

CTGGAN: Reliable Fetal Heart Rate Signal Generation Using GANs

Zichang Yu^{1,2}, Yating Hu^{1,2}, Yu Lu^{2,*}, Leya Li², Huilin Ge³, Xianghua Fu²

¹ College of Applied Science, Shenzhen University, Shenzhen, China

² College of Big Data and Internet, Shenzhen Technology University, Shenzhen, China

³ College of Automation, Jiangsu University Science and Technology, Zhenjiang, China

* Corresponding author. Email: lvyu@sztu.edu.cn

Abstract—Ensuring fetal health during pregnancy is critically dependent on precise Fetal Heart Rate (FHR) monitoring. A major challenge in this area is the limited availability of labeled FHR data, which poses a barrier to developing reliable automated analysis systems. To address this gap, our study introduces CTGGAN, a novel method employing Generative Adversarial Networks (GANs) to create synthetic, high-quality FHR signals. Specifically, CTGGAN integrates self-attention and residual modules within a Conditional GAN framework, fine-tuned to replicate the complex patterns characteristic of FHR data accurately. A notable feature of CTGGAN is its effective loss function, which combines Wasserstein distance with a gradient penalty to ensure training stability and enhance the authenticity of the generated signals. In performance metrics, Our method demonstrates the highest signal fidelity and distribution similarity, across five key measures: 0.215 maximum mean deviation (MMD), 0.012 sliced Wasserstein distance (SWD), 4.821 percent root mean square difference (PRD), 5.621 relative entropy (RE), and 0.614 Frechet distance (FD). This advancement in generating realistic FHR data with CTGGAN addresses critical issues like data insufficiency and class imbalance, thus advancing the field of prenatal healthcare technology. The code for CTGGAN is available at <https://github.com/ijcnn2024/CTGGAN>.

Index Terms—Generative Adversarial Networks, data augmentation, CTG, Fetal Heart Rate, Fetal Monitoring Systems

I. INTRODUCTION

Fetal Monitoring Systems can be considered as the Medical Decision-Making Systems for Cardiotocography (CTG). These systems play a crucial role in analyzing the data obtained from CTG, which includes fetal heart rate and maternal uterine contractions, to assist healthcare professionals in making informed decisions about maternal and fetal care during pregnancy and labor. These systems are designed to aid in the interpretation of CTG data, detect potential issues, and guide clinical interventions if necessary.

FHR monitoring provides critical insights into the fetus's physiological state and responses to stressors. Expert interpretation plays a key role in early detection of anomalies, which can prevent irreversible fetal damage. However, the subjective nature of FHR tracings interpretation introduces inter-observer variability among different experts, which poses potential risks to fetal health [1], [2]. Consequently, there is an increasing demand for objective and precise FHR analysis methods [3]–[5].

Despite deep learning's potential for automated FHR interpretation [6], [7], the scarcity of labeled FHR datasets for training robust models remains a significant challenge. For instance, the CTU-UHB database, one of the few available, contains only 552 recordings [8]. Addressing this challenge, generative models, particularly generative adversarial networks (GANs), have been explored for generating synthetic FHR signals that closely mimic real data [9]. This approach not only provides a nearly unlimited source of training data but also ensures patient confidentiality. However, it is worth noting that research in FHR signal generation using GANs is still in its early stages.

In response to these challenges, our work introduces CTGGAN, a novel Generative Adversarial Network specifically designed for the generation of synthetic FHR signals. Firstly, it employs conditional generative adversarial networks (CGANs) and incorporates label training, which enables our model to generate samples of various categories. Additionally, it utilizes 1D transposed convolutional layers, which excel at capturing the intricate patterns and high variability presented in FHR signals. Furthermore, CTGGAN incorporates state-of-the-art techniques, such as self-attention mechanisms and residual learning, into its architecture. The self-attention mechanism allows the model to focus on salient features in the data, thereby enhancing the quality and authenticity of the generated signals. On the other hand, residual learning facilitates the training of deeper networks by addressing the vanishing gradient problem, to improve model performance.

Another significant innovation in CTGGAN is its advanced loss function, which combines the Wasserstein distance with a gradient penalty. This not only enhances the stability of the training process but also ensures a higher diversity and quality in the generated data [10], [11]. Such a loss function is particularly effective in GANs, addressing common challenges like mode collapse and training instability. Our extensive experimental evaluations demonstrate that CTGGAN can generate FHR signals of superior fidelity and diversity, surpassing existing state-of-the-art methods [9]. This makes CTGGAN a powerful tool in addressing the critical issues of data scarcity and class imbalance in FHR research and paves the way for more advanced, data-driven approaches in prenatal care and monitoring.

II. RELATED WORK

The acquisition of large datasets is a cornerstone in enhancing machine learning and deep learning algorithms, particularly in imbalanced and scarce data scenarios. The evolution of deep learning has intensified the focus on data augmentation techniques across various domains including biosignals [12], images [13], text [14], and speech [15]. Data augmentation has undergone significant developments, notably:

Oversampling Methods: Widely used in computer vision, techniques such as random flipping, translation, cropping, and affine transformations have effectively addressed class imbalance [16]. In biosignal contexts like FHR, EEG, and sleep stages, overcoming class imbalance is essential for robust data generation and analysis [17].

Mathematical Methods: Clifford's team introduced a non-linear model for one-dimensional time sequences [18] and later, an artificial vector methodology [19]. These methods, while innovative, often result in uniformity, limiting diverse data exploration. Each simulated sample relies on specific equations, making varied sample generation complex [20], [21]. Recent semi-supervised methods like those proposed by Dabrowski et al. [22] aim to generate accurate and varied biosignal samples, surpassing traditional mathematical constraints.

Deep Learning Methods: Approaches like variational autoencoders (VAEs) and Generative Adversarial Networks (GANs) have emerged as effective for generative modeling. GANs, in particular, are renowned for their generative efficiency, inspiring enhancements such as Deep Convolutional GANs (DCGAN) [23], Wasserstein GANs [24], CycleGANs [25], SeqGANs [26], Conditional GANs (CGAN) [27], and Augmented CycleGANs [28]. These GAN variants have been successful in generating time series data for biosignals.

The application of GANs in FHR signal generation is less explored. Unlike other biomedical signals, FHR signals, characterized by complex morphologies and high variability, present unique challenges. To date, Zhang et al. [9] have reported on FHR signal generation, which is FHRGAN, noting gaps between generated and real FHR signals. Our proposed method also addresses this gap, using conditional adversarial networks combined with self-attention mechanisms and residual modules for high-quality synthetic FHR signal generation.

Additionally, work on synthetic CTG data generation using conditional GANs has been done by Ahmed et al. [29], who focused on generating feature records from a CTG feature dataset. In contrast, our approach, alongside FHRGAN, concentrates on generating complete FHR signals, offering a novel contribution to the field.

III. METHODS

A. Overview of CTGGAN

CTGGAN is an innovative adaptation of the classical Generative Adversarial Network (GAN) [30], specifically tailored for Fetal Heart Rate (FHR) signal analysis. GANs are a sophisticated class of AI algorithms used in unsupervised machine

learning, comprising two neural networks: the Generator (G) and the Discriminator (D). In a zero-sum game, G aims to generate data that deceives D , while D learns to distinguish between real and generated data.

The discriminator in CTGGAN plays a crucial role in assessing both the authenticity and label categories of $G(z)$ and real samples, which is integral to the adversarial training process. This not only augments the generator's and discriminator's capabilities but also ensures the generation of high-quality FHR signals that closely resemble real-world data.

CTGGAN is illustrated in Fig. 1. It uses two-dimensional transposed convolutions, our CTGGAN model employs one-dimensional transposed convolutions. This choice is tailored to the specific nature of FHR signals, allowing for more effective feature extraction and signal generation. Additionally, we have integrated a self-attention mechanism and residual modules into our model.

Compared to FHRGAN, CTGGAN is improved in two key aspects: (1) Architecture for label classification: Unlike FHRGAN, which employs auxiliary classifiers for signal label classification, CTGGAN adopts a conditional GAN architecture to generate signals with different labels. This strategic choice enhances overall coherence and effectiveness in label classification. In CTGGAN, the generator's input, z , integrates a noise vector (sourced from a uniform or Gaussian distribution) and the signal's label category (expressed through one-hot encoding). This combination allows G to generate signals influenced by the noise vector, while $G(z)$ strives to mimic the original data distribution closely. (2) Convolutional layer: Another notable feature of CTGGAN is its use of one-dimensional transposed convolutions combined with a self-attention mechanism and residual modules. This change is particularly effective for FHR signals, aiding in more efficient feature extraction and signal generation.

The self-attention mechanism enhances the model's ability to focus on salient features in the data, capturing long-range dependencies more effectively. Meanwhile, the residual modules ensure smooth training by allowing gradients to flow through the network, preventing the vanishing gradient problem and enhancing the overall learning capacity. Finally, in terms of the loss function, the Wasserstein distance and gradient penalty were introduced to enhance the training stability of GAN networks, and improvements were made to the them.

The objective function of a CTGGAN can be represented as:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

Where:

- \mathbb{E} denotes the expected value.
- x are real samples drawn from the true data distribution $p_{\text{data}}(x)$.
- z are noise samples drawn from the noise distribution $p_z(z)$.

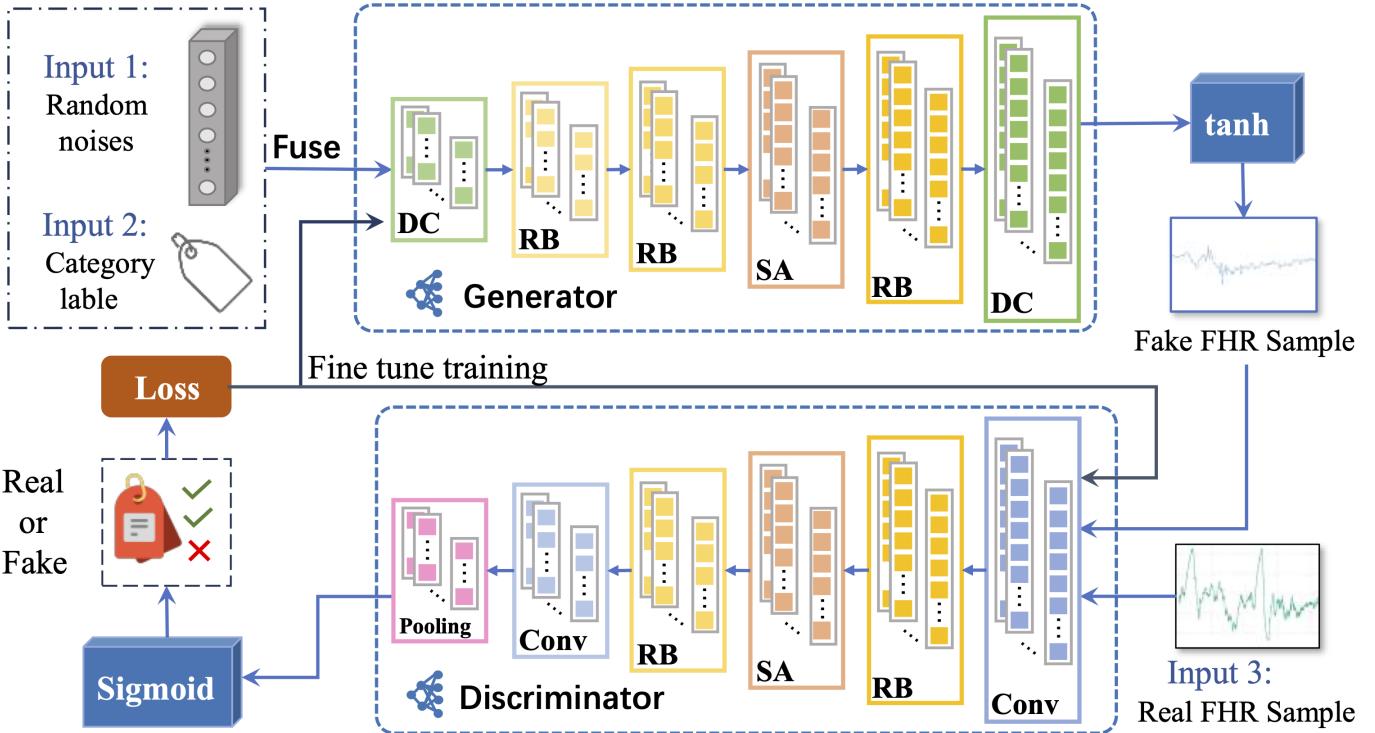


Fig. 1: Architecture of CTGGAN for FHR signal generation.

- $D(x)$ is the discriminator's estimate of the probability that a real data instance x is real.
- $G(z)$ represents the data generated by the generator from noise samples z .

B. Self-Attention and Residual Blocks

Recently, Zhang et al. [31] introduced the self-attention mechanism to GANs, and their model enables attention-driven, long-range dependency modeling for image generation tasks. We apply self-attention mechanism to one-dimensional FHR signal generation, which promises enhanced modeling of complex dependencies and improved performance. Our model also incorporates residual blocks, which can improve training efficiency. These designs are key to effectively capturing complex dependencies, thereby elevating overall model performance.

1) *Self-Attention Mechanism:* The Self-Attention mechanism within our model is strategically designed to enhance focus on specific regions of the input feature map. This design enables the network to effectively capture and process long-range dependencies, which are crucial for understanding complex data patterns. The inner workings of this self-attention block are comprehensively illustrated in Fig. 2.

In operation, the mechanism follows these steps:

- It begins by generating a query (q), key (k), and value (v) from the input feature map. This is achieved through the application of 1D convolution operations, allowing for a transformation of the input data into these three distinct components.

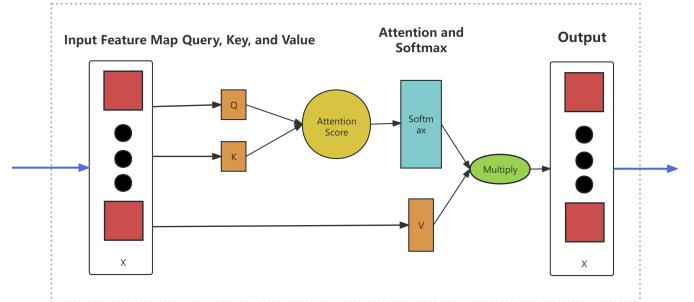


Fig. 2: Detailed Architecture of the Self-Attention Block in CTGGAN.

- Subsequently, attention weights are computed. This involves executing a matrix multiplication between the query and key, which is then followed by applying a softmax function. This process effectively assesses the degree of relevance or 'attention' that different parts of the input should receive.
- The final output is crafted by combining these attention weights with the value. This output is carefully scaled and then reintegrated with the original input. This fusion results in the enriched output of the self-attention layer, which now embodies both the original data and the additional insights gleaned from the self-attention process.

2) *Residual Blocks:* The Residual Blocks in the model consist of sequences of 1D transposed convolutions, batch

normalization, and activation functions. The structure of a Residual Block is as follows:

- Each block contains two main paths: the main convolutional path and a shortcut (residual) path.
- The main path includes a series of convolutions and non-linear activations.
- The residual path provides a shortcut for the input to flow through, bypassing the main convolutional operations.
- The outputs of both paths are added together to form the final output of the block.

The combination of self-attention and residual blocks allows the model to effectively learn complex patterns in the data, making it particularly suitable for tasks requiring high-level feature extraction and representation learning.

C. Wasserstein GAN with Gradient Penalty Term

A major limitation of conventional GANs is the instability of the discriminator during training. This instability often leads to the discriminator collapsing, recognizing only a narrow subset of modes in the data distribution as real. Consequently, this restriction on the discriminator's capabilities impacts the diversity of the generator's output. Recent advancements, particularly the WGAN-GP loss function with gradient penalty, have significantly enhanced training stability and sample quality [8]. The incorporation of gradient penalties in WGAN-GP effectively addresses issues like mode collapse and improves convergence properties.

The key contribution of Wasserstein GAN (WGAN) lies in its adoption of the Wasserstein distance as a loss metric, which distinguishes between real and generated data distributions. This metric replaces the Jensen-Shannon divergence used in traditional GANs, providing a more meaningful and smoother gradient landscape, thus enhancing training stability [32]. The Wasserstein distance W between two probability distributions, P_r (real data distribution) and P_g (generated data distribution), is defined as:

$$W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} \mathbb{E}_{(x,y) \sim \gamma} [\|x - y\|] \quad (2)$$

Where:

- P_r is the real data distribution.
- P_g is the generated data distribution.
- $\Pi(P_r, P_g)$ denotes the set of all joint distributions $\gamma(x, y)$ that have marginals P_r and P_g .
- x is a sample from P_r .
- y is a sample from P_g .
- $\|x - y\|$ represents the distance between samples x and y .

In training, the discriminator D minimizes the Wasserstein distance, fixed by θ_G^* :

$$L_D(\theta_D, \theta_G^*) = W(P_r, P_f) \quad (3)$$

Similarly, the generator G is trained to minimize the Wasserstein distance:

$$\begin{aligned} L_G(\theta_G, \theta^* D) &= \mathbb{E}_{x_f \sim P_f} [D_{\theta^* D}(x_f)] \\ &= \mathbb{E}_{x_f \sim P_f} [D_{\theta_D}(G_{\theta_G^*}(z))] \end{aligned} \quad (4)$$

Initially, WGAN applied weight clipping to the discriminator weights within $[-c, c]$ to enforce the K-Lipschitz constraint. This approach improved stability, however, sometimes led to suboptimal samples or convergence issues [8].

WGAN-GP resolves this by imposing Lipschitz continuity on the discriminator more effectively. It includes a gradient penalty in the loss function, as follows:

$$\begin{aligned} W(P_r, P_f) &= \bar{W}(P_r, P_f) \\ &\quad + \lambda \mathbb{E}_{\hat{x} \sim \hat{P}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \end{aligned} \quad (5)$$

Where \hat{x} represents samples along linear interpolations between P_r and P_f , and λ is a hyperparameter that balances the original WGAN loss and the gradient penalty.

In addition to these improvements, our training also incorporates spectral normalization for the discriminator and feature matching for the generator's loss. Spectral normalization stabilizes the discriminator by constraining its Lipschitz constant, enhancing overall training dynamics. Feature matching loss, meanwhile, encourages the generator to produce samples with feature distributions that closely match those of real data. This is achieved by minimizing the difference in feature representations between real and generated samples at an intermediate layer of the discriminator. These enhancements further stabilize GAN training and improve the fidelity and diversity of the generated samples.

The improvement of the loss function for the generator includes a feature matching component, which is mathematically represented as follows:

$$L_G = -\mathbb{E}_{\tilde{x} \sim P_g} [D(\tilde{x})] + \alpha \cdot \|F_r - F_g\|_2^2 \quad (6)$$

In this equation:

- L_G represents the loss function of the generator.
- \tilde{x} denotes the samples generated by the generator, following the distribution P_g .
- $D(\tilde{x})$ is the discriminator's output for the generated samples, indicating the probability of the sample being real.
- The term $\alpha \cdot \|F_r - F_g\|_2^2$ is the feature matching component, where:
 - α is a weighting factor that balances the importance of the feature matching component in the overall loss function.
 - $\|F_r - F_g\|_2^2$ is the squared L2 norm, which measures the distance between feature representations of real samples (F_r) and generated samples (F_g) extracted from an intermediate layer of the discriminator.

This feature matching term is designed to minimize the difference in feature representations between real and generated samples, guiding the generator to produce more realistic samples.

D. Detailed Design of the Network Structure

In our proposed model, we made specific modifications to the GAN. Unlike FHRGAN, which used an auxiliary classifier to discriminate the labels, we used a conditional

generative adversarial network to perform label classification and feedback training. Furthermore, we did not use a two-dimensional convolutional network, but a one-dimensional transposed convolutional network. This modification aids in better capturing the sequential dependencies in FHR signals. Additionally, we used smaller kernel size (1x4) and each transposed convolution layer was equipped with a ReLU activation and batch normalization, which helps in emphasizing the contrast between adjacent features. The design of our GAN model has been meticulously tailored to address the specific intricacies involved in FHR signal generation. Table I presents the detailed configuration of our model.

TABLE I: Architectural Details of the Generator and Discriminator for FHR Signal Data Augmentation. Note: LReLU is LeakyReLU with an alpha value of 0.2; BN: Batch Normalization; Conv denotes convolution; SA: Self-Attention Block; RB: Residual Block. FC: Fully Connected Layer.

Generator		
Layer	Structure	Output Shape
Input	-	(3, 100)
DC1	Conv1d-ReLu-BN	(256, 125)
RB2	(Conv1d-ReLu-BN)*2	(128, 250)
SA3	Self-Attention	(128, 250)
RB4	Conv1d-ReLu-BN	(64, 500)
RB5	Conv1d-ReLu-BN	(32, 1000)
DC6	Conv1d-Tanh	(1, 1000)

Discriminator		
Input	-	(3, 1000)
Conv1	Conv1d-LReLu	(32, 1000)
RB2	Conv1d-LReLu-BN	(64, 500)
SA3	Self-Attention	(64, 500)
RB4	Conv1d-LReLu-BN	(128, 250)
RB5	Conv1d-LReLu	(256, 125)
FC6	Pooling-Sigmoid	(1, 1)

IV. EXPERIMENT

A. Configurations

The model training and evaluation were performed on a workstation with 2 Intel Xeon Silver 4210 CPUs, 128GB DDR4 memory and Nvidia RTX 4090 GPU. The operating system is Windows 10. The implementations were based on Python 3.11.4, PyTorch 2.0, cuda 11.8, and other common deep learning libraries.

B. Dataset and Pre-processing

The CTU-UHB dataset [8] contains 552 CTG recordings selected from 9164 collected between 2010-2012. CTG recordings include FHR and uterine contraction signals sampled at 4Hz over a 90-minute period. The neonatal umbilical artery pH is a gold standard for evaluating fetal status [33]. pH values ≥ 7.15 are considered normal, while values ≤ 7.15 are deemed pathological, resulting in 447 normal and 105 pathological FHRs.

The study used 200 real FHR samples, comprising 100 normal and 100 pathological. A ratio of 8:2 was used for dividing these samples into training and testing sets. Data

pre-processing is crucial for noise reduction and enhancing the quality of fetal heart rate data, significantly improves performance in evaluation. The preprocessing involved the following steps:

- Discarding data where FHR was zero for > 15 s, or FHR was < 50 bpm or > 200 bpm.
- Replacing abnormalities with Hermite interpolation.

The impact of data preprocessing on FHR signals is depicted in Fig. 3.

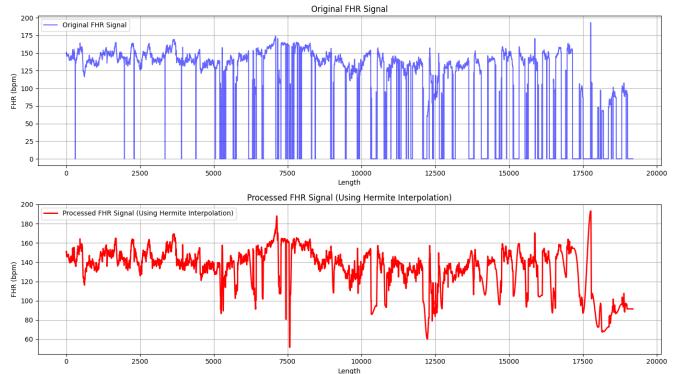


Fig. 3: Comparison of FHR Signals Before and After Data Preprocessing.

C. Training Process

We developed the proposed Wasserstein GAN with gradient penalty in PyTorch. The model's noise input vector was set to a length of 100. Training was conducted over 150 epochs using the ADAM optimizer, set at a learning rate of 0.00001. The gradient penalty coefficient was set to 10, and a batch size of 32 was maintained. After training, the generator model is capable of accepting random noise vectors and generating synthetic FHR signals corresponding to various category labels. The comprehensive training procedure of our model is outlined in Algorithm 1.

D. Evaluation Indicators

To evaluate the quality and diversity of the generated FHR signals, we adopted quantitative metrics, and we also conducted visual assessments by clinical experts.

1) *Qualitative Visual Inspection*: In the realm of GAN models, quantitative visual inspection of the synthesized data is regarded as an intuitive method [34]. We create visual representations to access the quantity and accuracy of these signals in comparison to the ground truth signals, followed by soliciting expert evaluations for signal assessment.

2) *Signal Fidelity*: To assess the quality of generated FHR signals, we introduce relative entropy(RE) and the Frechet distance(FD) metrics for similarity between real $x \sim P_{\text{data}}$ and generated $z \sim P_z$ FHR samples. Lower RE and FD values indicate better signal quality.

$$RE_{(x||z)} = \sum_{i \in \mathcal{X}} x_i \log \frac{X_i}{z_i} \quad (7)$$

Algorithm 1 CTGGAN Training Procedure

Require: Real FHR data samples from CTG dataset $x \sim P_{\text{data}}$ and latent vectors z of dimension 100

Ensure: Fully trained generator model G_h

- 1: Initialize gradient penalty coefficient $\lambda = 10$, epochs $n_{\text{epochs}} = 150$, batch size $m = 32$, critic-to-generator iteration ratio $n_{\text{critic}} = 5$, Adam parameters: $\alpha = 0.0001$, $\beta_1 = 0$, $\beta_2 = 0.9$. Initialize generator G and discriminator D parameters using a truncated normal distribution, variance 0.02
- 2: **for** $j \leftarrow 1$ to n_{epochs} **do**
- 3: **for** $t \leftarrow 1$ to n_{critic} **do**
- 4: **for** each m -sized batch **do**
- 5: Sample real FHR data $x \sim P_{\text{data}}$ and noise vectors $z \sim \mathcal{N}(0, 1)$
- 6: $\hat{x} \leftarrow G(z)$
- 7: Calculate discriminator loss L_D from real and generated data
- 8: Update discriminator D via Adam optimizer with α, β_1, β_2
- 9: **end for**
- 10: $\omega \leftarrow \text{Adam}(\nabla \frac{1}{m} \sum_{i=1}^m L_D, \omega, \alpha, \beta_1, \beta_2, \varepsilon)$
- 11: **end for**
- 12: Sample a batch of noise matrices $z^{(i)}_{i=1} \sim \mathcal{N}(0, 1)$
- 13: Determine generator loss L_G
- 14: Update generator G using Adam optimizer and parameters α, β_1, β_2
- 15: **end for**
- 16: **Output:** Store the trained CTGGAN model G_h

$$FD(x, z) = \min_{\alpha, \beta} \|D\|, D = \max_t \|X(\alpha(t)) - Z(\beta(t))\| \quad (8)$$

where D represents the Euclidean distance, which is the length of the sequence.

3) *Distribution Similarity*: Unlike FHRGAN, which aims for signal consistency, our model uses random seeds to produce signals with greater diversity and fidelity. We evaluate the generated and original distributions using three metrics: maximum mean deviation (MMD) [35], sliced Wasserstein distance (SWD) [11], and percent root mean square difference (PRD) [36]. Lower values indicate closer distribution similarity.

V. RESULT

A. Training performance

In this study, we embarked on training real data using various models, with an emphasis on assessing the performance of five GAN-based architectures: DCGAN, SeqGAN, CGAN, FHRGAN, and our CTGGAN model. The key evaluation metrics were convergence speed and stability oscillation in GANs. Using a dataset of 160 training samples, we analyzed the loss function trajectories of each model, as depicted in Figures 4 and 5.

The results indicate that FHRGAN and CTGGAN initiated with considerably lower loss values than DCGAN, SeqGAN, and CGAN, demonstrating smoother and more consistent

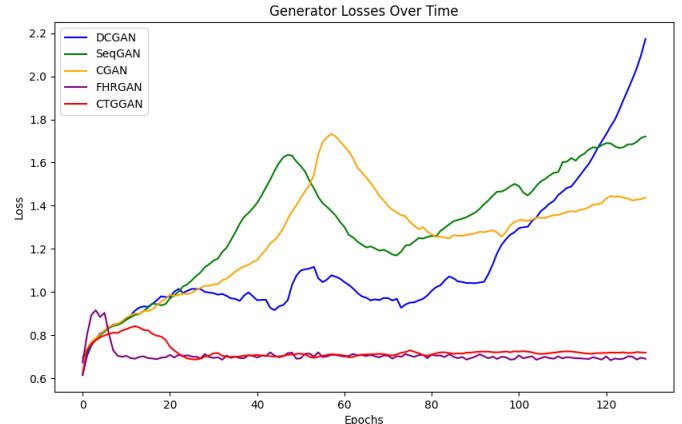


Fig. 4: Comparative Generator Loss across Different Models.

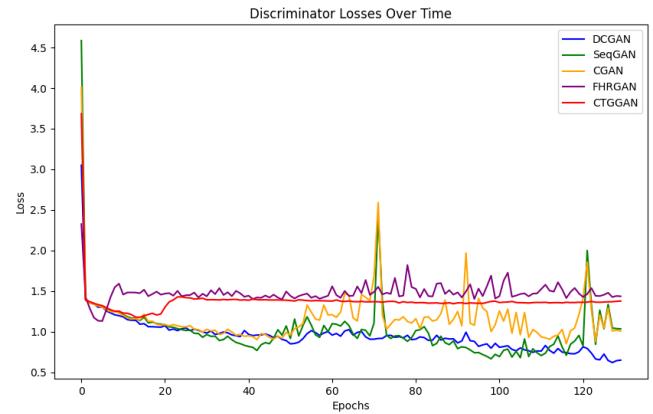


Fig. 5: Comparative Discriminator Loss across Different Models

initial performance over 30 epochs. This pattern highlighted common challenges in traditional GAN models like gradient vanishing and mode collapse, which hinder the generator's ability to produce varied samples or compel it to generate specific types to fool the discriminator. The use of Wasserstein distance in FHRGAN and CTGGAN contributed to their training stability.

We noted that DCGAN, SeqGAN, and CGAN exhibited the slowest convergence rates, indicating potential weaknesses in their training effectiveness, as further evidenced by our analysis of the generated signal samples. After 60 training epochs, their loss trajectories suggested modal breakdown, raising concerns about their suitability for FHR signal generation.

While FHRGAN showed rapid initial convergence, it experienced slight oscillations thereafter, indicating room for improvement in handling complex FHR signals. In contrast, CTGGAN demonstrated remarkable stability after a slower convergence, with losses of both generator and discriminator trending towards equilibrium, indicating effective model operation. In conclusion, our study suggests that while FHRGAN maintains high efficiency, CTGGAN exhibits stronger stability,

a vital attribute for successful GAN network training.

B. Evaluation of Signal Generation Fidelity

1) *Expert Assessment of Generated FHR Signals:* As depicted in Fig. 6, the FHR signal generated by our GAN model demonstrates a high degree of fidelity. This observation is further validated by specialist evaluations, underscoring the robustness of our model in generating accurate FHR signals.

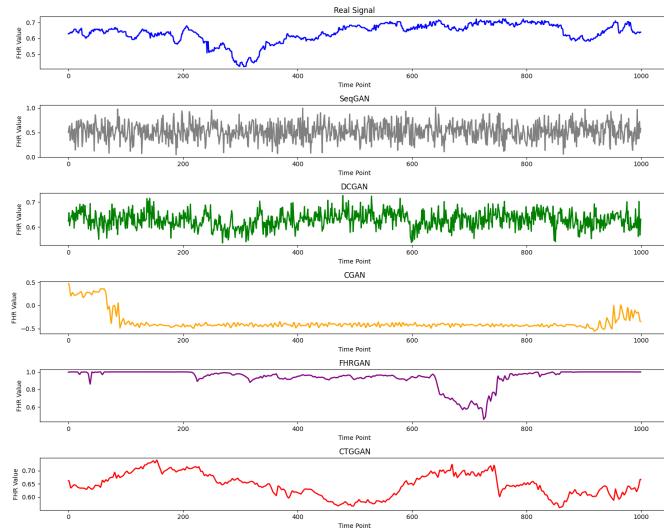


Fig. 6: Comparison of FHR Signals Generated by Various GAN Models.

2) *Quantitative Analysis: Reconstruction Error (RE) and Fidelity Distance (FD):* Due to the limited research in FHR signal generation, we adopted a comprehensive comparison approach similar to that of FHRGAN. Each model was applied to the test dataset, resulting in 60 sets of normal samples. Table II summarizes the average scores for RE and FD metrics. Our model notably outperforms others, achieving the highest scores in both RE and FD categories, indicating superior fidelity and detail in the generated signals.

TABLE II: Comparative Analysis of Quality Metrics for Normal FHR Signals Generated by Different Models.

Model	Testing data	RE	FD
SeqGAN		10.324	0.974
DCGAN		9.532	0.795
CGAN	60 samples	9.012	0.763
FHRGAN		7.059	0.682
CTGGAN (Ours)		5.621	0.614

C. Distribution Similarity

Table III presents a summary of the distribution similarity results, comparing real and synthesized FHR data across various GAN architectures. Notably, our model, CTGGAN, consistently achieves the lowest scores in all evaluated metrics, indicating a substantial enhancement in distribution similarity compared to other GANs. In contrast to the approach by

Zhang et al., who concentrated on signal generation for specific categories and labels, our methodology leveraged random seeds to foster greater diversity in signal generation. This approach has yielded signals that not only demonstrate enhanced diversity but also exhibit a closer similarity to real FHR data distributions, aligning with visual inspection.

TABLE III: Comparative Evaluation of Distribution Similarity Metrics for FHR Signals Generated by Different GAN Models.

Model	MMD	SWD	PRD
SeqGAN	6.527	0.541	10.635
DCGAN	5.267	0.254	10.431
CGAN	5.022	0.254	9.431
FHRGAN	1.975	0.086	5.437
CTGGAN (Ours)	0.215	0.012	4.821

D. Discussion

The results from our study validate the exceptional capability of our model in generating high-fidelity, diverse FHR signals. Expert evaluations have verified the morphological resemblance of these synthesized signals to actual FHR traces. Notably, our model outperforms others in achieving the lowest scores in RE, FD, MMD, SWD, and PRD metrics. This indicates a superior fidelity and closer distribution similarity to real FHR data compared to competing models. Collectively, both qualitative and quantitative assessments underscore our model's adeptness in replicating the essential visual and statistical characteristics of real FHR signals.

Nevertheless, certain limitations are present. Although our model effectively captures the overall variability, it falls short in emulating some highly specific characteristics of FHR signals. Additionally, the potential for further enhancing the model's generalization capabilities exists, particularly through the utilization of a larger, real-world dataset. Despite these challenges, our work signifies a promising stride in the application of GANs for FHR signal generation, holding substantial promise for augmenting automated FHR monitoring in the future.

VI. CONCLUSION

This paper presented CTGGAN, a novel model for synthetic FHR signal generation. Integrating a customized Self-Attention mechanism, Residual network architecture, and an optimized loss function, CTGGAN achieves the SOTA performance in producing FHR traces with high fidelity. Our extensive evaluations show that CTGGAN outshines existing GAN models in mimicking real FHR signal characteristics, offering superior sample quality and diversity while maintaining stable training. This research contributes significantly to overcoming data scarcity and class imbalance in FHR research and clinical practice. CTGGAN's ability to generate high-quality, diverse FHR data on demand enriches datasets for robust FHR monitoring systems and advances computer-aided FHR analysis, marking a notable step forward in the field.

REFERENCES

- [1] D Ayres-de Campos, J Bernardes, A Costa-Pereira, and L Pereira-Leite, "Inconsistencies in classification by experts of cardiotocograms and subsequent clinical decision," *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 106, no. 12, pp. 1307–1310, 1999.
- [2] Mark E O'Sullivan, Elizabeth C Considine, Mairead O'Riordan, William P Marnane, JM Rennie, and Geraldine B Boylan, "Challenges of developing robust AI for intrapartum fetal heart rate monitoring," *Frontiers in Artificial Intelligence*, vol. 4, pp. 765210, 2021.
- [3] Cang Chen, Weifang Xie, Zhiqi Cai, and Yu Lu, "Deep Learning for Cardiotocography Analysis: Challenges and Promising Advances," in *ICIC*. Springer, 2023, pp. 354–366.
- [4] Huanwen Liang and Yu Lu, "A CNN-RNN unified framework for intrapartum cardiotocograph classification," *Computer Methods and Programs in Biomedicine*, vol. 229, pp. 107300, 2023.
- [5] Giulio Steyde, Edoardo Spairani, Giovanni Magenes, and Maria G Signorini, "Fetal heart rate spectral analysis in raw signals and PRSA-derived curve: normal and pathological fetuses discrimination," *Medical & Biological Engineering & Computing*, vol. 62, no. 2, pp. 437–447, 2024.
- [6] Zhidong Zhao, Jiawei Zhu, Pengfei Jiao, Jinpeng Wang, Xiaohong Zhang, Ximiao Lu, and Yefei Zhang, "Hybrid-FHR: a multi-modal ai approach for automated fetal acidosis diagnosis," *BMC Medical Informatics and Decision Making*, vol. 24, no. 1, pp. 19, 2024.
- [7] Yu Lu, Huanwen Liang, Zichang Yu, and Xianghua Fu, "MT-1DCG: A Novel Model for Multivariate Time Series Classification," in *ICIC*. Springer, 2023, pp. 222–234.
- [8] Václav Chudáček, Jiří Spilka, Miroslav Burša, et al., "Open access intrapartum CTG database," *BMC Pregnancy and Childbirth*, vol. 14, pp. 1–12, 2014.
- [9] Yefei Zhang, Zhidong Zhao, Yanjun Deng, and Xiaohong Zhang, "FHRGAN: Generative adversarial networks for synthetic fetal heart rate signal generation in low-resource settings," *Information Sciences*, vol. 594, pp. 136–150, 2022.
- [10] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville, "Improved Training of Wasserstein GANs," *NeurIPS*, vol. 30, 2017.
- [11] Julien Rabin, Gabriel Peyré, Julie Delon, and Marc Bernot, "Wasserstein Barycenter and Its Application to Texture Mixing," in *SSVM 2011*. Springer, 2012, pp. 435–446.
- [12] Laurenz Berger, Max Haberbusch, and Francesco Moscato, "Generative adversarial networks in electrocardiogram synthesis: Recent developments and challenges," *Artificial Intelligence in Medicine*, p. 102632, 2023.
- [13] Beijing Chen, Xingwang Ju, Bin Xiao, Weiping Ding, Yuhui Zheng, and Victor Hugo C de Albuquerque, "Locally GAN-generated face detection based on an improved Xception," *Information Sciences*, vol. 572, pp. 16–28, 2021.
- [14] Lauri Juvela, Bajibabu Bollepalli, Junichi Yamagishi, and Paavo Alku, "Waveform Generation for Text-to-speech Synthesis Using Pitch-synchronous Multi-scale Generative Adversarial Networks," in *ICASSP*. IEEE, 2019, pp. 6915–6919.
- [15] Zhongjian Qi, Jun Sun, Jinzhao Qian, Jiajia Xu, and Shu Zhan, "PCCM-GAN: Photographic Text-to-Image Generation with Pyramid Contrastive Consistency Model," *Neurocomputing*, vol. 449, pp. 330–341, 2021.
- [16] Paula Branco, Luís Torgo, and Rita P Ribeiro, "A Survey of Predictive Modeling on Imbalanced Domains," *ACM Computing Surveys*, vol. 49, no. 2, pp. 1–50, 2016.
- [17] Le Yu, Peiwang Tang, Zhiguo Jiang, and Xianchao Zhang, "Denoise Enhanced Neural Network with Efficient Data Generation for Automatic Sleep Stage Classification of Class Imbalance," in *IJCNN*. IEEE, 2023, pp. 1–8.
- [18] GD Clifford and PE McSharry, "Generating 24-hour ECG, BP and respiratory signals with realistic linear and nonlinear clinical characteristics using a nonlinear model," in *Computers in Cardiology*. IEEE, 2004, pp. 709–712.
- [19] Gari D Clifford, Shamim Nemati, and Reza Sameni, "An artificial vector model for generating abnormal electrocardiographic rhythms," *Physiological Measurement*, vol. 31, no. 5, pp. 595, 2010.
- [20] N Jafarnia-Dabanloo, DC McLernon, H Zhang, A Ayatollahi, and V Johari-Majd, "A modified Zeeman model for producing HRV signals and its application to ECG signal generation," *Journal of Theoretical Biology*, vol. 244, no. 2, pp. 180–189, 2007.
- [21] Ebadollah Kheirati Roonizi and Reza Sameni, "Morphological modeling of cardiac signals based on signal decomposition," *Computers in Biology and Medicine*, vol. 43, no. 10, pp. 1453–1461, 2013.
- [22] Matthieu Dabrowski, José Mennesson, Jérôme Riedi, and Chaabane Djerafa, "Semi-supervised GAN with sparse ground truth as Boundary Conditions," in *IJCNN*. IEEE, 2023, pp. 1–9.
- [23] Alec Radford, Luke Metz, and Soumith Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," *arXiv preprint arXiv:1511.06434*, 2015.
- [24] Ming Zheng, Tong Li, Rui Zhu, Yahui Tang, Mingjing Tang, Leilei Lin, and Zifei Ma, "Conditional wasserstein generative adversarial network-gradient penalty-based approach to alleviating imbalanced data classification," *Information Sciences*, vol. 512, pp. 1009–1023, 2020.
- [25] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros, "Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks," in *ICCV*, 2017, pp. 2223–2232.
- [26] Lantao Yu, Weinan Zhang, Jun Wang, and Yong Yu, "SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient," in *AAAI*, 2017.
- [27] Mehdi Mirza and Simon Osindero, "Conditional Generative Adversarial Nets," *arXiv preprint arXiv:1411.1784*, 2014.
- [28] Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu, Tie-Yan Liu, and Wei-Ying Ma, "Dual Learning for Machine Translation," *NeurIPS*, vol. 29, 2016.
- [29] Halal Abdulrahman Ahmed, Juan A Nepomuceno, Belén Vega-Márquez, and Isabel A Nepomuceno-Chamorro, "Generating Synthetic Fetal Cardiotocography Data with Conditional Generative Adversarial Networks," in *SOCO*. Springer, 2023, pp. 111–120.
- [30] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.
- [31] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena, "Self-Attention Generative Adversarial Networks," in *ICML*. PMLR, 2019, pp. 7354–7363.
- [32] Martin Arjovsky, Soumith Chintala, and Léon Bottou, "Wasserstein Generative Adversarial Networks," in *ICML*. PMLR, 2017, pp. 214–223.
- [33] Claude Racinet, Paul Ouellet, Jonathan Muraskas, and Thierry Daboval, "Neonatal cord blood eucapnic pH: A potential biomarker predicting the need for transfer to the NICU," *Archives de Pédiatrie*, vol. 27, no. 1, pp. 6–11, 2020.
- [34] Ali Borji, "Pros and cons of GAN evaluation measures: New developments," *Computer Vision and Image Understanding*, vol. 215, pp. 103329, 2022.
- [35] Karsten M Borgwardt, Arthur Gretton, Malte J Rasch, Hans-Peter Kriegel, Bernhard Sch"olkopf, and Alex J Smola, "Integrating structured biological data by Kernel Maximum Mean Discrepancy," *Bioinformatics*, vol. 22, no. 14, pp. e49–e57, 2006.
- [36] Debapriya Hazra and Yung-Cheol Byun, "SynSigGAN: Generative Adversarial Networks for Synthetic Biomedical Signal Generation," *Biology*, vol. 9, no. 12, pp. 441, 2020.