



A comprehensive review on GANs for time-series signals

Da Zhang^{1,2,3} · Ming Ma⁴ · Likun Xia^{1,2,3,5} 

Received: 8 October 2020 / Accepted: 4 January 2022 / Published online: 28 January 2022
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

Abstract

During the last decade, deep learning (DL) techniques have demonstrated the capabilities in various applications with a large number of labeled samples. Unfortunately, it is normally difficult to obtain such large amounts of samples in practice. As one of the most promising research directions of data generation in DL, generative adversarial network (GAN) can process not only images but also time-series signals. Unfortunately, it is easy to lose the time-dependence information from the latter due to characteristics of GAN, which increases the challenge of signal generation. Besides, the existing evaluation methods cannot evaluate the performance of GAN comprehensively. Therefore, this paper summarizes the current work of time-series signals generation based on GAN and the existing evaluation methods of GAN. As compared to existing GAN-related review work, this paper claims four unique points: (1) we specify the difficulties of GAN for time-series generation, particularly for the biological signal generation with potential solutions; (2) we analyze drawbacks of existing evaluation methods, and propose feasible solutions; (3) some suggestions are provided for the further research of robust time-series signal generation, especially for biological signal generation; (4) we provide a preliminary experiment to demonstrate the effectiveness of GANs for time-series signals generation, particularly, electroencephalogram (EEG).

Keywords Deep learning · Generative adversarial networks · Time-series signals · Evaluation methods · Data generation · Generative model

1 Introduction

During the last decade, big data has been a hot topic arising with the explosive growth of informative data and the rapid development of network computing technology. It drives the further development of artificial intelligence (AI) [1, 2]; meanwhile, AI turns big data into knowledge or

productivity, appearing in various different aspects of our lives, including computing [3, 4], medicine [5, 6], education [7, 8] and so on. Machine learning (ML) [9, 10], considered as the foundation of AI, is a multidisciplinary subject that specializes in how computers simulate learning behaviors from human to acquire new knowledge or skills and reorganize existing knowledge structures to improve their performance continuously. Unfortunately, traditional ML or shallow network models such as supporter vector machine (SVM), multilayer perceptron (MLP) normally struggles to process or learn big data due to its high temporal and spatial complexity [11], resulting in poor performance of AI. As the part of ML, deep network or deep learning (DL) [12] has been developed to handle such data and achieved huge success in different domains including brain-computer interface (BCI) [13, 14], natural language processing (NLP) [15, 16], affective computing [17, 18], games [19, 20], and so on. However, such technology is heavily dependent on sample size, i.e., big data is necessary, and substantial computational efforts are needed to train models with huge capacity, whereas it normally takes a lot of efforts to obtain such amount of training samples, especially in medical and military fields [21]; on the other

Likun Xia and Da Zhang contributed equally to the work.

✉ Likun Xia
xlk@cnu.edu.cn

- ¹ College of Information Engineering, Capital Normal University, Beijing, China
- ² International Science and Technology Cooperation Base of Electronic System Reliability and Mathematical Interdisciplinary, Capital Normal University, Beijing, China
- ³ Laboratory of Neural Computing and Intelligent Perception, Capital Normal University, Beijing, China
- ⁴ Department of Computer Science, Winona State University, Winona, USA
- ⁵ Beijing Advanced Innovation Center for Imaging Theory and Technology, Capital Normal University, Beijing, China

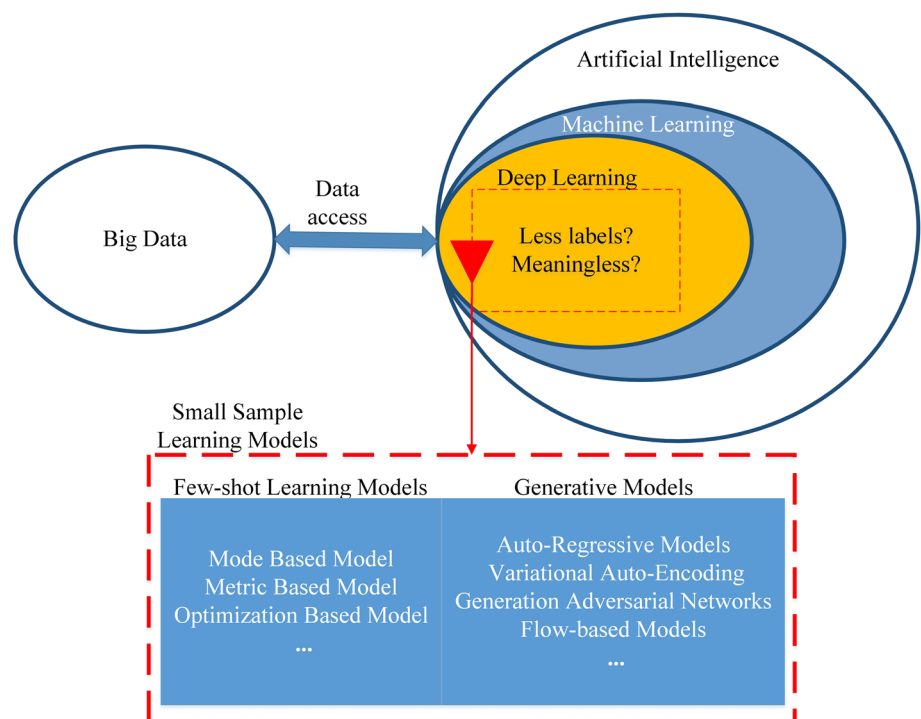
hand, the following two cases may still be considered insufficient amount of samples/small samples even if the size is large enough, i.e., big data: large amount of data but with small amount of targeting/expecting samples (meaningless/less features), for example, for identification of weak blood vessel in magnetic resonance angiography (MRA) images, due to its low proportion in such image, it is difficult for a ML-based model to be fully trained to identify such vessels even though the sample size is sufficient [22]; large amount of samples without labels or with small number of labels. Therefore, numerous methods have been developed on small sample size issue in another perspective including few-shot learning models [23–27] and generative models [28–30]. The former mainly consists of mode-based model [25], metric-based model [26], and optimization-based model [27], whereas the generative models generally encompass auto-regressive models, variational auto-encoding (VAE) [28], generative adversarial networks (GANs) [29] and flow-based models [30], as illustrated in Fig. 1.

Comparably, generative models have become more favorable for data/sample generation and have demonstrated their usefulness in various fields, particularly on natural samples such as images [31–37], video [28–40], music [41–43], speech [44, 45] and other research fields such as medical images [46–48], security [49, 50], wearing collocation [51, 52], time-series prediction [53], and anomaly detection [54–56].

Among them, GAN and its variants provide an alternative approach to sample generation and have demonstrated their capabilities of coping with not only images [54–58] but also time-series signals related applications such as music [42, 59–62], speech [63, 64] and biological signal. However, there are still more challenging to generate time-series signals than images due to three reasons: (1) most of GANs are built via the Full Connection (FC) layer and 2DCNN. These network layers are unable to model the changes of time-series signals, which results in the generation of fake sample; (2) in the task of time-series signal generation, the output data of the generator are discrete, which result in inability to calculate the gradient directly; (3) traditional GANs may fail to capture characteristics of time-series signals and the complex correlation between the related attributes, which results in the loss of the time-dependence of real signals.

There are a number of approaches that can further evaluate quality of the generated signals and GAN models. The approaches consist of two parts: (1) evaluation based on empiricism and (2) evaluation based on computation. Usually, both will be combined to get the evaluation results. Part (1) is very useful for visual/audio signals that require human-level interaction and evaluation, e.g., images [34, 35], music [59, 60] and speech [63]. So far they have achieved promising results. It is one necessary step since they can directly filter the ‘obvious’ signals for reducing both risks of misleading and evaluation time. For other signals such as biological signals, there are also

Fig. 1 A general picture of small sample methods for DL



temptations [65–69] but without human-level interaction, this is due to the characteristics of the signals. For part (2), at present, commonly used evaluation methods include the probability distribution measurement [70, 71], quality measurement [72–75], and classification measurement [76, 77] by calculating the similarity, clarity and diversity of generative samples, respectively. Unfortunately, there is not individual method that is capable of providing a comprehensive solution, i.e., GAN cannot be evaluated efficiently.

In this paper, we mainly focus on the GAN-related methods for time-series signal generation and corresponding evaluation methods. As compared to existing GAN-related review work [21, 78–84], this paper highlights three unique points: (1) we mainly summarize the applications of GAN in time-series signals in recent years, especially in the complex biological signals, which enables the development of BCI systems using less subjects and better personalization; (2) we specify difficulties and causes of GANs for time-series signal generation, which may help to improve effectiveness and robustness of GANs; (3) we analyze drawbacks of existing evaluation methods for the signals, e.g., they are lack of capabilities of automatic combination and decision-making, which may help to improve evaluation methods, so as to better generative models.

The rest of the paper is structured as follows: in Sect. 2, various GANs on time-series signals are introduced. Existing evaluation techniques for the GANs are highlighted in Sect. 3. In Sect. 4, GANs and corresponding evaluation techniques are discussed based on time-series signals and followed by suggestions in future research work. Conclusive remarks are provided in Sect. 5.

2 GANs on time-series signals

The GAN proposed by Goodfellow et al. [29] is a novel framework inspired by the minimax game theory for generating models through adversarial process estimation. Basically, it consists of two parts: a generator (G) and a discriminator (D). The latter performs adversarial training based on real samples and the former generates new samples by capturing the distribution of the real samples. GAN has been applied in various fields due to its capabilities of reliable model generation and enhancement. However, the standard GAN has drawbacks of training instability because of gradient disappearance and mode collapse [85–87]. A large number of GAN's variants have been developed in order to improve the stability, which are

implemented either by theory-based approaches such as changing the loss function [85–87], or frame (structure)-based approaches including interior structure [88, 89] or even adding the other conditions for input [90].

For the loss function changes, Mao et al. [91] proposed least squares GAN (LSGAN) by replacing least squares loss function to improve stability and converging speed. Martin et al. [85] theoretically analyzed the problem of the standard GAN, which was further highlighted by Wasserstein GAN (WGAN). WGAN solves mode collapse with Wasserstein Distance (WD) as the loss function. Moreover, WGAN applies weight clipping method to ensure Lipschitz restriction, but the method may cause gradient vanishing or gradient exploding. To improve the situation, Ishaan et al. presented WGAN-GP [92] by adding a Gradient Penalty (GP) into WGAN.

For interior structure changes, Alec et al. [88] proposed an unsupervised learning deep convolutional GAN (DCGAN) by integrating convolutional neural networks (CNN) [93] with GAN to achieve a more stable framework. Tero et al. [89] suggested a progressive growing GAN (PGGAN) using gradual growth approach to generate high-resolution samples. For condition setup, Mirza et al. [90] presented conditional GAN (CGAN) by adding conditional variables in random noise for input to solve training freedom. However, it is still unable to improve existing problem of traditional GAN.

The GANs discussed above mainly take images processing into consideration, but the fundamental structures provided can also be applied to other time-series signals with modification in various applications. Recently, there have been several models that can be applied to anomaly detection in the area of nonlinear systems. For example, in order to process time-series data, MAD-GAN [94] constructs the generator and discriminator of GAN into Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN). However, when the dataset is small, it can cause a larger discriminator to overfit easily, and a shallow generator cannot generate data that is realistic enough to defeat the discriminator. Similar to MAD-GAN, TAnoGAN [95] embeds LSTM into generator and discriminator models to process time-series data. The difference is that TAnoGAN uses different effective architectures to detect anomalies in small datasets, but the original formula for standard loss-fighting causes problems of gradient instability and schema collapse. Geiger et al. improve the problems by developing TadGAN [96] that combines Wasserstein Loss and Cycle Consistency to achieve the final minmax target. They use several new methods to calculate reconstruction error, as well as different methods to combine the reconstruction

error with the output of the discriminant to calculate the anomaly score. Moreover, there are some generative models working on natural signals and biological signals [97–99].

In this work, we mainly focus on time-series signal generation including natural signals (music, speech, text) and biological signals (EEG, EOG, ECG) since there are huge demands in extensive applications such as film-making, movie dubbing, music synthesis, and neuroscience. The following subsections introduce GAN and its variants based on the applications, respectively.

We categorize the GAN-based methods in time-series signals based on their application fields, which are introduced in the following subsections. GANs on music are introduced in Sect. 2.1. Section 2.2 presents the framework of GANs in speech enhancement. GANs on biological signals are introduced in Sect. 2.3.

2.1 Music

In recent years, there have been huge demands on music generation including films and videos [44–46]. However, such generation is inherently difficult because music contains a wide range of source information, and there are complex coherences between different sources; thus, it is difficult to develop such an effective mathematical model to deal with music signal [100–102]. GAN-based models have the capabilities for such music signal generation with supervised and even unsupervised learning techniques. They include continuous recurrent neural networks GAN (C-RNN-GAN) [59], sequence GAN (SeqGAN) [100], objective-reinforced GAN (ORGAN) [42], MuseGAN [101], WaveGAN [60], conditional WaveGAN (cWaveGAN) [61], GANSynth [46]. The following subsections discuss the techniques, respectively.

2.1.1 Continuous recurrent neural networks GAN

C-RNN-GAN [59] is a kind of continuous recurrent neural network (RNN) [102] that performs confrontation training on the basis of long short-term memory (LSTM) [103, 104]. It directly extracts whole sequences and then increases variety of generative samples by using feature matching approach. Moreover, it replaces the loss function L_G with standard G in order to avoid overfitting of D . Both loss functions L_D and L_G are defined as seen in Eqs. (1) and (2), respectively, where R is the representation from the

last layer before the final logistic classification layer in D , $z^{(i)}$ is a sequence of uniform random vectors in $[0,1]^k$, k indicates dimensionality from the random sequence, and $x^{(i)}$ is a sequence from the training samples.

$$L_G = \frac{1}{m} \sum_{i=1}^m \left(R(x^{(i)}) - R(G(z^{(i)})) \right)^2 \quad (1)$$

$$L_D = \frac{1}{m} \sum_{i=1}^m \left[-\log D(x^{(i)}) - \left(\log(1 - D(G(z^{(i)}))) \right) \right] \quad (2)$$

Typically, G is to maximize the error produced by the

D . When performing feature matching, the target becomes an internal representation that matches the actual samples at some level in the D . If either D or G is excessively powerful during training, it may cause model collapse. If it takes place, C-RNN-GAN will stop training the powerful party.

2.1.2 Sequence GAN

SeqGAN [100] generates time-series signals by combining GAN via policy gradient with reinforcement learning (RL) [105], where G is regarded as an agent in RL, state is the current produced tokens, and action is the next token for selecting G to generate sample. In the pre-training, sequence dataset was pretreated with maximum likelihood estimate (MLE) [104]. For each token, D uses Monte Carlo (MC) search [106] with a roll-out policy to explore a reward for the entire samples and return the current location. Algorithm 1 shows full details of SeqGAN, where the reward guides G to generate samples and guide D to be dynamically updated in order to further improve the generative model iteratively, as shown in Eq. (3):

$$\begin{aligned} Q_{D_\phi}^{G_\theta}(s = Y_{1:t-1}, a = y_t) \\ = \begin{cases} \frac{1}{N} \sum_{n=1}^N D_\phi(Y_{1:T}^n), Y_{1:T}^n \in MC^{G_\beta}(Y_{1:t}; N) & \text{for } t < T \\ D_\phi(Y_{1:t}) & \text{for } t = T \end{cases} \end{aligned} \quad (3)$$

where $Q_{D_\phi}^{G_\theta}$ is the action-value function that represents the reward at state s of taking action a , s is state, a is taking action, G_θ is following policy, G_β is the roll-out policy, and D_ϕ is discriminator policy.

Algorithm 1 Sequence Generative Adversarial Nets**Require:** following policy G_o ; discriminator policy D_ϕ ; Sequence dataset $S=\{X_{1:T}\}$

```

1: Initialize  $G_o, D_\phi$ .
2: Pre-train using MLE on  $S$ .
3: Generate fake sequence using  $G_o$ 
4: repeat
5:   for g-steps do
6:     Generate sequence  $Y_{1:T}=(y_1, y_2, \dots, y_T)$ 
7:     for  $t$  in  $1:T$  do
8:       Compute  $Q(a=y_t; s=Y_{1:t-1})$  by Equation (3).
9:     end for
10:    Update generator parameters via policy gradient
11:  end for
12:  for d-steps do
13:    Use current  $G_o$  to generate fake sequence and combine with real sequence
14:    Train discriminator  $D_\phi$ 
15:  end for
16: until SeqGAN converges

```

SeqGAN is one of the first extended GANs to generate music and has demonstrated its excellent performance. Unfortunately, during experiments issues on GAN such as gradient disappearance and model collapse still remain, so quality of the generative music can still be doubted.

2.1.3 Objective-reinforced GAN

ORGAN [42] combines GAN, expert reward with RL to improve quality of generative samples based on the framework of SeqGAN. The G of ORGAN is also regarded as an agent. However, ORGAN adds a reward function $R(Y_{1:T})$ defined for full length sequences to replace D_ϕ , and the action-value function is shown as Eq. (4). When given an incomplete sequence $Y_{1:t} = (y_1, y_2, \dots, y_t)$, also referred to a state, G produces an action a , along with the next token y_{t+1} . Moreover, to improve the stability of learning, WD is adopted to avoid problems of GAN convergence like ‘perfect discriminator.’

$$Q(Y_{1:t-1}, y_t) = \begin{cases} \frac{1}{N} \sum_{n=1 \dots N} R(Y_{1:T}^n), & \text{with} \\ Y_{1:T}^n \in MC^{G_\beta}(Y_{1:t}^n; N), & \text{if } t < T. \\ R(Y_{1:T}), & \text{if } t = T. \end{cases} \quad (4)$$

ORGAN can not only generate qualitative samples, but also retain essential information obtained from real samples. As compared to SeqGAN, it is capable of achieving better quality of samples and maintaining their diversities. Unfortunately, ORGAN has not integrated the choice of heuristic into the method to further improve its performance, e.g., an approach to choose a more effective loss function.

2.1.4 MuseGAN

MuseGAN [99] proposed a novel model of multi-generator and single-discriminator to generate multi-track polyphonic music signal. It divides the G into the shared temporary structure generator G_{temp} , the private temporary structure generator $G_{\text{temp},i}$, and the bar generator G_{bar} . Both G_{temp} and $G_{\text{temp},i}$ accept the noise of temporal dependency and inter-track, respectively. Finally, the outputs from G_{temp} and $G_{\text{temp},i}$ are concatenated into the G_{bar} which can generate piano-rolls sequentially. The generative process is shown as Eq. (5):

$$G(\bar{z}) = \left\{ G_{\text{bar},i} \left(z, G_{\text{temp}}(z_t)^{(t)}, z_i, G_{\text{temp},i}(z_{i,t})^{(t)} \right) \right\}_{i,t=1}^{M,T} \quad (5)$$

where i is inter-track, t indicates temporal dependency, M represents the number of tracks, and T is the number of bar.

Compared with other music GANs, MuseGAN can generate more unified rhythm and even clearer musical structure multi-track piano-rolls. However, esthetically, it still lags behind the level of human musicians.

2.1.5 WaveGAN

WaveGAN [60] is the first attempt at applying GANs to unsupervised audio generation. It is based on DCGAN [88] in terms of same parameters and numerical operations, but it has a different structure that is achieved by flattening 2D convolutions of DCGAN into 1D convolution (e.g., 4×4 2D convolution becomes 16×1 1D) for generating music as illustrated in Fig. 2.

Moreover, it adopts the WD as loss function to improve the stability. The function is shown in Eq. (6), where D_ω is

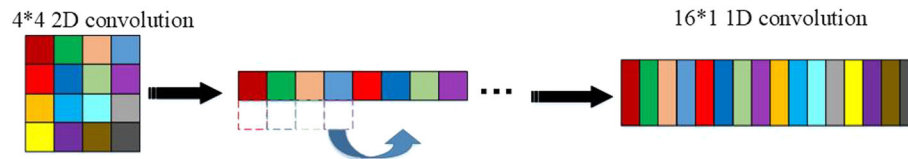


Fig. 2 Depiction of the transposed convolution operation from DCGAN to WaveGAN. (e.g., 4×4 2D convolution becomes 16×1 1D)

the 1-Lipschitz of WD, x represents real samples, and z is random noise.

$$V(D_\omega, G) = E_{x \sim P_x}[D_\omega(x)] - E_{z \sim P_z}[D_\omega(G(z))] \quad (6)$$

WaveGAN is fully parallelizable and can generate hours of audio in a few seconds [60]. As compared to other GANs for music, it is capable of generating more realistic music by using both waveform and spectrogram strategies. However, the generative music cannot be used in real applications such as movie dubbing because of the inflexibility and uncertainty of length of generative samples using unsupervised learning.

2.1.6 Conditional WaveGAN

As compared to WaveGAN discussed in Sect. 2.1.4, cWaveGAN [61] is designed based on supervised learning to generate audio samples. Basically, two hyper-parameter tuning methods are added as conditions in the system as illustrated in Fig. 3: (a) it embeds/concatenates conditioning information into an input feature vector; (b) it then scales hidden layers by initially mapping conditioning information to a scaling vector and then multiplying it with the input vector for both D and G .

With the conditions, cWaveGAN can generate music and audio as expected and be better controlled by human.

Unfortunately, the conditions added increase the training time.

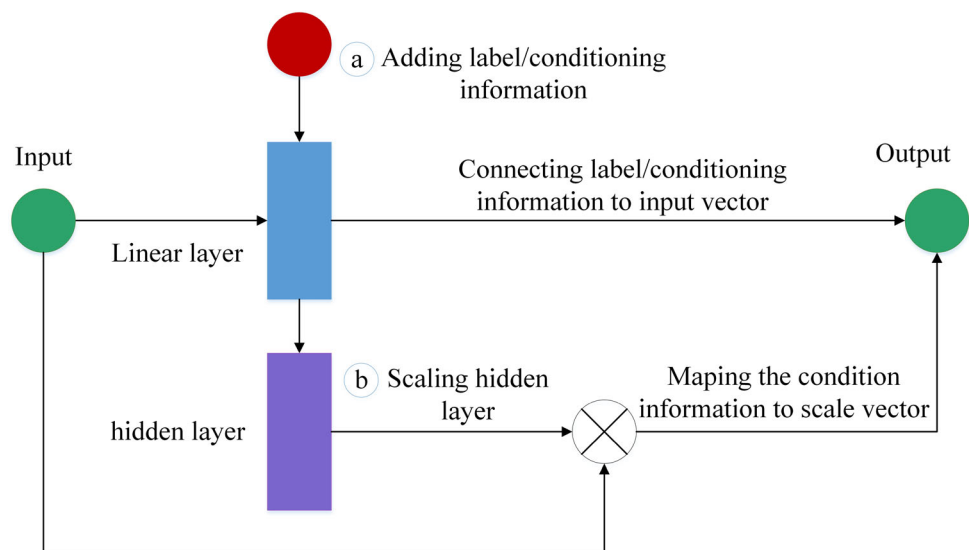
2.1.7 GANSynth

GANSynth [62] was developed based on structure of PGGAN [89] by gradually applying samples from a single vector to a complete music through convolution. Similar to WaveGAN [86], it is trained by progressive variants which increase dimensions of channel to generate more music in parallel. In this way, the speed of generating samples on modern GPUs is faster than other generative frameworks. The activation function x of G is also normalized as shown in Eq. (7), where C indicates the total number of channels, and n , h , w , and c refer to the batch, height, width, and channel dimensions, respectively.

$$x = x_{nhwc} / \left(\frac{1}{C} \sum_c x_{nhwc}^2 \right)^{0.5} \quad (7)$$

The GANSynth is able to generate music faster than other generative models, but it is only experimented based on NSynth dataset [41]; more datasets may be needed for evaluation. In addition, as GANSynth cannot capture the full data distribution well, mode collapse and issue of diversity still exist.

Fig. 3 cWaveGAN workflow: (a) adding conditioning information; (b) scaling hidden layers



2.1.8 Brief summary on various GANs for music

In this subsection, GANs discussed above for music generation are summarized in Table 1 regarding their individual advantages and disadvantages.

2.2 Speech

Speech processing is one of the most straightforward methods for human–computer interaction, but there is a lot of noise in the naturally collected speech, which affect processing results [85]. Many denoising methods [87, 107] can be used to process speech, but it may lose the important information. GAN and its variants can provide a novel approach to process speech for conserving the important information and denoising, e.g., Speech Enhancement GAN (SEGAN) [63] and Variational Auto-encoding Wasserstein GAN (VAWGAN) [64]. The following subsections introduce the techniques, respectively.

2.2.1 Speech enhancement GAN

SEGAN [63] provides an end-to-end structure, in which G and D share the same underlying structure. G is fully

convolutional, so that dense layer is not needed at all. This enforces the network to pay close attention to the time correlations in the input signal and throughout the whole layering process. Its input is fed into both the speech signals with noise and the latent representation; the output is an enhanced speech signal. With reference to CGAN [90], SEGAN adds the LSGAN [91] and replaces the loss function by binary coding to perform mapping and classification, respectively, i.e., in order to further stabilize the training and improve the quality of the generative samples. The loss functions of G and D can be found in Eqs. (8) and (9) accordingly, where λ is a parameter that controls the size of the L1 norm, c denotes the value that G and D to believe for fake samples, and \tilde{x} is the noisy signal which needs to be enhanced.

$$\begin{aligned} \min_D V_{LSGAN}(D) = & \frac{1}{2} E_{x, x_C \sim p_{\text{data}}(x, x_C)} \left[(D(x, x_C) - 1)^2 \right] \\ & + \frac{1}{2} E_{z \sim p_z(z), x_C \sim p_{\text{data}}(x_C)} \left[D(G(z, x_C), x_C)^2 \right] \end{aligned} \quad (8)$$

Table 1 Advantages and disadvantages of GANs on music generation

GANs' Variants	Advantages	Disadvantages
C-RNN-GAN [59]	The first application to generate classical music Generative music is more diverse in terms of the number and intensity of tones	Fail to pass human expert judgment Lack of diversity of generative samples
SeqGAN [100]	Reduce possibility of training issue by using gradient policy update of RL Produce more stable and effective samples as discriminator is pre-trained with supervised signals	Ability for decision-making is not directly proportional to the sample size The reward mechanism used causes time-consuming during training
ORGAN [42]	Able to maintain information and retain sample diversity Provide better results than either MLE or SeqGAN	Lack of a more effective loss function to stabilize training Lack of heuristic choice to further improve its performance
MuseGAN [99]	Propose the model of multiple generators and single discriminator Able to generate multi-track music Can be applied to AI cooperative music generation or music accompaniment	The generation process takes up a lot of GPU resources Fall behind the level of human musicians
WaveGAN [86]	The first application of GANs for unsupervised audio generation Fully parallelizable	Fail to generate multi-scale length of audio as required in films without setting conditions More noise is produced, especially in bird vocalizations
cWaveGAN [61]	Produce better results by setting conditional information and conditional scaling	The training time is increased comparably due to the conditioning methods Unexpected noise still exists Issue of unstable training process is still unsolved
GANSynth [62]	Comparably faster than other generative models with experiments Capable of generating entire sequence Keep harmonics from overlapping	Issue on mode collapse and diversity cannot be solved as expected Only applicable to specific data set

$$\min_G V_{LSGAN}(G) = \frac{1}{2} E_{\tilde{x} \sim p_{data}(\tilde{x}), z \sim p_z(z)} [(D(G(z, \tilde{x}), \tilde{x}) - 1)^2] + \lambda \|G(z, \tilde{x}) - x\|_1 \quad (9)$$

As compared with other models, SEGAN can reduce the operation speed of denoising waveform significantly [63], but there may be some high-frequency artifacts during speech enhancement [52].

2.2.2 Variational auto-encoding wasserstein GAN

VAWGAN [62] is a non-parallel voice convert (VC) system [107] combining VAE [28] with WGAN [108]. It is divided into three parts for implementing VC system: a conditional VAE (C-VAE) [107] reconstructs the tone, GAN is applied to improve the over-simplified C-VAE, and WGAN explicitly considers VC in the training objectives.

It is highlighted in Eq. (10), where q_ϕ is the expectation of z , α is a coefficient which emphasizes the loss of WGAN, p_t^* is distribution for real samples, and D_ψ is the discriminator.

$$J_{vawgan} = -D_{KL}(q_\phi(z_n|x_n)||p(z_n)) + E_{z \sim q_\phi}(z|x)[\log p_\theta(x|z, y)] + \alpha E_{x \sim p_t^*}[D_\psi(x)] - \alpha E_{z \sim q_\phi}(z|x)[D_\psi(\zeta_\theta(z, y_t))] \quad (10)$$

Moreover, in the system, VAE encodes the content z of the speech signal, and the decoder generates the converted target voice ζ_θ by giving a target speaker information y .

VAWGAN is the first GAN to implement a non-parallel VC and generate more realistic samples. However, the effectiveness of the conversion needs to be improved.

2.2.3 Brief summary on various GANs for speech generation

In the subsection, the GANs discussed above on speech generation are summarized in Table 2 in terms of individual advantages and disadvantages.

2.3 Biological signal

Biological signal can assist to detect physiological status of human being, and related research has become a hot topic in many fields such as medicine and computer science. Unlike computer vision or traditional medical imaging which requires thousands of data, in the EEG signal analysis, it is very rare and difficult obtain such amount of biological signals with significant pathological characteristics (such as epilepsy, depression, etc.) for effective training models. Therefore, it is meaningful to generate biological signals in this case. However, since biological signals are multi-channel with strong nonlinear, the generator may not capture the characteristics of these signals effectively. During the past few years, GAN and its variants have been applied to EEG generation to optimize the classification models, such as EEG-GAN [107], Recurrent GAN (RGAN) [65], Conditional WGAN (CWGAN) [109], and Class Conditional WGAN-GP (CC-WGAN-GP) [110]. The following subsections introduce the techniques, respectively.

2.3.1 EEG-GAN

As far as we know, EEG-GAN [107] is the first framework to apply GAN to EEG signal generation. It takes the structure of DCGAN and replaces its loss function by WGAN, which improves the stability and interpretability of the generative framework. The loss function is defined as seen in Eq. (11), where P_{data} and P_G represent the probability distribution of real samples and generative samples, respectively.

Table 2 Overview of GANs for speech generation

GANs Variants	Advantages	Disadvantages
SEGAN [63]	Provide an end-to-end structure for feature extraction Learn efficiently from various sources Faster enhancement process than other speech enhancement models	There may be some high-frequency artifacts in SEGAN during speech enhancement SEGAN is unable to generate long sequence of speech
VAWGAN [64]	The first GAN to implement a non-parallel VC and generate more realistic samples Does not require extra voice mapping as compared with other VC systems	The similarity of the converted voice is not improved The global variance (GV) from the output of VAW-GAN is not as good as that of the real data

$$W(P_x, P_G) \approx \max_{D \in 1\text{-Lipschitz}} \{E_{x \sim P_x}[D(x)] - E_{x \sim P_G}[D(x)]\} \quad (11)$$

In term of model training, EEG-GAN training starts with low resolution and gradually increases to all EEG signals in order to reduce the artifacts during sampling process. It generates EEG signals with similar frequency to real samples, which proves that GAN can be applied to EEG generation tasks. However, there are also high frequency artifacts in generative samples.

2.3.2 Recurrent GAN

RGAN [65] does not only effectively learn the characteristics of EEG signals, but also augment EEG signals. It replaces the full connection (FC) layer [111] in GAN by RNN [104]. This will help GAN to generate samples more accurately since RNN is able to capture time dependencies in the EEG signals. RGAN is presented as Eq. (12):

$$\min_G \max_D V(D, G) = E_{x \sim P_x}[\log D(x; \theta_D)] - E_{z \sim P_z}[\log(1 - D(G(z; \theta_G)))] \quad (12)$$

where θ_G is the parameter set of the G function, θ_D is the parameter set of the D function, z is random noise, and D classifies real samples and generative samples by minimizing the mean square error (MSE) [112, 113].

RGAN has shown its capability of generating EEG signals effectively, but it has been only evaluated with BCI2000 datasets [114]; thus, the generalization ability is doubted.

2.3.3 Conditional WGAN

CWGAN [109] generates EEG signals with WGAN-GP to improve the ability of EEG-based emotion recognition. Different from other frameworks of EEG generation, it generates differential entropy (DE) features that suite for EEG-based emotion recognition tasks [115]. To generate multiple emotion categories, an auxiliary label Y_r is fed into both D and G , where Y_r instructs the G to generate various kinds of samples and controls the kinds of generative samples in D . The proposed CWGAN can be formulated by Eq. (13):

$$\begin{aligned} \min_{\theta_G} \max_{\theta_D} L(X_r, X_g, Y_r) \\ = E_{x_r \sim X_r, y_r \sim Y_r} [D(x_r | y_r)] \\ - E_{x_g \sim X_g, y_r \sim Y_r} [D(x_g | y_r)] \\ - \lambda E_{\hat{x} \sim \hat{X}, y_r \sim Y_r} \left[(||\nabla_{\hat{x}|y_r} D(\hat{x} | y_r)||_2 - 1)^2 \right] \end{aligned} \quad (13)$$

where θ_g and θ_d represent the parameters of the G and D , λ is a hyper-parameter controlling the trade-off between original objective and gradient penalty, and \hat{x} denotes the data points sampled from the straight line between real distribution X_r and generative distribution X_g .

CWGAN sufficiently preserves the diversity of generative samples. However, CWGAN only generates DE instead of EEG signals. Therefore, it can only be used in the field of EEG signal classification.

2.3.4 Class conditional WGAN-GP

Similar to CWGAN, CC-WGAN-GP [90] generates high-quality multi-channel EEG signals based on WGAN-GP. The difference is that CC-WGAN-GP adds a classifier (C) to the D to improve the ability of GAN for EEG classification, as shown in Fig. 4. Specifically, both D/C share parameters ϕ_s at the lower layers (S), but have a unique discriminator branch ϕ_D and a classifier branch ϕ_C at the output of S .

In addition, the G generates samples χ_f conditioned on the class labels y_f that can be served as target labels of C . The loss function is presented as Eq. (14).

$$L_G(\phi_G, \phi_D^*, \phi_C^*, \phi_s^*) = E_{\chi_f \sim P_f} [D_{\phi_D^*}(G_{\phi_G}(z, y_f))] + L_C(\phi_G, \phi_C^*, \phi_s^*) \quad (14)$$

CC-WGAN-GP improves the classification ability of the model by training G and C with a unified loss. Meanwhile, it reduces the high frequency artifacts of the generative samples. However, the generative samples are only applicable to classification models.

2.3.5 Brief summary on various GANs for EEG generation

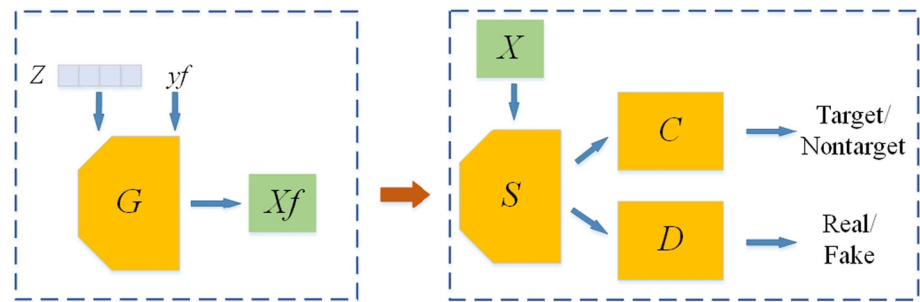
In this subsection, the GANs discussed above for EEG generation are summarized in Table 3 in terms of individual advantages and disadvantages.

2.4 A summary on GANs to process time-series signals

In this section, we will focus on analyzing the reasons that various GAN may fail. Moreover, we provide a comprehensive relationship of GAN and its variants in different applications.

2.4.1 Difficulties in processing time-series signals with GANs

Table 4 summarizes the difficulties of GANs in generating samples of time-series signals including music, voice and EEG.

Fig. 4 The structure of CC-WGAN-GP**Table 3** Overview of GANs on EEG generation

GANs' Variants	Advantages	Disadvantages
EEG-GAN [107]	First framework to generate EEG Generate EEG that is close to the real samples	Generate only single channel EEG signals There is no specific application for generating samples
RGAN [65]	Capable of capturing time dependencies in EEG signal Preserve the integrity of the information	Currently, RGAN is only evaluated with the BCI2000 dataset, thus the generalization ability is doubted
CWGAN [109]	Preserve the diversity of generative samples Improve the model's performance	Generate DE instead of EEG
CC-WGAN-GP [110]	Combine D with C to improve the classification ability Reduced the high frequency artifacts	The generative samples are only applicable to classification models

Table 4 Overview of GANs on time-series signals

Application	GANs variants	Difficulties of generation
Music	C-RNN-GAN [59], SeqGAN [100], ORGAN [42], MuseGAN [99], WaveGAN [86], cWaveGAN [61], GANSynth [62]	Music contains a wide range of information sources, and there are complex dependencies and mutual influences among the sources It is difficult to generate samples with musical characteristics and diversity During music generation, G cannot save features of the generated music (signal), which may lead to information loss and difficulties in subsequent music processing
Speech	SEGAN [63], VAWGAN [64]	Various environmental factors are involved in speech signals, including geographical reception distance and crowd, etc. Therefore, the influence from the factors must be taken into account in the process of speech signal processing The limitations of GAN may lead to some information missing in the real samples, which result in issues of uncertainty to the characteristics of context and confusion of order
EEG	EEG-GAN [107], RGAN [65], CWGAN [109], CC-WGAN-GP [110]	EEG is a multi-channel signal, and each channel contains complex information with stronger nonlinearity than other time-series signals, thus it is challenging for generative models to handle such signal accurately and robustly GAN cannot simultaneously generate multi-channel signals Since the signals are randomly generated, the corresponding position of EEG features cannot be guaranteed

2.4.2 A summary of GANs with time-series signals

It is known that the original GAN has drawbacks such as training instability and lack of diversity because of gradient disappearance and mode collapse. Its variants are developed and applied in time-series signal by changing either frames or theories in order to fit individual applications and improve the generation ability. Figure 5 summarizes the development of GANs for time-series signal generation. It is seen that the GANs (second row) were developed based on the original GAN and some GANs that were originally used for image processing (first row). The GANs in the first row are divided into frame based and theory based; the GANs in second row are categorized into application based in terms of music (yellow), speech (in brown) and EEG (in gray) accordingly.

Due to the network structure of traditional GAN, it is easy to destroy the time dependence information of time-series signals, which increases the difficulty of applying GAN to time-series signals. In recent years, many researchers have made attempts on different time-series signals. So far, GAN has demonstrated good performance in music and speech processing, and gradually applied to more complex biological signals. In 2016, researchers made a preliminary attempt to use relatively simple single track music signal in C-RNN-GAN. In the following year, SeqGAN and ORGAN being combined with reinforcement learning further improved music signal generation quality in multi-track long-time continuous music samples. In 2018, MuseGAN was designed with multiple generators

and discriminators to generate superimposed music signals. Since then, more generation frameworks have improved the generation efficiency in terms of the framework architecture and loss function, such as WaveGAN, cWaveGAN and GANsynth. On the other hand, GAN has been applied to the speech signal enhancement with more time dependence information. In 2017, SEGAN and VAWGAN were developed for speech enhancement. In 2019, GAN-TTS [116] generated higher fidelity speech signals than other models. Moreover, since 2018, several frameworks with GAN have been proposed for generating EEG signals, from single channel such as EEG-GAN to multi-channel such as RGAN, CWGAN, and CC-WGAN-GP. Figure 6 summarizes the trends of developing GANs for time-series signal generation against the publication years and publication sources of exiting works.

To further illustrate the performance of GAN in time-series signals, we select the results of representative models (GANSynth, VAW-GAN, and EEG-GAN) in music, speech and EEG signals, as shown in Fig. 7a, b, and c. In Fig. 7a, the illustrative results indicate that the real samples and generative samples in GANSynth have coherent waveforms that result in strong consistent colors for each harmonic. Figure 7b implies that the Global Variance (GV) of VAW-GAN may not generate better samples than real sample, but the higher values of which demonstrate its capability of diversity. In Fig. 7c, the generative/fake samples have similar waveform to the real samples in terms of the mean and standard deviation using EEG-GAN. The above results show that GAN and its variants can generate

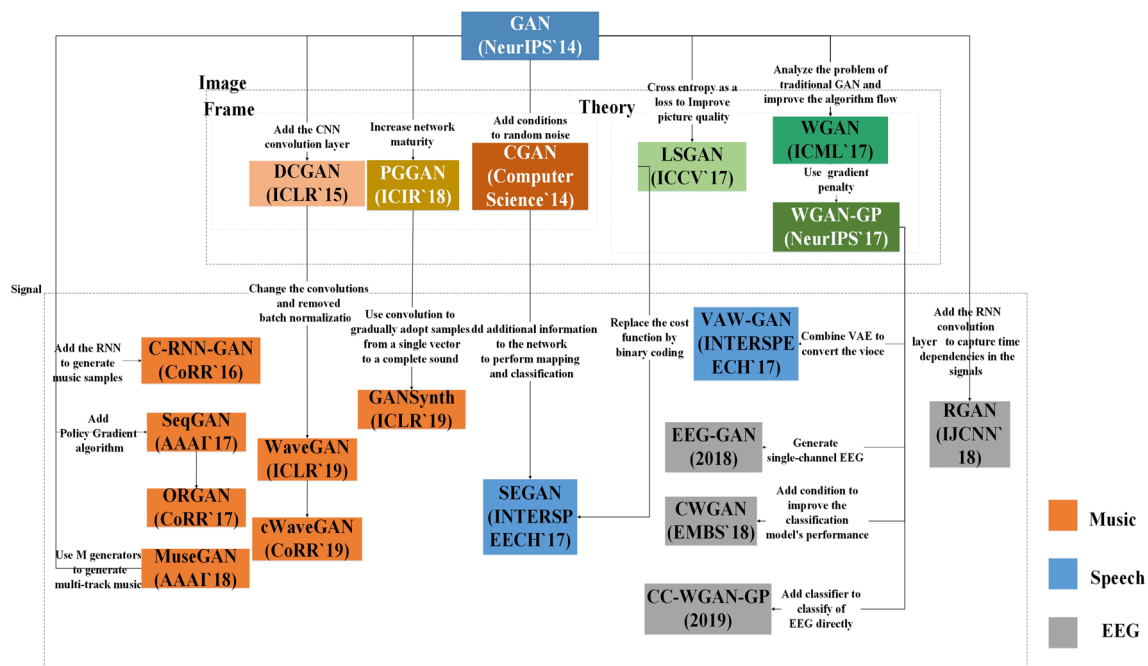
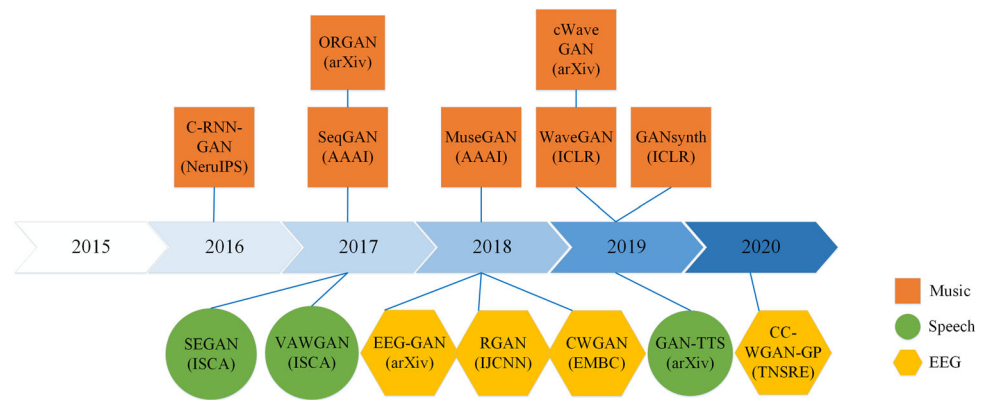


Fig. 5 Relationship between the GAN and its variants for time-series signal processing

Fig. 6 The trend of the GAN for time-series signals



similar or even better samples than real samples. They have opened up a new way for the development of time-series signals.

3 Evaluation methods

Evaluation and comparison of samples generated by various GANs have always been challenging [77, 116, 117]. Part of the reasons is that the existing evaluation methods may misjudge in some scenarios [118]. Therefore, evaluation of the samples still needs manual decision-making, which is time-consuming and may produce subjective results. Moreover, existing evaluation methods have some limitations, each of which only takes into account some indicators including information entropy, similarity, clarity, variety, and accuracy. Therefore, it is necessary to improve them by quantifying the GANs for time-series signals [77, 118] in terms of applicability and generalization ability.

3.1 Evaluation methods for GANs

In addition to subjectively evaluating the quality of the generated model, various evaluation methods/features have been used to evaluate the generative samples in terms of information entropy, clarity and variety, similarity, accuracy. Generally, entropy is obtained with indicators such as Kullback–Leibler Divergence (KLD) [119, 120] and Jensen–Shannon Divergence (JSD) [121]; clarity and variety are achievable with indicators like Inception Score (IS) [72], Fréchet Inception Distance (FID) [73] and Mode Score (MS) [118]; similarity can be obtained with distance-based indicators such as WD [122], Kernel MMD [74], and ED [75]; accuracy is achievable with 1-NN classifier [76]. More details on the methods are introduced in the following subsections.

3.1.1 Kullback–Leibler divergence

KLD [120], also called relative entropy, is equivalent to the difference between two probability distributions (P and Q) in cross-entropy as shown in Eq. (15), where P and Q are the probability distribution of real samples and generative samples accordingly. The cross-entropy is shown in Eq. (16).

$$KL(P||Q) = H(P) - H(P, Q) = \sum P(x) \log \frac{P(x)}{Q(x)} \quad (15)$$

$$H(P, Q) = \sum_x P(x) \cdot \log \left(\frac{1}{Q(x)} \right) \quad (16)$$

As the optimization and evaluation method for standard GAN, KLD is asymmetric, i.e., $KLD(P||Q) \neq KLD(Q||P)$, which implies that it could only evaluate the information entropy in one perspective. So KLD has uncertainty.

3.1.2 Jensen–Shannon divergence

JSD [121] is proposed to improve KLD as defined in Eq. (17). It calculates both two distributions the KLD (P , Q) and the mean of two distributions.

$$JS(P||Q) = \frac{1}{2} KL \left(P || \frac{P+Q}{2} \right) + \frac{1}{2} KL \left(Q || \frac{P+Q}{2} \right) \quad (17)$$

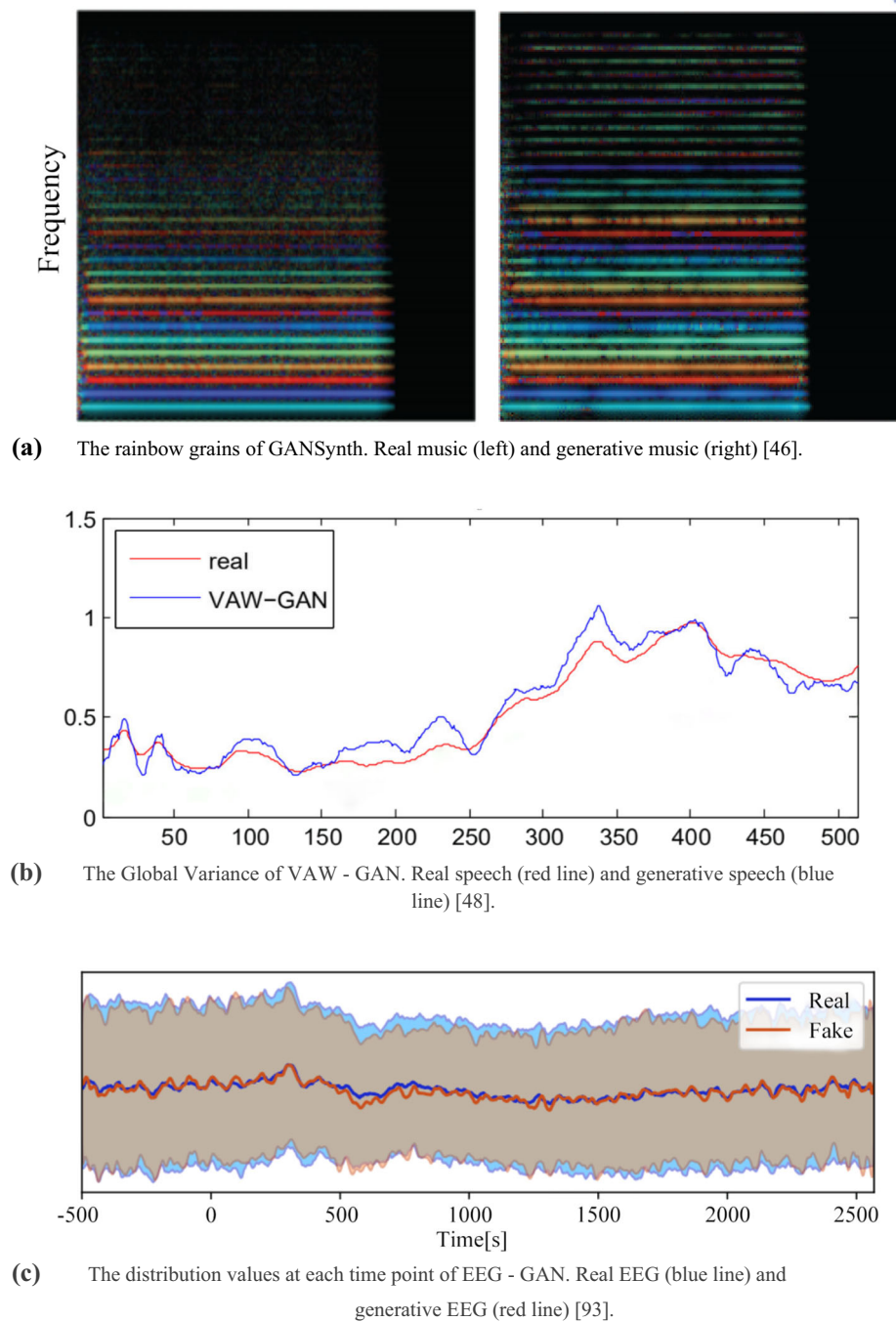
Compared with KLD, there is a problem in the measurement of KLD and JSD: if the distributions of P and Q do not overlap or overlap a little, then KLD and JSD become meaningless. Therefore, WD is recommended.

3.1.3 Wasserstein distance

WD [122], also known as earth-mover distance (EMD), is to measure the distance between two distributions. It is defined in Eq. (18):

$$W(P_r, P_g) = \inf_{\gamma \sim \prod (P_r, P_g)} E_{(x^r, x^g) \sim \gamma} [||x^r - x^g||] \quad (18)$$

Fig. 7 Some results of representative GANs on time-series signals



where $\Pi(P_r, P_g)$ denotes the set of all joint distributions (i.e., probabilistic couplings), γ is each possible joint distribution, and $\|x_r - x_g\|$ denotes the base distance between two samples.

WD can not only solve the problem of unstable training, but also provide a reliable indicator during training process. Moreover, compared with KLD and JSD, it produces superior smoothness, so as to theoretically solve the problem of gradient disappearance. However, calculating WD becomes more complicated when the number of

samples is increasing; thus, it is difficult to achieve expected results in reality.

3.1.4 Inception score

IS [72] originated from Inception Net [123] was initially applied to evaluate the clarity and diversity of the generative image. IS as defined in Eq. (19) calculates the KLD between $p(y|x)$ and $p(y)$, where $p(y|x)$ denotes distribution

of the label y by x obtained by prediction, and $p(y) = \int_x p(y|x) dp_g$.

$$IS(P_g) = e^{E_{x \sim p_g} [D_{KL}(p(y|x)||p(y))]} \quad (19)$$

If clarity in generative sample x is large enough, it may be classified easily, i.e., the entropy in $p(y|x)$ is low. Moreover, if variety in the generative sample is large enough, it should be evenly distributed across all categories, i.e., the entropy in $p(y)$ is high.

IS is able to quantify the diversity and clarity of generative samples. However, it is highly sensitive to noise and entirely relies on the final probability of classifier.

3.1.5 Fréchet inception distance

FID [73] approximates the Inception activation values of real samples and generative samples as Gaussian distribution [124], so the difference between the two samples can be obtained with the calculation of their mean and variance. The lower the FID indicates the higher the quality of the samples and the better the diversity. It is expressed in Eq. (20):

$$FID(x, g) = \|\mu_x - \mu_g\|_2^2 + Tr\left(C_x + C_g - 2(C_x C_g)^{\frac{1}{2}}\right) \quad (20)$$

where x is a real sample and g is a generative sample, μ_x and μ_g are means of real samples and generative samples, C_x and C_g are sample covariance, and the Tr represents the sum of the elements on the diagonal of the matrix.

Compared with IS, FID is more sensitive to model collapse and more robust to noise. However, neither FID nor IS can detect the overfitting of GAN.

3.1.6 Mode score

MS [118] makes an improvement based on the defect that IS does not consider the relationship between the generative samples and the real samples, as shown in Eq. (21). Compared with IS, MS calculates the KLD of the real samples distribution $p(y)$ and the sample label distribution $p(y^*)$, which could be closer to the real sample distribution.

$$MS(P_g) = e^{E_{x \sim P_g} [KL(p_M(y|x)||p_M(y)) - KL(p_M(y)||p_M(y^*))]} \quad (21)$$

MS cannot only evaluate the quality and diversity of generative samples, but also pay more attention to the similarity between real samples and generative samples. However, MS cannot detect the over fitting phenomenon in the training process of GAN.

3.1.7 Euclidean distance

ED [75] evaluates the similarity between the generative samples and the real samples. In the optimal case, the minimum ED_{\min} between two samples should be equivalent to the minimum distance distribution. It is defined as Eq. (22), where ED is the Euclidean distance between two distribution $A(x_1, x_2, \dots, x_n)$ and $B(y_1, y_2, \dots, y_n)$, x_i and y_i represent the characteristics of A and B, respectively.

$$ED = \sqrt{(x_n - y_n)^2 + \dots + (x_2 - x_1)^2 + (y_2 - y_1)^2} \\ = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (22)$$

ED is the simplest method to evaluate the similarity of samples. However, when the samples have multiple attributes, it tends to produce random fluctuations of varying sizes resulting in inaccurate results.

3.1.8 Kernel maximum mean discrepancy

Kernel MMD [74, 116] measures the dissimilarity between real samples P_r and generative samples P_g by giving some fixed kernel function k as defined as Eq. (23), where X_r indicates real samples and X_g indicates generative samples, they are sampled from P_r and P_g , and MMD can be calculated by the expected approximation between two distributions.

$$MMD^2(P_r, P_g) = E_{\substack{X_r, X'_r \sim P_r, \\ X_g, X'_g \sim P_g}} [k(X_r, X'_r) - 2k(X_r, X_g) + k(X_g, X'_g)] \quad (23)$$

The two distributions have equal mean values for the corresponding outputs on the function k , which indicates that the distributions can be considered as the same one.

The sample complexity and computational complexity of Kernel MMD are relatively low. However, the evaluation results may be biased.

3.1.9 1-Nearest neighbor accuracy

1-NN classifier [76] was used to evaluate whether the two distributions were identical in the two-sample test, as defined in Eq. (24), where S_r and S_g denote real samples and generative samples, respectively.

$$S_r \sim P_r^n \text{ and } S_g \sim P_g^m, \text{ with } |S_r| = |S_g| \quad (24)$$

When S_g matches S_r , the optimal scores can be obtained. Different from the most commonly used measure of accuracy, 1-NN classifier correlates the accuracy with Leave-One-Out (LOO) accuracy under certain condition [125]. It not only has all the advantages of other indicators,

but also can detect the mode collapse. However, the training process of 1-NN classifier is complex and consumes a lot of memory.

3.2 Summary of evaluation methods

Existing evaluation methods evaluate the quality of generative samples according to the indicators including information entropy, similarity, clarity, diversity, and accuracy. Among them, KLD and JSD are used for evaluating information entropy; however, if the two distributions are far apart, or no overlap, they can be meaningless. WD, SWD, ED, and the Kernel MMD have been used to evaluate the similarity of generative samples. Among the four methods, WD has the highest time complexity, while Kernel MMD has the lowest one. In terms of image clarity and diversity, IS, FID and MS are able to provide better results, but they have only been applied to image evaluation. Finally, 1-NN classifier is capable of classifying samples easily with excellent accuracy. Table 5 summarizes the evaluation methods from different perspectives.

4 Preliminary experiment

In this section, we analyzed the performance of WGAN-GP to generate EEG signals on a publicly available dataset Bi2015a [127]. The dataset includes EEG recordings of 50 subjects playing a visual P300 BCI video game called “Brain Invaders.” It recorded EEG signals from 32 channels in the 10–10 international system at sampling frequency of 512 Hz. After preliminary screening, we deleted the data with corrupted channels, so 35 subjects were selected; meanwhile, compared with the characteristics of P300 among all 32 channels, it is found that C3 has the most significant characteristics. Therefore, data of C3 channel from 35 subjects were used in the experiment. In this study, two different networks were performed to accordingly construct the generative models: (1) CNN, and (2) LSTM. The models were then evaluated using two evaluation metrics including SWD and MS, respectively. The former is used to evaluate the similarity of the model, and the latter is used to evaluate its diversity. The results are shown in Table 6.

Table 5 Comparison of various evaluation methods

Perspectives	Method	Advantages	Disadvantages
Information entropy	KLD [120]	Low complexity Easy to calculate	No gradient available Mutability Results are meaningless when the distributions of between samples Do not overlap
	JSD [121]	Perform Evaluation of information entropy better than KLD	No gradient available Mutability Results are meaningless when the distributions of between samples do not overlap
Clarity and variety	IS [72]	Highly sensitive to noise	Insufficient generalization ability
	MS [118]	Capable of evaluating both clarity and variety Pay more attention to the similarity than IS	Entirely relies on the final probability of classifier Cannot detect the over fitting phenomenon
	FID [73]	More sensitive to model collapse	Has difficulty in capturing subtle details
		Better discriminant power, robustness and efficiency than IS	Cannot describe the spatial relationship of sample features
Similarity	WD [122]	Suitable for multi-attribute samples WD performs well in feature space	High time complexity
	Kernel MMD [126]	Low sample complexity Low time complexity	Results can be biased
	ED [75]	More intuitive response to sample differences	Accuracy decreases as the increment of sample attributes
Accuracy	1-NN accuracy [76]	Better robustness and efficiency than other methods	The training process is much more complicated Computational cost is high Consume a lot of memory

Table 6 The quantitative evaluation results of generative EEG by different frameworks

Framework	Network	SWD	MS
WGAN-GP($G_D D_D$)	CNN	3.37	12.25
WGAN-GP($G_L D_L$)	LSTM	1.90	3.95

A higher MS score in bold indicates better diversity, and a lower SWD score in bold indicates greater similarity

According to the evaluation results, the SWD values of WGAN-GP ($G_D D_D$) and WGAN-GP ($G_L D_L$) were 3.37 and 1.90, respectively, indicating that the generated samples from WGAN-GP ($G_L D_L$) are more similar to the real EEG. In addition, the MS values of WGAN-GP ($G_D D_D$) and WGAN-GP ($G_L D_L$) are 12.25 and 3.95, respectively. The result indicates that WGAN-GP ($G_D D_D$) can generate more diverse samples than WGAN-GP ($G_L D_L$).

We further evaluated the quality of the generative EEG from a visual perspective. As illustrated in Figs. 8 and 9, it is observed that the range of the EEG signals generated by WGAN-GP ($G_D D_D$) is obviously inconsistent with that of the real EEG. In Fig. 9, the range of EEG signals generated by WGAN-GP ($G_L D_L$) is closer to the range of real EEG. Therefore, considering the similarity, diversity and visual observation, WGAN-GP ($G_L D_L$) has demonstrated significant performance for EEG signal generation.

5 Discussion and future work

During the last few years, GAN and its variants have been achieving massive success in terms of generative models, for not only images generation [55–58] but also time-series signal generation [42–49]. Unfortunately, it is difficult to

apply to all time-series signals, particularly on biological signal such as EEG signals. This may be due to the following reasons:

- (1) As compared with natural time-series signals such as music and audio signals, EEG signal generation and processing are much more complex due to its unreadable information hidden in multi-channel. Multi-channel may cause G to fail to generate specific features at one particular channel, whereas unreadable information reduces possibility of detection of neural correlation and activities. Moreover, manual intervention is not efficient enough to achieve convincing evaluation results. Therefore, it is essential to have a generalized and robust GAN for EEG generation and evaluation.
- (2) Since the standard GAN needs a series of sampling and convolution operations, it can destroy the continuity of the time-series signal. It also causes the loss of information in the signal, resulting in D unable to guide G to improve the generative ability through discrete information. Furthermore, issues like the gradient disappearance and mode collapse affect the quality of the generation since Nash equilibrium is hard to be achieved [73, 119]. Therefore, it is important to choose either more suitable loss functions or the framework.
- (3) Quality of generative models is still doubted. There are various evaluation methods (e.g., information entropy, clarity, variety, similarity, and accuracy [72–76]) for evaluating the GANs and generative models. However, none of them is capable of fully representing the quality with confidence; in other words, each of them only focuses on individual perspectives, so the GANs cannot be evaluated

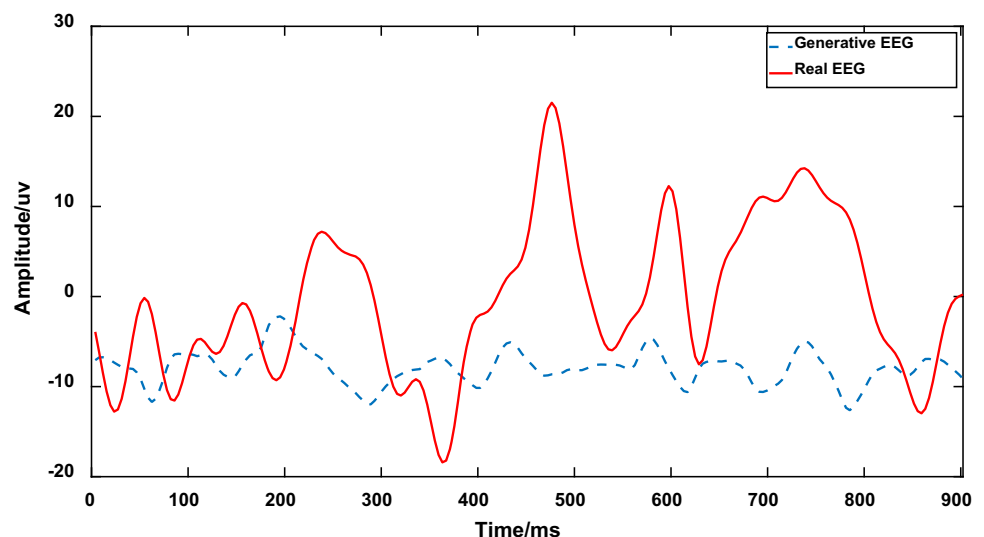
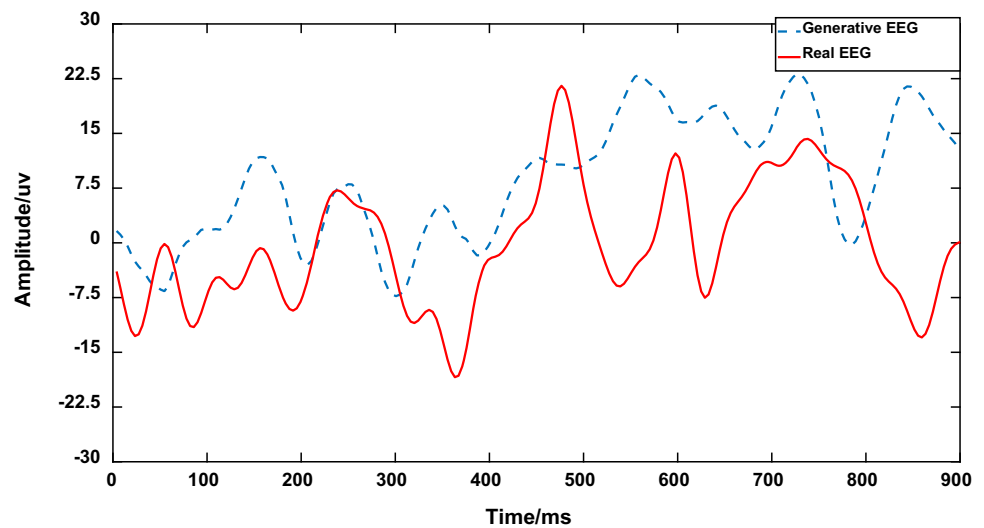
Fig. 8 The generative EEG of WGAN-GP($G_D D_D$)

Fig. 9 The generative EEG of WGAN-GP($G_L D_L$)

adequately. Unfortunately, they may raise some questions:

- (a) Manual selection of indicators by the user. This approach is subjective and time-consuming; quality of the generative models is heavily dependent on individual preference; if this is the case, one questions will be how it can be implemented automatically.
- (b) Lack of clear decision-making criteria. When the indicators are providing conflict solutions, it is almost impossible for users to make decision and provide weights manually; if this is the case, one question will be which one or more should be given more attention than others;
- (c) Lack of human expertise intervention for EEG signals. It has been demonstrated that existing indicators are capable of performing image and audio signal generation better than EEG signal, because their outcomes can be evaluated by human experts based on visual and auditory recognition, but this is not the case for EEG signal; one question will be how the generative EEG signals can be evaluated by neuroscientists even though information is hidden.

Based on the questions addressed above, potential further work can be proposed:

- (1) Development of a novel framework for EEG signal generation to handle robustness and stability. It is able to reduce the possibility of gradient disappearance and curse of multi-dimensions. This may be achieved by designing a novel loss function [91, 92, 108] such as adding appropriate super-parameters to improve stability, changing other

networks (e.g., LSTM, RNN) [101, 102], and even changing the random noise [127].

- (2) Development of a more intelligent and adaptive evaluation system to provide robust evaluation solution. This can be implemented with existing and/or modified unsupervised learning techniques such as cluster learning [128, 129] and Ricci Flow method [130], so that individual indicator will be assigned weights automatically and adaptively based on the contribution in the system.

6 Conclusion

This paper provides a comprehensive review on time-series signal generation with GAN-related framework and evaluation methods for the generated signals. Various state-of-the-art GANs are introduced and analyzed on how they are modified to fit in different applications such as music, speech and EEG signals. The techniques developed mainly focus on issues of unstable training on mode collapse and gradient disappearance. Moreover, existing evaluation methods are lack of clear decision-making criteria, especially when different indicators have provided conflict solutions. Compared with current review papers, four unique findings are addressed: (1) few related studies specifically focus on time-series signals generation due to limitation of GAN for time-series signals; (2) we clearly point out the difficulties of traditional GAN for time-series signals; (3) existing evaluation methods may fail to provide representative and comprehensive results on GANs due to lack of correlation between individual method and capability of automation on decision-making criteria. Based on the findings, we propose possible solutions to improve quality of GANs and a more intelligent and adaptive

evaluation system; (4) we demonstrate the effectiveness of GANs by generating EEG signals using a Bi2015a dataset, followed by quality checking using two evaluation metrics including SWD and MS. Above all, in spite of the defects in GANs on time-series signal generation, they have still gained great prospects in future research in various applications.

Acknowledgements This work is supported by Beijing Natural Science Foundation (4202011), National Natural Science Foundation of China (61572076, 61772351) and Key Research Grant of Academy for Multidisciplinary Studies of CNU (JCKXYJY2019018).

Declarations

Conflict of interest The authors declare that they have no conflicts of interest to this work.

References

- Ma Y, Wang Z, Yang H, Yang L (2020) Artificial intelligence applications in the development of autonomous vehicles: a survey. *IEEE CAA J Autom Sin* 11(3):315–329
- Xu LD, Duan L (2019) Big data for cyber physical systems in industry 4.0: a survey. *Enterp IS* 13(2):148–169
- Montes N, Rosillo N, Mora MC, Hilario L (2021) A novel real-time matlab/simulink/LEGO EV3 Platform for academic use in robotics and computer science. *Sensors* 21(3):1006
- Shhadat I, Al-bataineh B, Hayajneh A, Al-Sharif ZA (2020) The use of machine learning techniques to advance the detection and classification of unknown malware - sciencedirect. *Procedia Comput Sci* 170:917–922
- He H, Liu B, Luo H, Zhang T, Jiang J (2020) Big data and artificial intelligence discover novel drugs targeting proteins without 3D structure and overcome the undruggable targets. *Stroke Vasc Neurol* 5(4):381–387
- Karl E, Michel R (2018) Big data and artificial intelligence for diagnostic decision support in atypical dementia. *Nervenarzt* 89(8):875–884
- Huang C, Chen D, Guo W (2019) Innovation in methodology of education: big data and artificial intelligence. In: *Processing of national conference on computer science technology and education* 49–60
- Gordon SG, Robert H, Slavko Z (2020) Computer science meets education: natural language processing for automatic grading of open-ended questions in eBooks. *J Educ Comput Res* 58(3):1227–1255
- Bishop CM, Nasrabadi NM (2007) *Pattern recognition and machine learning*. J Electron Imaging 16(4):049901
- Trentin E, Schwenker F, El Gayar N, Abbas HM (2018) Off the mainstream: advances in neural networks and machine learning for pattern recognition. *Neural Process Lett* 48(2):643–648
- Rao DJ (2019) *Big data and artificial intelligence*. Hoboken, New Jersey
- LeCun Y, Bengio Y, Hinton GE (2015) Deep learning. *Nature* 521(7553):436–444
- Zhang X, Yao L, Wang X, Monaghan JJM, Mcalpine D, Zhang Y (2020) A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers. *J Neural Eng* 18(3):1–44
- Sakhavi S, Guan C, Yan S (2018) Learning Temporal Information for Brain-Computer Interface Using Convolutional Neural Networks. *IEEE Transac Neural Netw Learn Syst* 29(11):5619–5629
- Shao Y, Hardmeier C, Tiedemann J, Nivre J (2017) Character-based joint segmentation and POS tagging for chinese using bidirectional RNN-CRF. In: *Processing of international joint conference on natural language processing (IJCNLP)* pp. 173–183
- Devlin J, Chang MW, Lee K, Toutanova K (2019) BERT: pre-training of deep bidirectional transformers for language understanding. In: *Processing of the North American chapter of the association for computational linguistics – human language technologies (NAACL-HLT)* pp 4171–4186
- Zhang D, Yao L, Zhang X, Wang S, Boots R (2018) Cascade and parallel convolutional recurrent neural networks on EEG-based intention recognition for brain computer interface. In: *Proceedings of the 32nd AAAI conference on artificial intelligence (AAAI)* pp. 1703–1710
- Song T, Zheng W, Song P, Cui Z (2020) EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans Affect Comput* 11(3):532–541
- Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D (2016) Mastering the game of Go with deep neural networks and tree search. *Nature* 7587(2016):484–489
- Moudik J, Roman N (2016) Determining player skill in the game of go with deep neural networks. In: *Proceedings of theory and practice of natural computing* pp. 188–195
- Kazemini S, Baur C, Kuijper A, van Ginneken B, Navab N, Albarqouni S, Mukhopadhyay A (2018) GANs for medical image analysis. *arXiv preprint arXiv:1809.06222*
- Zhang H, Xia L, Song R, Yang J, Hao H, Liu J, Zhao Y (2020) Cerebrovascular segmentation in MRA via reverse edge attention network. In: *Processing of medical image computing & computer assisted intervention (MICCAI)* pp. 66–75
- Higy B, Bell P (2018) Few-shot learning with attention-based sequence-to-sequence models. *arXiv preprint arXiv:1811.03519*
- Kadam S, Vaidya V (2018) Review and analysis of zero, one and few shot learning approaches. In: *Processing of the 18th international conference on intelligent systems design and applications (ISDA'18)* pp. 100–112
- Santoro A, Bartunov S, Botvinick M, Wierstra D, Lillicrap TP (2016) One-shot learning with memory-augmented neural networks. *arXiv preprint arXiv:1605.06065*
- Munkhdalai T, Yu H (2017) Meta networks. In: *Processing of the 34th international conference on machine learning (ICML'17)* pp. 2554–2563
- Ravi S, Larochelle H (2017) Optimization as a model for few-shot learning. In: *Processing of the 5th international conference on learning representations (ICLR'17)* pp. 24–26
- Kingma DP, Welling M (2014) Auto-encoding variational bayes. In: *Processing of the 2nd international conference on learning representations (ICLR'14)* pp.1–14
- Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville AC, Bengio Y (2014) Generative adversarial nets. In: *Processing of the 27th neural information processing systems: annual conference on neural information processing systems (NeurIPS'14)* pp. 2672–2680
- Kingma DP, Dhariwal P (2018) Glow: generative flow with invertible 1x1 convolutions. In: *Processing of the 31st neural information processing systems: annual conference on neural information processing systems (NeurIPS'18)* pp. 10236–10245
- Ajakan H, Germain P, Larochelle H, Laviolette F, Marchand M (2014) Domain-adversarial neural networks. *arXiv preprint arXiv:1412.4446*

32. Bousmalis K, Silberman N, Dohan D, Erhan D, Krishnan D (2017) Unsupervised pixel-level domain adaptation with generative adversarial networks. In: *Processing of IEEE conference on computer vision and pattern recognition (CVPR'17)* pp. 95–104
33. Kim T, Cha M, Kim H, Lee JK, Kim J (2017) Learning to discover cross-domain relations with generative adversarial networks. In: *Proceedings of the 34th international conference on machine learning (ICML'17)* pp. 1857–1865
34. Ledig C, Theis L, Huszar F, Caballero J, Cunningham A, Acosta A, Aitken AP, Tejani A, Totz J, Wang Z, Shi W (2017) Photo-realistic single image super-resolution using a generative adversarial network. In: *Processing of the IEEE conference on computer vision and pattern recognition (CVPR'17)* pp. 105–114
35. Wang C, Xu C, Wang C, Tao D (2018) Perceptual adversarial networks for image-to-image transformation. *IEEE Trans Image Process* 27(8):4066–4079
36. Wu J, Zhang C, Xue T, Freeman B, Tenenbaum J (2016) Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. In: *Processing of the 29th neural information processing systems: annual conference on neural information processing systems (NeurIPS'16)* pp. 82–90
37. Zhu JY, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Processing of the IEEE international conference on computer vision (ICCV'17)* pp. 2242–2251
38. Tulyakov S, Ming-Yu L, Yang X, Kautz J (2018) Mocogan: decomposing motion and content for video generation. In: *Processing of IEEE international conference on computer vision (ICCV'18)* pp. 1526–1535
39. Vondrick C, Pirsiavash H, Torralba A (2016) Generating videos with scene dynamics. In: *Processing of the 29th neural information processing systems: annual conference on neural information processing systems (NeurIPS'16)* pp. 613–621
40. Walker J, Marino K, Gupta A, Hebert M (2017) The pose knows: video forecasting by generating pose futures. In: *Processing of the IEEE international conference on computer vision (ICCV'17)* pp. 3352–3361
41. van den Oord A, Dieleman S, Zen H, Simonyan K, Vinyals O, Graves A, Kalchbrenner N, Senior AW, Kavukcuoglu K (2016) WaveNet: a generative model for raw audio. In: *Processing of the 9th isca speech synthesis workshop, Sunnyvale (ISCA'16)* pp. 125–140
42. Guimaraes GL, Sanchez-Lengeling B, Farias PLC, Aspuru-Guzik A (2017) Objective-reinforced generative adversarial networks (ORGAN) for sequence generation models. *arXiv preprint arXiv:1705.10843*
43. Lee S, Hwang U, Min S, Yoon S (2017) A SeqGAN for polyphonic music generation. *arXiv preprint arXiv:1710.11418*
44. Hsu CC, Hwang HT, Wu YC, Tsao Y, Wang HM (2017) Voice conversion from unaligned corpora using variational autoencoding wasserstein generative adversarial networks. In: *Processing of the 18th annual conference of the international speech communication association (INTERSPEECH'17)* pp. 3364–3368
45. Lin K, Li D, He X, Sun MT, Zhang Z (2017) Adversarial ranking for language generation. In: *Processing of the 30th neural information processing systems: annual conference on neural information processing systems (NeurIPS'17)* pp. 3155–3165
46. Dai W, Dong N, Wang Z, Liang X, Zhang H, Xing EP (2018) SCAN: structure correcting adversarial network for organ segmentation in chest X-ray. In: *Processing of the medical image analysis-and-multimodal learning for clinical decision support - 4th international workshop (DLMIA'18), and 8th international workshop (ML-CDS'18), Held in Conjunction with MICCAI* pp. 263–273
47. Mardani M, Gong E, Cheng JY, Vasanawala S, Zaharchuk G, Alley MT, Thakur N, Han S, Dally WJ, Pauly JM, Xing L (2017) Deep generative adversarial networks for compressed sensing automates MRI. *arXiv preprint arXiv:1706.00051*
48. Yang D, Xiong T, Xu D, Huang Q, Liu D, Zhou SK, Xu Z, Park JH, Chen M, Tran TD, Chin SP, Metaxas DN, Comaniciu D (2017) Automatic vertebra labeling in large-scale 3d CT using deep image-to-image network with message passing and sparsity regularization. In: *Processing of the 25th international conference information processing in medical imaging (IPMI'17)* pp. 633–644
49. Shi H, Dong J, Wang W, Qian Y, Zhang X (2017) SSGAN: secure steganography based on generative adversarial networks. In: *Processing of the advances in multimedia information processing - 18th Pacific-rim conference on multimedia (PCM)* pp. 534–544
50. Volkhonskiy D, Nazarov I, Borisenko B, Burnaev E (2017) Steganographic generative adversarial networks. *arXiv preprint arXiv:1703.05502*
51. Liu LL, Zhang HJ, Xu XF, Zhang Z, Yan SC (2020) Collocating clothes with generative adversarial networks cosupervised by categories and attributes: a multidiscriminator framework. *IEEE Trans Neural Netw Learn Syst* 31(9):3540–3554
52. Zhang HJ, Wang XH, Liu LL, Zhou DL, Zhang Z (2020) WarpClothingOut: a stepwise framework for clothes translation from the human body to tiled images. *IEEE Multimed* 27(4):58–68
53. Pano-Azucena AD, Tlelo-Cuautle E, Ovilla-Martinez B, Gerardo de la Fraga L, Li R (2020) Pipeline FPGA-based implementations of ANNs for the Prediction of up to 600-steps-ahead of chaotic time series. *J Circuits Syst Comput* 30(9):1142–1153
54. Brock A, Donahue J, Simonyan K (2019) Large scale GAN training for high fidelity natural image synthesis. In: *Processing of the 7th international conference on learning representations (ICLR'19)*
55. Huang SW, Lin CT, Chen SP, Wu YY, Hsu PH, Lai SH (2018) AugGAN: cross domain adaptation with GAN-based data augmentation. In: *Processing of the 15th European conference computer vision (ECCV'18)* pp. 731–744
56. Wang X, Yu K, Wu S, Gu J, Liu Y, Dong C, Qiao Y, Loy CC (2018) ESRGAN: enhanced super-resolution generative adversarial networks. In: *Processing of the 15th European conference computer vision (ECCV'18)* pp. 63–79
57. Shaham TR, Dekel T, Michaeli T (2019) SinGAN: learning a generative model from a single natural image. In: *Processing of the IEEE international conference on computer vision (ICCV'19)* pp. 1–11
58. Zhu JY, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Processing of the IEEE international conference on computer vision (ICCV'17)* pp. 2242–2251
59. Mogren O (2016) C-RNN-GAN: continuous recurrent neural networks with adversarial training. *arXiv preprint arXiv:1611.09904*
60. Donahue C, McAuley JJ, Puckette MS (2019) Adversarial audio synthesis. In: *Processing of the 7th international conference on learning representations (ICLR'19)*
61. Lee CY, Toffy A, Jung GJ, Han WJ (2018) Conditional wavegan. *arXiv preprint arXiv:1809.10636*
62. Engel JH, Agrawal KK, Chen S, Gulrajani I, Donahue C, Roberts A (2019) Gansynth: adversarial neural audio synthesis. In: *Processing of the 7th international conference on learning representations (ICLR'19)*

63. Pascual S, Bonafonte A, Serra J (2017) SEGAN: speech enhancement generative adversarial network. In: Processing of the 18th annual conference of the international speech communication association (ISCA'17) pp. 3642–3646
64. Hsu CC, Hwang HT, Wu YC, Tsao Y, Wang HM (2017) Voice conversion from unaligned corpora using variational autoencoding wasserstein generative adversarial networks. In: Processing of the 18th annual conference of the international speech communication association (ISCA'17) pp. 3364–3368
65. Abdelfattah SM, Abdelrahman GM, Wang M (2018) Augmenting the size of EEG datasets using generative adversarial networks. In: Processing of the international joint conference on neural networks (IJCNN'18) pp. 1–6
66. Karimian N, Guo Z, Tehranipoor MM, Forte D (2017) Highly reliable key generation from electrocardiogram (ECG). *IEEE Trans Biomed Eng* 64(6):1400–1411
67. Palazzo S, Spampinato C, Kavasidis I, Giordano D, Shah M (2017) Generative adversarial networks conditioned by brain signals. In: Processing of the IEEE international conference on computer vision (ICCV'17) pp. 3430–3438
68. Wang F, Zhong S, Peng J, Jiang J, Liu Y (2018) Data augmentation for eeg-based emotion recognition with deep convolutional neural networks. In: Processing of the 24th international conference multimedia modeling (MMM'18) pp. 82–93
69. Mabu S, Fujita K, Kuremoto T (2019) Disaster area detection from synthetic aperture radar images using convolutional autoencoder and one-class SVM. *J Robot Netw Artif Life* 6(1):48–51
70. Lopez-Paz D, Oquab M (2017) Revisiting classifier two-sample tests for GAN evaluation and causal discovery. In: Processing of the 5th international conference on learning representations (ICLR'17)
71. Ma T (2018) Generalization and equilibrium in generative adversarial nets (GANs) (invited talk). In: Proceedings of the 50th Annual ACM SIGACT symposium on theory of computing (STOC'18) p. 2
72. Barratt S, Sharma R (2018) A note on the inception score. *arXiv preprint arXiv:1801.01973*.
73. Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S (2017) GANs trained by a two time-scale update rule converge to a local nash equilibrium. In: Processing of the 30th neural information processing systems: annual conference on neural information processing systems (NeurIPS'17) pp. 6626–6637
74. Slimene A, Zagrouba E (2013) Kernel maximum mean discrepancy for region merging approach. In: Processing of the 15th international conference computer analysis of images and patterns (CAIP'13) pp. 475–482
75. Wang L, Zhang Y, Feng J (2005) On the Euclidean distance of images. *IEEE Trans Pattern Anal Mach Intell* 27(8):1334–1339
76. Govindarajan M, Chandrasekaran RM (2010) Evaluation of k-nearest neighbor classifier performance for direct marketing. *Expert Syst Appl* 37(1):253–258
77. Shmelkov K, Schmid C, Alahari K (2018) How good is my gan?. In: Processing of the 15th european conference computer vision (ECCV'18) pp. 218–234
78. Wang K, Gou C, Duan Y, Lin Y, Zheng X, Wang F-Y (2017) Generative adversarial networks: introduction and outlook. *IEEE CAA J Autom Sin* 4(4):588–598
79. Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B, Bharath AA (2018) Generative adversarial networks: an overview. *IEEE Signal Process Mag* 35(1):53–65
80. Hong Y, Hwang U, Yoo J, Yoon S (2019) How generative adversarial networks and their variants work: an overview. *ACM Comput Surv* 52(1):10:1–10:43
81. Wang Z, She Q, Ward TE (2019) Generative adversarial networks: a survey and taxonomy. *arXiv preprint arXiv:1906.01529*
82. Hitawala S (2018) Comparative study on generative adversarial networks. *arXiv preprint arXiv:1801.04271*
83. Zamorski M, Zdobylak A, Zieba M, Swiatek J (2019) Generative adversarial networks: recent developments. In: Processing of the 18th international conference artificial intelligence and soft computing (ICAISC'19) pp. 248–258
84. Pan Z, Weijie Yu, Yi X, Khan A, Yuan F, Zheng Y (2019) Recent progress on generative adversarial networks (GANs): a survey. *IEEE Access* 7(2019):36322–36333
85. Engel JH, Resnick C, Roberts A, Dieleman S, Norouzi M, Eck D, Simonyan K (2017) Neural audio synthesis of musical notes with WaveNet autoencoders. In: Processing of the 34th international conference on machine learning (ICML'17) pp. 1068–1077
86. Hendrix GG, Carbonell JG (1981) A tutorial on natural-language processing. In: Processing of the ACM 1981 annual conference pp. 4–8
87. Wu Z, Li H (2013) Voice conversion and spoofing attack on speaker verification systems. In: Processing of the IEEE Asia-Pacific signal and information processing association annual summit and conference (APSIPA'13) pp. 1–9
88. Radford A, Metz L, Chintala S (2016) Unsupervised representation learning with deep convolutional generative adversarial networks. In: Processing of the 4th international conference on learning representations (ICLR'16)
89. Karras T, Aila T, Laine S, Lehtinen J (2018) Progressive growing of GANs for improved quality, stability, and variation. In: Processing of the 6th international conference on learning representations (ICLR'18)
90. Mirza M, Osindero S (2014) Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*
91. Mao X, Li Q, Xie H, Lau RYK, Wang Z, Smolley SP (2017) Least squares generative adversarial networks. In: Processing of the IEEE international conference on computer vision (ICCV'17) pp. 2813–2821
92. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC (2017) Improved training of wasserstein GANs. In: Processing of the 30th neural information processing systems: annual conference on neural information processing systems (NeurIPS'17) pp. 5767–5777
93. Dosovitskiy A, Springenberg JT, Brox T (2015) Learning to generate chairs with convolutional neural networks. In: Processing of the IEEE conference on computer vision and pattern recognition (CVPR'15) pp. 1538–1546
94. Li D, Chen D, Shi L, Jin B, Goh J, Ng SK (2019) MAD-GAN: multivariate anomaly detection for time series data with generative adversarial networks. In: Processing of the international conference on artificial neural networks (ICANN'19) pp. 703–716
95. Bashar MA, Nayak R (2020) TAnoGAN: time series anomaly detection with generative Adversarial networks. *arXiv:2008.09567v1*
96. Geiger A, Liu D, Alnegheimish S, Cuesta-Infanteet A, Veeramachaneni K (2020) TadGAN: time series anomaly detection using generative adversarial networks. In: IEEE international conference on big data 978-1-7281-6251
97. Schlegl T, Seebock P, Waldstein SM, Schmidt-Erfurth U, Langs G (2017) Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), 10265 LNCS:146–147

98. Zenati H, Foo CS, Lecouat B, Manek G, Chandrasekhar VR (2018) Efficient gan-based anomaly detection. arXiv preprint [arXiv:1802.06222](https://arxiv.org/abs/1802.06222)
99. Akcay S, Atapour-Abarghouei A, Breckon TP (2019) GANomaly: semi-supervised anomaly detection via adversarial training. In: Processing of Asian conference on computer vision (ACCV'19) pp. 622–63
100. Yu L, Zhang W, Wang J, Yu Y (2017) SeqGAN: sequence generative adversarial nets with policy gradient. In: Processing of the 31th AAAI conference on artificial intelligence (AAAI'17) pp. 2852–2858
101. Dong HW, Hsiao WY, Yang LC, Yang YH (2018) MuseGAN: multi-track sequential generative adversarial networks for symbolic music generation and accompaniment. In: Processing of the 32th AAAI conference on artificial intelligence (AAAI'18) pp. 34–41
102. Martin E, Cundy C (2017) Parallelizing linear recurrent neural nets over sequence length. arXiv preprint [arXiv:1709.04057](https://arxiv.org/abs/1709.04057)
103. Huang Y, Huang X, Cai Q (2018) Music generation based on convolution-LSTM. *Comput Inf Sci* 11(3):50–56
104. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780
105. Sutton RS, McAllester DA, Singh SP, Mansour Y (1999) Policy gradient methods for reinforcement learning with function approximation. In: Processing of the 30th neural information processing systems: annual conference on neural information processing systems (NeurIPS'99) pp. 1057–1063
106. Falkowski BJ (2020) Maximum likelihood estimates and a Kernel k-means iterative algorithm for normal mixtures. In: Processing of the IEEE industrial electronics society (IECON'20) pp. 2115–2118
107. Hartmann KG, Schirmer RT, Ball T (2018) EEG-GAN: generative adversarial networks for electroencephalographic (EEG) brain signals. arXiv preprint [arXiv:1806.01875](https://arxiv.org/abs/1806.01875)
108. Arjovsky M, Chintala S, Bottou L (2017) Wasserstein GAN. arXiv preprint [arXiv:1701.07875](https://arxiv.org/abs/1701.07875)
109. Luo Y, Lu BL (2018) EEG data augmentation for emotion recognition using a conditional wasserstein GAN. In: Processing of the 13th IEEE engineering in medicine and biology society (EMBS) pp. 2535–2538
110. Che T, Li Y, Jacob AP, Panwar S, Rad P, Jung TP, Huang Y (2019) Modeling EEG data distribution with a Wasserstein Generative Adversarial Network to predict RSVP Events. arXiv preprint [arXiv:1911.04379](https://arxiv.org/abs/1911.04379)
111. Fiesler E, Caulfield HJ, Choudry A, Ryan JP (1990) Maximum capacity topologies for fully connected layered neural networks with bidirectional connections. In: Processing of the international joint conference on neural networks (IJCNN'90) pp. 827–831
112. Kingma DP, Ba J (2015) Adam: a method for stochastic optimization. In: Processing of the 3th international conference on learning representations (ICLR'15)
113. Weiss A, Yeredor A (2019) A maximum likelihood-based minimum mean square error separation and estimation of stationary Gaussian sources from noisy mixtures. *IEEE Trans Signal Process* 67(19):5032–5045
114. Schalk G, McFarland DJ, Hinterberger T, Birbaumer N, Wolpaw JR (2004) BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans Biomed Eng* 51(6):1034–1043
115. Sun S, Zhou J (2014) A review of adaptive feature extraction and classification methods for eeg-based brain-computer interfaces. In: Processing of the international joint conference on neural networks (IJCNN'14) pp. 1746–1753
116. Bińkowski M, Donahue J, Dieleman S, Clark A, Elsen E, Casagrande N, Cobo LC, Simonyan K (2020) High fidelity speech synthesis with adversarial networks. arXiv preprint [arXiv:1377.11646](https://arxiv.org/abs/1377.11646)
117. Theis L, van den Oord A, Bethge M (2016) A note on the evaluation of generative models. In: Processing of the 4th international conference on learning representations (ICLR'16)
118. Xu Q, Huang G, Yuan Y, Guo C, Sun Y, Wu F, Weinberger KQ (2018) An empirical study on evaluation metrics of generative adversarial networks. arXiv preprint [arXiv:1806.07755](https://arxiv.org/abs/1806.07755)
119. Lucic M, Kurach K, Michalski M, Gelly S, Bousquet O (2017) Are gans created equal? A large-scale study. arXiv preprint [arXiv:1711.10337](https://arxiv.org/abs/1711.10337)
120. Onishi K, Imai H (1997) Voronoi diagram in statistical parametric space by kullback-leibler divergence. In: Processing of the 13th annual symposium on computational geometry (2017) pp. 463–465
121. Weinberg GV, Glenny VG (2016) Optimal rayleigh approximation of the k-distribution via the kullback-leibler divergence. *IEEE Signal Process Lett* 23(8):1067–1070
122. Otto F, Westdickenberg M (2005) Eulerian calculus for the contraction in the wasserstein distance. *SIAM J Math Anal* 37(4):1227–1255
123. Szegedy C, Liu W, Jia Y, Sermanet P, Reed SE, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Processing of the IEEE conference on computer vision and pattern recognition (CVPR'15) pp. 1–9
124. Wang Z, Lyu S, Liu L (2019) Learnable markov chain monte carlo sampling methods for lattice Gaussian distribution. *IEEE Access* 7(2019):87494–87503
125. Kowalczyk A (2000) Sparsity of data representation of optimal kernel machine and leave-one-out estimator. In: Processing of the 31th neural information processing systems: annual conference on neural information processing systems (NeurIPS'99) pp. 252–258
126. Reddi SJ, Ramdas A, Póczos B, Singh A, Wasserman LA (2014) Kernel MMD, the median heuristic and distance correlation in high dimensions. arXiv preprint [arXiv:1406.2083](https://arxiv.org/abs/1406.2083)
127. Korczowski L, Cederhout M, Andreev A, Cattani G, Congedo M (2019) Brain Invaders calibration-less P300-based BCI with modulation of flash duration Dataset (bi2015a). [Research Report] GIPSA-lab. hal-02172347
128. Gurumurthy S, Kiran Sarvadevabhatla R, Venkatesh Babu R (2017) DeLiGAN: generative adversarial networks for diverse and limited data. In: Processing of the IEEE conference on computer vision and pattern recognition (CVPR'17) pp. 4941–4949
129. Bengio Y, Li W (2017) Mode regularized generative adversarial networks. In: Processing of the 5th international conference on learning representations (ICLR'17)
130. Buzano R, Rupflin M (2015) Smooth long-time existence of Harmonic Ricci Flow on surfaces. *J Lond Math Soc* 95(1):277–304

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.