these datasets are usually highly unbalanced towards the healthy control group. Datasets that represent very specific pathologies, rare diseases, or uncommon arrhythmic conditions, are usually extremely small or even totally missing.

For instance, one of the most used databases for atrial fibrillation (AF) detection is AF Classification from a Short Single Lead ECG

Recording: The PhysioNet/Computing in Cardiology Challenge 2017 (CinC2017) [26,27]. Despite being a large dataset, it is highly unbalanced: AF is present in about 9% of the recordings only. Another example is the PTB Diagnostic ECG Database (PTBDB) [27,28], the most popular database for myocardial infarction (MI) detection: in this case, the MI group is around 3 times larger than the healthy control group leading to severe diagnosis bias.

ECG data can be collected from scratch, with new ad hoc measurements, or by applying for secondary usage of already existing recordings. In both cases, several challenges may be encountered and need to be faced.

2.1. Dataset building challenges

Dataset building refers to the process that includes data collection design, data acquisition, and data labeling procedures.

The main causes of data scarcity derive from challenges arising during data collection design and data acquisition. In fact, due to time, economic, organizational, and bureaucratic constraints, data collection often occurs within a single clinical center. This results in a restricted perspective and limited coverage of the population and hence, low dataset heterogeneity, leading to several possible biases (e.g. demographic, geographic).

A particularly critical case is the collection of ECG recordings from patients affected by rare diseases. According to the European Union, a *rare disease* is a medical condition that affects a small number of individuals in the population, specifically no more than 1 person in 2000 [29]. Examples of rare diseases in cardiac fields are most of the inherited arrhythmogenic diseases (IADs) like long QT syndrome (1:2000), Brugada syndrome (1:2000), short QT syndrome (2.7:100,000), catecholaminergic polymorphic ventricular tachycardia (1:10,000), and arrhythmogenic cardiomyopathy (1:5000). Due to the limited access to patient populations, the amount of cases that can be studied in a restricted geographic area is quite limited. To the best of our knowledge, for these rare diseases, no ECG datasets are publicly available to date. However, studying rare diseases is of great relevance since, despite their low prevalence, they are often associated with significant morbidity and mortality. The lack of understanding of rare diseases can lead to delayed diagnosis, misdiagnosis, and inadequate treatment. Undeniably, the only way of gathering knowledge from a sufficiently representative dataset of rare diseases implies collaboration among different hospitals and different countries and, therefore, data sharing.

Other issues need to be faced in case of certain pathological conditions that may be intermittent, e.g. paroxysmal AF, making it difficult to capture the abnormal rhythm on an ECG. In this case, the Holter monitor or other wearable devices need to be used to allow continuous recording and be able to identify intermittent events. On the other hand, data quality may be limited when compared to the standard 12-leads ECG recording.

Clearly, in case of life-threatening conditions (e.g. during ventricular fibrillation (VF)) data acquisition problems are further amplified since the priority is obviously on saving the patient's life, not on collecting ECG data. Indeed, only a few open-access databases (i.e. MIT-BIH Arrhythmia Database (MITDB) [27,30], MIT-BIH Malignant Ventricular Ectopy Arrhythmia Database (VFDB) [27,31] and Creighton University Ventricular Tachyarrhythmia Database (CUDB) [27,32]) with ventricular tachycardia or fibrillation events are available. All of them are small in size containing 48, 22, and 35 short ECG recordings, respectively, and an insufficient number of critical arrhythmias. As pointed out also by [33], this hinders the improvement of algorithms for solving tasks related to these arrhythmias.

Another arduous process, which is essential for supervised ML approaches, is data labeling. ECG signal needs to be opportunely labeled, i.e. specific ECG clinical patterns need to be properly identified. This
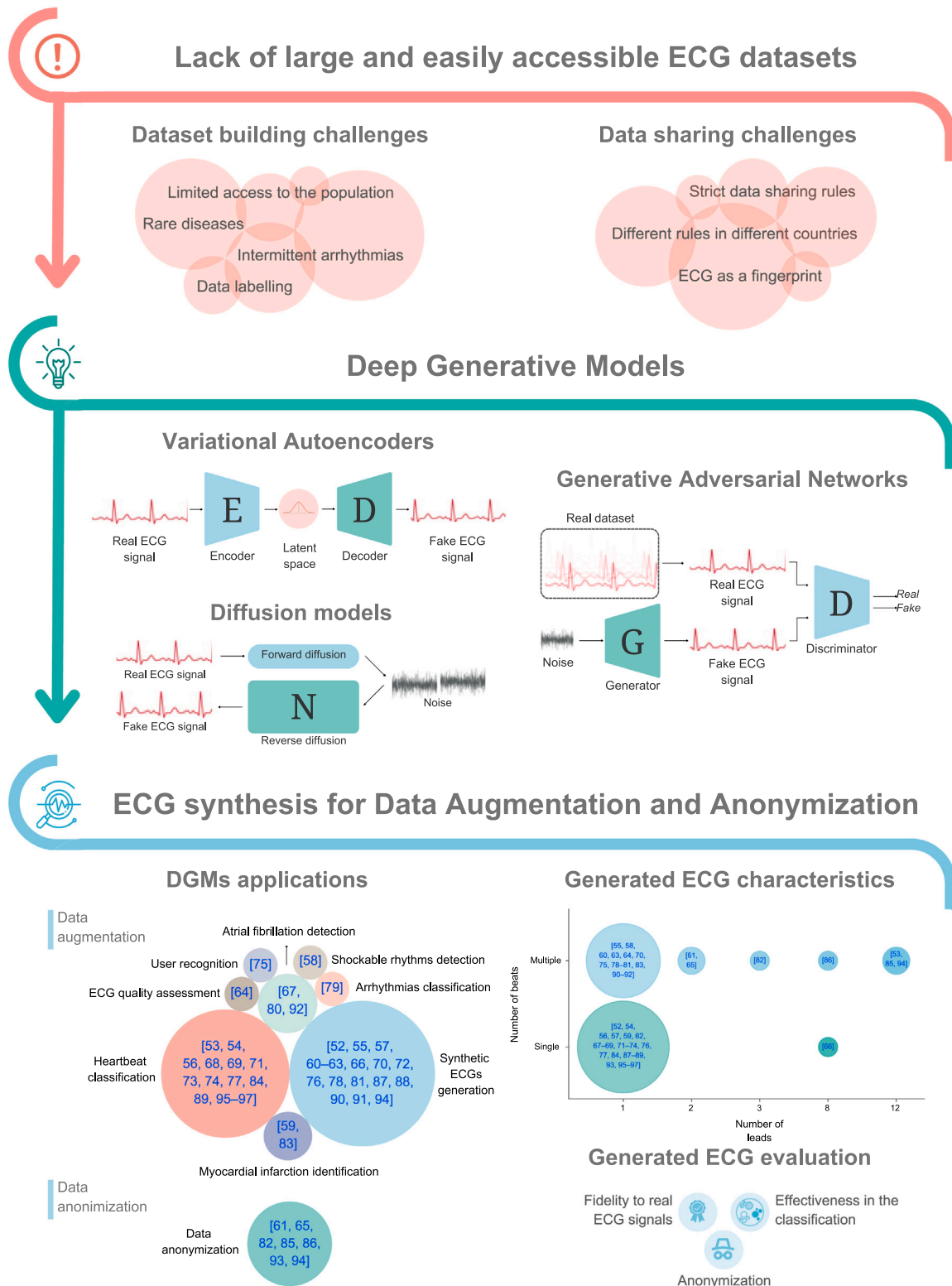
Fig. 1. Graphical summary. First, we deeply analyze the problem of low data availability, then we investigate the DGMs as a tool for ECG synthesis and, finally, we explore and evaluate the potential and limitations of the application of DGMs in solving this specific problem.

process is time-consuming and highly prone to human errors. Inaccurate labeling can lead to sub-optimal algorithm performances due to the erroneous knowledge provided during the training phase. Moreover, for certain less common pathological conditions, e.g. genetic diseases, there is a need for specific expertise in interpreting specific ECG patterns, that a general cardiologist may not have. The time-consuming

nature of this procedure is another huge issue, especially in the case of secondary usage of ECG recordings collected in the past by clinical centers. Indeed, these data are usually not labeled because it is not necessary for clinical practice. Therefore, all these ECG recordings need to be labeled a posteriori, requiring a considerable amount of time and increasing the likelihood of errors. Clearly, the problem is

further amplified in the case of long ECG recording (e.g. 24 h Holter recording).

Softwares for automatic labeling (e.g. the ones that are embedded in certain ECG recording devices) are in principle available but they cannot completely replace clinicians' expertise. To guarantee good quality labels, these softwares should be used only as a support and the clinician should verify the correctness of the assigned labels.

Depending on the specific clinical task, different types of labels may be needed: heartbeats and rhythm identification, and cardiac disease diagnosis are the main classification tasks in cardiology. Given the time-consuming nature of the procedure, it may be hard to find open-access datasets containing accurate labels for the specific disturbance under analysis, further reducing the amount of usable data for a specific task. Moreover, certain diseases or disturbances may require specific annotations on the ECG signal morphology (e.g. Brugada syndrome, early repolarization syndrome).

Unfortunately, the ECG labels may change also based on the clinical guidelines and rules adopted. This lack of agreement can make it challenging to merge different datasets in a homogeneous way to train a single model. For instance, in the context of heartbeat classification, the Association for the Advancement of Medical Instrumentation (AAMI) standard [34] is the most common and is adopted worldwide. It classifies heartbeats into five categories: normal beats (N), ventricular ectopic beats (V), supraventricular ectopic beats (S), fusion beats (F), and unknown beats (Q). However, there is also another widely adopted heartbeats classification system: the one exploited in the MITDB [27, 30]. It consists of 15 classes that can also be reassigned to the 5 AAMI super-classes, following AAMI recommendations [34,35].

2.2. Data sharing challenges

In order to get access to an existing ECG dataset for secondary usage or to build a new ad hoc one it is required to obtain the approval of the local research ethics committee and to comply with the local laws or recommendations related to personal data protection and management. Research must conform to recognized ethical standards that include respecting the dignity, rights, safety, and well-being of the people who take part in the study. Everything gets more complicated if the research involves importing data from a country that obeys different regulations.

The Health Insurance Portability and Accountability Act (HIPAA) [36] in the United States and the General Data Protection Regulation (GDPR) [37] in the European Union are two key regulations that set guidelines for protecting patients' privacy and data sharing. Switzerland is introducing a new legislation in 2023, the new Federal Act on Data Protection (nFADP) [38], to better protect its citizens' data and to ease the compatibility problems of Swiss law with European law, and in particular with the GDPR.

In the research ethical proposal, it is necessary not only to describe the project in detail but also regulatory aspects, safety and quality control, and data protection and management, including data handling and coding.

In fact, personal data protection and confidentiality are strictly associated with the terms *anonymization* and *de-identification*, which are frequently erroneously used as equivalent terms. As reported and deeply explored in the scoping review [39], these terms are often interpreted and used in different ways leading to confusion and poor understanding of the actual privacy level of the data. According to the GDPR *anonymization* leads to an irreversible state. More precisely, information can be considered anonymous if it "does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable" [37]. On the other hand, the terms *pseudonymization* and *de-identification*, defined respectively in the GDPR and in the HIPAA, are related to the concept of making the data less identifiable, for example by removing the so-called protected health information (PHI) (e.g. names, email address, age if over 89, etc.).

Data *anonymization* is an attractive solution for researchers to comply with data protection law since it makes data sharing much easier. Indeed, GDPR does not apply to anonymized data.

With certain categories of medical data, e.g. genomic data, it is clear that *anonymization* does not coincide with *de-identification* since the data themselves contain individual identifiers.

However, when researchers plan to work with biosignals, e.g. with ECG signals, a question may arise: is it enough to remove the associated PHI to ensure data anonymization?

According to several recent publications, biosignals like electrooculogram (EOG), electroencephalogram (EEG), and even ECG can be used as biometrics for identification purposes as well as the more popular anatomical and behavioral characteristics like fingerprint, face, iris, and gait [40].

Indeed, ECG analysis, in both frequency and time domains, can reveal patterns and properties that are specific to the individual, enabling subject identification. As pointed out in [41], ECG derives its uniqueness from the DNA. Indeed, heart geometrical characteristics, such as size, shape, position, orientation, etc., are typical of the individual and influence ECG morphology. Moreover, the latter is also affected by individual attributes like age, gender, weight, and pregnancy [41]. As reported in [42], the first attempt to use the ECG signal as a biometric dates back to 1977 [43]. In this work, the authors exploited 10 features extracted from a few heartbeats and based on the main ECG fiducial points. They proved that by producing a reference feature vector for each individual, these simple features enable the subject identification with quite high performances. These results have then been confirmed by other studies and the interest in using the ECG as an authentication factor has grown over the years. As reported in [40], several studies have been published in this field in recent years reaching close to 100% of accuracy. They proved to be able to solve the authentication task also with a single heartbeat and by extracting various types of features, not only based on ECG fiducial points but also on time and/or frequency analysis of the ECG. It is worth pointing out that some of these works take advantage of ECG fingerprint-like attitude to develop security systems for the medical field itself. [44,45] used ECG data to access patients' personal health records and confidential information. [46] instead developed a system for wearable device security.

The identification task may be more difficult in the presence of bad signal quality and/or different acquisition conditions (e.g. emotion, task, activity, pathological conditions, time). Nevertheless, some studies [47,48] already demonstrated that it is possible to identify the individual also in these complicated conditions.

Hence, removing the associated PHI is definitely not sufficient to anonymize ECG recordings. Other approaches need to be explored in order to share ECG data with the scientific community while respecting individual rights to privacy.

3. DGM for ECG synthesis

Deep generative models may represent a potential solution to the above-mentioned problems as they can be exploited both to augment and anonymize ECG datasets.

In this section we will describe the main types of DGMs and their peculiarity. The following sections, instead, will focus on the main DGMs applications found in literature and useful for the specific problem, on the characteristics of the generated ECG signals, and on their evaluation.

DGMs are unsupervised ML approaches that try to approximate the underlying distribution of the training data to be able to generate realistic examples that could belong to the learned data distribution. The most commonly used DGMs for synthetic ECG generation are variational autoencoders (VAE), generative adversarial networks (GAN) and diffusion models (DM), firstly introduced in [49–51], respectively. Some authors also attempted to combine GANs with VAEs or with concepts derived from mathematical models, proposing hybrid generative models.
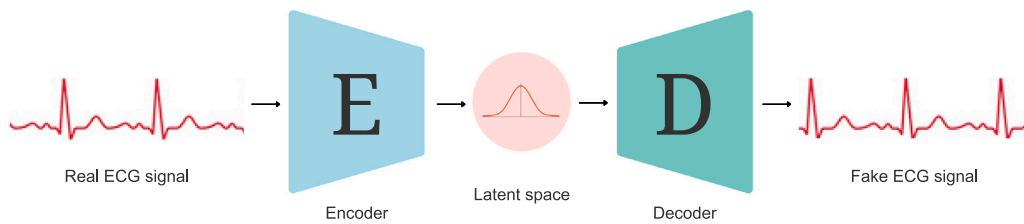
Fig. 2. VAE general scheme.

Fig. 3. GAN general scheme.

3.1. Variational autoencoders

Traditional autoencoders (AEs) are ML models composed of two networks: the Encoder and the Decoder. The Encoder reduces the input signal in a lower dimensional set of variables by extracting some latent features characterizing the input. Then, the Decoder tries to reconstruct the original signal from its encoded version.

VAEs differ from AEs because the input signal is not encoded into a single set of variables but into a distribution of sets of variables. This allows the Decoder to reconstruct new signals by sampling this distribution and, in this case, to generate new different ECGs.

A general scheme of a VAE is provided in Fig. 2. Examples of how VAEs can be adopted for synthetic ECG signal generation can be found in [52–55].

3.2. Generative adversarial networks

GANs are DL models that consist of two networks, namely Generator and Discriminator, that work as antagonists. The ultimate goal is to generate synthetic data that is difficult to distinguish from real data. The Generator, starting from vectors of noise provided in input, produces synthetic data that are as close as possible to the real data of the training set and tries to fool the Discriminator. The Discriminator, on the other hand, takes in input both real data from the original dataset and fake data generated by the Generator and tries to distinguish between them.

During training, the Generator learns to generate increasingly realistic data by continually adjusting its parameters to produce outputs that fool the Discriminator. Meanwhile, the Discriminator gets better at distinguishing between real and fake data. Due to the so-called adversarial training, the main challenge that can be faced when exploiting GANs is the training itself.

A general scheme of GAN is provided in Fig. 3.

Many kinds of GAN have been implemented through the years and applied for ECG data generation.

The simplest one is called vanilla GAN (VGAN) and it is based on fully connected layers [56–58].

Recurrent GANs (RGANs) were developed for time-series generation and therefore they were the first exploited to generate synthetic ECG signals [59–68].

Deep convolutional GANs (DCGANs) are based on convolution layers and are among the most suitable approaches for high-resolution image generation. Despite this, several works exploited DCGANs for ECG signals generation, directly [69–78] or indirectly by representing 1D input data in a 2D domain (e.g. time-frequency representation) [79–83].

Conditional GANs (CGANs) and auxiliary classifier GANs (ACGANs) have been introduced to deal with multi-labeled data (e.g. different types of arrhythmia) [54,57,58,69,72,84]. In both cases, the GAN is conditioned by the data labels: the Generator produces synthetic data for each label and the Discriminator distinguishes between synthetic and real data taking into account also class labels. The difference is that the Discriminator of ACGAN has a dual task: to recognize the nature of the data (synthetic or real) and to identify the correct class (e.g. arrhythmia class).

Finally, some other variants of GANs have been proposed or simply applied for ECG signal generation. As example, we refer to personalized GAN (PGAN) [77], progressive GAN (PGAN) [76], generative feature matching network (GFMN) [78] and sequence generative adversarial net optimized by policy gradient (SeqGAN-PG) [70].

Fig. 4 presents an example of a real 12-leads normal sinus rhythm (NSR) ECG signal compared to the 12-leads ECG signal generated with the GAN proposed in [85].

3.3. Diffusion models

DMs are probabilistic generative models that include two multi-step processes: the forward and the reverse diffusion. The forward diffusion process is fixed and consists of progressively adding noise to the input signal until it turns into noise. The reverse diffusion, instead, exploits a
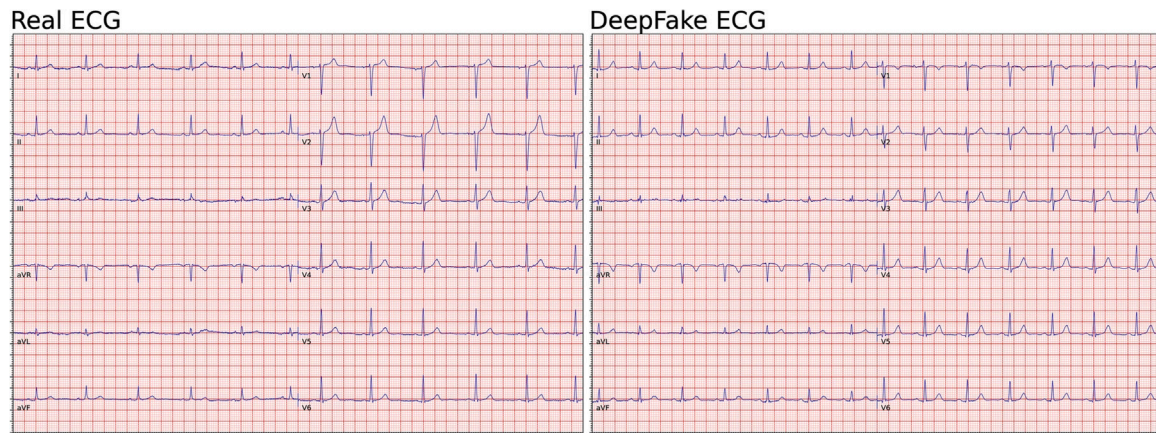
Fig. 4. Comparison of a real (left) 12-leads NSR ECG signal and a fake (right) 12-leads ECG signal generated with the GAN proposed in [85]. This figure was created by V. Thambawita et al. 09 November 2021, Deepfake electrocardiograms using generative adversarial networks are the beginning of the end for privacy issues in medicine, Scientific reports, vol. 11, no. 1. This figure is licensed under a Creative Commons Attribution 4.0 International Licence. To view a copy of this license, visit http://creativecommons.org/licences/by/4.0/.
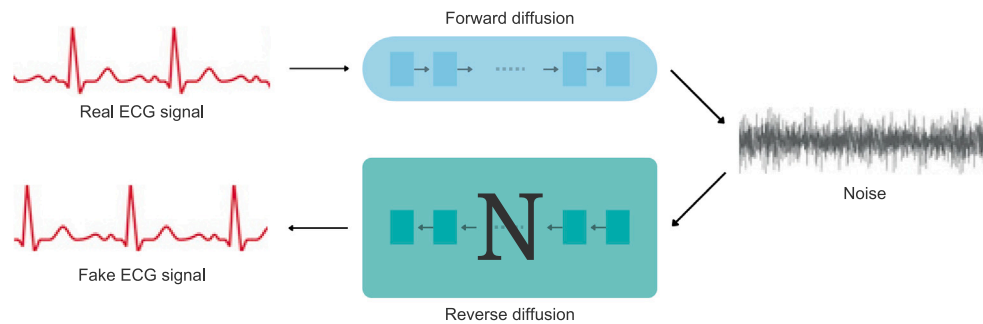

Fig. 5. DM general scheme.

trainable neural network to gradually remove the noise and reconstruct the signal. The synthetic signals are then generated by applying the trained neural network to simple noise. A general scheme of a DM is provided in Fig. 5. Few very recent works [86–88] attempted to exploit DMs for synthetic ECG signal generation. As with GANs, also DMs allow to generate data conditioned to a certain label [86,87]. The approaches proposed in both these works seem to be able to outperform GAN-based competitors. In [88], instead, the authors report that a GAN-based baseline was consistently superior to the approach they proposed. Certainly, more research is needed to understand DMs possibilities in ECG application.

3.4. Hybrid models

In [54,55] the authors tried to compare GANs and VAEs to assess which approach was the best but failed to identify a clear superiority. VAEs excel at discovering latent features within the data but, thanks to adversarial training, GANs tend to produce higher-quality synthetic data. In order to take advantage of the strengths of both approaches, it has been attempted to combine AE or VAE and GAN in hybrid generative models [84,85,89–91].

Other authors exploited techniques and concepts derived from mathematical modeling together with GANs in order to produce data constrained by physiological boundaries derived from cardiac models [67,73,74,92,93]. The authors of [94], instead, recently proposed a conditional generative model that combines a VAE, an autoregressive transformer-based model, and a GAN to generate synthetic ECG signals from text records.

4. DGM applications

DGMs have shown great potential in generating large amounts of realistic synthetic ECG signals. They have been exploited for two main applications: *data augmentation* and *data anonymization*. Both of them can be useful in solving the problem of low availability and accessibility of ECG data highlighted in this work. The distribution of works adopting DGMs for these two main applications is represented in Fig. 6.

4.1. Data augmentation

It appears clear from Fig. 6 that DGMs have been studied and adopted mainly for data augmentation purposes, to overcome data scarcity and/or class unbalancing problems. In most of the cases DGMs have been exploited in the context of a specific classification task with the final goal of improving the performance of the downstream classifier. A large number of studies addressed the problem of insufficient training samples of different classes and high class unbalance affecting the MIT-BIH Arrhythmia [30] database in the context of heartbeats or arrhythmia classification. [79], for instance, generated synthetic ECG segments with 7 different GANs, one for each heart rhythm (i.e. atrial fibrillation, atrial flutter, ventricular bigeminy, NSR, paced rhythm, and ventricular trigeminy), to augment the training set for a downstream arrhythmia classifier. They obtained an increase of F1-score of about 1% with all the 4 classifier architectures they tested. The improvement was particularly pronounced for the atrial flutter class, for which the number of original samples was very low and the F1 score increased
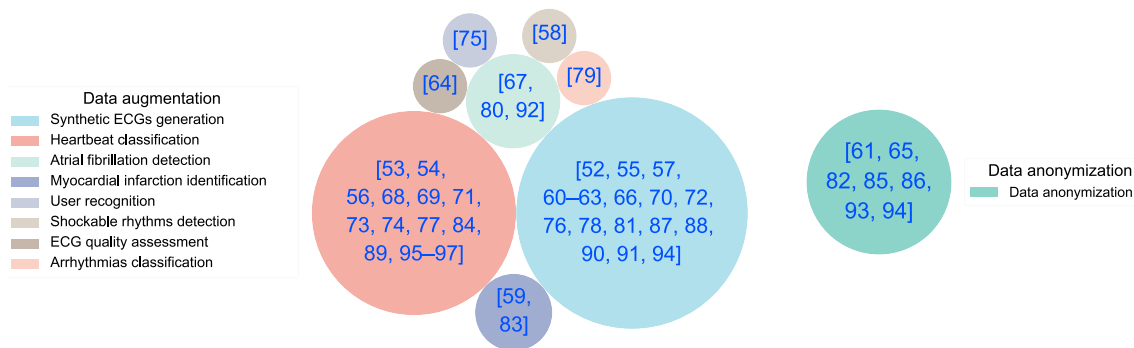
Fig. 6. Tasks for which DGMs have been exploited for data augmentation purposes, i.e. synthetic ECGs generation [52,55,57,60–63,66,70,72,76,78,81,87,88,90,91,94], heartbeats classification [53,54,56,68,69,71,73,74,77,84,89,95–97], AF detection [67,80,92], MI identification [59,83], user recognition [75], shockable rhythms detection [58], ECG quality assessment [64] and arrhythmias classification [79], or for data anonymization purposes [61,65,82,85,86,93,94].

by about 9%. In [83], the authors exploited a GAN to overcome the problem of data unbalancing toward the MI class in a MI identification task. Thanks to the generation of synthetic NSR ECG signals they managed to increase the number of training samples by 150% while keeping the dataset balanced. This resulted in an increase of the F1 score from 90.87% to 99.24%, where the baseline was obtained by balancing the dataset by undersampling the MI class. The authors of [64], instead, did not focus on cardiac diagnosis but on an automatic ECG quality assessment algorithm to provide an accurate initial screening of ECG quality. They exploited a conditional GAN to generate ECG samples for the "unacceptable" class and they managed to increase the accuracy of the ECG quality assessment task by about 2%. DGMs have been applied also to other downstream tasks including AF identification, shockable rhythms detection, and user recognition. The adoption of this tool for various classification tasks suggests how potentially it can be exploited to generate a more proper training set for any kind of classification task.

A considerable number of works, indeed, focused on data generation per se, regardless of any specific classification task, generating "standalone" synthetic ECG datasets. For instance, [55] introduced a DGM to generate ECG signals that could be used to train automated medical-aided diagnosis algorithms requiring a large volume of labeled clinical data and it was one of the first works presenting a DGM independently of the downstream task. [62], instead, proposed a GAN able to generate different kinds of bio-signals, i.e. ECG, EEG, electromyogram (EMG), and photoplethysmogram (PPG). Recently, [94] presented a DGM able to produce 12-leads ECG signals based on clinical text records.

4.2. Data anonymization

As mentioned above, one of the main barriers to ECG data sharing is its fingerprint-like attitude. The signal itself contains sensitive information, therefore it cannot be easily anonymized. DGMs and especially GANs, can be exploited to anonymize ECG signals: by creating a new set of synthetic data from a real dataset only general characteristics are derived from the entire set of subjects. Therefore, it is extremely unlikely to find in the generated dataset one ECG with exactly the same identifying characteristics of a single ECG belonging to a real subject of the original dataset.

Despite DGMs have been widely exploited for anonymizing electronic health records or clinical tabular data [98–104], medical images [105,106], and speech data used for remote health monitoring [107], their adoption in the ECG field is still quite limited.

A few very recent works explored the possibility of using the proposed GANs for ECG data anonymization, as shown in Fig. 6. The authors of [86], for instance, evaluated the possibility of exploiting a DM to generate a digital twin of a dataset, reflecting the same characteristics of the original one, but anonymized and thus easily shareable. To the best of our knowledge, the architecture proposed in [93] is the

first designed specifically for the ECG de-identification task. Its aim is to generate, for each real ECG, a new ECG by removing the subject identity information while maintaining domain-specific characteristics. This is made possible by the addition of two elements: the first one aims to verify if the physiological characteristics of the ECG are maintained, while the second one, called Verificator, has to assess if the original and the generated signals belong to the same subject.

5. Generated ECG characteristics

The synthetic ECGs generated with DGM can have different characteristics (i.e. number of leads, number of beats, type of heartbeat/rhythm reproduced) which may have an impact on their usability in the final task. These characteristics depend on both the model and the starting dataset. Indeed, the data generated with the DGMs developed so far replicate the same structure of the real data used to train the model, in terms of number of leads and beats. It means that, if the model is trained with single lead ECG recordings segmented beat-by-beat, then it will be able to generate only single lead ECG heartbeats.

Fig. 7 shows the distribution of studies of the literature depending on the characteristics of generated ECGs in terms of number of leads and beats.

As can be noted from Fig. 7, most of the works exploiting DGMs produce single-lead ECGs. This is probably due to the easiness of the approach but also to the fact that the majority of the open-access ECG datasets contain only one or two leads. Few recent works explored the possibility of generating multiple leads, adopting two different approaches. The first consists in training a model for each lead independently. The second one consists in training one single model by combining all the desired leads in a matrix. In this way, the model is able to learn the inter-dependency of the channels and preserve it in the generated leads. Therefore, the data generated with the second approach are expected to be more realistic: all the channels are generated simultaneously and their synchronization and dependency should be preserved.

Hence, DGMs are able to reproduce either single or multiple leads, according to the specific needs of the final task. Clearly, it would be more conservative to always perform data augmentation on all the available leads. Officially, only the standard 12-leads ECG is recognized as a diagnostic tool in clinics. In the field of AI, instead, researchers often try to develop algorithms (e.g. classifiers for diagnosis) able to obtain the same results with fewer leads, to reduce and optimize the information needed to solve the task. Therefore, in these cases, it could be sufficient to reproduce even a single lead.

The DGMs proposed in the literature have demonstrated their ability to produce both single and multiple heartbeats. The latter is essential for applications requiring the analysis of series of beats like rhythm analysis. DGMs have also shown to be able to generate ECG signals up to one-minute length [75,78].
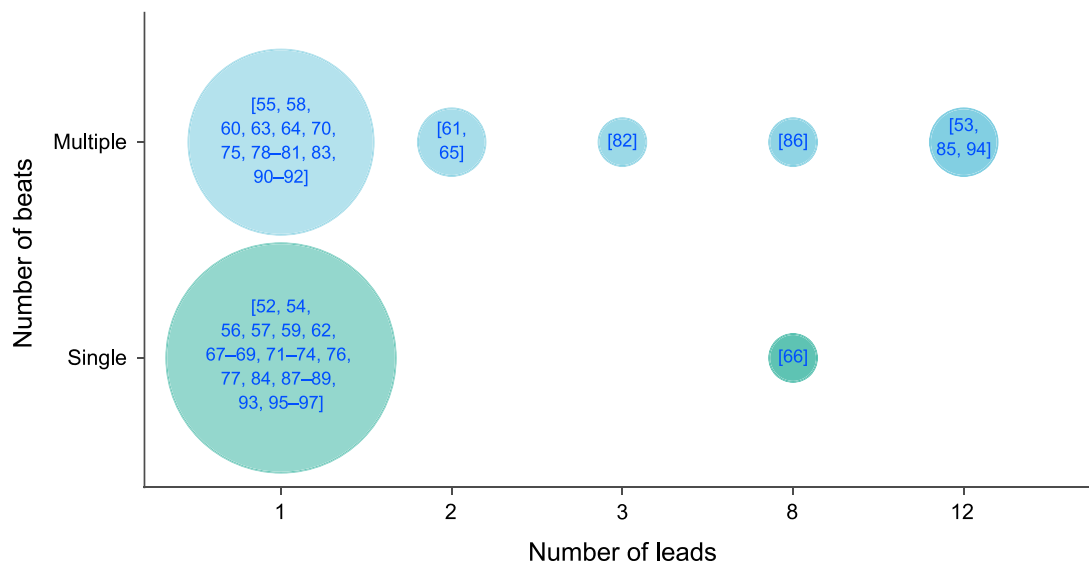
Fig. 7. Characteristics of the ECGs generated with DGMs in literature by single beat, i.e. single-lead [52,54,56,57,59,62,67–69,71–74,76,77,84,87–89,93,95–97], eight-leads [66] and by multiple-beats, i.e., single-lead [55,58,60,63,64,70,75,78–81,83,90–92], two-leads [61,65], three-leads [82], eight-leads [86], twelve-leads [53,85,94].

It should be highlighted that the length of the generated ECG is also limited by the training dataset: if the original dataset contains only single heartbeats, the model will not be able to generate more than one heartbeat at a time. Combining the obtained beats in a sequence may not result in a realistic ECG signal, as the generative algorithm was never exposed to characteristics over longer timescales, like heart rate variability.

There are different ways to produce multiple ECG classes (e.g. type of heartbeat, rhythm). The simplest and most common approach is to use an independent model for each class. However, the introduction of conditional GANs and auxiliary classifier GANs has opened the opportunity to generate multiple types of heartbeats or rhythms with the same model, by conditioning it with the desired class.

6. Generated ECG evaluation

After generating synthetic ECGs, either for data augmentation or data anonymization purposes, how can we evaluate the goodness of the generated data?

Unfortunately, there is no single answer to this crucial question. Depending on the final application three main aspects can be evaluated. Whether the end goal is data augmentation and/or anonymization, it is always important to assess the realness of the generated ECG signals which means the fidelity to the real ECG signals of the original dataset. In a second case, where synthetic ECGs generation is directly related to a specific classification task, it is also useful to examine the effect of the generated data in the classification task itself. Finally, when DGMs are adopted for data anonymization purposes, it is essential to verify that generated data are actually anonymized and privacy is preserved.

6.1. Fidelity to real ECG signals

Several criteria have been proposed in the literature to evaluate the realness of the generated ECG signals.

The most simple and intuitive criterion is the visual inspection of the signals [59,82,86,87,90–92,94,96]. It consists of visually assessing generated synthetic signal by evaluating its morphology, rhythmicity, or other peculiar characteristics.

This is a subjective and qualitative criterion and it may be quite time-consuming. Moreover, to enhance reliability the inspection of synthetic data should be always performed by expert clinicians. An expert cardiologist, indeed, is able to conduct a 360-degree evaluation of the signal by assessing every typical feature and the relationship between them, especially in case of pathological conditions.

Another visual method has been adopted by [78], by evaluating the real-vs-fake classification performance of a cardiologist with a set of real and generated data. This approach makes the evaluation less qualitative and less subjective but it should be conducted with due attention. In fact, the evaluation should be performed by an expert cardiologist, employing an unknown proportion of real and fake signals. The signals should be presented to the clinician in a random order and without giving feedback, e.g. by telling the correct answer during the classification. Moreover, subjectivity should be reduced by repeating the classification with other expert cardiologists. The main problem remains the time needed for classifying all the signals. Given the large quantity of data that is usually generated, this is typically adopted just as a supplementary evaluation applied to a restricted sample of synthetic data.

In [65], the author exploited a real-vs-fake classifier that acts as an external discriminator, and its performance is used to evaluate the closeness of the generated data to the original ones. It may be seen as an attempt to automate the real-vs-fake classification performed by the clinician.

This approach considerably reduces the time required for assessing the quality of generated signals but, on the other hand, the final evaluation also depends on the classifier adopted.

In the literature, quantitative metrics have also been proposed, to evaluate the similarity and fidelity of the generated signals to the training ones. These metrics can be divided into two groups.

The first one includes all the metrics that evaluate the statistical distribution of the signal features (e.g. the amplitude), regardless of the time order. This group comprehends: maximum mean difference (MMD) based metrics, relative entropy (RE) based metrics, earth mover's distance (EMD), statistics of features distribution (SFD) extracted from the ECG and visualization using T-distributed stochastic neighbor embedding (TDSNE). For instance, in [63,66,85] the authors tried to evaluate if the typical ECG characteristics were preserved in the generated signals. They calculated the amplitude, duration, or axis of some ECG waves or fiducial points (e.g. RR interval, QT interval, their correlation, R amplitude, QRS axis, etc.) to assess the plausibility of these values and/or their similarity with the respective distributions in the real data.

The second group comprehends metrics that compare pairs of signals, taking into account the temporal order. It includes Euclidean

distance (ED), mean absolute error (MAE), structural similarity index (SSI), mean square error (MSE) based metrics, time warping (TW) based metrics, Frechet distance (FD) based metrics and Chamfer distance (CD). In most cases, the comparison is carried out between each pair of real-fake signals. A few other works [69,95], instead, obtained a reference template of a single heartbeat by averaging the real ones and then compared each real or fake heartbeat with the template.

There is a lack of consensus regarding the metrics to be used as a reference to compare different models or synthetic datasets. Moreover, even if some of these metrics have been adopted in multiple works, they are often implemented in different ways, making the studies' comparison impossible. Another downside is the lack of reference values or thresholds to define if a signal is close enough to the real signals and thus realistic enough.

6.2. Effectiveness in classification

When synthetic ECGs generation is directly related to a specific classification task, the accuracy of the classifier can be exploited as an indirect measure to evaluate the goodness of the generated data. For data augmentation applications, this is the most used quantitative measure and it consists of evaluating the increment in the downstream classifier performance due to the addition of the generated ECG signals in the training set. This is a very simple way to assess the effectiveness of the data augmentation technique.

Similarly, for data anonymization applications, a downstream classifier can be exploited to verify if the performance obtained with the anonymized dataset is comparable to the ones obtained with the original dataset [86,93]. The aim is to assess if the ECG characteristics needed for the classification task are preserved in the anonymized dataset.

However, in both cases, these metrics can be adopted only if the generation of synthetic data is directly associated with and adopted for a classification task. Another drawback is that this evaluation criterion is strictly dependent on the specific classification task and on the classifier adopted, making it quite difficult to compare different works. Moreover, in data augmentation applications, the original real dataset quality has naturally a huge influence on the baseline performance of the classifier (i.e. when trained without synthetic data) and hence on the possible improvement.

6.3. Anonymization

When DGMs are adopted for ECG anonymization, there is a need to assess not only that the new dataset has preserved the real data characteristics but also that generated data are actually anonymized and privacy is secured. The assessment of the anonymization level reached can be extremely complicated. Only three [61,65,93] of the works exploiting DGMs for ECG data anonymization provided an evaluation metric for this purpose. In [61,65] a so-called "membership inference attack" on the synthetic dataset is performed. In brief, this is a privacy attack that aims to determine whether a specific data instance was used in the training set of a machine learning model.

In [93] the authors assess the effectiveness of the anonymization based on the accuracy of the Verificator. If it is close to the random guessing (i.e. 50%) it means that it cannot recognize that the two recordings belong to the same subject and thus the generated signal can be considered anonymized.

7. Discussion and conclusions

The application of AI techniques in cardiac clinical research is too often slowed down or even stopped, because of the lack of high-quality ECG datasets, and because of the limited number of samples in existing datasets. This is a major constraint.

Dataset building and even dataset access for secondary usage, are complicated processes that present several obstacles. Rarity and/or paroxysmal behavior of certain pathology obviously limits the amount of ECG data that can be collected in a single area. Data protection and ethical regulations, although defending and securing human rights, too often become barriers and decelerate open science progress. Data sharing is obstructed by regulation requirements, that may differ from one country to another, and sometimes even from one hospital to another. The fact that the ECG signal can be considered a human fingerprint and cannot be easily anonymized further complicates data sharing. Fostering a culture of open science and aligning regulations to favor wider collaboration of different clinical centers around the world is needed.

The DGM's ability to produce a larger synthetic dataset that shares only general characteristics with the original dataset while removing subjects' identity information could certainly support both the building of large datasets and ECG data sharing. Synthetic datasets have managed to boost AI algorithms' performance in various ECG classification tasks. The amount of data that DGMs can generate is potentially unlimited. To the best of our knowledge, there are no studies focusing on measuring the maximal possible size of a generated dataset. Since DGMs can only resemble the original dataset, an upper limit could be set by the specific downstream algorithm trained on these data and, in particular, by the reaching of a plateau in its performance.

These techniques have been originally conceived in the computer vision domain, and consequentially have been firstly exploited in the medical field, for generating images [108,109]. Afterward, they have been applied in the more challenging domain of biosignals. We deeply analyzed DGM applications to generate ECG signals, but they have also found utility with other biosignals, such as EEG [110–113], EMG [114], PPG [115].

DGMs present several practical advantages. They can be adopted to enhance the performance of ML algorithms whenever class imbalance or data scarcity is encountered, by generating synthetic samples for the less representative classes or for all of them according to need. Potentially, each cardiological center could create an anonymized version of its own real dataset and openly share useful information while ensuring the privacy of patients. It should be highlighted that the generation of new synthetic data would not increase the workload for clinicians in the data labeling phase since ECG segments can be generated together with the corresponding label.

Especially in the case of rare diseases, e.g. Brugada syndrome, the sharing of anonymized data coming from different centers could allow to produce a more representative dataset. The knowledge that could be extracted with AI techniques from a large representative dataset could not only increase algorithm performance but also result beneficial from a strictly clinical perspective. Indeed, a larger-scale analysis could allow to extract more useful and reliable information than the one that a single center can extract from a restricted number of recordings.

Naturally, DGMs are not the only possible solution to the challenges identified. For class unbalancing problems many popular data balancing techniques such as random oversampling, synthetic minority oversampling technique (SMOTE), adaptive synthetic (ADASYN), can be employed, even if DGMs have proven to be more performing [56, 67,80]. When large datasets are available but are not suitable for the specific task required, transfer learning techniques [116] can be exploited to train a model on a different task and then adjust the pre-trained model to the desired task. Semi-supervised or unsupervised learning approaches – e.g. promising recently proposed transformer-based architectures – can also be alternative solutions when a high number of unlabeled data are available. They can be exploited to learn data properties and extract representative features that can be used to solve classification tasks with less data. However, there is still a need for more research, to clearly understand if such approaches can be a valid alternative to data generation. Federated learning [117] is another possible approach to exploit sensitive data while preserving

privacy. The model can be locally trained, and only the parameters are exported. A new model is then created by combining multicenter information. This approach has technical limitations and is not certainly focused on building an openly accessible dataset, but on creating stronger algorithms.

The main limitation, common to all generative methods, is the lack of a standard quantitative and task-independent criterion to evaluate the realness of the generated data. There is indeed a huge need to promote a fair comparison between different algorithms and methodologies.

Another gap is represented by the lack of a robust anonymization assessment metric. There is an absolute need for privacy-secure data sharing, but only a few attempts have been found in the literature in this sense.

Another limitation may be given by the original dataset used to train the DGM. If it is too small and at the same time very heterogeneous, the model may not be able to learn the underlying data properties. Clearly, if a pathological condition is absent in the original dataset, the DGM will not be able to produce the related specific patterns in the generated data.

From a more technical point of view, it should be pointed out that, despite being the most used DGM, GANs are quite challenging to train because of problems such as training instability, slow convergence, and mode collapse [118].

All these considerations highlight the need for further research.

The first step should be the establishment of a standardized quality evaluation criterion. Probably, the adoption of a set of metrics evaluating the signals from different perspectives (e.g. ECG morphology, rhythmicity, presence of noise, etc.) would be beneficial, attempting to mimic the 360-degree visual inspection performed by the clinician.

Similarly, for anonymization applications, a standard criterion able to assess if subject privacy is actually guaranteed needs to be further explored. The concurrent development of advanced and performing techniques for easily and accurately identifying individuals from their ECG must be also kept in mind.

Given the recent popularity and successful application of DM in various fields and the first attempts to exploit them also for biosignals generation, it would be worth exploring these techniques further.

In summary, we believe that DGMs can certainly support from a technical point of view the creation of large anonymized datasets. Although, several challenges are on the way. A cultural change that could favor open science concepts adoption and standardization is also absolutely needed.

CRediT authorship contribution statement

Giuliana Monachino: Performed the literature analysis, Designed and structured the paper. **Beatrice Zanchi:** Performed the literature analysis, Designed and structured the paper. **Luigi Fiorillo:** Critically revised the paper. **Giulio Conte:** Critically revised the paper. **Angelo Auricchio:** Critically revised the paper. **Athina Tzovara:** Critically revised the paper. **Francesca Dalia Faraci:** Designed and structured the paper.

Declaration of competing interest

None Declared

Acknowledgments

References

[1] A. Holden, M. Poole, J. Tucker, Reconstructing the heart, Chaos Solitons Fractals 5 (3–4) (1995) 691–704.

[2] M. Boyett, A. Holden, I. Kodama, R. Suzuki, H. Zhang, Atrial modulation of sinoatrial pacemaker rate, Chaos Solitons Fractals 5 (3–4) (1995) 425–438.

[3] A. Holden, V. Biktashev, Computational biology of propagation in excitable media models of cardiac tissue, Chaos Solitons Fractals 13 (8) (2002) 1643–1658.

[4] M. Poole, A. Holden, J. Tucker, Hierarchical reconstructions of cardiac tissue, Chaos Solitons Fractals 13 (8) (2002) 1581–1612.

[5] S.R. Gois, M.A. Savi, An analysis of heart rhythm dynamics using a three-coupled oscillator model, Chaos Solitons Fractals 41 (5) (2009) 2553–2565.

[6] B. van der Pol, J. Van der Mark, The Heartbeat Considered As a Relaxation-Oscillation, and an Electrical Model of the Heart, Verlag nicht ermittelbar, 1929.

[7] K. Grudziński, J.J. Żebrowski, Modeling cardiac pacemakers with relaxation oscillators, Physica A 336 (1–2) (2004) 153–162.

[8] M.A. Quiroz-Juárez, J.A. Rosales-Juárez, O. Jiménez-Ramírez, R. Vázquez-Medina, J.L. Aragón, ECG patient simulator based on mathematical models, Sensors 22 (15) (2022) 5714.

[9] E.K. Roonizi, R. Sameni, Morphological modeling of cardiac signals based on signal decomposition, Comput. Biol. Med. 43 (10) (2013) 1453–1461.

[10] S. Parvaneh, M. Pashna, Electrocardiogram synthesis using a Gaussian combination model (GCM), in: 2007 Computers in Cardiology, IEEE, 2007, pp. 621–624.

[11] R. Dubois, P. Roussel, M. Vaglio, F. Extramiana, F. Badilini, P. Maison-Blanche, G. Dreyfus, Efficient modeling of ECG waves for morphology tracking, in: 2009 36th Annual Computers in Cardiology Conference (CinC), IEEE, 2009, pp. 313–316.

[12] M.A. Awal, S.S. Mostafa, M. Ahmad, M.A. Alahe, M.A. Rashid, A.Z. Kouzani, M.P. Mahmud, Design and optimization of ECG modeling for generating different cardiac dysrhythmias, Sensors 21 (5) (2021) 1638.

[13] Z. Li, M. Ma, ECG modeling with DFG, in: 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, IEEE, 2006, pp. 2691–2694.

[14] P.E. McSharry, G.D. Clifford, L. Tarassenko, L.A. Smith, A dynamical model for generating synthetic electrocardiogram signals, IEEE Trans. Biomed. Eng. 50 (3) (2003) 289–294.

[15] P. Tiam Kapen, S.U. Kouam Kouam, G. Tchuen, A comparative study between normal electrocardiogram signal and those of some cardiac arrhythmias based on McSharry mathematical model, Aust. Phys. Eng. Sci. Med. 42 (2) (2019) 511–528.

[16] N. Jafarnia-Dabanloo, D. McLernon, H. Zhang, A. Ayatollahi, V. Johari-Majd, A modified Zeeman model for producing HRV signals and its application to ECG signal generation, J. Theoret. Biol. 244 (2) (2007) 180–189.

[17] R.M. Evaristo, A.M. Batista, R.L. Viana, K.C. Iarosz, J.D. Szezech Jr., M.F. de Godoy, Mathematical model with autoregressive process for electrocardiogram signals, Commun. Nonlinear Sci. Numer. Simul. 57 (2018) 415–421.

[18] R. Sameni, G.D. Clifford, C. Jutten, M.B. Shamsollahi, Multichannel ECG and noise modeling: Application to maternal and fetal ECG signals, EURASIP J. Adv. Signal Process. 2007 (2007) 1–14.

[19] P. Cao, X. Li, K. Mao, F. Lu, G. Ning, L. Fang, Q. Pan, A novel data augmentation method to enhance deep neural networks for detection of atrial fibrillation, Biomed. Signal Process. Control 56 (2020) 101675.

[20] N. Nonaka, J. Seita, RandECG: Data augmentation for deep neural network based ECG classification, 2021, 2N1IS2a02, in 35 (2021).

[21] N. Nonaka, J. Seita, Data augmentation for electrocardiogram classification with deep neural network, 2020, arXiv preprint arXiv:2009.04398.

[22] N. Nonaka, J. Seita, Electrocardiogram classification by modified EfficientNet with data augmentation, in: 2020 Computing in Cardiology, IEEE, 2020, pp. 1–4.

[23] J. An, R.E. Gregg, S. Borhani, Effective data augmentation, filters, and automation techniques for automatic 12-lead ECG classification using deep residual neural networks, in: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2022, pp. 1283–1287.

[24] C.G. Turhan, H.S. Bilge, Recent trends in deep generative models: a review, in: 2018 3rd International Conference on Computer Science and Engineering (UBMK), 2018, pp. 574–579.

[25] N. Norori, Q. Hu, F.M. Aellen, F.D. Faraci, A. Tzovara, Addressing bias in big data and AI for health care: A call for open science, Patterns 2 (10) (2021) 100347.

[26] G. Clifford, C. Liu, B. Moody, L.-w. Lehman, I. Silva, Q. Li, A. Johnson, R. Mark, AF classification from a short single lead ECG recording: the physionet computing in cardiology challenge 2017, 2017, pp. 1–4, http://dx.doi.org/10.22489/CinC.2017.065-469.

[27] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals, Circulation 101 (June 23) (2000 (June 13)) e215–e220, Circulation Electronic Pages: http://circ.ahajournals.org/content/101/23/e215.full PMID:1085218.

[28] R. Bousseljot, D. Kreiseler, A. Schnabel, Nutzung der EKG-signaldatenbank CARDIODAT der PTB über das internet, Biomed. Tech. / Biomed. Eng. 40 (1995).

[29] EU research on rare diseases, 2023, https://research-and-innovation.ec.europa.eu/research-area/health/rare-diseases_en. Accessed: 2023-04-26.

[30] G. Moody, R. Mark, The impact of the MIT-bih arrhythmia database, IEEE Eng. Med. Biol. Mag.: Q. Mag. Eng. Med. Biol. Soc. 20 (2001) 45–50.

[31] G. S.D., Development and analysis of a ventricular fibrillation detector, 1986.

[32] F.M. Nolle, F.K. Badura, J.M. Catlett, R.W. Bowser, M.H. Sketch, CREI-GARD, a new concept in computerized arrhythmia monitoring systems, Comput. Cardiol. (1986).

[33] M. Pelter, M. Carey, S. Al-Zaiti, J. Zegre-Hemsey, C. Sommargren, L. Isola, P. Prasad, D. Mortara, F. Badilini, An annotated ventricular tachycardia (VT) alarm database: Toward a uniform standard for optimizing automated VT identification in hospitalized patients, Ann. Noninvasive Electrocardiol.: Off. J. Int. Soc. Holter Noninvasive Electrocardiol. Inc 28 (2023) e13054.

[34] E. ANSI-AAMI, Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms assoc. Adv., Med. Instrum. Arlington, VA (1998).

[35] E.J. da S. Luz, W.R. Schwartz, G. Cámara-Chávez, D. Menotti, ECG-based heartbeat classification for arrhythmia detection: A survey, Comput. Methods Programs Biomed. 127 (2016) 144–164.

[36] U.D. of Labor Employee Benefits Security Administration, The health insurance portability and accountability act (hipaa), 2004, http://purl.fdlp.gov/GPO/gpo10291. Accessed: 2023-05-03.

[37] European Parliament and Council of the European Union, Regulation (EU) 2016/679 of the European Parliament and of the Council, 2016.

[38] CH, New federal act on data protection (nFADP), 2023, https://www.kmu.admin.ch/kmu/en/home/facts-and-trends/digitization/data-protection/new-federal-act-on-data-protection-nfadp.html. Accessed: 2023-05-05.

[39] R. Chevrier, V. Foufi, C. Gaudet-Blavignac, A. Robert, C. Lovis, Use and understanding of anonymization and de-identification in the biomedical literature: Scoping review, J. Med. Internet Res. 21 (2019).

[40] H. Alrubaish, N. Saqib, Your vital signs as your password? in: M. Sarfraz (Ed.), Recent Advances in Biometrics, IntechOpen, Rijeka, 2022, ch. 10.

[41] T.M.C. Pereira, R.C. Conceição, V. Sencadas, R. Sebastião, Biometric recognition: A systematic review on electrocardiogram data acquisition methods, Sensors 23 (3) (2023).

[42] K. Revett, F. Deravi, K. Sirlantzis, Biosignals for user authentication - towards cognitive biometrics? in: 2010 International Conference on Emerging Security Technologies, 2010, pp. 71–76.

[43] G.E. Forsen, M. Nelson, R.J. Staron, Personal attributes authentication techniques, 1977.

[44] P. Huang, B. Li, L. Guo, Z. Jin, Y. Chen, A robust and reusable ECG-based authentication and data encryption scheme for ehealth systems, 2016, pp. 1–6, http://dx.doi.org/10.1109/GLOCOM.2016.7841541.

[45] B. Sivaranjani, N. Radha, Securing patient's confidential information using ECG steganography, in: 2017 2nd International Conference on Communication and Electronics Systems (ICCES), 2017, pp. 540–544.

[46] S. Yin, C. Bae, S. Kim, J.-s. Seo, Designing ECG-based physical unclonable function for security of wearable devices, in: Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, Vol. 2017, 2017, pp. 3509–3512.

[47] M. Komeili, W. Louis, N. Armanfard, D. Hatzinakos, On evaluating human recognition using electrocardiogram signals: From rest to exercise, in: 2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 2016, pp. 1–4.

[48] A. Ghazarian, J. Zheng, H. El-Askary, H. Chu, G. Fu, C. Rakovski, Increased risks of re-identification for patients posed by deep learning-based ECG identification algorithms, in: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 1969–1975.

[49] D.P. Kingma, M. Welling, Auto-encoding variational bayes, 2013, arXiv preprint arXiv:1312.6114.

[50] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets. proceedings of the 27th international conference on neural information processing systems-volume 2. NIPS'14, 2014.

[51] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, S. Ganguli, Deep unsupervised learning using nonequilibrium thermodynamics, in: F. Bach, D. Blei (Eds.), Proceedings of the 32nd International Conference on Machine Learning, in: Proceedings of Machine Learning Research, vol. 37, PMLR, Lille, France, 2015, pp. 2256–2265.

[52] V. Kuznetsov, V. Moskalenko, N.Y. Zolotykh, Electrocardiogram generation and feature extraction using a variational autoencoder, 2020, arXiv preprint arXiv:2002.00254.

[53] H. Liu, Z. Zhao, X. Chen, R. Yu, Q. She, Using the VQ-VAE to improve the recognition of abnormalities in short-duration 12-lead electrocardiogram records, Comput. Methods Programs Biomed. 196 (2020) 105639.

[54] Y. Xia, W. Wang, K. Wang, ECG signal generation based on conditional generative models, Biomed. Signal Process. Control 82 (2023) 104587.

[55] F. Zhu, F. Ye, Y. Fu, Q. Liu, B. Shen, Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network, Sci. Rep. 9 (1) (2019) 1–11.

[56] A.M. Shaker, M. Tantawi, H.A. Shedeed, M.F. Tolba, Generalization of convolutional neural networks for ECG classification using generative adversarial networks, IEEE Access 8 (2020) 35592–35605.

[57] Z. Zhou, X. Zhai, C. Tin, Fully automatic electrocardiogram classification system based on generative adversarial network with auxiliary classifier, Expert Syst. Appl. 174 (2021) 114809.

[58] K. Dahal, M.H. Ali, A hybrid GAN-based DL approach for the automatic detection of shockable rhythms in AED for solving imbalanced data problems, Electronics 12 (1) (2023).

[59] S. Haradal, H. Hayashi, S. Uchida, Biosignal data augmentation based on generative adversarial networks, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2018, pp. 368–371.

[60] A.M. Delaney, E. Brophy, T.E. Ward, Synthesis of realistic ECG using generative adversarial networks, 2019, arXiv preprint arXiv:1909.09150.

[61] E. Brophy, Synthesis of dependent multichannel ECG using generative adversarial networks, in: Proceedings of the 29th ACM International Conference on Information & Knowledge Management, 2020, pp. 3229–3232.

[62] D. Hazra, Y.-C. Byun, SynSigGAN: Generative adversarial networks for synthetic biomedical signal generation, Biology 9 (12) (2020) 441.

[63] S. Dasgupta, S. Das, U. Bhattacharya, CardioGAN: An attention-based generative adversarial network for generation of electrocardiograms, in: 2020 25th International Conference on Pattern Recognition (ICPR), IEEE, 2021, pp. 3193–3200.

[64] X. Zhou, X. Zhu, K. Nakamura, M. Noro, Electrocardiogram quality assessment with a generalized deep learning model assisted by conditional generative adversarial networks, Life 11 (10) (2021) 1013.

[65] E. Brophy, M. De Vos, G. Boylan, T. Ward, Multivariate generative adversarial networks and their loss functions for synthesis of multichannel ecgs, IEEE Access 9 (2021) 158936–158945.

[66] Y.-H. Zhang, S. Babaeizadeh, Synthesis of standard 12-lead electrocardiograms using two-dimensional generative adversarial networks, J. Electrocardiol. 69 (2021) 6–14.

[67] R. Banerjee, A. Ghose, Synthesis of realistic ecg waveforms using a composite generative adversarial network for classification of atrial fibrillation, in: 2021 29th European Signal Processing Conference (EUSIPCO), IEEE, 2021, pp. 1145–1149.

[68] J. Qin, F. Gao, Z. Wang, L. Liu, C. Ji, Arrhythmia detection based on WGAN-GP and SE-ResNet1D, Electronics 11 (21) (2022) 3427.

[69] P. Wang, B. Hou, S. Shao, R. Yan, ECG arrhythmias detection using auxiliary classifier generative adversarial network and residual network, IEEE Access 7 (2019) 100910–100922.

[70] F. Ye, F. Zhu, Y. Fu, B. Shen, ECG generation with sequence generative adversarial nets optimized by policy gradient, IEEE Access 7 (2019) 159369–159378.

[71] T. Golany, G. Lavee, S.T. Yarden, K. Radinsky, Improving ECG classification using generative adversarial networks, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, 2020, pp. 13280–13285.

[72] D. Nankani, R.D. Baruah, Investigating deep convolution conditional GANs for electrocardiogram generation, in: 2020 International Joint Conference on Neural Networks (IJCNN), IEEE, 2020, pp. 1–8.

[73] T. Golany, K. Radinsky, D. Freedman, SimGANs: Simulator-based generative adversarial networks for ECG synthesis to improve deep ECG classification, in: International Conference on Machine Learning, PMLR, 2020, pp. 3597–3606.

[74] T. Golany, D. Freedman, K. Radinsky, Ecg ode-gan: Learning ordinary differential equations of ecg dynamics via generative adversarial learning, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35, 2021, pp. 134–141.

[75] M.-G. Kim, S.B. Pan, A study on user recognition using the generated synthetic electrocardiogram signal, Sensors 21 (5) (2021) 1887.

[76] H. Yang, J. Liu, L. Zhang, Y. Li, H. Zhang, Proegan-ms: A progressive growing generative adversarial networks for electrocardiogram generation, IEEE Access 9 (2021) 52089–52100.

[77] T. Golany, K. Radinsky, PGANs: Personalized generative adversarial networks for ECG synthesis to improve patient-specific deep ECG classification, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 557–564.

[78] F. Cao, A. Budhota, H. Chen, K.S. Rajput, Feature matching based ECG generative network for arrhythmia event augmentation, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2020, pp. 296–299.

[79] T. Lan, Q. Hu, X. Liu, K. He, C. Yang, Arrhythmias classification using short-time Fourier transform and GAN based data augmentation, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2020, pp. 308–311.

[80] F.N. Hatamian, N. Ravikumar, S. Vesal, F.P. Kemeth, M. Struck, A. Maier, The effect of data augmentation on classification of atrial fibrillation in short single-lead ECG signals using deep neural networks, in: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2020, pp. 1264–1268.

[81] N. Wulan, W. Wang, P. Sun, K. Wang, Y. Xia, H. Zhang, Generating electrocardiogram signals by deep learning, Neurocomputing 404 (2020) 122–136.

[82] E. Piacentino, A. Guarner, C. Angulo, Generating synthetic ecgs using gans for anonymizing healthcare data, Electronics 10 (4) (2021) 389.

[83] W. Li, Y.M. Tang, K.M. Yu, S. To, SLC-GAN: An automated myocardial infarction detection model based on generative adversarial networks and convolutional neural networks with single-lead electrocardiogram synthesis, Inform. Sci. 589 (2022) 738–750.

[84] C. Du, P.X. Liu, M. Zheng, Classification of imbalanced electrocardiosignal data using convolutional neural network, Comput. Methods Programs Biomed. 214 (2022) 106483.

[85] V. Thambawita, J.L. Isaksen, S.A. Hicks, J. Ghouse, G. Ahlberg, A. Linneberg, N. Grarup, C. Ellervik, M.S. Olesen, T. Hansen, et al., DeepFake electrocardiograms using generative adversarial networks are the beginning of the end for privacy issues in medicine, Sci. Rep. 11 (1) (2021) 1–8.

[86] J. Alcaraz, N. Strodthoff, Diffusion-based conditional ECG generation with structured state space models, Comput. Biol. Med. 163 (2023) 107115.

[87] N. Neifar, A. Ben-Hamadou, A. Mdhaffar, M. Jmaiel, DiffECG: A generalized probabilistic diffusion model for ECG signals synthesis, 2023, arXiv preprint arXiv:2306.01875.

[88] E. Adib, A. Fernandez, F. Afghah, J. Prevost, Synthetic ECG signal generation using probabilistic diffusion models, IEEE Access PP (2023) 1.

[89] K.F. Hossain, S.A. Kamran, A. Tavakkoli, L. Pan, X. Ma, S. Rajasegarar, C. Karmaker, ECG-adv-GAN: Detecting ECG adversarial examples with conditional generative adversarial networks, in: 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, 2021, pp. 50–56.

[90] H.-C. Seo, G.-W. Yoon, S. Joo, G.-B. Nam, Multiple electrocardiogram generator with single-lead electrocardiogram, Comput. Methods Programs Biomed. 221 (2022) 106858.

[91] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, C. Fookes, Generalized generative deep learning models for biosignal synthesis and modality transfer, IEEE J. Biomed. Health Inf. (2022).

[92] Z. Xiong, M.K. Stiles, A.M. Gillis, J. Zhao, Enhancing the detection of atrial fibrillation from wearable sensors with neural style transfer and convolutional recurrent networks, Comput. Biol. Med. 146 (2022) 105551.

[93] S. Jafarlou, A.M. Rahmani, N. Dutt, S. Rahimi Moosavi, ECG Biosignal Deidentification Using Conditional Generative Adversarial Networks, Vol. 2022, 2022.

[94] H. Chung, J. Kim, J.-M. Kwon, K.-H. Jeon, M.S. Lee, E. Choi, Text-to-ECG: 12-lead electrocardiogram synthesis conditioned on clinical text reports, in: ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023, pp. 1–5.

[95] Y. Xia, Y. Xu, P. Chen, J. Zhang, Y. Zhang, Generative adversarial network with transformer generator for boosting ECG classification, Biomed. Signal Process. Control 80 (2023) 104276.

[96] T.H. Rafi, Y.W. Ko, HeartNet: Self multihead attention mechanism via convolutional network with adversarial data synthesis for ECG-based arrhythmia classification, IEEE Access 10 (2022) 100501–100512.

[97] M.S. Islam, M.N. Islam, N. Hashim, M. Rashid, B.S. Bari, F. Al Farid, New hybrid deep learning approach using BiGRU-BiLSTM and multilayered dilated CNN to detect arrhythmia, IEEE Access 10 (2022) 58081–58096.

[98] E. Choi, S. Biswal, B.A. Malin, J.D. Duke, W.F. Stewart, J. Sun, Generating multi-label discrete patient records using generative adversarial networks, in: Machine Learning in Health Care, 2017.

[99] A. Torfi, E.A. Fox, CorGAN: Correlation-capturing convolutional generative adversarial networks for generating synthetic healthcare records, in: The Florida AI Research Society, 2020.

[100] D. Lee, H. Yu, X. Jiang, D. Rogith, M. Gudala, M. Tejani, Q. Zhang, L. Xiong, Generating sequential electronic health records using dual adversarial autoencoder, J. Am. Med. Inform. Assoc. 27 (9) (2020) 1411–1419.

[101] J. Yoon, L. Drumright, M. Schaar, Anonymization through data synthesis using generative adversarial networks (ADS-gan), IEEE J. Biomed. Health Inf. PP (2020) 1.

[102] N. Park, M. Mohammadi, K. Gorde, S. Jajodia, Data synthesis based on generative adversarial networks, 2018.

[103] L. Gonzalez-Abril, C. Angulo, J.A. Ortega, J.L. López-Guerra, Generative adversarial networks for anonymized healthcare of lung cancer patients, Electronics (2021).

[104] H. Bae, D. Jung, H.-S. Choi, S. Yoon, Anomigan: Generative adversarial networks for anonymizing private medical data, Pac. Symp. Biocomput. Pac. Symp. Biocomput. 25 (2020) 563–574.

[105] B. Kim, J. Dolz, P.-M. Jodoin, C. Desrosiers, Privacy-net: An adversarial approach for identity-obfuscated segmentation, 2019.

[106] H.-C. Shin, N.A. Tenenholtz, J.K. Rogers, C.G. Schwarz, M.L. Senjem, J.L. Gunter, K.P. Andriole, M. Michalski, Medical image synthesis for data augmentation and anonymization using generative adversarial networks, in: A. Gooya, O. Goksel, I. Oguz, N. Burgos (Eds.), Simulation and Synthesis in Medical Imaging, Springer International Publishing, Cham, 2018, pp. 1–11.

[107] K. Vatanparvar, V. Nathan, E. Nemati, M.M. Rahman, J. Kuang, Adapting to noise in speech obfuscation by audio profiling using generative models for passive health monitoring, 2020 (2020) 5700–5704.

[108] T. Wang, Y. Lei, Y. Fu, W.J. Curran, T. xing Liu, X. Yang, Medical imaging synthesis using deep learning and its clinical applications: A review, Medical Physics (2020) arXiv:.

[109] M. Devi, S. Kamal, Review of medical image synthesis using GAN techniques, ITM Web Conf. 37 (2021) 01005.

[110] S.M. Abdelfattah, G.M. Abdelrahman, M. Wang, Augmenting the size of EEG datasets using generative adversarial networks, in: 2018 International Joint Conference on Neural Networks (IJCNN), 2018, pp. 1–6.

[111] N.K. Nik Aznan, A. Atapour-Abarghouei, S. Bonner, J.D. Connolly, N. Al Moubayed, T.P. Breckon, Simulating brain signals: Creating synthetic EEG data via neural-based generative models for improved SSVEP classification, in: 2019 International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1–8.

[112] K. Rasheed, J. Qadir, T.J. O'Brien, L. Kuhlmann, A. Razi, A generative model to synthesize EEG data for epileptic seizure prediction, IEEE Trans. Neural Syst. Rehabil. Eng. 29 (2021) 2322–2332.

[113] Y. Luo, L.-Z. Zhu, Z.-Y. Wan, B.-L. Lu, Data augmentation for enhancing EEG-based emotion recognition with deep generative models, J. Neural Eng. 17 (5) (2020) 056021.

[114] V. Mendez, C. Lhoste, S. Micera, EMG data augmentation for grasp classification using generative adversarial networks, in: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022, pp. 3619–3622.

[115] K. Vo, E.K. Naeini, A. Naderi, D. Jilani, A.-M. Rahmani, N.D. Dutt, H. Cao, P2E-WGAN: ECG waveform synthesis from PPG with conditional wasserstein generative adversarial networks, in: Proceedings of the 36th Annual ACM Symposium on Applied Computing, 2021.

[116] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (2010) 1345–1359.

[117] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, Y. Gao, A survey on federated learning, Knowl.-Based Syst. 216 (2021) 106775.

[118] J. Langr, V. Bok, GANs in Action: Deep Learning with Generative Adversarial Networks, Manning, 2019.