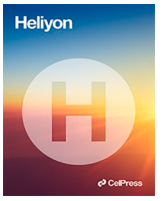
Heliyon 10 (2024) e34145

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Prediction method of human defecation based on informer audio data augmentation and improved residual network

Tie Zhang [a,*], Cong Hong [a], Yanbiao Zou [a], Jun Zhao [b]

[a] School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou, 510640, China
[b] China Rehabilitation Research Center, Beijing, 100000, China

ARTICLE INFO

Keywords:
Defecation prediction
Bowel sounds
Multi-domain features
Residual neural networks
Audio data augmentation
Timing signal prediction

ABSTRACT

Defecation care for disabled patients is a major challenge in health management. Traditional post-defecation treatment will bring physical injury and negative emotions to patients, while existing pre-defecation forecasting care methods are physically intrusive. On the basis of exploring the mechanism of defecation intention generation, and based on the characteristic analysis and clinical application of bowel sounds, it is found that the generation of desire to defecate and bowel sounds are correlated to a certain extent. Therefore, a deep learning-based bowel sound recognition method is proposed for human defecation prediction. The wavelet domain based Wiener filter is used to filter the bowel sound data to reduce other noise. Statistical analysis, fast Fourier transform and wavelet packet transform are used to extract the integrated features of bowel sound in time, frequency and time-frequency domain. In particular, an audio signal expansion data algorithm based on the Informer model is proposed to solve the problem of poor generalization of the training model caused by the difficulty of collecting bowel sound in reality. An improved one-dimensional residual network model (1D-IResNet) for defecation classification prediction is designed based on multi-domain features. The experimental results show that the proposed bowel sound augmentation strategy can effectively improve the data sample size and increase the sample diversity. Under the augmented dataset, the training speed of the 1D-IResNet model is accelerated, and the classification accuracy reaches 90.54 %, the F1 score reaches 83.88 %, which achieves a relatively good classification stability while maintaining a high classification index.

1. Introduction

In recent years, the need for health management and long-term care for patients with disabilities has become increasingly prominent. Among them, the defecation care of disabled patients has become a major challenge. For disabled patients with limited mobility, the commonly used method of defecation care is post-defecation care, which involves the use of diapers or anal pouches to collect feces, and then cleaning the feces after defecation, as well as cleansing the private parts of the human body. However, in most cases, due to the lack of timely cleaning of feces, which makes feces and skin longer direct contact, ultimately leading to itching and even pain, increasing the risk of local skin breakage and infection, and further deterioration of physical health [1]. In the long run, on the one hand, the patient's health is seriously impaired and prone to depression, shame and other negative emotions, and on the other

* Corresponding author.
E-mail address: merobot@scut.edu.cn (T. Zhang).

https://doi.org/10.1016/j.heliyon.2024.e34145
Received 26 November 2023; Received in revised form 16 June 2024; Accepted 4 July 2024
Available online 6 July 2024

hand, it increases the cost of nursing care and treatment [2].

Therefore, in order to solve the above difficulties, pre-defecation prediction has become a defecation care modality with great potential. Zan et al. [3,4] proposed to use an artificial anal sphincter system to extract the feature vectors of rectal pressure signals using a wavelet packet analysis algorithm, and construct a rectal sensory prediction model for defecation pattern recognition based on a support vector mechanism. However, this method requires surgical implantation of the artificial anal sphincter into the human body, and there exists wound infection due to wire connection or pain due to battery replacement during operation, which is highly invasive and not applicable to routine nursing care. Saitoh et al. [5] proposed a pre-defecation prediction method that extracts features from the log data of the nursing record system and discriminates between defecation or no defecation. However, the accuracy of the prediction results of this method is low and it is not real-time and cannot be accurate within a certain point in time. Yabunaka et al. [6] proposed to assess the rectal fecal storage status of healthy adults by pocket ultrasonography to define the normal desire for rectal defecation. However, the validity of pocket ultrasound imaging in this method is more dependent on operator expertise, limiting feasibility and translation into routine practice.

Based on the above pre-defecation prediction study, this paper tries to find a human-related physiological parameter with high correlation with defecation and simple access, and then use deep learning algorithms to mine deep information about it, and use the features for defecation-intention discrimination. Based on this idea, this paper firstly explores the mechanism of defecation reflex and defecation-intention generation.

The desire to defecate occurs when feces in the intestines move through the intestines toward the rectum, stimulating the nerve endings around the rectum and anus, which causes contractions of the intestinal wall and sphincter muscles, creating the desire to defecate and a feeling of pressure. Specifically, in the large intestine, feces move toward the colon, sigmoid colon, and rectum. When the feces reaches the rectum, the rectum is sufficiently stretched, and this process stimulates the rectal wall and sphincter, producing the desire to defecate [7,8].

Studies have pointed out that the human body is accompanied by specific intestinal motility patterns prior to defecation, and there are subtle changes in indicators of intestinal physiological parameters. The colon is capable of propelling intestinal contents distally, and approximately 90 % of the entire intestinal transport time is through the colon. Colonic propagation pressure waves (CPPWs) are important determinants of intestinal propulsion and defecation. Alterations in the characteristics of CPPWs may be an important marker of defecation [9–11]. In addition, it has been pointed out that one of the roles of high-amplitude propagation contractions (HAPCs) among the multiple modes of colonic motility in humans is to help the intestinal contents to pass through the colon to form stool, and that HAPCs can be used as one of the criteria for determining defecation [11,12].

However, most of the physiological indicators of intestinal motility in the above studies are obtained from built-in sensors, some of which required surgical operations. Such operations are not only intrusive to the human body, causing discomfort and pain to the patients, but also fail to provide long-term and non-destructive real-time monitoring and prediction of human defecation.

Meanwhile, Bowel sounds (BS) have attracted the attention of researchers as a physiological indicator of intestinal motility. BS usually refer to the "rumbling" sound emanating from the intestines, which is caused by the movement of intestinal contents (i.e., feces) driven by intestinal motility, and can reflect the motility and physiological status of the intestines well [13]. BS has the characteristics of non-invasive, low cost of data detection, simple access, etc., which has great medical value.

BS are widely used in clinical diagnosis for the assessment of intestinal obstruction, intestinal motility, and gastrointestinal disorders. Ching et al. [14] used an electronic stethoscope to collect BS from the abdomen of patients. By analyzing the sound duration, interval between sounds, dominant frequency and peak frequency of BS, they proved that BS can be used to assist in the diagnosis of intestinal obstruction, and the difference in the sound characteristics of intestinal noises between large intestinal obstruction and small intestinal obstruction may also help to determine the site of obstruction. Kim et al. [15] performed continuous intestinal monitoring and estimation of intestinal motility by extracting acoustic features of BS from each segment as input vectors to a back propagation (BP) neural network model. Du et al. [16] achieved a positive diagnosis of Irritable Bowel Syndrome using BS analysis by extracting acoustic features from BS of patients with Irritable Bowel Syndrome and healthy normal subjects and using the features as inputs to a machine learning model.

The mechanism of BS and the mechanism of defecation reflex are correlated to a certain extent, and the special intestinal movements of the human body before defecation can be captured and expressed by BS, thus avoiding the use of invasive devices on the human body. Research in clinical medicine has further confirmed that BS contains a great deal of regular information, which is of great importance in interpreting and assessing intestinal function. Therefore, this paper concludes that it is feasible and scientific to use BS as an indicator of defecation monitoring, combined with deep learning algorithms, and using computers instead of human beings to explore its potential characteristics and use it for defecation prediction.

However, BS collected using stethoscope device often have noise pollution, and if the original signal is directly used, it may cause some signal energy parameters and audio characteristics for analysis to be incorrectly calculated, resulting in misclassification, so the noise removal process is critical. Existing BS denoising algorithms are generally based on wavelet transforms (WT), Ranta et al. [17] proposed a smooth non-smooth filter based on iterative WT, which is used without additional parameters but has the disadvantage that it tends to distort detected BS. Hadjileontiadis [18] used wavelet domain fractal dimension filter to improve the distortion in the detection of BS, but still suffered from insufficient noise reduction. Hadjileontiadis [19] later proposed the use of higher order statistics instead of wavelets, and the use of iterative kurtosis detector (IKD) for BS detection and analysis. However, the computational cost of the iterative processing scheme is a major drawback. Dimoulas [20] et al. proposed a wavelet domain based Wiener filtering algorithm (WDWF) for BS enhancement, which introduces the frequency dependent parametric Wiener filtering technique of discrete WT and wavelet packet, which is more adaptive in terms of scanning resolution and thresholding rules, and has been successfully developed and tested for BS denoising. Considering the performance advantages of multiple filtering algorithms, this paper decides to combine

the WDWF technique to realize the enhancement and filter denoising of BS.

For recorded BS signals, there are many variations in the shape of the signal curves in the time and frequency domains, however, interpreting BS by hand is complicated, which not only requires difficult-to-obtain expertise, but also is time-consuming and laborious. Therefore, it is necessary to use computers to assist in analyzing BS characteristics, and commonly used empirical characterization metrics include sound duration, energy, power, amplitude, dominant frequency, and silence duration [21–23]. Yin [24] et al. extracted a total of 420 time-domain and frequency-domain features including the total number of shaped peaks, the length of neighboring peak intervals, the relative amplitude and kurtosis of the peaks, and the average relative intensity, and utilized a BP neural network to classify the digestive state. Qiao [25] et al. proposed the use of binaural features to describe the amplitude and time-delay relationship of BS, specifically including short-time energy, Mel cepstrum coefficient, amplitude ratio, and phase difference, and used them for BS detection. Based on the above study, it can be seen that the BS signal is characterized by asymmetry, weak signal, large individual differences and randomness, etc., and comprehensive features including time domain, frequency domain and time-frequency domain must be considered comprehensively when feature extraction and analysis are performed on it.

As a result, this paper proposes a deep learning method based BS recognition classification for human pre-defecation prediction. In this study, a wavelet domain based Wiener filter is used to filter the BS data collected using an electronic stethoscope to extract the BS signals from the overlapping signals. In addition, in order to fully explore the complex intrinsic characteristics of the BS signal, comprehensive features including time domain, frequency domain and time-frequency domain are considered in this study for feature extraction and analysis. Finally, based on the multi-domain features, an improved one-dimensional residual network model(1D-IResNet) is designed for defecation classification prediction.

In particular, during the BS data collection process, patients had fewer defecation per day, the frequency of defecation is more fixed, and some subjects had irregular defecation, which made it difficult to collect BS data with a defecation during the time period being measured. Based on the above factors, the collected dataset presents an overall small sample size and unequal positive and negative sample sizes, and it is necessary to augment and expand the training dataset before proceeding with the construction of the classification prediction model.

For audio data, commonly used augmentation strategies are categorized into 1D time-domain waveform augmentation and 2D time-frequency map augmentation. The commonly used methods for 1D time-domain waveform augmentation include time stretching, pitch shifting, dynamic range compression, audio shifting, and adding background noise, etc. [26,27], and there are a large number of

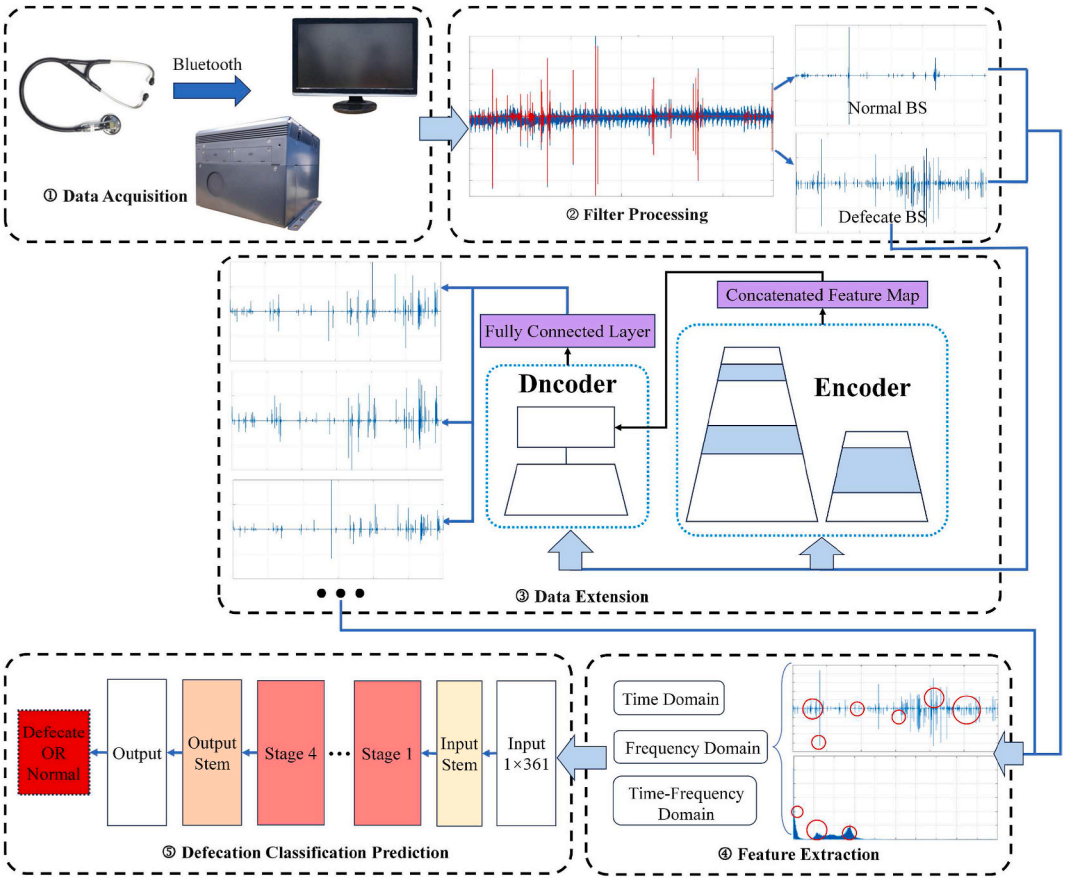


Fig. 1. Research steps framework of this paper.

standard libraries that provide operation interfaces. These operations can quickly and easily obtain a large amount of data, but the specific parameter selection varies depending on the dataset, and the augmentation effect is not the same. 2D time-frequency map augmentation, on the other hand, is to convert audio samples into Mel spectrograms by spectral transformation, and to augment Mel spectrograms with image data. Traditional image augmentation methods include simple random cropping, horizontal flipping, random rotation, and affine transformation, etc., and there are also other special augmentation methods, for example, Yazgaç [28] proposed to augment Mel spectrograms based on fractional differentiation of Mel scales, and Bahmei [29] and others used deep convolutional generative adversarial networks structure to generate excellent Mel spectrogram samples with high scalability. However, the data samples obtained from 2D time-frequency map augmentation are often not efficiently converted to audio data, and their augmentation results are usually directly used as training datasets for image classification models. For this reason, this paper proposes a way of expanding audio signal data based on the Informer temporal signal prediction model, where a single raw BS data is taken as input, and the model will learn to fit the input signal distribution and output the fitting result in the form of prediction, thus obtaining the augmented dataset based on the original data.

The specific step-by-step block diagram of this study is shown in Fig. 1.

1. Electronic stethoscope to capture BS signals before defecation and during normal.
2. WDWF for captured BSs.
3. BS data augmentation based on Informer temporal signal prediction model.
4. Time domain, frequency domain and time-frequency domain feature extraction of BS signals.
5. defecation intention classification prediction model based on multi-domain features of BS and improved 1D residual network.

The main contributions of this study are.

- Based on the BS signal data, an improved residual neural network prediction algorithm is proposed to predict the defecation intention for human, and the intrinsic correlation between BS and defecation reflex is experimentally verified, which provides a new approach for solving the problem of defecation care for disabled patients.
- In order to solve the problem of small sample size and unequal positive and negative samples in the BS dataset, a way of expanding audio signal data based on the Informer temporal signal prediction model is proposed, which learns to fit the distribution of the input signals and outputs the fitting results in the form of a prediction, so as to obtain an expanded data set based on the original data, and solve the problem of generalized training model due to the difficulty of collecting BS data in reality.

The remainder of this paper is organized as follows:

Section II will introduce the noninvasive acquisition system for BS signal acquisition; Section III will introduce the method and process of processing BS signals, including filtering processing and multidomain feature extraction for the acquired BSs; Section IV introduces the data augmentation algorithm for BS based on the Informer temporal signal prediction model; Section V introduces the improved 1D residual neural network model, and proposes an algorithm based on modeling of BS signal feature modeling; Section VI presents the analysis of experimental results to verify the effectiveness of data augmentation and the feasibility of the prediction neural network; conclusions are presented in Section VII.

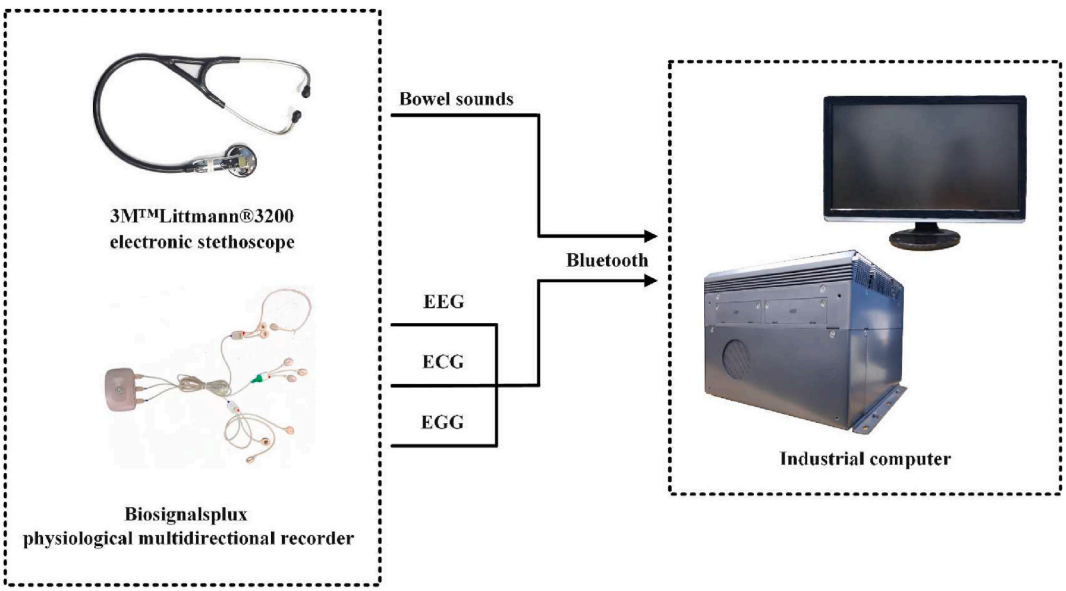


Fig. 2. Physiological data acquisition system.

2. BS acquisition system and BS dataset

2.1. Physiological data acquisition system

In this study, a physiological data collection system for human defecation prediction is established. As shown in Fig. 2, the system mainly includes a signal acquisition module and an industrial computer module, which can realize the monitoring and storage of BS signals and other human physiological signals. The signal acquisition module uses 3M™ Littmann® 3200 electronic stethoscope to collect human BS signals and Biosignalsplux physiological multidetector recorder to collect other human physiological parameters including electroencephalogram (EEG), electrocardiogram (ECG) and gastric electrogram (EGG) signals, respectively, and then transmits these physiological data to the industrial computer module via Bluetooth interface. The industrial computer module displays the collected signals of various types in real time and stores them. Of these, only the BS signal data collected by the electronic stethoscope is studied in this paper, and the rest of the physiological parameters are used for future follow-up studies.

2.2. BS dataset sources and collection details

The 3M™ Littmann® 3200 electronic stethoscope used in this study has a sampling frequency of 4000 Hz, and an analysis of the frequency characteristics of the BS shows that only about 0.5 % of the signal energy is located above 1000 Hz, only about 2 % is located above 500 Hz, and almost all of the BS energy is located in the range of 100–500 Hz [30]. Therefore, according to the Nyquist sampling theorem, the electronic stethoscope used retains the information in the original BS signal intact.

Fig. 3 shows a schematic diagram of the location of BS collection. The three common locations for BS collection are the upper right (ascending colon), upper left (descending colon) and lower left (sigmoid colon) of the abdomen, as shown in Fig. 3(a). Among them, the sigmoid colon is connected to the rectum, and when the feces in the intestines move toward the sigmoid colon and rectum, they stimulate the nerve endings around the rectum and anus, which causes contraction of the intestinal wall and sphincter muscles, generating the desire to defecate and a sense of pressure [7]. Therefore, in this study, the stethoscope fixed in the left lower quadrant of the abdomen is the most appropriate to collect BS.

In the process of data acquisition, the pickup component of the electronic stethoscope is fixed on the left lower abdomen of the human body with medical tape, as shown in Fig. 3(b). Data acquisition is done as much as possible in a ward or empty room without loud noise. The subject is asked to adopt the supine position, minimize body movement, maintain uniform breathing, and stay awake, and try not to communicate verbally with the subject during the process, except for asking if he/she had a desire to defecate. The data is sent via Bluetooth to the computer, where it is received and recorded using Littmann StethAssist software. The data recording duration was 60 s. On the one hand, it can ensure that the collection time range includes the key time periods of bowel sounds, and capture as many bowel sounds as possible; on the other hand, it can ensure the reliability and stability of the data samples, and avoid the samples to be affected by noise, interference, or transient changes; and also it is a reasonable trade-off between the realistic cost of data processing and storage, and the feasibility of sample collection. The data with strong defecation intention within 10 min before defecation is labeled as label 1(with defecation intention), and the rest is labeled as label 0 (without defecation intention). At the same time, in order to avoid selection bias caused by choosing a specific population as the sample, we recruited volunteers from different hospitals and schools, and the composition of the population included not only disabled patients but also normal young people. On the other hand, in order to avoid introducing the confounding variable of gastrointestinal diseases, we did not include people with gastrointestinal diseases when recruiting volunteers, and ensured the validity of the results of the study by controlling a single variable.

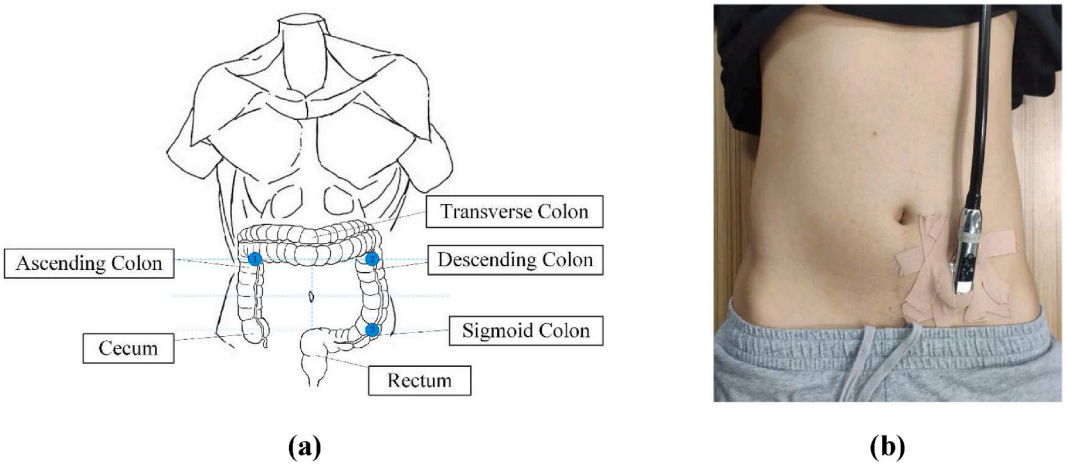


Fig. 3. Schematic diagram of BS collection location. (a) diagram of the structure of the human large intestine bowel and location of BS collection; (b) BS collection in the lower left position of the abdomen using an electronic stethoscope.

3. BS signal processing method and process

The processing of BS signals mainly includes filtering denoising and feature extraction. In this study, a WDWF is used to denoise the BSs, and a 6-band discrete wavelet transform (DWT) decomposition topology is realized. After that, we extract the features of pure BS signal in time domain, frequency domain and time frequency domain, and then construct the multi-domain feature set of BS.

3.1. WDWF based noise reduction filtering for BS

It has been shown that BS, as a non-stationary short-duration acoustic signal associated with intestinal contractile activity, exhibit three basic patterns: (a) stationary intestinal motility; (b) contractions that are explosive and of short duration; and (c) clusters of contractions with regular release slopes and durations [31]. Based on such information, the BS signal is usually defined as a mixture of five modes: bowel silent period, bowel explosive period, bowel regular sustained contraction, respiratory heartbeat noise contamination and environmental background noise contamination [32]. And it has been pointed out that the frequency range of BS is limited, with almost all signal power spectral densities lying between 100 and 500 Hz and only a small number of signals with power spectral densities higher than 500 Hz and above [30].

According to the signal pattern and sound characteristics of BS, it is known that the BS are mostly sudden "bursts" and regular continuous contractions, and some conventional filtering algorithms are difficult to eliminate the incidental heart sounds, lung sounds and environmental noise. Existing denoising algorithms for BS are generally based on WT, and among which wavelet domain based on Wiener filtering has a stronger adaptability in terms of scanning resolution and thresholding rules, and has been successfully developed and tested for BS denoising. Comparing the performance of the above filters, the WDWF is chosen to be used in this study [24].

In order to ensure that the filtered BS are not distorted and to reduce unnecessary memory computational overhead, a 5-level decomposition tree, which is common in audio signal applications, is chosen to realize a 6-band decomposition topology. And the wavelet analytic synthesis topology also provides perceptual criteria about the key bands of human auditory attributes to achieve robust enhancement with reduced artifacts. In this study, the $WDWF_{II\text{-}6}$ algorithm for BS denoising is implemented and tested to realize a 6-band DWT decomposition topology. The basic operational characteristics of the $WDWF_{II\text{-}6}$ denoising algorithm are briefly described below.

3.1.1. Parametric Wiener filter

In the time domain, the BS signal mixed with background noise, the pure uncontaminated BS signal and the background noise signal can be expressed as $x(i)$, $s(i)$, and $n(i)$, respectively, and the spectral representations of the three signals in the frequency domain are obtained by analyzing and estimating them using the short-time Fourier transform(STFT), which are $X(k)$, $S(k)$, and $N(k)$, respectively, and the spectral relationship among the three can be expressed by the following equation (1):

$$S(k) = X(k) - N(k) \tag{1}$$

The solution to the BS denoising problem is to formulate a filter $H(k)$, which most closely approximates the spectral decimation operation described in equation (1).

Assuming that the noise signal is a smooth stochastic process, an estimate of the guttural noise can be obtained by applying the Fourier transform to the quiet period of the available signal $N_{FP}(k)$. Based on the above reasoning, by multiplying the filter algorithm $H(k)$ with the spectrum of the BS signal mixed with background noise $X(k)$, an estimate of the pure guttural signal $S\sim(k)$ can be obtained, which is given by equation (2):

$$S\sim(k) = X(k) - N_{FP}(k) = H(k)\cdot X(k) \tag{2}$$

Further, by relying on frequency and simulating the perception of broadband noise by the human ear, a frequency-dependent parametric Wiener filter (FDPWF) can be implemented. The mathematical expression for the FDPWF transfer function H_{FDPWF} is given below:

$$H_{FDPWF}(k) = \begin{cases} \left\{1 - c\cdot A(k)\cdot\left[\frac{|N_{FP}(k)|}{|X_T(k)|}\right]^a\right\}^b, if\ c\cdot A(k)\cdot\left[\frac{|N_{FP}(k)|}{|X_T(k)|}\right]^a \le 1 \\ 0, otherwise \end{cases} \tag{3}$$

Eq. (3) where a, b, and c are the real-valued parameters of the filter, $X_T(k)$ is the spectrum of the windowed signal of duration T, $N_{FP}(k)$ is the Fourier transform of the noise footprint, and the parameters $A(k)$ use a noise perception criterion related to the band center frequency and human auditory attributes.

3.1.2. Wiener filters based on wavelet domains

Next, in order to implement parametric Wiener filtering in the wavelet domain, the WT needs to be used instead of the STFT as an analytical estimation method for the spectrum. After WT, $X_k(w)$ is the WT coefficients of the kth band (band $w = 0, 1, \ldots, W_{Xk} - 1$), $N_{FP-k}(w)$ is the noise footprint wavelet coefficients of the kth band, and the FDPWF will produce a unique transfer function $H_k(w)$ in the wavelet domain for each kth frequency band:

$$H_k(w)=\left\{\begin{array}{l}\left\{1-c\cdot A_{kw}\cdot\left[\frac{\langle N_{FP-k}(w)\rangle}{\langle X_k(w)\rangle}\right]^a\right\}^b, if\ c\cdot A_{kw}\cdot\left[\frac{\langle N_{FP-k}(w)\rangle}{\langle X_k(w)\rangle}\right]^a\le 1\\ 0, otherwise\end{array}\right\} \tag{4}$$

where the partial expressions within Eq. (4) are Eq. (5), Eq. (6) and Eq. (7), available at the kth band:

$$\langle N_{FP-k}(w)\rangle=\frac{1}{W_{Nk}}\sqrt{\sum_{w=0}^{W_{Nk}-1}[N_{FP}(w)]^2} \tag{5}$$

$$\langle X_k(w)\rangle=\frac{1}{W_{Xk}}\sqrt{\sum_{w=0}^{W_{Xk}-1}[X(w)]^2} \tag{6}$$

$$A_{kw}=\left\{\begin{array}{l}1, for\ f_{c-k}\le 500\\ \frac{f_{c-k}}{500}, for\ f_{c-k}>500\end{array}\right\} \tag{7}$$

In Eq. (7), the parameter f_{c-k} is the center frequency of the corresponding kth frequency band, which is the geometric mean of the frequency limit of the frequency band, and A_{kw} is the perceptual parameter of wavelet estimation.

Applying the above processing of the WT coefficients of the kth band, the estimate of the filter coefficients for BS $S_k\sim(w)$ is obtained by multiplying the filter transfer function $H_k(w)$ with the spectrum of the incoming mixed background noise BS signal $X(k)$, as in Eq. (8) below:

$$S_k\sim(w)=H_k\cdot X_k(w) \tag{8}$$

Further integration yields Eq. (9):

$$S\sim(k,w)=H_{WDWF}(k,w)\cdot X(k,w), w=0,1,\ldots,W_{Xk}-1 \tag{9}$$

Wherein, the filter of the BS denoising algorithm WDWF_{Π} is represented by the following equation:

$$\begin{array}{l}H_{WDWF|\Pi}(k,w)=H_{bn|\Pi}(k,w)\cdot H_{dn|\Pi}(k,w)\\ where,\\ H_{bn|\Pi}(k,w)=\left\{c\cdot A_{kw}\cdot P_{nFP;a}(k)\le P_{x;a}(k,w)|_{\Pi}\right\}\\ H_{dn|\Pi}(k,w)=\left\{1-c\cdot A_{kw}\cdot\left[\frac{P_{nFP;a}(k)}{P_{x;a}(k,w)|_{\Pi}}\right]^b\right\}\end{array} \tag{10}$$

In Eq. (10), the noise estimate $P_{nFP;a}(k)$ is the average of all a-power amplitude noise estimates $P_{nFP;a}(k,w)$, provided by the kth band wavelet coefficients of the noise footprint, the specific expression is Eq. (11) below:

$$P_{nFP;a}(k)=\frac{1}{W_{Nk}}\sum_{w=0}^{W_{Nk}-1}|N_{FP}(k,w)|^2 \tag{11}$$

where the power estimate $P_{x;a}(k,w)|_{\Pi}$ is provided by equation (12), d is the memory parameter, and d_{PS} is the varying memory term:

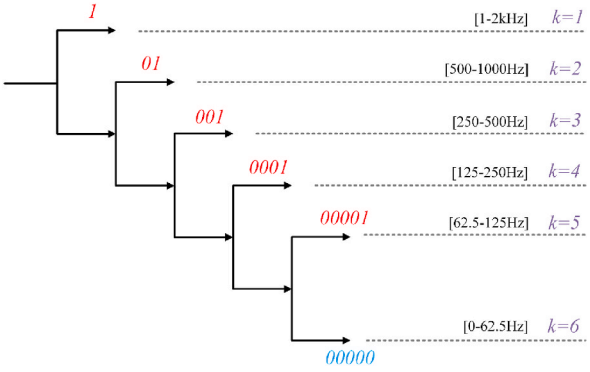


Fig. 4. 6-band DWT decomposition tree for WDWF.

$$P_{x;a}(k,w)|_{\Pi} = \left\{ \begin{array}{l} |X(k,w)|^a, w = 0 \\ d \cdot |X(k,w)|^a + (1-d) \cdot P_{s\sim;a}(k,w-1), w = 1,2,\ldots,W_{Xk}-1 \end{array} \right\}$$

where, (12)

$$P_{s\sim;a}(k,w) = d_{PS} \cdot |S\sim(k,w)|^a + (1-d_{PS}) \cdot P_{s\sim;a}(k,w-1), w = 1,2,\ldots,W_{Xk}-1, d_{PS}\to 1$$

$$P_{s\sim;a}(k,w) = |S\sim(k,w)|^a, d_{PS} \cong 1$$

3.1.3. *WDWF realized by 6-band DWT*

When performing WT, there are usually two types of decomposition: discrete DWT and wavelet packet (WP). However, WP decomposition topology requires more computational cost and computational time, in comparison, DWT has minimal signal distortion and compromised computational cost although it loses a small portion of performance accuracy. Therefore, the standard 6-level DWT is chosen as the denoising process in this study, and the decomposition scheme of this 6-band WDWF is shown in Fig. 4. At a given sampling frequency of 4 kHz, the entire range of 0–2 kHz is divided into a classical octave analysis from 62.5 Hz to 2 kHz.

In this study, 6-band DWT (II) type WDWF is used with parameter configurations chosen as: $a = 2, b = 1, c = 1$, and $d = 0.1$. The perceptual parameter A_k is also chosen to be used through which the robustness of the denoising algorithm is enhanced.

By using 6-band DWT (II) WDWF to filter the original BS, the heart sounds, lung sounds and environmental background noise can be significantly attenuated. Fig. 5 shows the difference between the time domain and frequency domain before and after filtering the BS signal.

3.2. *Multi-domain feature extraction of BS signal*

On the one hand, BS signals contain a large amount of information with large individual variability, and if rely only on the human ear auscultation of a professional physician for analysis, this not only requires strong professional knowledge, but also makes it difficult to ensure that the information obtained can fully represent the characteristics of BS. On the other hand, BS are characterized by asymmetry, weak signals, high randomness, and wide dynamic range, and their amplitude levels and frequency components have significant variations. When analyzing BSs, if only consider a single linear relationship or a single feature domain, will unintentionally give up some important feature information, and it is difficult to explore the intrinsic information of the data in depth.

Therefore, this study uses a rather complex combination of feature parameters extracted from the analysis of time, frequency and joint time-frequency signal to complete the task of defecation intention prediction. Specifically, we summarize the 23 time-domain features $T_1, T_2, T_3, \ldots, T_{23}$ in Table 1, the 16 frequency-domain features $F_1, F_2, F_3, \ldots, F_{16}$ in Table 2 and the 2 time-frequency-

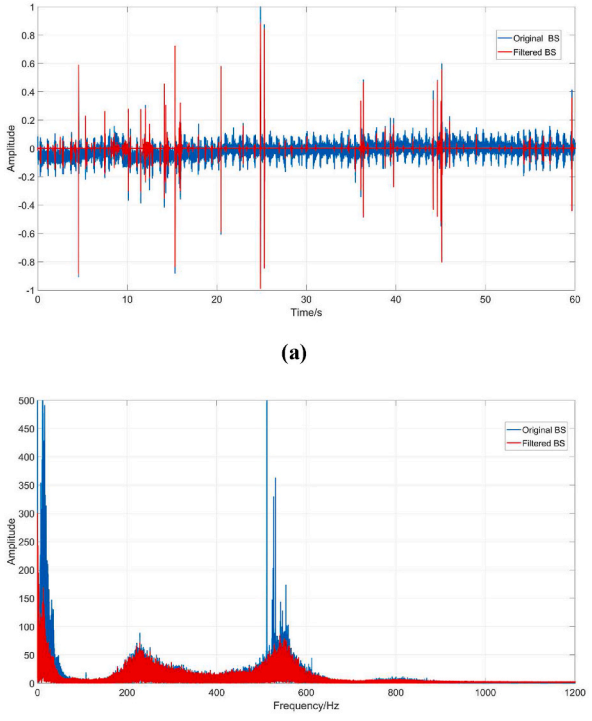


Fig. 5. Waveform and spectrogram of BS signal before and after filtering. (a) waveform diagram; (b) spectrogram.

Table 1
Time domain features expressions.

Feature Name	Feature Expression	Feature Connotation
Max	$T_1 = max\langle x(i)\rangle$	Intensity
Min	$T_2 = min\langle x(i)\rangle$	Intensity
Mean	$T_3 = \frac{1}{N}\sum_{i=1}^{N} x(i)$	Shift
Variance	$T_4 = \frac{1}{N}\sum_{i=1}^{N} [x(i) - T_3]^2$	Degree of dispersion
Standard deviation	$T_5 = \sqrt{T_4}$	Stability
Mean absolute deviation	$T_6 = \frac{1}{N}\sum_{i=1}^{N} \lvert x(i) - T_3\rvert$	Degree of dispersion
Root mean square	$T_7 = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x(i)^2}$	Stability
Average difference	$T_8 = \frac{1}{N}\sum_{i=1}^{N} [x(i) - T_3]$	Amplitude scale
Absolute energy	$T_9 = \sum_{i=1}^{N} x(i)^2$	Energy distribution
Peak to peak distance	$T_{10} = T_1 - T_2$	Periodicity
Sum of absolute differences	$T_{11} = \sum_{i=0}^{N-1} \lvert x(i+1) - x(i)\rvert$	Intensity of change
Shannon Entropy	$T_{12} = -\sum_{i=1}^{N} p(x(i)) \cdot \log_2 p(x(i))$	Uncertainties
Area under the curve	$T_{13} = \sum_{i=0}^{N-1} [(i+1) - (i)] \cdot \frac{x(i+1) + x(i)}{2}$	Energy distribution
Autocorrelation	$T_{14} = \frac{\frac{1}{N-k}\sum_{i=k+1}^{N} [x(i) - T_3][x(i-k) - T_3]}{\frac{1}{N}\sum_{i=1}^{N} [x(i) - T_3]^2}$	Periodicity
Signal center point	$T_{15} = \frac{\sum_{i=0}^{N} i \cdot x(i)^2}{\sum_{i=0}^{N} x(i)^2}$	Shift
Neighbor peaks	$T_{16} =$ *Number of peaks*	Repeatable
Signal distance	$T_{17} = \sum_{i=0}^{N-1} \sqrt{1 + [x(i+1) - x(i)]^2}$	Difference in point in time
Total energy	$T_{18} = \frac{\sum_{i=0}^{N} x(i)^2}{N}$	Overall strength and power
Zero crossing rate	$T_{19} =$ *Number of signal reversals*	Rate of change
Skewness	$T_{20} = \frac{1}{N}\sum_{i=1}^{N} \left[\frac{x(i) - T_3}{T_5}\right]^3$	Symmetry
kurtosis	$T_{21} = \frac{1}{N}\sum_{i=1}^{N} \left[\frac{x(i) - T_3}{T_5}\right]^4$	Peak distribution pattern
Number of positive/negative turning points	$T_{22} =$ *Number of PTP* $T_{23} =$ *Number of NTP*	Waveform characteristics

domain features TF_1, TF_2 in Table 3 into a feature ensemble $\{T_1, T_2, \ldots, T_{23}, F_1, F_2, \ldots, F_{16}, TF_1, TF_2\}$ and use it as a single piece of data in the deep learning training process. The collected BS signal can be expressed as $x(i), i = 1, 2, \cdots, N$, where N is the number of signal data points, here N is 240000.

(1) Time domain features

BS mostly show explosive contractions, which are manifested in the signal waveform graph as a sudden and large change in the signal amplitude level. In order to capture the characteristic information of explosive BS signals, a total of 23 time-domain features of BS are extracted by the method of statistical analysis using the TSFEL library [33], which are denoted as $T_1, T_2, T_3, \ldots, T_{23}$, respectively, as shown in Table 1.

(2) Frequency domain features

From the previous analysis, it is clear that the frequency range of BS is more limited, with almost all of the signal power spectral densities located between 100 and 500 Hz and significant variations in the frequency components. Based on this characteristic, it is necessary to characterize its frequency domain. Through the fast Fourier transform (FFT), the BS signal is changed from the time domain to the frequency domain, which can be expressed as $X(k), k = 1, 2, \cdots, K$, K is the number of spectral lines, f_k is the frequency value of the kth spectral line, $S(f_k)$ is the power spectral density within the f_k, and from this, a number of eigenvalues within the frequency domain are calculated, which are expressed as $F_1, F_2, F_3, \ldots, F_{16}$, and are shown in Table 2.

(3) Time-frequency domain features

One limitation of the time domain and frequency domain feature extraction methods is that it is difficult to observe the frequency change of the non-stationary component of the signal over time, because the information is only calculated from one domain, so important features with high resolution are discarded.

Table 2
Frequency domain features expressions.

Feature Name	Feature Expression	Feature Connotation
Max power spectrum	$F_1 = MAX\langle S(f_k)\rangle, 0 \le f_k \le f_{\max}$	Main frequency components
Max frequency	$F_2 = MAX\langle f_k\rangle, 0 \le f_k \le f_{\max}$	High frequency and bandwidth
Median frequency	$F_3 = f_m$, $which \int_0^{f_m} S(f)df = 0.5 \int S(f)df$	Spectral characteristics
Spectral centroid	$F_4 = \frac{\int f \cdot S(f)df}{\int S(f)df}$	The center of the frequency
Power bandwidth	$F_5 = \frac{f_h - f_l}{F_4}$ f_h is high frequency -3dB point, f_l is the low frequency -3dB point	Bandwidth and energy
Spectral distance	$F_6 = \int [S_1(f) - S_2(f)]^2 df$	Similarity between frequency
Spectral entropy	$F_7 = -\sum_{k=1}^{K} P(f_k) \cdot log_2 P(f_k)$, $which\ P(f_k) = \frac{S(f_k)}{\int S(f_k)df}$	Uncertainty in frequency
Spectral spread	$F_8 = \sqrt{\frac{\int (f - F_4)^2 \cdot S(f)df}{\int S(f)df}}$	Bandwidth
Spectral skewness	$F_9 = \frac{\frac{\int (f - F_4)^3 \cdot S(f)df}{\int S(f)df}}{{F_8}^3}$	Degree of skewness of the frequency
Spectral kurtosis	$F_{10} = \frac{\frac{\int (f - F_4)^4 \cdot S(f)df}{\int S(f)df}}{{F_8}^4}$	Degree of kurtosis in the frequency
Spectral roll-off	$F_{11} = f_{off}$, $which \int_0^{f_{off}} S(f)df = 0.95 \int S(f)df$	Energy decreases with frequency
Spectral roll-on	$F_{12} = f_{on}$, $which \int_0^{f_{on}} S(f)df = 1.15 \int S(f)df$	Energy rises with frequency
Fundamental frequency	$F_{13} = f_a$, $which\ |X(a)| = max\langle X(k)\rangle$	Harmonic analysis
Human range energy	$F_{14} = \frac{\int_{0.6Hz}^{2.5Hz} S(f)df}{\int S(f)df}$	Body energy range
Spectral slope	$F_{15} = K \cdot log_2\left(\frac{H_{high}}{H_{low}}\right)$ H_{high}/ H_{low} is the amplitude at high/low frequencies, K is the scaling factor	Filter characteristics and frequency response
Spectral variation	$F_{16} = 1 - \frac{\int_1^{f_K} f^2 \cdot S_1(f) \cdot S_2(f)df}{\sqrt{\left[\int_1^{f_K/2} f^2 \cdot S_1(f)df\right]\left[\int_{f_K/2}^{f_K} f^2 \cdot S_2(f)df\right]}}$ $S_1(f)$ is $(1 \sim \frac{f_K}{2})$ of the spectrum, $S_2(f)$ is $(\frac{f_K}{2} \sim f_K)$ of the spectrum	Spectral structure and frequency response

Table 3
Time-frequency domain features expressions.

Feature Name	Feature Expression	Feature Connotation
Wavelet energy	$TF_1 = \int |X_k(w)|^2 dw$	Time-frequency distribution properties
Wavelet entropy	$TF_2 = -\sum_{w=0}^{W_{xk}-1} d_k(w) \cdot log_2 d_k(w)$, $which\ d_k(w) = \frac{X_k(w)}{\int X_k(w)dw}$	Stochastic characterization at different scales

In order to further improve the frequency information of BS signal observed in a small range, WT is selected to extract the time-frequency domain features of BS. The sliding window length, $L = 20000$, $X_k(w)$ is the WT coefficients (band $w = 0, 1, \ldots, W_{Xk} - 1$) for the kth band. The extracted features are wavelet energy and wavelet entropy, denoted as TF_1, TF_2, respectively, as shown in Table 3. On the premise that multiple features have been extracted in the time and frequency domains, we use multimodal feature fusion to process the features in the time-frequency domain, which does not introduce more complex calculations, but also makes it possible to characterize the bowel sound signals in a more integrated and comprehensive way, and to use the complementary information to further enhance the overall analysis.

4. BS data augmentation algorithm based on informer model

Deep learning models are data-driven classification algorithms that require a training set with a large data sample capacity and relatively balanced data categories for parameter optimization, so that the optimally trained model can have higher classification performance and better generalization.

However, in the realistic BS data collection process, due to the low number of daily defecation of patients, the frequency of defecation is relatively fixed, and some subjects have irregular defecation, it is difficult to collect BS data with defecation intention in the measured period. Based on the above factors, the collected dataset presents an overall small sample size and an imbalance between positive and negative sample sizes. Such a raw dataset as a training set will cause the model classification to be more biased towards categories with more sample sizes, which will result in training prone to overfitting and weakened model generalization.

To this end, this paper proposes a way of expanding audio signal data based on the Informer temporal signal prediction model, where a single raw BS data is taken as an input, and the model will learn to fit the input signal distribution and output the fitting result in the form of a prediction, thus obtaining an augmented dataset based on the original data.

4.1. Long series time series forecasting model — informer

Informer is proposed in order to solve the relatively basic but very important problem—long sequence time-series forecasting [34]. which is able to effectively capture the precise long-term dependency coupling between the outputs and inputs with high predictive capability. Informer improves on Transformer [35] and focuses on solving the original model's problems of high time complexity, high memory usage and architectural limitations.

As shown in Fig. 6, the Informer consists of an embedding coding block, an Encoder component, a Decoder component, a fully connected layer, and connections between them. The embedding coding block performs embedding projection of the input BS sequence, which mainly consists of position coding and timestamp coding, and then passes the coded sequence to the Encoder and Decoder. The Encoder component consists of two coding stacks of varying sizes, each stacked by multiple Muti-head *ProbSparse* Self-attention blocks and Self-attention Distilling blocks, and the Encoder extracts the internal linkage features of the BS coding sequence and passes them as outputs into the Decoder. The Decoder component consists of a Masked Muti-head *ProbSparse* Self-attention block and a Muti-head Attention block, where the Decoder extracts the features of the BS sequence to be predicted and fuses them with the incoming features from the Encoder. The fully connected layer receives the set of features passed in by the Decoder and directly predicts the output target BS sequence.

Notable among them are the following three particular innovations of Informer.

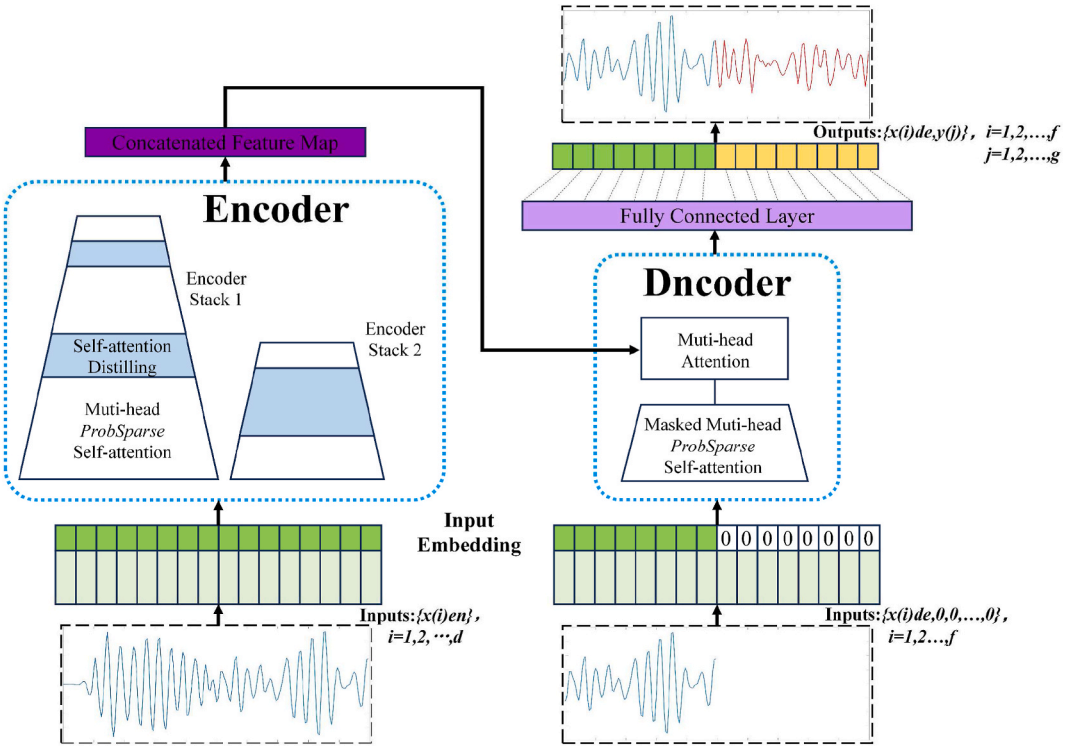


Fig. 6. Informer's fitting prediction process for BS signals.

- Probabilistic sparse self-attention mechanism (*ProbSparse* Self-attention). This mechanism addresses the higher memory footprint and higher cost of quadratic dot product computation
- Self-attention Distilling tightens the model size and reduces the number of model weight matrices without loss of self-attention performance.
- Generative style decoder. Regardless of training or prediction, Decoder's input sequence is divided into two parts: a known sequence before the time point to be predicted and a sequence of "0" after the time point to be predicted, and Decoder's output goal is to complete the "0" in the sequence with the prediction result.

4.2. *Informer application for data augmentation of BS*

The idea of using Informer to implement fitted augmentation for BS comes from a common audio augmentation strategy— Time masking. Time masking is the process of masking or replacing one or more segments of the original signal with a silent signal. However, Time masking is generally used randomly, and it is easy to erase important parts of the original signal, which results in the loss of important information in the sample, even though a new sample is obtained. Instead, a fitting prediction method is used, which learns by fitting a distribution to a part of the signal and replaces the original part with the predicted output. Because the replaced part is the fitting learning of the original part, and most of the information is still retained after the replacement while it is not a simple repetition of the original part, so as to realize the data augmentation and expansion without loss of information.

Based on this idea, this study chose to use the Informer model. Informer learns to fit the distribution and internal characteristics of a given whole sequence segment and outputs the sequence values for a number of unknown points subsequent to a location point within the sequence segment in the final prediction result.

As shown in Fig. 7, taking a small segment of BS signal of length *C* as an example, when the signal comes into the model, a segment of BS sequence of length *D* will be intercepted as the input to the Encoder (green), which consists of data points (blue) and signal values (gray). Then, the Encoder end input sequence of length *F* and the target prediction sequence of length *G* are taken to form the input of the decoder, where the former part also contains data points and signal values, and the latter part consists of data points (blue) and assigned "0" signal values (white).

As shown in Fig. 8, the BS input sequence is subjected to a positional embedding operation through the embedding coding module before it enters the Encoder, which mainly consists of local timestamps and aligning dimensions, with the aim of obtaining the local timing information of the input sequence and aligning it dimensionally. In the next step, the Encoder receives the input sequence where the primary coding stack (stack1) receives the entire input sequence while the secondary coding stack (stack2) takes half of the input and the second coding stack throws away a layer of self-attention module and self-attention distilling module. The input sequence is summed by multiple *ProbSparse* self-attention modules to obtain the features representation, and the output dimensions of the two coding stacks are aligned and summed to obtain the final set of features representations. The *ProbSparse* self-attention mechanism

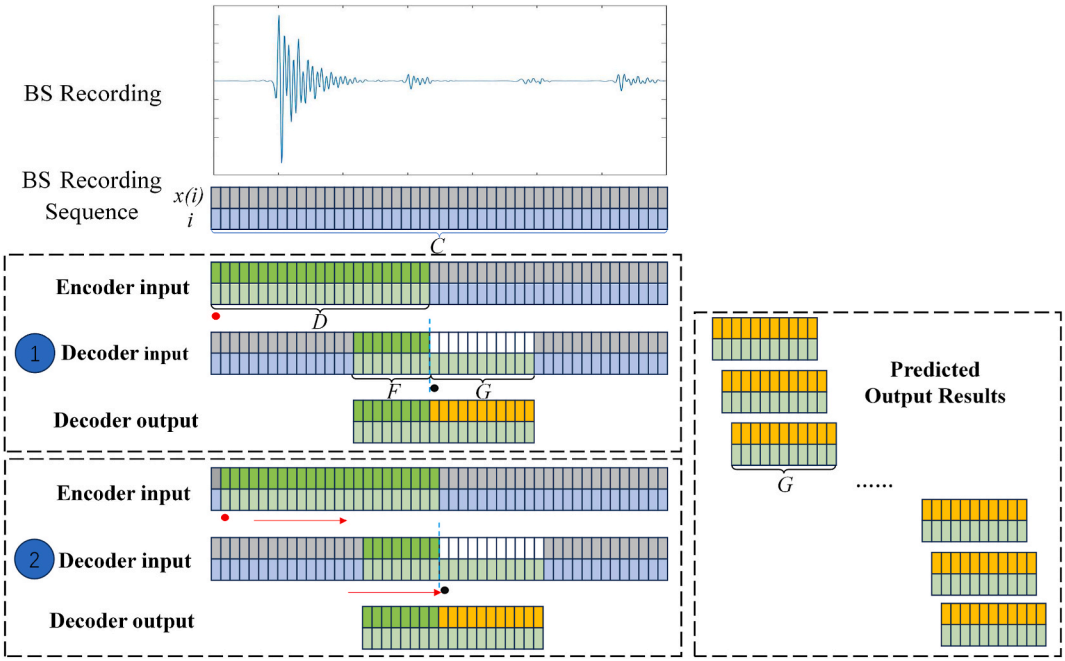


Fig. 7. Example of segmentation of BS signal input and the BS sequence of predicted output results, with black dots marking the target prediction points and red dots marking the start of the sequence. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

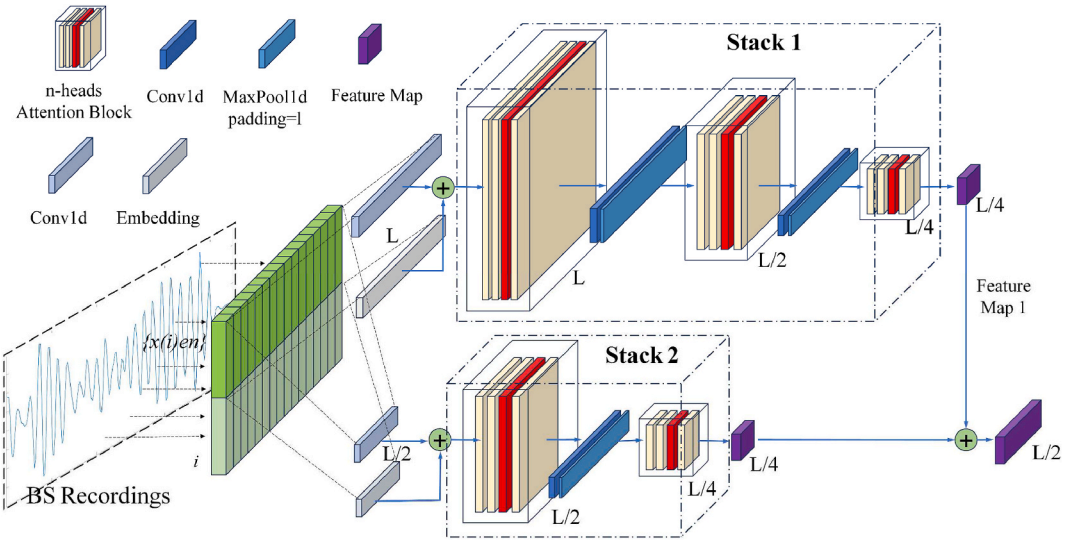


Fig. 8. Self-attentive features extraction of BS signal by Encoder.

utilizes sparse matrices to replace the original attention matrix, dramatically reducing the arithmetic demand while maintaining good performance, while the self-attention distillation module can tighten the size of the model without losing self-attention performance.

Fig. 9 illustrates the dynamic decoding process of the Decoder, where the BS input sequence also requires the positional embedding operation of the embedding coding module before entering the Decoder. Immediately after that, the Decoder receives the input sequence, interacts with the set of feature representations input from the Encoder via the Multi-head Attention module, and passes the set of features into the fully connected layer, and finally directly predicts the output target sequence (yellow part in Fig. 9).

As a result, the predicted signal value at the target prediction point about the subsequent G data points is obtained. After that, each time the sequence start point will be shifted one place backward and keep the value of each length of the signal sequence segmentation unchanged, repeat the above process, and so on, and finally obtain the predicted sequences within the BS signal of length $C - D$ about G points after each position point, as shown in Fig. 7. These sequences are arranged in the order of the target prediction points, and a complete predicted BS signal of length $C - D$ is obtained by piecing together multiple sequences with G as the interval length. By performing the same operation, G predicted BS signals are obtained.

The purpose of applying the Informer model in this paper is to perform data augmentation while trying to avoid the signal data obtained from augmentation being too similar to the original signal data. For such consideration, the discontinuity generated from the end point of the generated signal to the next point when filling in the blanks is within the acceptable range, because a certain degree of discontinuity is rather beneficial to promote the diversity of the augmented data. Based on the power of the Informer model, only 2–3 epochs of iterative training are carried out in the actual application, and better broadening results can be obtained.

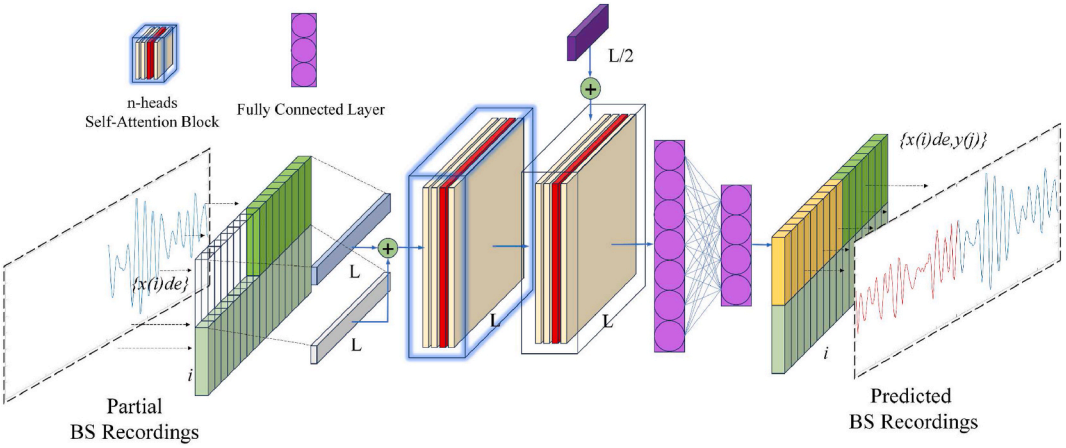


Fig. 9. Dynamic decoding process of the Decoder.

5. Defecation intention prediction algorithm based on BS features

The defecation intention classification prediction model based on BS multidomain features and an improved 1D residual network (1D-IResNet) is shown in Fig. 10. A sequence of BS signals is filtered and processed to extract appropriate time domain, frequency domain and time-frequency domain features constructed as a multi-domain feature set, which is used as an input to the training process and testing process of the 1D-IResNet in order to generate the optimal weights for outputting global prediction results.

The modified residual neural network used in this study is different from the conventional residual neural network [36]. One of its most important improvements is a minor modification to the path B of the down sampling block, as shown in Fig. 11. The 1×1 convolution in path B of the original down sampling block ignores 3/4 of the input feature representation. Based on experience, a 1×2 average pooling layer with a step size of 2 is added before the convolution, and its step size is changed to 1. After the modification, the down sampling block does not ignore the feature information, which works well in practice and has little impact on the computational cost.

The convolutional layer is the backbone part of 1D-IResNet. The convolutional layer uses a 1D convolutional kernel to extract features from the input data to obtain a feature map of the BS information. After each convolutional operation, a batch norm and ReLU operation is performed on the output feature map to prevent multiple convolutional layers from degenerating into one. The input data after batch norm not only retains the original features, but also accelerates the training. ReLU activation function can increase the nonlinear fitting ability of the network, which is conducive to the learning of the deep network.

The final output layer of the network is a fully connected layer, where the BS signal feature representation obtained by convolution is compressed to one dimension by a fully connected layer after global average pooling, and finally the binary classification output is achieved by a Sigmoid function layer (0:no defecation, 1:defecation).

6. Experimental results and analysis

All data processing algorithms for this experiment are compiled on PyCharm Community Edition software, and the deep learning algorithms are implemented based on TensorFlow and Pytorch deep learning libraries. All experiments are conducted on Windows 10 server with Intel(R) Xeon(R) CPU, NVIDIA GeForce GTX 2080 (11 GB).

6.1. *Description of experimental data and ethical approval*

A total of 29 volunteers are recruited for this study, and a total of 378 audio data of BS signals with a single duration of 60s are collected. Some of the data had been labeled as special cases at the time of data collection because the volunteers had substantial limb movement or coughing. This type of data is excluded from subsequent use, and a total of 223 sets of data are actually included in the experiment, of which 153 sets are labeled as no defecation and 70 sets are labeled as having defecation. Volunteers included disabled

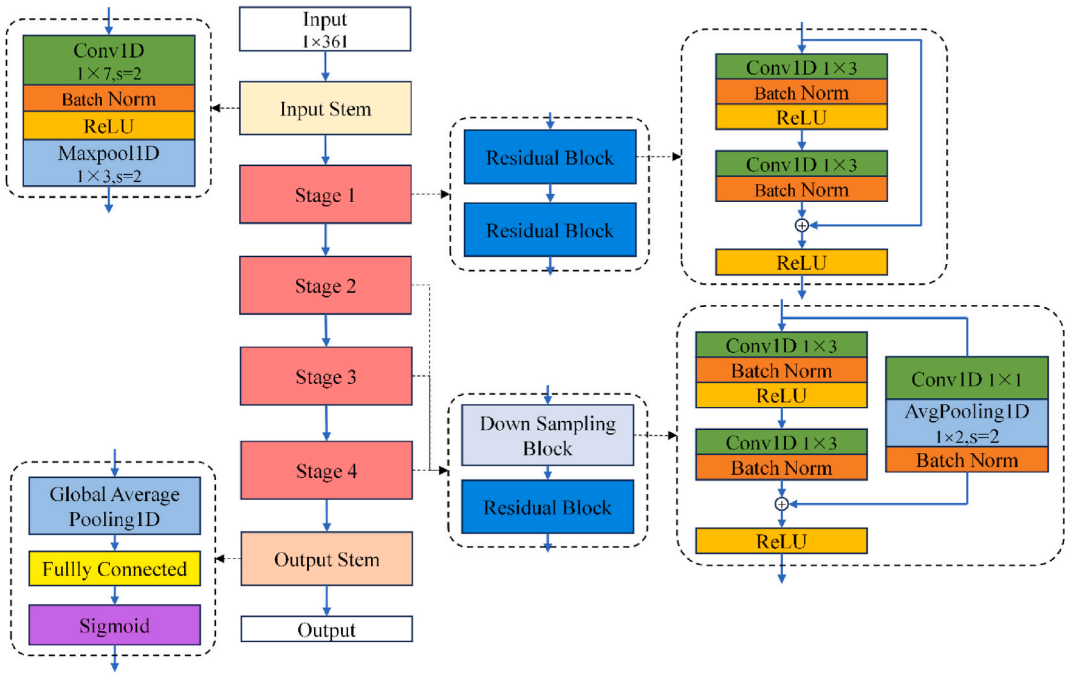


Fig. 10. 1D-IResNet defecation intention classification model.

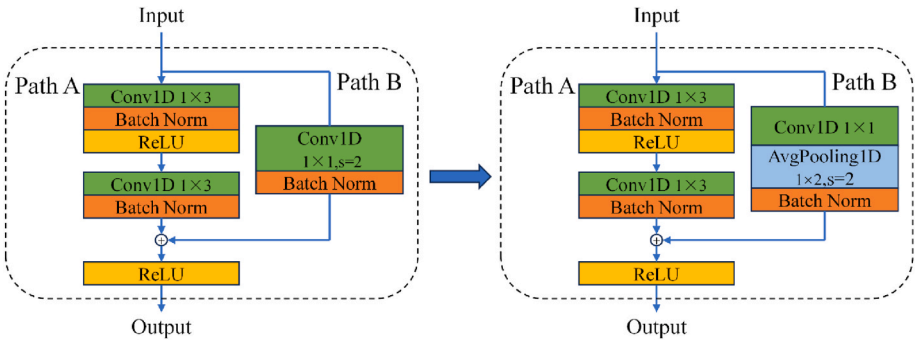


Fig. 11. Modifications to the Down sampling block in the improved residual network.

patients from Beijing Boai Hospital, disabled patients from the Third Affiliated Hospital of Sun Yat-sen University in Guangzhou, and some master's degree students from South China University of Technology, with an age distribution ranging from 23 to 80 years old. The recruitment period for this study begins on June 1, 2021 and ends on May 30, 2022. This study was reviewed and approved by

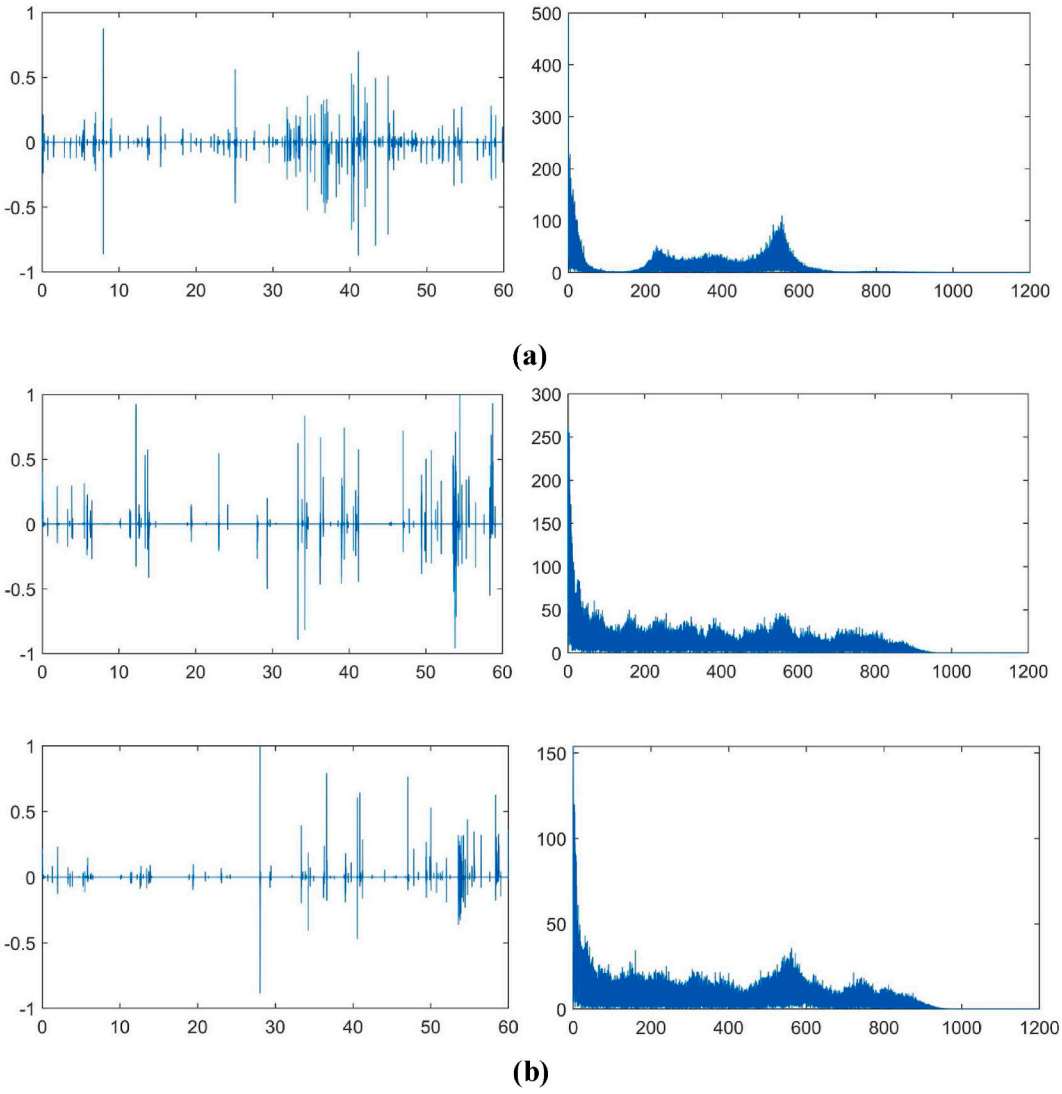


Fig. 12. Waveforms and spectra of original BS, routine augmented BS and Informer augmented BS. (a) the original BS; (b) the routine augmentation of BS; (c) the Informer augmentation of BS.

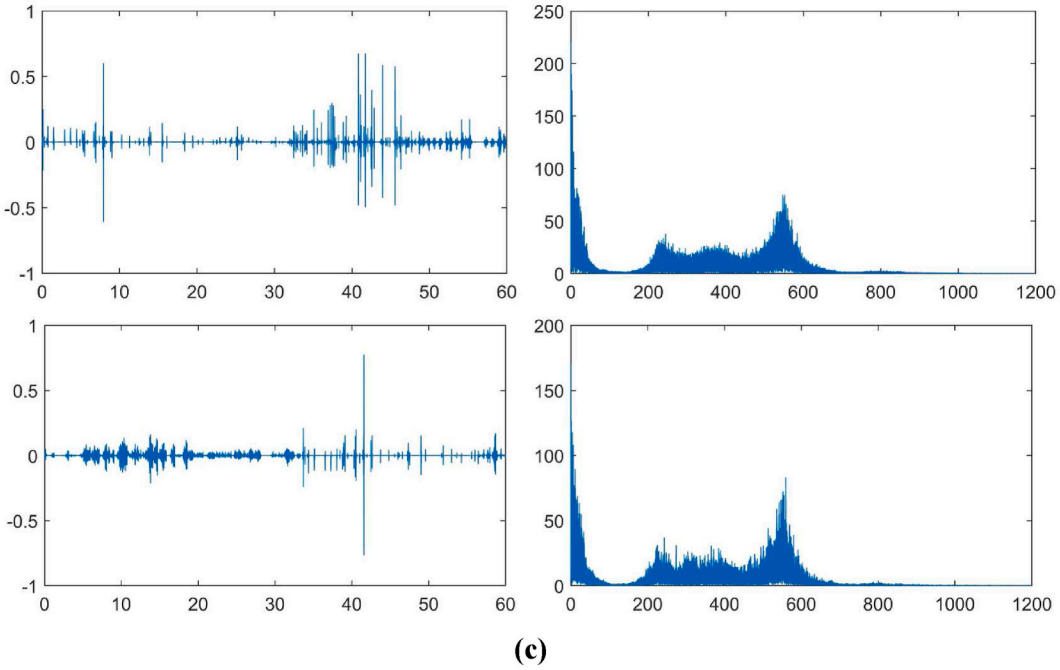


Fig. 12. (*continued*).

[The Medical Ethics Committee of China Rehabilitation Research Center], with the approval number: [2020-085-1]. All participants volunteered to participate and are fully aware of the content and data flow of the study before participation. The group of disabled patients signed a written informed consent and the group of postgraduate students were given a verbal informed consent witnessed by their advisors, and no minors were involved in this study. The test set used to validate the accuracy of the model in the subsequent experiments was derived from bowel sound data from different sources or populations, and the source of this portion of the data and itself was not modified throughout the experiments.

6.2. *Comparison of the effectiveness of audio data augmentation methods*

In order to investigate the effect of the Informer data augmentation algorithm for audio signals proposed in this paper, this paper also used conventional audio data augmentation algorithms to process the same BS signals, and used the Pearson correlation coefficient index to initially quantitatively analyze the effect of BS augmentation. Time masking, time stretching, pitch shifting, segment transversal shifting and polarity reversal are mixed as the conventional audio augmentation, and three to five of them are randomly

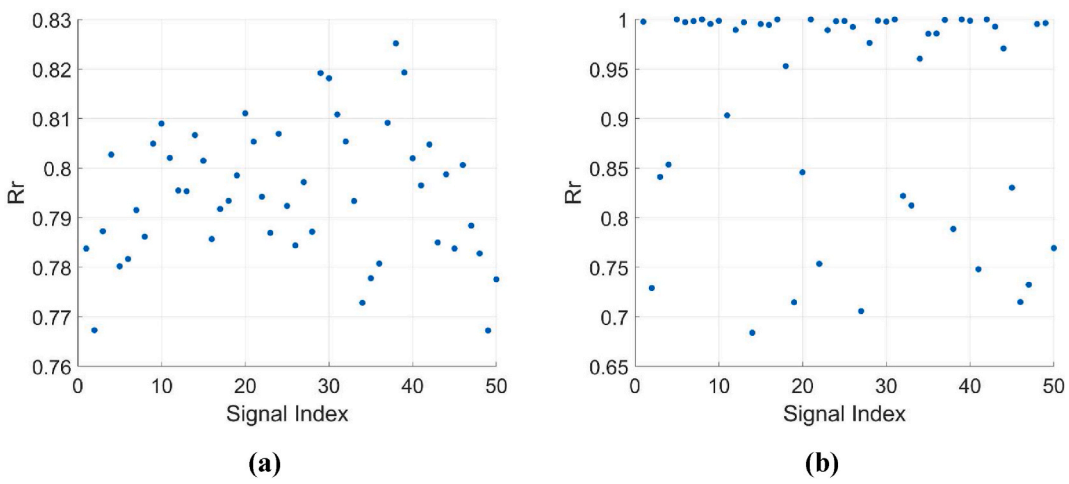


Fig. 13. Scatterplot of Pearson's correlation coefficient Rr in the frequency domain for the original BS signal and the augmented BS signal. (a) Rr of Informer augmentation algorithm; (b) Rr of routine augmentation algorithm.

used to augment the original BS, which are finally saved in the form of audio.

In order to compare the BS augmentation effect of the two algorithms more intuitively under the visual perception, some of the results of the original BS signals and the BS signals generated by the two methods are shown in Fig. 12 (filtered by the 6-band DWT (II) type WDWF filter and low-pass filter).

As can be seen in Fig. 12-(a) and Fig. 12-(b), the BS signal obtained by the conventional augmentation algorithm has a high spectral similarity to the original bowel tone, but its waveform profile in the time domain range differs significantly from the original value. On the other hand, as shown in Fig. 12-(c), the waveforms of the BS obtained by the Informer broadening algorithm are more diverse, and have the single burst and continuous multiple bursts of BS, but their spectrum are still significantly different from the original values.

Pearson's Correlation Coefficient is commonly used in the field of signal processing as a means of assessing the similarity of time-series signal data, and the larger the absolute value of the correlation coefficient between two sets of signals, the stronger the linear correlation. By calculating the Pearson's correlation coefficient Rr in the frequency domain for the original BS signal and the augmented BS signal, the correlation can initially reflect the effect of BS augmentation. The Pearson's Correlation Coefficient Rr is calculated for a single original BS signal and 50 BS signals obtained by the two broadening algorithms respectively, and the correlation is presented in the form of a scatter plot as shown in Fig. 13.

From the scatterplot, it can be seen that the correlation between the BS signals obtained by the Informer algorithm and the original signals is strong, and the Rr values are more concentrated, and all of them are located in the interval of 0.76–0.83, which belongs to the strong correlation. This indicates that rather than having a serious impact on the quality of the augmentation data, the discontinuities created when filling in the gaps have contributed to the diversity of the augmentation data. The correlation between the BS signals obtained by the routine algorithm and the original signals spanned a wide range, with the Rr values ranging from moderate correlation to very strong correlation, and some of the data had very strong correlation, with the Rr values reaching more than 0.95.

From the preliminary analysis of the scatterplot, it can be seen that the Informer augmentation algorithm has the performance comparable to the routine augmentation algorithm, and the correlation of the generated data is more concentrated. This can initially verify the effectiveness of the generalization algorithm used in this paper.

6.3. *Experiments on defecation prediction based on BS features*

(1) Experiments on Defecation Prediction With Different Augmentation Algorithms

In the defecation prediction task, the ultimate goal is to classify BS labeled as having a defecation and BS without a defecation, and to use the appearance of BS as a signal of an impending defecation and to provide advance warning. In order to validate the effectiveness of the proposed BS classification algorithm based on multi-domain features of BS and improved 1D residual network, and to further test the performance of the BS data augmentation algorithm based on the Informer model. In this part, the original BS dataset (O set), the balanced BS dataset after censoring negative samples (Set B), the Informer augmented BS dataset (I set), and the routine augmented BS dataset (R set) are used as the training set for the defecation classification model (1D-IResNet), respectively, and the prediction results of 1D-IResNet are examined and analyzed on the test set.

The original dataset consisted of 223 signals, of which, 153 signals are labeled as no defecation and 70 signals are labeled as defecation, with a positive to negative sample ratio of approximately 1:2. We use 80 % of the data for model training and 20 % for testing the accuracy of the model. And 75 % of the training data is divided into training set and the remaining 25 % is divided into validation set. The training set is used to optimize the model parameters. The validation set is used to check the convergence of the model during training and to adjust the hyperparameters. The test set is used to evaluate the performance of the model after training. The data are randomly selected and the ratio of training set, validation set and test set is 6:2:2. With this division ratio, 80 % of the data

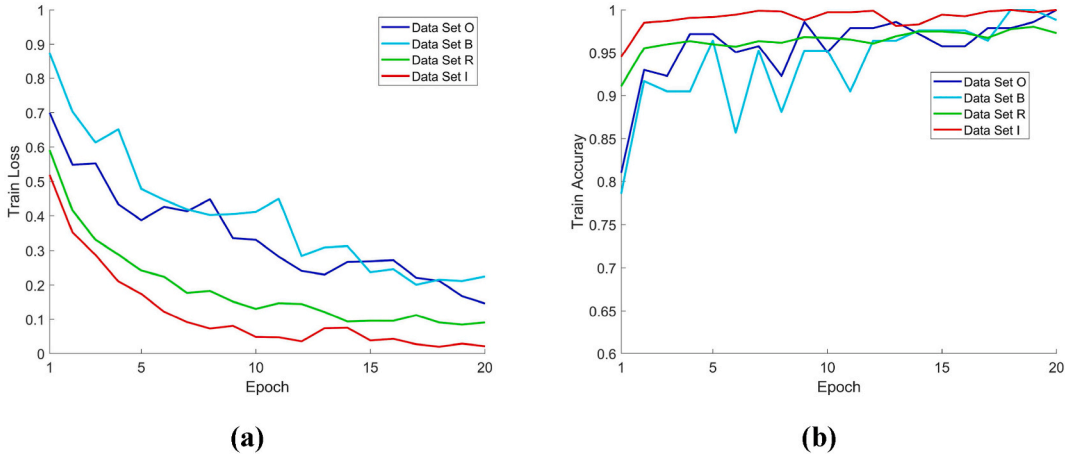


Fig. 14. Training loss and accuracy for O-set, B-set, R-set and I-set. (a) Training loss; (b) Training accuracy.

as training set provides enough samples for training and tuning the model while the remaining 20 % of the data as test set is sufficient for evaluating the performance of the model.

In addition, we randomly eliminate the signals without defecation intention in the original BS dataset according to the ratio of 1:1, so that the positive and negative sample sizes are balanced. The balanced BS dataset has a total of 140 signals, which are divided into training, validation, and testing sets in the ratio of 6:2:2. Again, the divisions of the data are chosen randomly.

In the data augmentation stage, 13 BS signals with defecation and 12 BS signals without defecation are randomly selected from the original dataset in this study, and these 25 signals are augmented using the Informer augmentation algorithm and the routine augmentation algorithm, respectively, and 50 child augmented BS signals are obtained using each BS as a parent template. When dividing the model dataset, 68 BS signals are randomly selected from the original dataset, and together with the 1250 child BS signals obtained from the generalization and 25 parent BS signals, they are used as the training set for the defecation intention classification model (1D-IResNet), and the remaining 130 original BS signals are used as the test set.

The network weights of the 1D-IResNet network are randomly initialized, the loss is calculated using the binary cross-entropy loss function, the initial learning rate is = 0.0005 during training, and the learning rate dynamic attenuation strategy is used when the parameter is updated. A minibatch size of 32 is chosen for each training phase for the I set and R set with large data sample sizes, while a minibatch size of 16 is chosen for the O set and B set with small sample sizes. the training process is 20 epochs for both.

Fig. 14 shows the model training loss and accuracy results of 1D-IResNet in each of the four datasets. Fig. 14-(a) shows that the training loss of the prediction model decreases dramatically in the initial 1–10 epochs; at the 20th epoch, the training loss on the O and B sets converges to 0.2, the training loss on the R set converges to 0.1, and the training loss on the I set converges to 0.05. Fig. 14-(b) further shows that in the initial 1–3 epochs, the accuracy of the prediction model varies greatly, and in the 10th epochs, the accuracy is already greater than 0.9, with large fluctuations in the accuracy on the O- and B-sets; and it has basically stabilized between 0.95 and 1 in the 20th epochs, with no significant changes in the fluctuations. Therefore, combining the loss function and accuracy of the model, the model has reached a converged state at 20 epochs.

Fig. 15-(a) and Fig. 15-(b) shows the model receiver operating characteristic curves and areas of 1D-IResNet on the test set of 4 datasets. As can be seen from the results, the AUC values on all four datasets exceed 0.9, indicating that 1D-IResNet has high accuracy and prediction ability on small, unbalanced and large sample datasets. However, the effect of the dataset's volume size on the model's classification ability is stronger than the dataset's imbalance.

We statistically analyzed the classification accuracy, precision, recall and F1-score of the model, and used them to evaluate the classification performance of the model in detail. Their calculation formulas are as follows:

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \tag{13}$$

$$Pre = \frac{TP}{TP + FP} \times 100\% \tag{14}$$

$$Rec = \frac{TP}{TP + FN} \times 100\% \tag{15}$$

$$F1 = \frac{2P \cdot R}{P + R} \times 100\% \tag{16}$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

In order to reduce the effect of random initialization of network weights on the analysis results, in this study, the network is trained

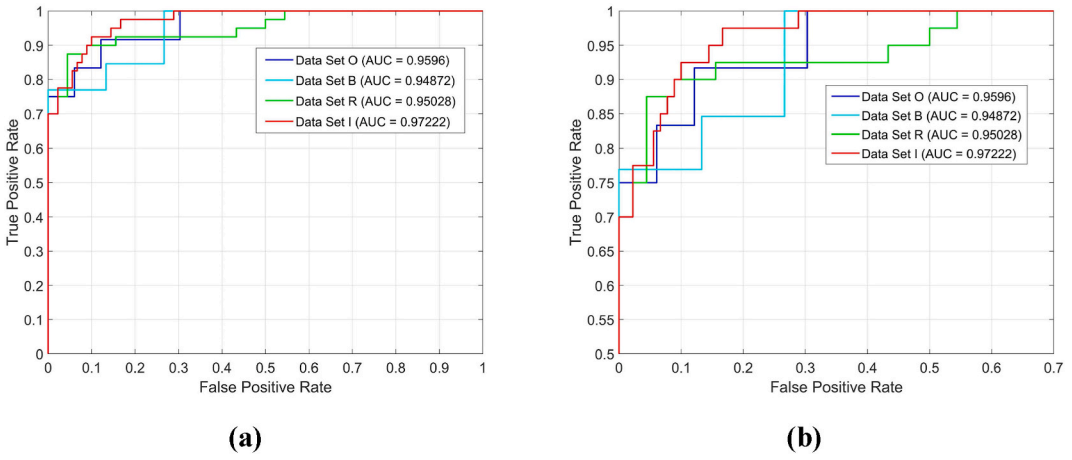


Fig. 15. Receiver operating characteristic curve and area (AUC-ROC) for O-set, B-set, R-set and I-set.

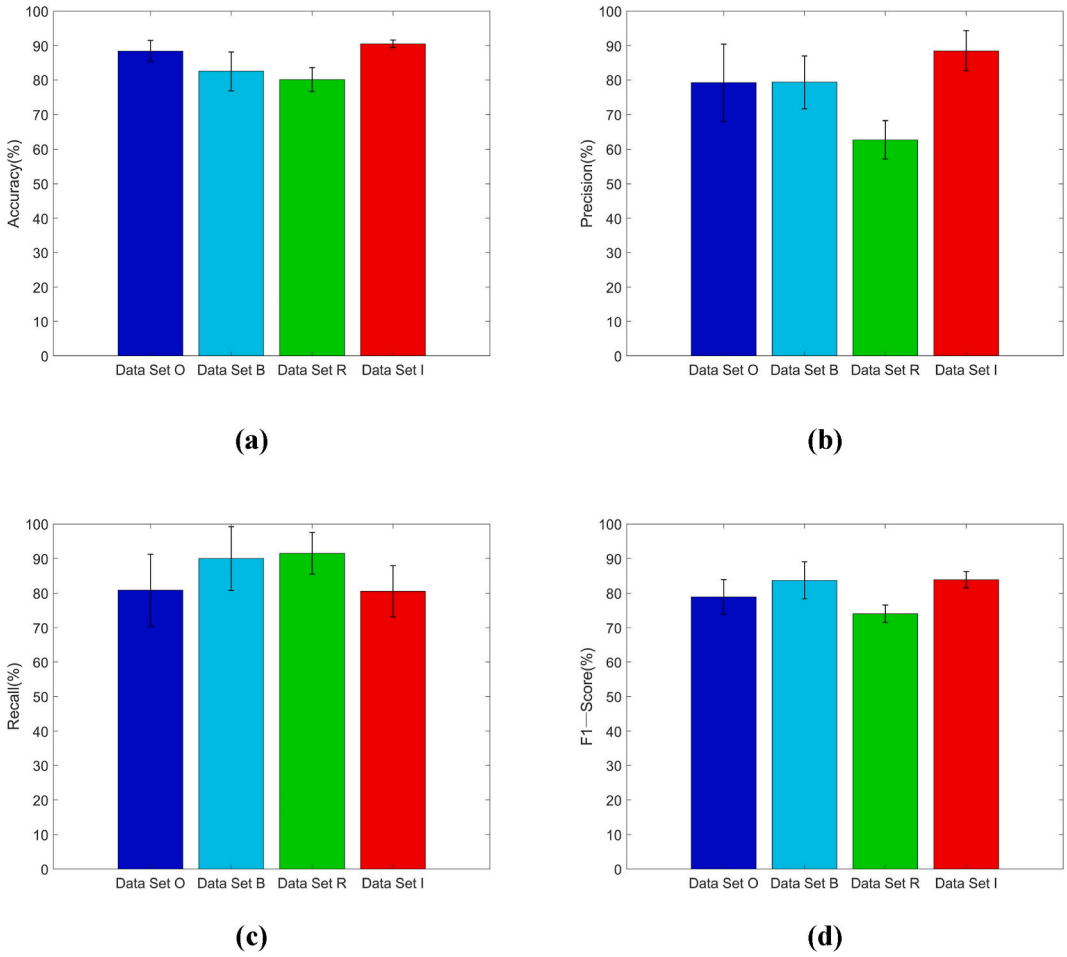


Fig. 16. Accuracy metrics for classification models trained on the four datasets. (a) accuracy; (b) precision; (c) recall; (d) F1-score.

Table 4
Means and standard deviations of the indicators obtained from the four datasets tested on the 1D-IResNet model.

Datasets	Acc(%)	Pre(%)	Rec(%)	F1(%)
O set	88.44 ± 3.11	79.22 ± 11.25	80.83 ± 10.43	78.91 ± 5.00
B set	82.50 ± 5.70	79.36 ± 7.68	90.00 ± 9.21	83.70 ± 5.31
R set	80.08 ± 3.49	62.68 ± 5.70	91.50 ± 6.03	74.02 ± 2.45
I set	**90.54 ± 1.15**	**88.51 ± 5.86**	80.50 ± 7.43	**83.88 ± 2.34**

10 times under each of four different datasets and the mean values of classification accuracy, precision, recall and F1-score on the test set are counted and the results are shown in Fig. 16 and Table 4.

From Fig. 16(a–d) and Table 5, it can be seen that the 1D-IResNet model has good classification accuracy and high F1 score in the original dataset, reaching 88.44 % and 78.91 %, respectively. At the same time, its precision and recall rates are also above 79 %. This indicates that the model is more accurate in the classification process and achieves a more balanced state between misclassifying negative cases and omitting true positive cases. The model can still have a better classification effect under small sample size.

The balanced dataset after removing some negative samples by 50 % has a mediocre performance, although the recall and F1 score have improved and the model is biased towards capturing real positive examples, the overall classification accuracy has dropped to 82.50 %. This performance is attributed to the small size of the original dataset, and although removing some of the negative samples will make the dataset balanced in terms of positive and negative samples, it also results in a training set that is too small, which affects the model's generalization ability and classification accuracy.

On the other hand, the BS dataset obtained by the conventional enhancement algorithm performs poorly after model training. Both the accuracy and F1 score decreased compared to the original dataset, with the accuracy dropping to 80.08 % and the F1 score only 74.02 %. And from the 62.68 % precision and 91.50 % recall, it can be seen that the model has been heavily biased towards capturing real positive examples, which leads to more serious misclassification cases. The reason for this problem is that the BS obtained by the routine augmentation algorithm are too similar to the original BS in the frequency domain, and a small portion of the data is even no different from the original data. Such a result leads to a reduction in the diversity of the training dataset, and the model repeatedly learns features from the same type of data several times during the training process, which makes it difficult to capture the main categorization features and results in the overfitting. From the analysis of Pearson's correlation coefficient, it can be seen that most of the Rr coefficients between the BS obtained by the routine augmentation algorithm and the original BS are more than 95 %, and although their correlation is very strong, the extremely strong correlation also brings the risk of easy overfitting.

On the contrary, the performance of the Informer augmentation algorithm after the model training, its classification accuracy and F1-score have been improved, the accuracy is improved to 90.54 %, F1 score is improved to 83.88 %, and it still has a better accuracy stability while maintaining a higher classification index, and its standard deviation is only 1.15 and 2.34. At the same time, the precision is also increased to 88.51 %, the recall remains above 80 %, indicating that although the model is more inclined to reduce the misjudgment of negative cases, but still maintains a better ability to capture the real positive cases, and the difference between the two is within 10 % in did not appear imbalance. It shows that after the Informer augmentation algorithm, the diversity of the BSs dataset is guaranteed, and the positive and negative samples in the training set are balanced, so that the model will not prioritize learning only one class of sample features during the training process. At the same time, because of the expansion of the training set samples, the convergence speed of the model has been improved, and the model can converge in only 10 epochs in multiple trainings, and also has a good performance on the test set.

(2) Experiments on Defecation Prediction With Different Deep Learning Algorithms

To further validate the performance advantages of the defecation prediction algorithms proposed in this paper, several representative classification algorithms are selected for comparison in this study, which include residual neural network (1D-ResNet), k-nearest neighbor algorithm (KNN), support vector machine (SVM) and BP neural network (BP). Using the same training strategy for each algorithm, each is trained 10 times under the balanced dataset obtained using the Informer data augmentation algorithm and their accuracy, precision, recall and F1-scores on the test set are counted respectively, and the results obtained are shown in Fig. 17 and Table 5.

From Fig. 17(a–d) and Tables 5 and it can be seen that among the five models, the 1D-IResNet model has the best classification effect, and its accuracy and F1-score are much higher than those of SVM, KNN, and BP, and maintains a relatively good balance between reducing the number of negative cases of miscarriage of justice and reducing the number of true positive cases of omission. Meanwhile, compared with the original 1D-ResNet, the improved residual block has a stronger field of view for feature extraction, which can improve the feature learning ability of the network and obtain better classification ability. The test results show that the proposed 1D-IResNet defecation prediction algorithm is efficient, It has a high detection accuracy while also having a high checking and recall rate, and a good balance in the classification of positive and negative cases, demonstrating the superiority of the text-proposed algorithm.

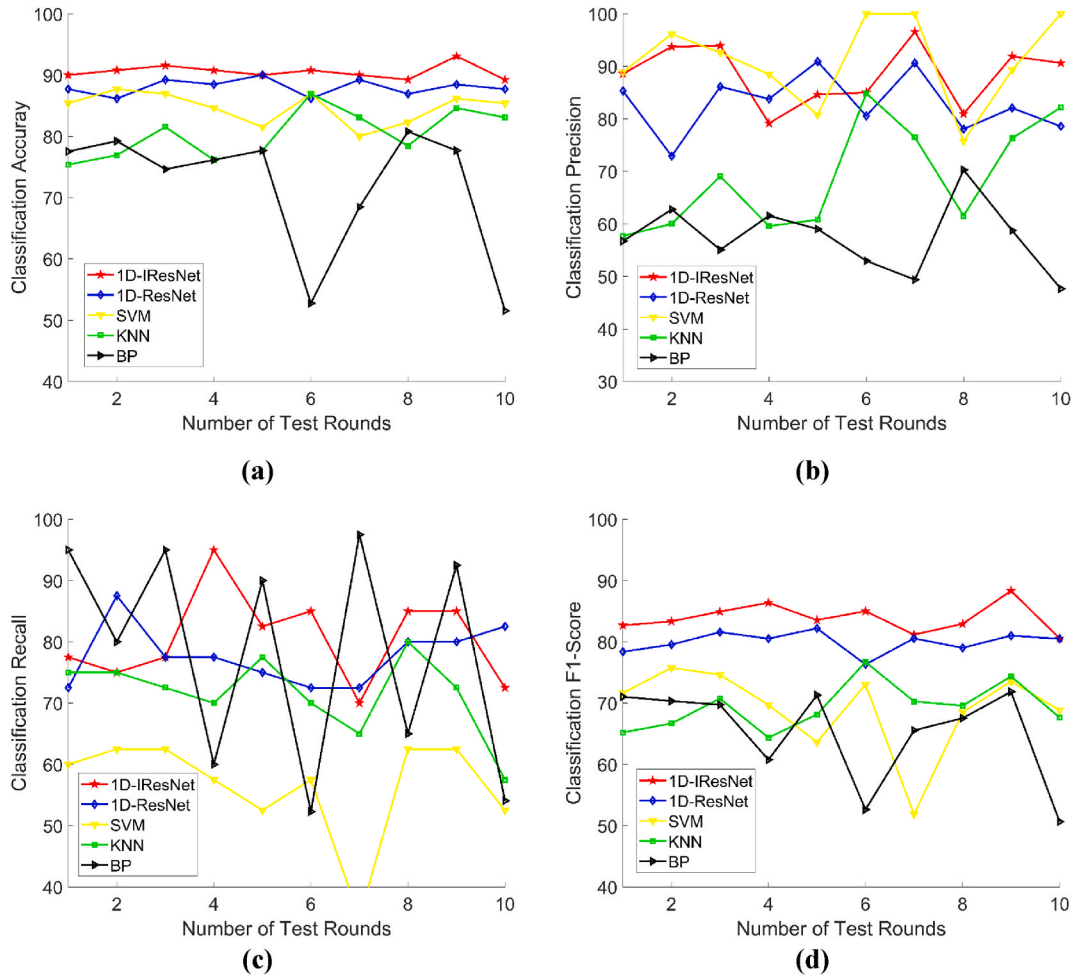


Fig. 17. Results of ten tests of 1D-IResNet, 1D-ResNet, SVM, KNN and BP on I set. (a) accuracy; (b) precision; (c) recall; (d) F1-score.

Table 5
Means and criteria of 1D-IResNet, 1D-ResNet, SVM, KNN and BP for testing each metric on the I set.

Algorithm	Acc(%)	Pre(%)	Rec(%)	F1(%)
1D-IResNet	90.54 ± 1.15	88.51 ± 5.86	80.50 ± 7.43	83.88 ± 2.34
1D-ResNet	88.00 ± 1.32	82.89 ± 5.66	77.75 ± 4.92	79.96 ± 1.72
SVM	84.69 ± 2.58	91.19 ± 8.32	56.50 ± 8.51	69.11 ± 7.03
KNN	80.38 ± 3.98	68.84 ± 10.29	71.50 ± 6.48	69.37 ± 3.88
BP	71.64 ± 10.80	57.41 ± 6.68	78.13 ± 18.39	65.13 ± 7.87

7. Conclusion

There are fewer studies on the perception of willingness to defecate. As shown in Table 6, Most of them use an invasive approach to defecation intention monitoring, and the correlation between bowel pressure changes and defecation reflexes is used for perception studies. Invasive defecation intention perception has higher recognition accuracy compared to non-invasive. However, it is difficult to achieve long-term stable real-time recognition with invasive methods, and the invasive nature of the device makes it difficult to ensure safety and comfort during use.

In this paper, a human defecation prediction method based on Informer audio data augmentation and improved residual neural network is proposed to address the challenges of defecation care for disabled patients. We use bowel sounds as a feature source and use deep learning to analyze the multi-domain features of bowel sounds for defecation intention. As an accurate, safe and practical non-invasive defecation intention detection method, its classification accuracy reaches 90.54 % and F1-score reaches 83.88 %, with its standard deviation only 1.15 and 2.34, meanwhile, the check-accuracy rate improves to 88.51 % and the recall rate stays above 80 %, which is able to capture the positive cases better while maintaining high prediction accuracy of the samples, and the accuracy of

Table 6
Comparison of relevant defecation intention perception studies.

Intrusive/non-intrusive	Signal Source	Research Methodology	Accuracy	Year
Intrusive	Rectal pressure	HAPC feature recognition based on wavelet packet analysis and SVM	94.7 %	2012 [3]
	Rectal pressure	Defecation intention based on optimal wavelet packet and SVM	95.45 %	2013 [37]
	Rectal pressure	Modeling fecal mass estimation based on pressure sensors	92 %	2020 [38]
	Rectal pressure	Modeling defecation intention perception based on pressure similarity	95 %	2022 [39]
Non-invasive	Defecation records	Identifying defecation-related features in nursing logs based on random forest algorithm	85 %	2018 [5]
	Bowel sound	Prediction method of human defecation based on Informer audio data augmentation and improved residual network	90.54 ± 1.15 %	/

defecation intention recognition is comparable to that of invasive. The method enables caregivers to provide better care by anticipating the defecation needs of disabled patients in advance of their defecation, and is more applicable to routine community care as the collection of BS is non-invasive and does not cause discomfort and pain to the using patients.

The innovations and contributions of this paper mainly include.

- We propose the possibility of human defecation prediction based on BS signals, and construct a dataset with a link between defecation and BS;
- Based on the BS signal data, we propose a human defecation prediction algorithm based on the improved residual neural network, and experimentally validate the intrinsic correlation between the BS signals and defecation reflexes, which offers a new idea to solve the problem of defecation care for the disabled;
- In order to solve the problem that the overall sample size of the BS dataset is small and the positive and negative sample sizes are not equal, an audio signal augmentation algorithm based on the Informer temporal signal prediction model is proposed, which effectively solves the problem of poor generalization of the training model due to the difficulty of collecting BS data in reality.

Although the defecation intention prediction method in this study achieved better classification results, it still has some limitations. Our current research work is in the beginning stage, and there are still directions and space for improvement. Although we have solved the problem of model generalizability by means of data augmentation at this stage, the problem of limited sample size and diversity of the dataset due to the difficulty of collecting bowel sounds in reality still exists. This limitation affects the improvement of the accuracy of the experimental results, as well as the ability to generalize the model to a wider range of people or different scenarios. In order to solve this problem, we are actively applying for new ethical approvals, and in our future work, we will reach cooperation with more hospitals to expand the scope of volunteer recruitment as much as possible. We will also try to extend the collection duration as much as possible with the consent of the volunteers, and try to incorporate information such as individual defecation habits and feces quality into the feature set.

On the other hand, due to the limited dataset at this stage, we have to sacrifice computational complexity for the generalizability and accuracy of the classification model. However, this limitation will increase the cost and complexity of deploying and running the model in real-time applications in resource-limited environments. Currently, the computational complexity and cost are mainly reflected in the data augmentation process of the Informer model. As the dataset slowly expands to a certain level, we will abandon the data augmentation process and the computation results will be greatly reduced, and the hardware requirements will be reduced as well. Meanwhile, in the future, we will plan to reduce the computational requirements and running time of the classification model through model compression techniques (e.g., pruning, quantization, distillation, etc.) to improve efficiency and reduce costs. And we will try to export them into lightweight and embeddable classification models to better adapt to micro or mobile devices.

In addition, our understanding of bowel sounds is still partially lacking at this stage, and the way of applying their multidomain features has not been studied deeply enough. For the eigenvalues extracted in the time, frequency, and time-frequency domains, we currently use multimodal feature fusion to integrate the features and further enhance the overall analysis by utilizing the complementary information without introducing too many complex calculations. In our future work, for features such as wavelet energy and wavelet entropy, which have correlation and connectivity between adjacent time and adjacent frequency, we will try to introduce the direction of combining applications such as time-frequency correlation analysis or time-frequency connectivity analysis, to reveal the complex dynamic relationship between the time-frequency features of the bowel sounds, in order to further improve our study.

Funding statement

This work is supported by the National Major Science and Technology Project of China [grant numbers 2020YFC2007600].

Data availability statement

We created a dataset for this study. Since further research is in progress, we cannot publish the dataset right now. As a result, we are temporarily unable to deposit data associated with our study into a publicly available repository.

Ethics declarations

This study was reviewed and approved by [The Medical Ethics Committee of China Rehabilitation Research Center], with the approval number: [2020-085-1].

All participants/patients (or their proxies/legal guardians) provided informed consent to participate in the study.

CRediT authorship contribution statement

Tie Zhang: Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Cong Hong:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Yanbiao Zou:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Formal analysis, Data curation. **Jun Zhao:** Writing – original draft, Investigation, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Banharak, et al., Prevention and care for incontinence-associated dermatitis among older adults: a systematic review, J. Multidiscip. Healthc. 14 (2021) 2983–3004.
[2] Y. Mugita, et al., Assessing absorbent products' effectiveness for the prevention and management of incontinence-associated dermatitis caused by urinary, faecal or double adult incontinence: a systematic review, J. Tissue Viability 30 (4) (2021) 599–607.
[3] P. Zan, et al., Study on reconstruction of rectal sensation based on wavelet packet analysis and SVM strategy, J. Med. Eng. Technol. 36 (4) (2012) 205–209.
[4] P. Zan, J. Zhao, L.H. Yang, Research on biomechanical compatibility for the artificial anal sphincter based on rectal perception function reconstruction, IET Sci. Meas. Technol. 9 (8) (2015) 921–927.
[5] T. Saitoh, et al., Excretion prediction using nursing record system log data, in: 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 2018. Nara, JAPAN.
[6] K. Yabunaka, et al., Assessment of rectal feces storage condition by a point-of-care pocket-size ultrasound device for healthy adult subjects: a preliminary study, Drug discoveries & therapeutics 12 (1) (2018) 42–46.
[7] N.G. Kock, J. Kewenter, T. Sundin, Studies on the defecation relfex in man, Scand. J. Gastroenterol. 7 (8) (1972) 689–693.
[8] A. Shafik, F. El-Sibai, I. Ahmed, Parasympathetic extrinsic reflex: role in defecation mechanism, World J. Surg. 26 (6) (2002) 737–741.
[9] G. Bassotti, et al., Colonic mass movements in idiopathic chronic constipation, Gut 29 (9) (1988) 1173–1179.
[10] P.G. Dinning, et al., Abnormal predefecatory colonic motor patterns define constipation in obstructed defecation, Gastroenterology 127 (1) (2004) 49–56.
[11] P.G. Dinning, M.M. Szczesniak, I.J. Cook, Proximal colonic propagating pressure waves sequences and their relationship with movements of content in the proximal human colon, Neuro Gastroenterol. Motil. 20 (5) (2008) 512–520.
[12] M. Camilleri, et al., American Neurogastroenterology and Motility Society consensus statement on intraluminal measurement of gastrointestinal and colonic motility in clinical practice, Neuro Gastroenterol. Motil. 20 (12) (2008) 1269–1282.
[13] S.N. Saito, et al., Generation mechanisms of bowel sounds by simultaneous measurements of X-ray fluoroscopy and bowel sounds, in: 43rd Annual International Conference of the IEEE-Engineering-In-Medicine-And-Biology-Society (IEEE EMBC), Electr Network, 2021.
[14] S.S. Ching, Y.K. Tan, Spectral analysis of bowel sounds in intestinal obstruction using an electronic stethoscope, World J. Gastroenterol. 18 (33) (2012) 4585–4592.
[15] K.S. Kim, J.H. Seo, C.G. Song, Non-invasive algorithm for bowel motility estimation using a back-propagation neural network model of bowel sounds, Biomed. Eng. Online 10 (2011).
[16] X.H. Du, et al., Noninvasive diagnosis of irritable bowel syndrome via bowel sound features: proof of concept, Clin. Transl. Gastroenterol. 10 (2019).
[17] R. Ranta, et al., Wavelet-based bowel sounds denoising, segmentation and characterization, in: 23rd Annual International Conference of the IEEE-Engineering-In-Medicine-And-Biology-Society, Istanbul, Turkey, 2001.
[18] L.J. Hadjileontiadis, L.T. Rekanos, Ieee, Enhancement of explosive bowel sounds using kurtosis-based filtering, in: 25th Annual International Conference of the IEEE-Engineering-In-Medicine-And-Biology-Society, Cancun, MEXICO, 2003.
[19] L.J. Hadjileontiadis, Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding - Part I: Methodology, IEEE Trans. Biomed. Eng. 52 (6) (2005) 1143–1148.
[20] C. Dimoulas, et al., Novel wavelet domain Wiener filtering de-noising techniques: application to bowel sounds captured by means of abdominal surface vibrations, Biomed. Signal Process Control 1 (2) (2006) 177–218.
[21] D. Bray, et al., Assessing motility through abdominal sound monitoring, in: Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 6, `Magnificent Milestones and Emerging Opportunities in Medical Engineering' (Cat. No.97CH36136), 1997, pp. 2398–2400.
[22] B.L. Craine, M.L. Silpa, C.J. O'Toole, Enterotachogram analysis to distinguish irritable bowel syndrome from Crohn's disease, Dig. Dis. Sci. 46 (9) (2001) 1974–1979.
[23] B.L. Craine, M.L. Silpa, C.J. O'Toole, Two-dimensional positional mapping of gastrointestinal sounds in control and functional bowel syndrome patients, Dig. Dis. Sci. 47 (6) (2002) 1290–1296.
[24] Y. Yin, et al., Bowel sound based digestion state recognition using artificial neural network, in: 11th IEEE Annual Biomedical Circuits and Systems Conference (BioCAS), 2015. Atlanta, GA.
[25] Y.Z. Qiao, L. Wang, X.P. Tao, A bowel sound detection method based on a novel non-speech body sound sensing device, in: 45th Annual International IEEE-Computer-Society Computers, Software, and Applications Conference (COMPSAC), Electr Network, 2021.

[26] J. Salamon, J.P. Bello, Deep convolutional neural networks and data augmentation for environmental sound classification, IEEE Signal Process. Lett. 24 (3) (2017) 279–283.
[27] W. Shengyun, et al., A comparison on data augmentation methods based on deep learning for audio classification, J. Phys. Conf. 1453 (2020) 012085.
[28] B.G. Yazgac, M. Kirci, Fractional-order calculus-based data augmentation methods for environmental sound classification with deep learning, Fractal and Fractional 6 (10) (2022).
[29] B. Bahmei, E. Birmingham, S. Arzanpour, CNN-RNN and data augmentation using deep convolutional generative adversarial network for environmental sound classification, IEEE Signal Process. Lett. 29 (2022) 682–686.
[30] R. Ranta, et al., Digestive activity evaluation by multichannel abdominal sounds analysis, IEEE Trans. Biomed. Eng. 57 (6) (2010) 1507–1519.
[31] B.L. Bardakjian, The Gastrointestinal System, 1999.
[32] C. Dimoulas, et al., Bowel-sound pattern analysis using wavelets and neural networks with application to long-term, unsupervised, gastrointestinal motility monitoring, Expert Syst. Appl. 34 (1) (2008) 26–41.
[33] M. Barandas, et al., TSFEL: time series feature extraction library, SoftwareX 11 (2020).
[34] H.Y. Zhou, et al., Informer: beyond efficient transformer for long sequence time-series forecasting, in: 35th AAAI Conference on Artificial Intelligence/33rd Conference on Innovative Applications of Artificial Intelligence/11th Symposium on Educational Advances in Artificial Intelligence, Electr Network, 2021.
[35] A. Vaswani, et al., Attention is all you need, in: NIPS, 2017.
[36] T. He, et al., Bag of tricks for image classification with convolutional neural networks, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
[37] E.Y. Jiang, et al., Rectal sensation function rebuilding based on optimal wavelet packet and support vector machine, IET Sci. Meas. Technol. 7 (3) (2013) 139–144.
[38] Z.R. Zhou, et al., Design and evaluation of puborectalis-like artificial anal sphincter that replicates rectal perception, Artif. Organs 44 (7) (2020) E300–E312.
[39] D. Han, et al., Design and evaluation of perception reconstruction with sensor system for artificial anal sphincter based on vector similarity, Artif. Organs 46 (12) (2022) 2391–2399.