

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

# Filling of Fetal Heart Rate Signal: Diffusion model based on Dimension Construction and Period Segmentation

Zhixin Zhou, Zhidong Zhao, *Member, IEEE*, Yefei Zhang, Yanjun Deng, Shuai Wang, Huaming Wu, *Senior Member, IEEE* and Pengfei Jiao, *Member, IEEE*

**Abstract**—Fetal Heart Rate (FHR) signal is widely used in doppler fetal heart monitors. However, incomplete FHR signals reduce the effectiveness of fetal heart rate monitoring. Filling missing data is a key technique to improve FHR quality, but existing filling algorithms lack consideration the correlation of FHR signals. Therefore, we focus on two correlations related to FHR, and propose a filling algorithm called Diffusion model for Missing Data Imputation in FHR (DMDI-FHR). Firstly, we construct the Dual-dimensional Sample Construction (DDSC) method that finds two FHR signals with maximizing the difference to form correlations between them. Secondly, the Multi-period Decomposition (MPD) method is introduced to obtain the internal correlation of FHR signal. Finally, DMDI-FHR algorithm controls the filling process based on the diffusion model. Experimental results demonstrate the performance of DMDI-FHR algorithm, which provides an effective way to improve the quality of FHR signal.

**Index Terms**—Consumer electronic, fetal heart rate, missing data, deep learning, diffusion model

## I. INTRODUCTION

MEDICAL monitors are common consumer electronic devices [1]. They have been widely used in recent years because of the advantages of non-invasive, real-time and low cost [2]. These devices are mainly used to collect different types of medical data [3]. In gynecology, doppler monitors can record Fetal Heart Rate (FHR) signals to reflect the state of fetus in utero [4]. With the development of computer technology, many hospitals and scholars try to analysis FHR signals and output auxiliary diagnosis results of fetal state in doppler fetal heart monitors.

For example, Dash et al. [5] proposed a novel FHR signal classification algorithm based on generative models and Bayesian theory, which incorporates an expert system and SVM to build the classifier. Fuentealba et al. [6] found that time and frequency variations in FHR signals are related to fetal status and

proposed a new method. The new method extracts time and frequency domains, which decomposes the FHR signal into several parts and computes the Time-Varying Autoregressive spectrum. Alsaggaf et al. [7] combines three existing algorithms in an organized way to increase accuracy in detecting fetal hypoxia.

Most smart electronic devices need continuous data [8]. However, missing data may occur in fetal monitoring. Missing data can reduce the variability of FHR signal, which may impact the performance of FHR classification algorithms.

Thus, continuous FHR signal is a prerequisite for fetal auxiliary diagnosis. Some scholars have done some studies on filling in FHR signals. Przybyła et al. [9] propose an embedded space to rebuild the missing parts of FHR. In rebuilding process, they calculate the missing data through K-Nearest Neighbor (KNN). Feng et al. [10] attempted to rebuild the missing data in FHR signals based on Gaussian process (GP).

Recent years, scholars have proposed more filling algorithms. In 2013, Oikonomou et al. [11] presented an algorithm to rebuild the FHR missing with two steps. In first step, compute an estimate of missing data by empirical dictionary. These computed values are updated in the second step. Repeat these two steps to get the best filling results. In 2017, Frigo et al. [12] reduced the computational requirements of FHR missing data filling by quantitatively comparing five regression techniques to enable online operation at 4 Hz. In 2017, Feng et al. [13] observed some hidden connections between FHR and uterine signals. They combined FHR and uterine signals to fill missing part by GP. In 2021, Wang et al. [14] proposed the Optimal Transport Based on KNN (KOT) algorithm combining optimal decision trees and KNN, which assumes that FHR signals of the same type have the same distribution. In 2021, Quirk et al. [15] found that FHR signals are correlated in time and space and proposed a Bayesian based on GP algorithm. In 2023, Shapira et al. [16] observed a positive effect of baseline on filling FHR missing data. They

This work was supported in part by the National Natural Science Foundation of China (Grant No.62071162, 62071327 and 62301205) and Zhejiang Provincial Natural Science Foundation of China (Grant No. LDT23F01012F01 and LDT23F01015F01).

(Corresponding author: zhaozd@hdu.edu.cn (Zhidong Zhao))

Z. Zhou, Y. Zhang and Y. Deng are with the College of Electronics and Information Engineering, Hangzhou Dianzi University, Hangzhou 310000,

China. (e-mail: zhouzx@hdu.edu.cn; zhangyf\_valora@hdu.edu.cn; yanjund@hdu.edu.cn).

Z. Zhao, S. Wang and P. Jiao are with the School of Cyberspace, Hangzhou Dianzi University, Hangzhou 310000, China. (e-mail: zhaozd@hdu.edu.cn; shuaiwang.tai@gmail.com; pjiao@hdu.edu.cn).

H. Wu is with the Center for Applied Mathematics, Tianjin University, Tianjin, 300072, China. (e-mail: whming@tju.edu.cn).

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

proposed automatic regression two-sided modified asymmetric least squares method as a baseline calculation algorithm. In 2023, Zhang et al. [17] found the length of missing values is strongly related to the difficulty of rebuilding FHR values.

However, these studies have two problems.

- Deep learning algorithms are good at mining correlations in multi-dimensional data. But most FHR studies are based on single dimension signals.
- These studies disregard the multiple hidden periods within FHR signal.

To solve the above two problems, we build a missing data filling algorithm with dimension construction and period segmentation. Dual-dimensional Sample Construction (DDSC) and Multi-period decomposition (MPD) methods are combined to realize the dimension augmentation and decompose periods of FHR signals. Specifically, DDSC selects two FHR signals with the maximum difference to create a new sample. MPD decomposes the distributions within FHR signals. Diffusion model for Missing Data Imputation in FHR (DMDI-FHR) can achieve data filling. The main contributions of this paper can be summarized as follows:

- We propose a sample construction method DDSC to search for two FHR signals with the largest difference to form dual-dimensional samples, which expands the correlation between FHR signals.
- We propose an MPD method to decompose the FHR signal into multiple periods and discuss the value of period diversity to missing data filling task.
- We construct an FHR missing data filling algorithm DMDI-FHR, which based on DDSC, MPD and diffusion model. The DMDI-FHR algorithm achieves good performance on filling task.

## II. RELATED WORK

Continuous signals are the basis of deep learning studies. Missing data is a prevalent problem in time series. We summarize the studies on filling missing data for time series.

### A. Statistics-based algorithm

Statistics-based algorithms utilize statistical measures to fill missing data. Based on statistical elements, Autoregressive Integrated Moving Average (ARIMA) model was introduced [18]. ARIMA combines trend, seasonal, and periodic information in time series with statistical tools [19-20]. However, ARIMA-based methods only consider common characteristics of samples, while ignoring unique attributes of each sample.

### B. Machine learning-based algorithm

Machine learning-based algorithms can fit original distribution of dataset. For each missing data, these algorithms generate a corresponding filling value.

For instance, Mir et al. [21] proposed an Imputation by Feature Importance (IBFI) algorithm for reconstructing missing

data. IBFI classifies three missing categories to calculate missing data. Zheng et al. [22] proposed a KNN-based method that combines multidimensional temporal and spatial information to capture short-term fluctuations and fit long-term trends. Li et al. [23] developed a matrix containing long-term and short-term correlations of signals. This matrix reduces interpolation errors in both random and continuous missing data cases.

Machine learning-based algorithms require manual construction of feature engineering. However, it is challenging to construct feature engineering in large amounts of data.

### C. Deep learning-based algorithm

Deep learning has ability to independently extract high-dimensional information in missing data filling tasks.

**Neural network-based algorithms:** Che et al. [24] introduced a novel algorithm GRU-D. GRU-D incorporates two types of missing patterns (masking and time intervals) to improve filling results. Li et al. [25] discovered that different missing patterns exhibit temporal and spatial correlations. Based on this observation, they proposed a spatiotemporal filling algorithm named LSTM-AEs. The experimental results demonstrate that LSTM-AEs performs well for filling missing data in spatiotemporal scenarios. Ma et al. [26] viewed the nonlinear dependence of time series as a challenge, and they propose a new loss function to analyze historical states of nonlinear dependence. The new loss function enables filling algorithm to boost quality of univariate and multivariable filled data. Ma et al. [27] believed that reducing interpolation value errors is crucial, as downstream deep learning tasks can amplify these errors. To address this concern, they classify time series while filling missing values.

**GAN-based algorithms:** Guo et al. [28] focused on multi-variate time series, they incorporate multi-channel convolutions into GANs to generate high-precision data distributions. Oh et al. [29] identified that neural networks still have limitations when filling complex multivariate time series. They use GANs and bidirectional recursive mechanism to fit potential distribution through a new attention layer. The new layer can capture the weighted correlation of the entire series to improve the accuracy of downstream tasks. Tang et al. [30] introduced adversarial strategy to address the disregarding global temporal dynamics in filling missing data. The adversarial strategy to improve the global temporal distributions to fit overall distribution of data.

In summary, deep learning-based algorithms employ the strong fitting ability to fill missing data directly. Neural network-based algorithms tend to improve recurrent networks, such as RNNs. However, the performance of recurrent network still needs to be improved in modeling long time series. GAN-based algorithms exhibit good performance for various missing lengths, while training is unstable.

## III. METHOD AND MATERIAL

The DMDI-FHR algorithm focuses on the correlation within

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

and between FHR signals. The architecture of DMDI-FHR is presented in Fig. 1. First, DDSC method selects two FHR signals with maximum differences from the same label dataset, creating a new dual-dimensional sample. Next, the MPD method decompose the most different parts in dual-dimensional sample, which considers the influence of multiple distributions within FHR signal. Finally, we introduce diffusion model to FHR missing data filling tasks.

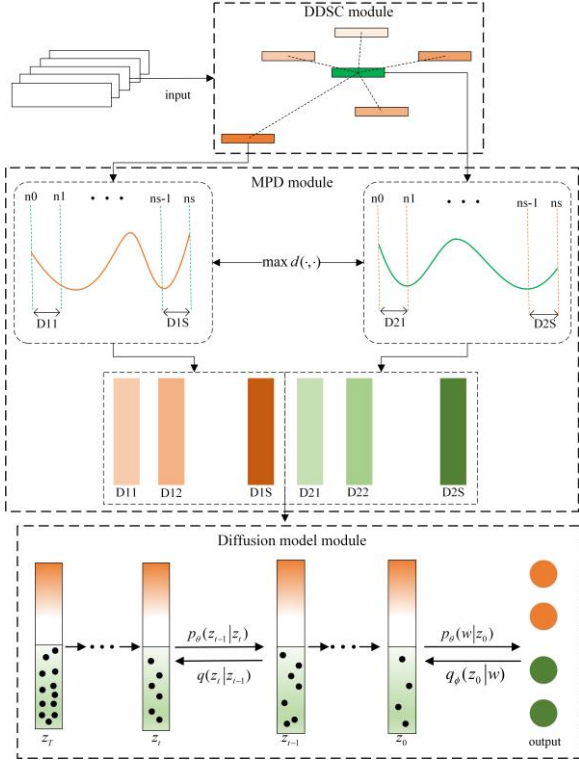


Fig. 1. The architecture of DMDI-FHR algorithm.

#### A. Dual-dimensional Sample Construction

One challenge of FHR missing data filling is ignoring the correlation between FHR signals. To tackle this challenge, we try to construct dual-dimensional samples by adding two FHR signals together. In this section, we propose a DDSC method to select an additional FHR signal as an extra dimension to an existing FHR signal.

Deep learning-based algorithms can achieve classification results by identifying commonalities in data [31]. In other words, the data with the same label exists differences. The distribution of data with same label is not uniform. This can also be explained from the signal perspective: abnormal FHR signals can be considered abnormal from multiple perspectives. There are some differences between the two FHR signals. Therefore, it is necessary to calculate the difference between two FHR signals with the same label.

The amplitude shift, stretch, and linear drift are the factors that affect in calculating the similarity of FHR signals. These factors make it challenging to determine the relative

displacement of two FHR signals.

Therefore, we propose DDSC method based on Dynamic Time Warping (DTW). In DDSC, FHR signals with the same label are adjusted in time domain through non-linear timing alignment to calculate their similarity, as shown in Fig. 2.

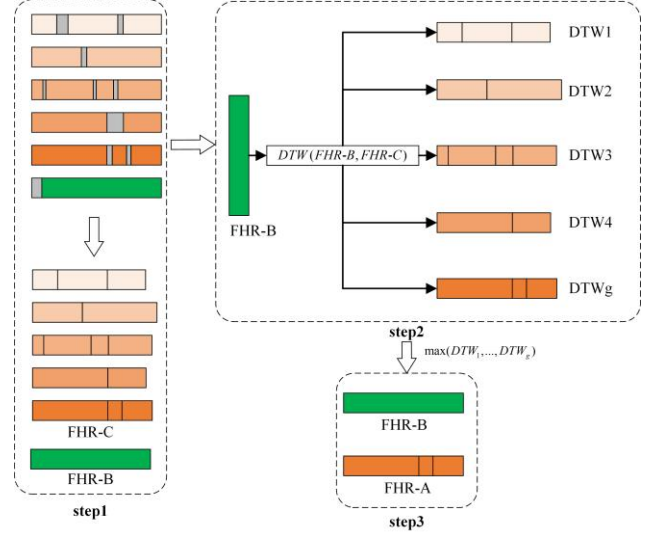


Fig. 2. DDSC technology roadmap.

The DDSC method is suitable for comparing the similarity of long distance and multi-period time series. Assuming that the number of FHR signals with the same label is  $N$  and their length (including missing part) is  $L$ , they can be expressed as:

$$FHR_{missing} = \{FHR^1, FHR^2, \dots, FHR^N\} \quad (1)$$

We randomly select an FHR signal from same label dataset, denoted as FHR-B. Next, we use DTW to measure its similarity to other signals, denoted as FHR-C. Then, we select an FHR signal with the greatest dissimilarity from FHR-C as the new dimension of FHR-B. The detailed steps are as follows:

Step 1: First, we delete missing parts of FHR signals to obtain  $FHR_{delete\_missing}$ :

$$FHR_{delete\_missing} = \{F^1, F^2, \dots, F^N\} \quad (2)$$

Step 2: We select an FHR signal as the base signal from  $FHR_{delete\_missing}$ . This can be expressed mathematically as:

$$\forall F^i \in FHR_{delete\_missing} = \begin{Bmatrix} f_1^1 & f_2^1 & \dots & f_{L_1}^1 \\ f_1^2 & f_2^2 & \dots & f_{L_2}^2 \\ \vdots & \vdots & \ddots & \vdots \\ f_1^N & f_2^N & \dots & f_{L_N}^N \end{Bmatrix} \quad (3)$$

$F^i$  can be denoted as FHR-B:

$$F^i = (f_1^i, f_2^i, \dots, f_{L_i}^i) \rightarrow FHR-B \quad (4)$$

where  $f$  represents the discrete signal inside the FHR,  $i \in \{1, 2, 3, \dots, N\}$ ,  $L_i \in \{1, 2, 3, \dots, L\}$ .

FHR-C is a set of  $N-1$  FHR signals. Selecting a candidate FHR signal  $F^j$  in FHR-C, assuming its length is  $L_j$ , where  $L_j$  is a variable and have different values in different FHR signals.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

This can be expressed as:

$$FHR-C = (F^1, F^2, \dots, F^{i-1}, F^{i+1}, \dots, F^N) \quad (5)$$

$$\forall F^j \in FHR - C, j = (1, 2, \dots, i-1, i+1, \dots, N) \quad (6)$$

$$F^j = (f_1^j, f_2^j, \dots, f_{L_j}^j), j = (1, 2, \dots, i-1, i+1, \dots, N) \quad (7)$$

Then,  $d(F^i, F^j)$  represents the distance between  $F^i$  and  $F^j$ . The distance calculation as follows:

$$d(F^i, F^j) = (F^i, F^j)^2 \quad (8)$$

Finally, we define the distance between any two points in the two FHR signals as a path denoted as  $W$ , which has  $K$  elements.  $w_k$  indicates alignment relationship between two FHR signals at the  $k$ -th path point.

$$W = w_1, w_2, \dots, w_K, \max(i, j) \leq K < i + j - 1 \quad (9)$$

The similarity of these two FHR signals is as follows:

$$DTW^{i,j} = \min \left( \sqrt{\sum_{k=1}^K w_k / K} \right) \quad (10)$$

Step 3: Repeat the step1 and step2 to calculate DTW values of all FHR signals in FHR-C and  $F^i$ :

$$DTW_{F^i} = (DTW^{i,1}, \dots, DTW^{i,i-1}, DTW^{i,i+1}, \dots, DTW^{i,N}) \quad (11)$$

Selecting the FHR signal  $F_{add}^i$  corresponding to the maximum value in the Eq. (11) as the new dimension of  $F^i$ :

$$DDSC_{F^i} = (F^i, F_{add}^i) \quad (12)$$

Step 4: Repeat step1 to step3 for all FHR signals with the same label, and finally form  $N$  two-dimensional samples.

### B. Multi-period Decomposition

Most FHR missing data filling studies select the single sample as the smallest unit. However, these studies neglect the periodicity of single FHR signal. For instance, acceleration and deceleration are periodic features within the FHR signal. Periodicity in time series have impact on downstream tasks [32]. Specifically, multiple periods in training data increases data variability, which may make it more difficult to fit the data distribution. Nevertheless, the periodicity also has a positive effect to enhance the diversity of generated data. Thus, the key to exploiting the internal periodicity is to separate these periods. However, an inescapable problem is that many periods in data cannot be accurately marked.

Periods refer to the different distribution within signals. Previous studies have attempted to uncover the hidden properties of signals through statistical tools. These studies aim to represent the all signals using a single tool. However, these tools overlook the multiple distributions within a signal. These multiple distributions can significantly affect the accuracy of statistical results. Therefore, we propose MPD method. This method adopts the concept of approximating the internal period using internal signal differences [33] and further extends this

concept to decompose dual-dimensional samples.

The number of periods and boundaries of each period are typically not labeled in FHR signals, which makes it a bivariate problem. We link the period distribution to the length of training data. Consequently, the bivariate problem (period and bound) become a univariate problem (period) by fixing length of training data. Multiple periods in a signal are equivalent to multiple distributions within signal. Thus, separating the periodicity requires maximizing the difference within the signals. We need to decompose the FHR into  $S$  parts with the largest difference.

The difference between two distributions can be measured using Relative Entropy (RE). Based on the principle of maximum entropy, we can identify the  $S$  least similar parts in the FHR signal.

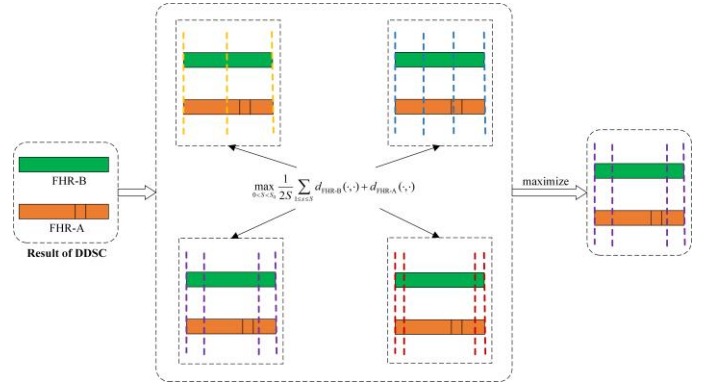


Fig. 3. MPD technology roadmap.

In addition, we consider the least similar parts in dual-dimensional FHR samples, with the aim of maximizing the differentiation of the dual-dimensional FHR sample. The corresponding mathematical expression is shown as follows:

$$\max_{0 < S < S_0} \frac{1}{2S} \sum_{1 \leq s \leq S} d_{FHR-B}(\cdot, \cdot) + d_{FHR-A}(\cdot, \cdot) \quad (13)$$

where  $d$  represents a distance metric,  $S$  represents number of periods, while  $S_0$  is a hyperparameter that prevents excessive splitting (prevent insufficient training data). In this paper, we use RE divergence as distance metric  $d(\cdot, \cdot)$ . If algorithm can learn the distribution from cases with maximum differences, it will have better filling ability.

### C. Diffusion Model Primary Network

Compared with other time series signals, FHR signals have more complex physiological characteristics and interference information. Diffusion model have powerful data generation capabilities and can be used for various generation-related tasks. By adding noise and then denoising, diffusion model can avoid the influence of noise on the original data. This feature of the diffusion model can be used to fill FHR missing data.

#### 1) Diffusion process

The diffusion process refers to gradually adding noise to  $x_0$



> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

until it reaches  $x_T$ . The initial signal  $x_0 \sim q(x)$  is gradually added to  $q(x_{1:T}|x_0)$ , where in  $x_1, \dots, x_T$  are same scale as  $x_0$ . The entire diffusion process is calculated based on Markov chain. As time  $t$  increases,  $x_t$  becomes an independent Gaussian distribution. The standard deviation of noise is determined by  $x_t$  and time  $t$ .

$$q(x_t | x_{t-1}) = N(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t I) \quad (14)$$

$$q(x_{1:T} | x_0) = \prod_{t=1}^T q(x_t | x_{t-1}) \quad (15)$$

where  $\beta_t$  and  $\sqrt{1 - \beta_t} x_{t-1}$  are the standard deviation and mean of noise.  $\beta_t I$  is the variance, where  $\alpha_t = 1 - \beta_t$  and  $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$ . Based on Markov process, the data distribution at any given time can be obtained gradually through recursion.

$$q(x_t | x_0) = N(x_t; \sqrt{\bar{\alpha}_t} x_0, (1 - \bar{\alpha}_t) I) \quad (16)$$

where  $q(x_t|x_0)$  represents the data distribution at any given time, which can be determined based on initial  $x_0$  and standard deviation of noise  $\beta_t$ .

## 2) Reverse process

The reverse process is the inverse of diffusion process. It refers to the process of recovering the original data distribution  $x_0$  from the Gaussian noise  $x_T$ . This process requires building a deep learning network to estimate the distribution, as shown in Eqs. (17) and (18).

$$p_\theta(x_{0:T}) = p(x_T) \prod_{t=1}^T p_\theta(x_{t-1} | x_t) \quad (17)$$

$$p_\theta(x_{t-1} | x_t) = N(x_{t-1}; \mu_\theta(x_t, t), \sum_\theta(x_t, t)) \quad (18)$$

where  $p_\theta(x_{0:T})$  is the reverse distribution, and  $p_\theta(x_{t-1} | x_t)$  needs to be obtained through deep learning network training. The conditional probability  $q(x_{t-1} | x_t, x_0)$  in diffusion process can be obtained using the Bayesian process:

$$q(x_{t-1} | x_t, x_0) = N(x_{t-1}; \mu_\theta(x_t, x_0), \beta_t I) \quad (19)$$

From Eq. (20), it can be inferred that given  $x_t$  and  $x_0$ , the data distribution at time  $x_{t-1}$  can be calculated.

$$x_0 = \frac{1}{\sqrt{\bar{\alpha}_t}} (x_t - \sqrt{1 - \bar{\alpha}_t} \varepsilon_t) \quad (20)$$

According to the above deduction, for  $q(x_{t-1} | x_t, x_0)$  to guide  $p_\theta(x_{t-1} | x_t)$  in training, it is necessary to determine the objective function  $-\log p_\theta(x_0)$ , as shown in Eq. (21).

$$L_{\text{simple}}(\theta) := E_{\varepsilon, x_0, \varepsilon} [\|\varepsilon - \varepsilon_\theta(\sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \varepsilon, t)\|^2] \quad (21)$$

where  $\varepsilon$  is an unknown random variable and  $\varepsilon_\theta$  is the specific objective form of defined network.

## 3) Network in diffusion model

To enhance FHR filling ability, we improve the deep learning network to recover more realistic distribution of FHR signals.

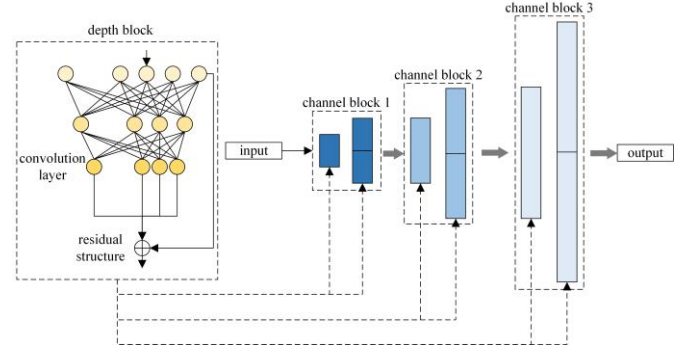


Fig. 4. Network of diffusion model.

The structure of network is shown in Fig. 4. The structure consists of two parts, the depth block and the channel block. It is driven by depth blocks based on residual structure and convolution and channel blocks in series. The depth block acquires high-dimensional features of network through multiple convolutional layers. Depth blocks prevent features from losing control by virtue of residual structure. Firstly, the input signal is passed through a depth block, features of different scales are extracted under convolution structure, and features of different convolution scales are aggregated through residual structure. Depth blocks are followed by connected channel blocks. The concatenation of multiple channel blocks continuously deepens features extracted by diffusion model.

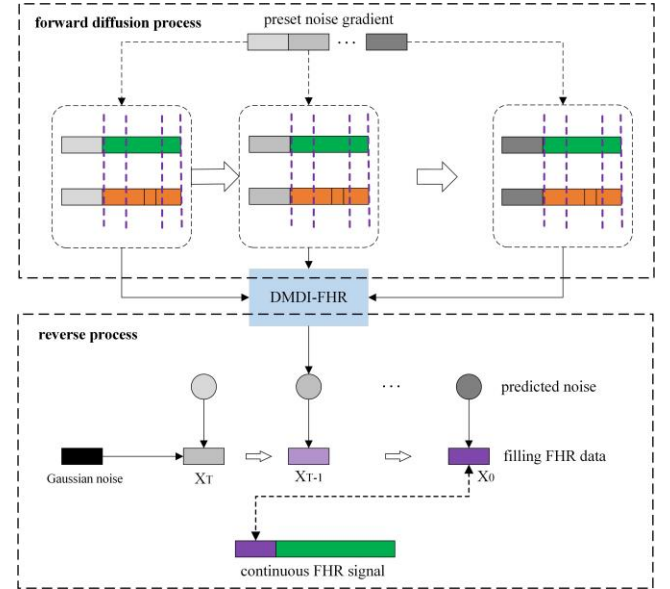


Fig. 5. FHR missing data filling process.

As shown in Fig. 5, the diffusion process of FHR signals refers to the process of gradually adding noise to FHR signals to an independent Gaussian distribution. The added noise with preset mean value and the train data has the same scale. The reverse process reconstructs FHR signals from Gaussian noise under the condition of noise.

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

#### IV. EXPERIMENTS

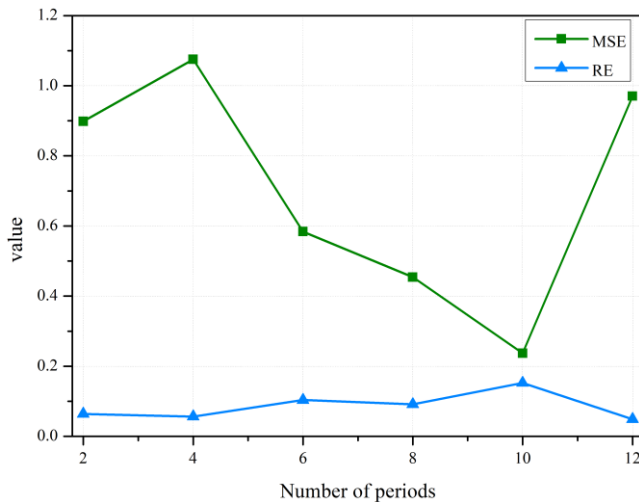
##### A. Dataset and experiments settings

The CTU-UHB dataset used in this study is from Physionet, an internationally authoritative open database of physiological signals. This dataset was collected by the Obstetrics Research Team at Brno University Hospital in Czech [34]. The dataset comprises 552 FHR records. The batch size, initial learning rate, and epochs for the experiment were set to 16,  $1e-5$ , and 20, respectively.

This section aims to assess the performance of DMDI-FHR in FHR missing data filling. First, the role of the DDSC and MPD was verified by ablation experiments. Following this, we verify the effect of DMDI-FHR algorithm on different missing rates. We also assess the impact of data filling on data augmentation.

##### B. Algorithm verification experiment

To confirm the effectiveness of the MPD method in processing FHR signals, we design an experiment that compared the quality of filled data at different decomposition scales by RE and MSE (Mean Squared Error). The results are presented in Fig. 6.



**Fig. 6.** Comparison of filling quality at different periods.

Two conclusions can be drawn from Fig. 6. Firstly, MPD method can affect the effect of data padding. DMDI-FHR has different filling ability under different periods. The maximum difference of MSE is 4.54 times. This verifies that MPD method can directly affect the quality of imputation of FHR data. Second, there is an effective synergistic relationship between data filling quality and periodic separation. The extreme values of the MSE in Fig. 6 are synergistic with those of the RE curve, reaching their extreme values at number of periods of 10.

In addition, we examined the impact of the DDSC and MPD methods on the data filling effect in DMDI-FHR. To this end, we conducted ablation experiments with the gap length of 60, and Table I presents the data filling capability of different

combinations within DMDI-FHR via MSE and MAE (Mean Absolute Error).

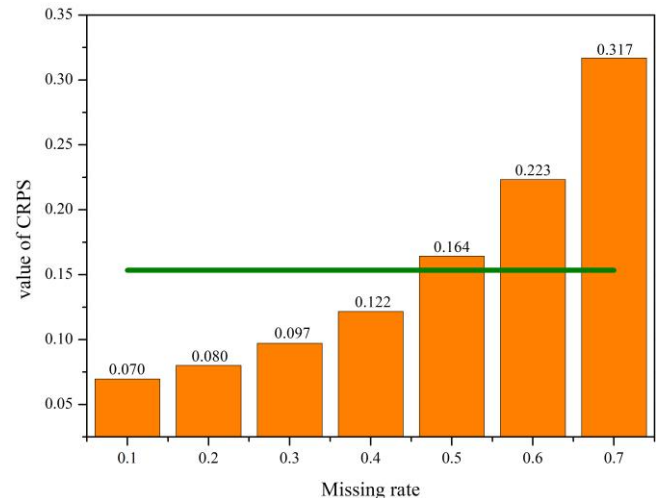
TABLE I

ABLATION EXPERIMENT OF DDSC AND MPD			
DDSC	MPD	MSE	MAE
√		5.466	4.558
	√	2.459	1.763
√	√	0.237	0.393

By adding the DDSC and MPD methods, the quality of FHR missing data filling can be improved to different degrees, with MSE and MAE reduced to 0.237 and 0.393. DDSC has a less impact on improving FHR data filling quality. The MSE and MAE indicators using DDSC alone were 3.007 and 2.795 higher than those using MPD alone, respectively. The reason for this is that the new dimensions of the dual-dimensional FHR sample come from data with the same label, and their variability is less than the variability within the FHR signal. In other words, MPD method can effectively decompose the multi distributions in FHR signal. In general, DDSC and MPD methods have positive effects on DMDI-FHR algorithm.

##### C. Data filling performance comparison experiment.

In Fig. 7, we compare the performance of DMDI-FHR at missing rates from 0.1 to 0.7 using CRPS (Continuous Ranked Probability Score) indicator.



**Fig. 7.** Experiment with different missing rates.

DMDI-FHR algorithm has good filling ability for different missing rates. Fig. 7 shows that DMDI-FHR algorithm reaches a minimum CRPS when the missing rate is 0.1, creating the most similar distribution. It is worth noting that 0.1 is a low missing rate. The filling ability of DMDI-FHR algorithm at higher missing rate can better reflect its performance. As shown in Fig. 7, DMDI-FHR algorithm can still perform valid data filling at missing rates are 0.5 and 0.6. Compared to the mean value (solid green line), the DMDI-FHR algorithm increases

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

CRPS by only 0.011 and 0.070 at missing rates are 0.5 and 0.6. Thus, DMDI-FHR algorithm has good filling ability when the missing rate is 0.1 to 0.6. The CRPS indicator increase rapidly when the missing rate is 0.7, which indicated that excessive missing rate is still an important factor affecting the filling effect.

In order to analyze the usability of DMDI-FHR algorithm more objectively, we add the more typical FHR classification research in recent years for comparative test. The experimental results are shown in Table II.

TABLE II  
COMPARISON OF PREVIOUS WORKS

Author	Algorithm	MSE
Frigo,2017,[13]	On-Line Regression	26.729
Feng, 2017,[14]	GP, without UA	1.357
Wang,2021,[15]	KOT (gap length:30)	0.476
Quirk, 2021,[16]	Bayesian based on GP	0.360
Xiang,2022,[35]	alternative interpolation	18.429
DMDI-FHR,2023	DMDI-FHR (gap length:20)	0.206

Table II shows a comparison between DMDI-FHR and related FHR filling studies. Each value represents the best performance achieved by the corresponding algorithm. As shown in Table III, the missing data filling accuracy of DMDI-FHR is higher than other five algorithms. Only KOT, Bayesian based on GP and DMDI-FHR algorithm achieve better filling accuracy. However, it is worth noting that KOT involves feature engineering. KOT improves the quality of filling data by manually improved KNN. Bayesian based on GP improves the filling accuracy by controlling Bayesian processes. On-Line Regression algorithm fills FHR missing data through regression algorithms, but its performance is worse than machine learning and deep learning algorithms. The results show that DMDI-FHR is effective with FHR missing data filling.

#### D. FHR usability experiment

Most downstream tasks require data augmentation due to the class imbalance of CTU-UHB dataset. Some data augmentation studies focus on fitting the entire FHR signals, while disregarding the impact of filling quality to data augmentation. We consider the impact of filling quality on data augmentation. Consequently, we used measures such as SSD (Sum of Squared Deviations) and PRD (Percentage Relative Difference) to assess the quality of Synthetic Minority Over-sampling Technique (SMOTE) [36] after DMDI-FHR application. Given the significant variance in the absolute values of these indicators, we present comparison results in percentages, with the actual values provided for reference.

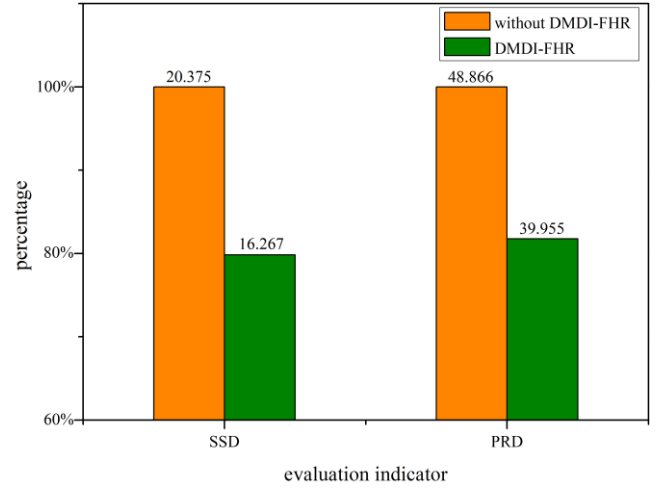


Fig. 8. The effect of data filling on data augmentation.

Fig. 8 demonstrates that SSD, MAD, and PRD indicators are significantly reduced after applying DMDI-FHR, thereby verifying the effectiveness of FHR data filling using DMDI-FHR. Specifically, each value improves when DMDI-FHR is added during the data augmentation, indicating an improvement in the quality of generated data. About SSD, the score decreases from 20.375 to 16.267, representing a decrease of 20.2%. The PRD decreased from 48.866 to 39.955, indicating a decrease of 18.2%. Overall, DMDI-FHR exhibited improved results across all indicators, further demonstrating the impact of data filling on data augmentation.

#### V. CONCLUSION

To support the development of consumer FHR monitoring electronics, we developed a new FHR missing data filling algorithm DMDI-FHR to improve data quality for fetal status assisted diagnosis. This algorithm combines two FHR signals as a new sample to achieve accurate feature extraction between dual-dimensions. Additionally, MPD method performs blind period decomposition of the FHR signal to enhance the independence of training data. After that, the missing part of FHR signal is generated from the noise by diffusion model network. Experimental results show that DMDI-FHR algorithm has a good effect in filling FHR missing data. In the future, it will be interesting to extend our work to real-time missing data filling.

#### REFERENCES

- [1] Wang X and Wu Y., "Fog-Assisted Internet of Medical Things for Smart Healthcare," *IEEE Transactions on Consumer Electronics*, vol. 69, no. 3, pp. 391-399, Aug. 2023.
- [2] Olokodana L et al., "EZcap: A Novel Wearable for Real-Time Automated Seizure Detection From EEG Signals," *IEEE Transactions on Consumer Electronics*, vol. 67, no. 2, pp. 166-175, Aug. 2021.
- [3] Ray P et al., "Implementation and Performance Analysis of Interoperable and Heterogeneous IoT-Edge Gateway for Pervasive Wellness Care," *IEEE*

> REPLACE THIS LINE WITH YOUR MANUSCRIPT ID NUMBER (DOUBLE-CLICK HERE TO EDIT) <

- Transactions on Consumer Electronics*, vol. 65, no. 4, pp. 464-473, Sep. 2019.
- [4] Hamelmann P *et al.*, "Doppler Ultrasound Technology for Fetal Heart Rate Monitoring: A Review," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, pp. 99, Sep. 2019.
- [5] Dash S, Quirk J G, and Djurić P M, "Fetal heart rate classification using generative models," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 11, pp. 2796-2805, Jun. 2014.
- [6] Fuentealba P, Illanes A, and Ortmeier F, "Cardiotocographic signal feature extraction through ceemdan and time-varying autoregressive spectral-based analysis for fetal welfare assessment," *IEEE Access*, vol. 7, pp. 159754-159772, Oct. 2019.
- [7] Alsagoff W *et al.*, "Predicting fetal hypoxia using common spatial pattern and machine learning from cardiotocography signals," *Applied Acoustics*, vol. 167, pp. 107429, Oct. 2020.
- [8] Pescador F and Mohanty P., "Machine Learning for Smart Electronic Systems," *IEEE Transactions on Consumer Electronics*, vol. 67, no. 4, pp. 224-225, Nov. 2021.
- [9] Przybyła T *et al.*, "A recovery of FHR signal in the Embedded Space," in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing.*, 2014, pp. 563-566.
- [10] Feng G, Quirk J G, and Djurić P M, "Recovery of missing samples in fetal heart rate recordings with Gaussian processes," in *IEEE European Signal Processing Conference.*, 2017, pp. 261-265.
- [11] Oikonomou V P *et al.*, "An adaptive method for the recovery of missing samples from FHR time series," in *IEEE Computer-Based Medical Systems.*, 2013, pp. 337-342.
- [12] Frigo G and Giorgi G, "Comparative evaluation of on-line missing data regression techniques in intrapartum FHR measurements," in *2017 IEEE International Instrumentation and Measurement Technology Conference.*, 2017, pp. 1-6.
- [13] Feng G *et al.*, "Recovery of missing samples in fetal heart rate recordings with Gaussian processes," in *2017 25th European Signal Processing Conference*, pp. 261-265.
- [14] Wang C *et al.*, "Imputation Method for Fetal Heart Rate Signal Evaluation Based on Optimal Transport Theory," *SN Computer Science*, vol. 2, pp. 1-12, Jun. 2021.
- [15] Quirk JG *et al.*, "Estimation of consecutively missed samples in fetal heart rate recordings," in *2020 28th European Signal Processing Conference*, pp. 1080-1084.
- [16] Shapira R *et al.*, "Double-sided asymmetric method for automated fetal heart rate baseline calculation," *Physical and Engineering Sciences in Medicine*, pp. 1-12, Mar. 2023.
- [17] Zhang Y *et al.*, "Reconstruction of missing samples in antepartum and intrapartum FHR measurements via mini-batch-based minimized sparse dictionary learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 276-288, Jan. 2022.
- [18] Velicer W F, and Colby S M, "A comparison of missing-data procedures for ARIMA time-series analysis," *Educational and Psychological Measurement*, vol. 65, no. 4, pp. 596-615, Aug. 2005.
- [19] Gómez V, Maravall A, and Peña D, "Missing observations in ARIMA models: Skipping approach versus additive outlier approach," *Journal of Econometrics*, vol. 88, no. 2, pp. 341-363, Feb. 1999.
- [20] Li Y *et al.*, "SO-slope entropy coupled with SVMD: A novel adaptive feature extraction method for ship-radiated noise," *Ocean Engineering*, vol. 280, pp. 114677, July. 2023.
- [21] Mir A *et al.*, "Imputation by feature importance (IBFI): A methodology to envelop machine learning method for imputing missing patterns in time series data," *PloS one*, vol. 17, no. 1, pp. e0262131, Aug. 2022.
- [22] Zheng L *et al.*, "A tensor-based K-nearest neighbors method for traffic speed prediction under data missing," *Transportmetrica B: Transport Dynamics*, vol. 8, no. 1, pp. 182-199, Jan. 2020.
- [23] Li L *et al.*, "Missing data estimation method for time series data in structure health monitoring systems by probability principal component analysis," *Advances in Engineering Software*, vol. 149, pp. 6085, Nov. 2018.
- [24] Che Z *et al.*, "Recurrent neural networks for multivariate time series with missing values," *Scientific reports*, vol. 8, no. 1, pp. 182-199, Apr. 2020.
- [25] Li D *et al.*, "Smoothed LSTM-AE: A spatio-temporal deep model for multiple time-series missing imputation," *Neurocomputing*, vol. 411, pp. 351-363, Oct. 2020.
- [26] Ma Q *et al.*, "End-to-end incomplete time-series modeling from linear memory of latent variables," *IEEE Transactions on Cybernetics*, vol. 50, no. 12, pp. 4908-4920, Apr. 2019.
- [27] Ma X *et al.*, "Co-regularized nonnegative matrix factorization for evolving community detection in dynamic networks," *Information Sciences*, vol. 528, pp. 265-279, Aug. 2020.
- [28] Guo Z, Wan Y, and Ye H, "A data imputation method for multivariate time series based on generative adversarial network," *Neurocomputing*, vol. 360, pp. 185-197, Sep. 2019.
- [29] Oh E *et al.*, "STING: Self-attention based Time-series Imputation Networks using GAN," in *IEEE International Conference on Data Mining (ICDM).*, 2021, pp. 1264-1269.
- [30] Ma X *et al.*, "Joint modeling of local and global temporal dynamics for multivariate time series forecasting with missing values," in *Proceedings of the AAAI Conference on Artificial Intelligence.*, 2020, pp. 5956-5963.
- [31] Wu W and Ma X, "Network-based structural learning nonnegative matrix factorization algorithm for clustering of scRNA-seq data," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, no. 1, pp. 566-575, Mar. 2022.
- [32] Ma X, Sun P, and Qin G, "Identifying condition-specific modules by clustering multiple networks," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 121, pp. 208216, Sep. 2018.
- [33] Du Y *et al.*, "Adarnn: Adaptive learning and forecasting of time series," in *2021 30th ACM international conference on information & knowledge management*, pp. 402-411.
- [34] Chudáčková V *et al.*, "Open access intrapartum CTG database," *BMC pregnancy and childbirth*, vol. 14, pp. 1-12, Jan. 2014.
- [35] Xiang J *et al.*, "Digital signal extraction approach for cardiotocography image," *Computer Methods and Programs in Biomedicine*, vol. 225, pp. 107089, Mar. 2022.
- [36] Chawla N *et al.*, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321-357, Jun. 2002.