A Human Defecation Prediction Method Based on Multi-Domain Features and Improved Support Vector Machine

Lin Li [1],[†] , Yuwei Ke [1],[†] , Tie Zhang [1],[*],[†] , Jun Zhao [2] and Zequan Huang [1]

[1] School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510641, China
[2] China Rehabilitation Research Center, Beijing 100068, China
[*] Correspondence: merobot@scut.edu.cn
[†] These authors contributed equally to this work.

Abstract: The difficulty of defecation seriously affects the quality of life of the bedridden elderly. To solve the problem that it is difficult to know the defecation time of the bedridden elderly, this paper proposed a human pre-defecation prediction method based on multi-domain features and improved support vector machine (SVM) using bowel sound as the original signal. The method includes three stages: multi-domain features extraction, feature optimization, and defecation prediction. In the stage of multi-domain features extraction, statistical analysis, fast Fourier transform (FFT), and wavelet packet transform are used to extract feature information in the time domain, frequency domain, and time-frequency domain. The symmetry of the bowel sound signal in the time domain, frequency domain, and time-frequency domain will change when the human has the urge to defecate. In the feature optimization stage, the Fisher Score (FS) algorithm is introduced to select meaningful and sensitive features according to the importance of each feature, aiming to remove redundant information and improve computational efficiency. In the stage of defecation prediction, SVM is optimized by the gray wolf optimization (GWO) algorithm to realize human defecation prediction. Finally, experimental analysis of the bowel sound data collected during the study is carried out. The experimental result shows that the proposed method could achieve an accuracy of 92.86% in defecation prediction, which proves the effectiveness of the proposed method.

Keywords: bowel sound; feature extraction; gray wolf optimization; healthcare; support vector machine

1. Introduction

With the increasing number of elderly [1], defecation care for the long-term bedridden elderly has become a pressing social issue. The defecation care of the long-term bedridden elderly is often accompanied by "dirty, messy and smelly" problems [2], which not only make the caregivers miserable but also make it difficult to guarantee the privacy and dignity of the elderly.

The existing methods of defecation care mainly include post-defecation treatment and pre-defecation warning. At present, the most commonly used nursing methods are post-defecation treatment, such as caregivers nursing, using diapers, using anal bags and using intelligent nursing robots, etc., but this kind of method often causes the skin to come into contact with excrement [3], which easily causes skin damage or even skin diseases in the elderly, and is not comfortable. Although methods based on pre-defecation warning can effectively solve the above problems, the existing methods are rare and difficult to be used in long-term care. For example, Zan et al. [4] proposed a method for predicting bowel intention based on rectal pressure signals monitored by biological parameter telemetry capsules. However, one capsule can only be monitored for 24 h. Long-term monitoring requires the long-term use of telemetry capsules, which is potentially risky and costly. Therefore, there is an urgent need for a new, cost-effective, long-term, and simple method

for predicting before defecation to improve the quality of life of the disabled elderly. In this paper, based on the study of physiological signals related to human defecation, a machine learning algorithm is used to identify the intention of defecation, and a defecation prediction method is proposed.

The physiological signal of the human body reflects the electrical activities of specific body parts [5], which is closely related to the health status and life process of the human body. Therefore, this information is a valuable data source for disease detection, rehabilitation, and treatment [6]. At present, physiological signals that have been widely studied in the field of healthcare include electromyography (EMG), electroencephalogram (EEG), electrocardiogram (ECG), and electrooculography (EOG). Among them, the EMG signal is a typical clinical recording method for diagnosing and monitoring neuromuscular behavior, which can be used to identify muscle injuries [7], neuroprosthetic control [8], and motor intention decoding [9], etc. EEG is the electrical activity of the brain and is widely used in sleep state recognition [10] epilepsy detection [11], and emotion recognition [12]. ECG, as the main means to detect the electrical activity of the heart, is an important and harmless means to predict and diagnose cardiovascular diseases [13,14]. EOG can detect eye movements, and can be used for ophthalmic diagnosis [15] and sleep status monitoring [16]. Inspired by these studies, we set out to investigate the physiological signals associated with defecation in humans.

A medical study shows that the human defecation intention is generated by the feedback loop formed by the pressure receptors distributed around the periphery of the rectum, nerve tissue, and brain [17]. Therefore, when the feces entering the rectum reaches a certain volume, the human body will have an obvious defecation intention. In addition, studies have shown that defecation is one of the important evaluation criteria reflecting gastrointestinal motility [18]. Based on the above studies, we can speculate that there is a high correlation between gastrointestinal status and human defecation activities, so it is feasible to identify defecation intention by monitoring human gastrointestinal status. Bowel sound (BS) is the sound produced when intestinal contractions push liquid and gas through different parts of the intestine during digestion [19]. It is a simple and effective physiological signal for evaluating the state of the gastrointestinal tract. Therefore, BS is used to monitor gastrointestinal status. In addition, the rectum is normally empty [20], but when feces stored in the colon are pushed into the rectum, the defecation center sends defecation signals through the efferent nerves. We hypothesized that the process of feces pushing into the rectum will be accompanied by more frequent BS, and the difference in the number of feces stored in the bowel may also lead to differences in the quality of BS.

However, manual interpretation of BS is complex, not only requires expertise acquired over many years of training, but is also time-consuming and labor-intensive. Therefore, computer-aided methods are needed to automatically identify and analyze BS, reduce human burden, and reduce errors caused by fatigue and internal variability [21]. The existing research techniques in BS can be divided into two categories: statistics-based methods and machine-learning-based methods. Iterative kurtosis-based detection (IKD) [22] is one of the classical statistics-based algorithms. It locates each BS event through point-by-point estimation of the kurtosis value of BS record points. The points with kurtosis values greater than the threshold are attributed to the existence of nearby BS events. However, IKD algorithms are acausal and cannot be implemented in real-time. On the other hand, with the development of artificial intelligence technology, some advanced machine learning technologies have been continuously used in BS clinical research, and have been widely used in the diagnosis and identification of intestinal diseases. Machine learning techniques conduct BS signal studies in a data-driven manner and are often more effective than statistical-based methods. In 2008, Dimoulas et al. [23] used wavelet feature extraction and multilayer perceptron (MLP) network classifier to analyze BS, which has high accuracy. In 2011, Kim et al. [24] used back propagation(BP) network to analyze BS, and their results showed that there is a good correlation between BS and colonic transit. In 2014, Ulusar et al. [25] proposed a method based on Naive Bayesian (NB) algorithm to confirm

the recovery of bowel function in patients after major abdominal surgery by monitoring BS. In 2018, Liu et al. [26] proposed a BS monitoring method based on Mel frequency cepstrum coefficient (MFCC) features and long-short term memory (LSTM) neural networks. This method can accurately distinguish BS and noises, and can be used for long-term detection of gastrointestinal motility. In 2018, Yin et al. [27] improved the BS monitoring system and proposed a BS recognition system based on Support Vector Machine (SVM) to identify intestinal motility events. In 2019, Du et al. [28] proposed the feasibility of applying BS to the diagnosis of irritable bowel syndrome based on a logistic regression algorithm.

It can be seen that applying machine learning technology to BS signal analysis provides a promising and effective method for gastrointestinal status-related diagnosis. However, there is no relevant research on the application of BS to defecation prediction. In addition, in previous studies, the distinction between types is mostly based on typical acoustic features, including spectral centroid, sub-band normalized energy, value of envelope thoracic coefficient, etc. [29], which are all single-domain features. Since BS has the characteristics of asymmetry, weak signal, strong background noise, large individual differences, and strong randomness [25], feature extraction based on single domain tends to ignore the information of signals in other domains, and it is difficult to fully reflect the inherent characteristics of the signal [30,31]. Moreover, in practical applications, the accuracy and generalization ability of the method need to be guaranteed, but the high computational complexity will hinder the applicability of the method in real-time continuous processing. Most previous studies have failed to balance accuracy and computational complexity. To solve the above problems, this paper proposes a new defecation prediction method based on multi-domain features and gray wolf optimizer-based support vector machine (GWO-SVM). This method is different from all previous methods. First of all, we extract the multi-domain features of BS, and select the features that are sensitive to classification. As far as we know, this is the first time in BS research. In addition, in feature extraction in time-frequency domain, we use wavelet packet transform and entropy instead of discrete wavelet transform coefficients, which can retain important information in high-frequency components. Finally, when we use SVM as the prediction model, we use the gray wolf optimization (GWO) algorithm for optimization, which can improve the recognition accuracy and generalization ability of the model. The method proposed in this paper is an innovative application of machine learning in the field of health care, and provides a promising and effective method for defecation care of the bedridden elderly. Since the monitoring of BS is non-invasive in vitro, it can avoid pain and inconvenience when applied to the human body. The proposed method has high recognition accuracy, which can meet the detection function of actual use, assist the bedridden elderly to get timely defecation care, and reduce the nursing cost and burden.

The research framework of the proposed method is shown in Figure 1, and the details are as follows. We firstly filter and denoise the collected BS signals, and then use three methods (statistical analysis, fast Fourier transform (FFT), and wavelet packet transform) to extract the features from multi-domain aspects (time domain, frequency domain, and time-frequency domain) to construct high-dimensional datasets. We believe that the multi-domain characteristics of normal human BS are distributed in a centrosymmetric arrangement, and this symmetry will change after adding the human BS signals. Therefore, the Fisher Score (FS) algorithm is used for feature selection of multi-domain features, and sensitive features with large changes in symmetry and meaningful for classification are selected to improve computational efficiency. After that, SVM which has a strong generalization ability in small datasets and nonlinear data classification is used to establish the defecation prediction model. In addition, the GWO algorithm is used to optimize SVM to get better accuracy of defecation prediction. Finally, experimental analysis is carried out based on the BS data collected in the experiment to verify the feasibility of defecation prediction based on the BS signal. The rest of this paper is organized as follows. Section 2 introduces the BS collection system and dataset. Section 3 introduces the processing methods of the BS signals and specifically introduces the process of multi-domain features extraction. Section 4 introduces the implementation process of GWO algorithm to optimize

SVM classifier. Section 5 verifies the effectiveness of the proposed method through an experimental case, and the conclusion is shown in Section 6.
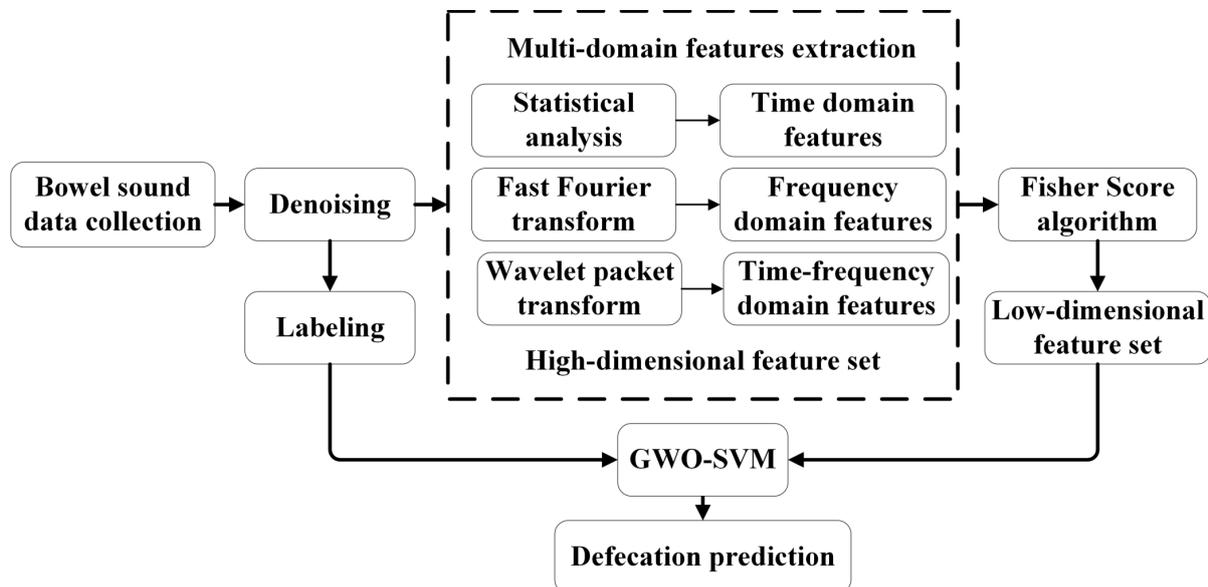

Figure 1. Research framework.

2. Bowel Sound Collection System and Dataset

2.1. Bowel Sound Collection System

To realize the monitoring of BS data and other physiological parameter signals, we build a BS collection system, as shown in Figure 2. The system includes a signal acquisition module, a data receiving module, and a real-time display module. The signal acquisition module includes a 3M™Littmann® 3200 electronic stethoscope for BS acquisition and a Biosignalsplux physiological multipurpose recorder for EEG, EGG, ECG, and other physiological parameter signals acquisition. The data receiving module adopts the industrial computer KMDA-3921, connected with the signal acquisition module through Bluetooth. The real-time acquisition module can display the physiological signals collected in real-time. Among them, the Biosignalsplux physiological multipurpose recorder in the signal acquisition module will be used for subsequent research.
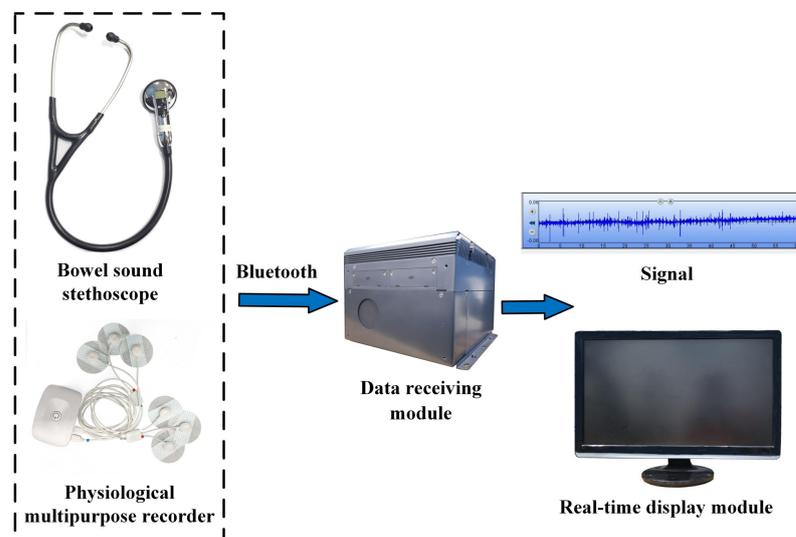

Figure 2. BS collection system.

2.2. Bowel Sound Data Acquisition

The sampling frequency of the BS stethoscope is 4000 Hz, which is the built-in sampling frequency of the sensor. Since the typical BS frequency range is between 50 Hz and 1500 Hz, according to the Nyquist sampling theory, when the sampling frequency is greater than double the highest frequency in the signal, the digital signal after sampling can completely retain the information of original signal. Therefore, the sampling frequency of 4000 Hz can completely retain the information of the BS signal. Since BS is best collected in the membrane-type auscultation mode (amplifying the 20–2000 Hz sound and strengthening the 100–500 Hz sound), the auscultation mode of the stethoscope is set to the membrane-type.

Figure 3 shows a schematic diagram of BS data collection locations. It can be seen that the common collection locations of BS are the upper right, upper left, and lower left positions of the abdomen. The lower left part of the abdomen is the sigmoid colon, which is connected to the rectum. When the feces stored in the sigmoid colon are large or the feces are pushed to the rectum from the sigmoid colon, humans will have a more obvious defecation intention. Therefore, we believe that it is the most suitable to fix the stethoscope in the lower left quadrant of the abdomen to collect BS in this paper.
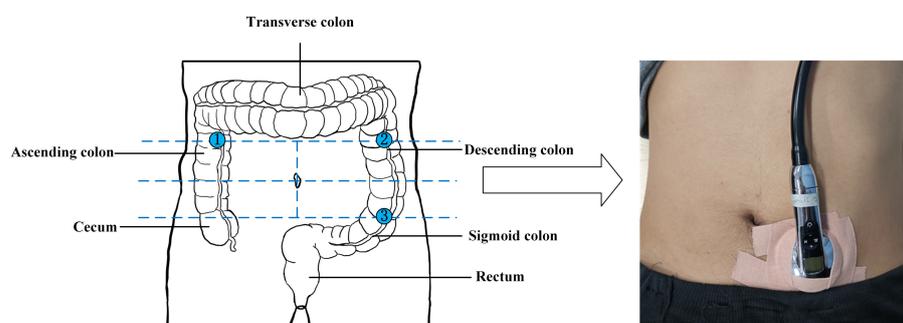

Figure 3. Schematic diagram of BS data collection.

Before BS data collection, it is necessary to shave off the excess hair where the BS stethoscope is located and apply a scrub to remove dead skin. During the collection process, to reduce unnecessary human error, the stethoscope is fixed on the lower left quadrant of the volunteer's abdomen with medical tape. The volunteers are asked to take a supine position, minimize body movement, try to breathe evenly, avoid speaking, and stay awake. The duration of each data point is 1 min. After data collection, the BS data within 10 min before defecation are manually marked as label 1 (intention to defecate), and the rest are marked as label 0 (no intention to defecate).

3. The Processing Methods of the Bowel Sound Signals

3.1. Filtering Method for Noise Reduction

BS is usually regarded as a non-stationary short-term signal with sudden characteristics, which is easily polluted by noise caused by activities such as heartbeat, breathing, and exercise. Therefore, the bowel sound signal needs to be filtered for noise reduction before extracting features. A recent study confirmed that the largest part of the power spectral density of BS is between 100 Hz and 500 Hz, and the power spectral density above 1000 Hz only accounts for a very small part [32]. Therefore, we mainly retain the information of the BS signal from 100 Hz to 1000 Hz in this paper.

By filtering the original BS signal with a second-order Butterworth-type high-pass filter with a cut-off frequency of 100 Hz, a low-pass filter with a cut-off frequency of 1000 Hz, and a notch filter of 50 Hz, the noise caused by breathing, heartbeat, electromagnetic interference, and other activities can be effectively reduced [25]. Figure 4 shows the time domain and frequency domain comparison diagram of the BS signal before and after filtering. It can be seen that the signal before filtering is noisy and contains more noise

below 100 Hz. After filtering, the noise is effectively suppressed. In addition, as shown by the arrow in Figure 4, the location of BS can be more clearly seen in the filtered signal.
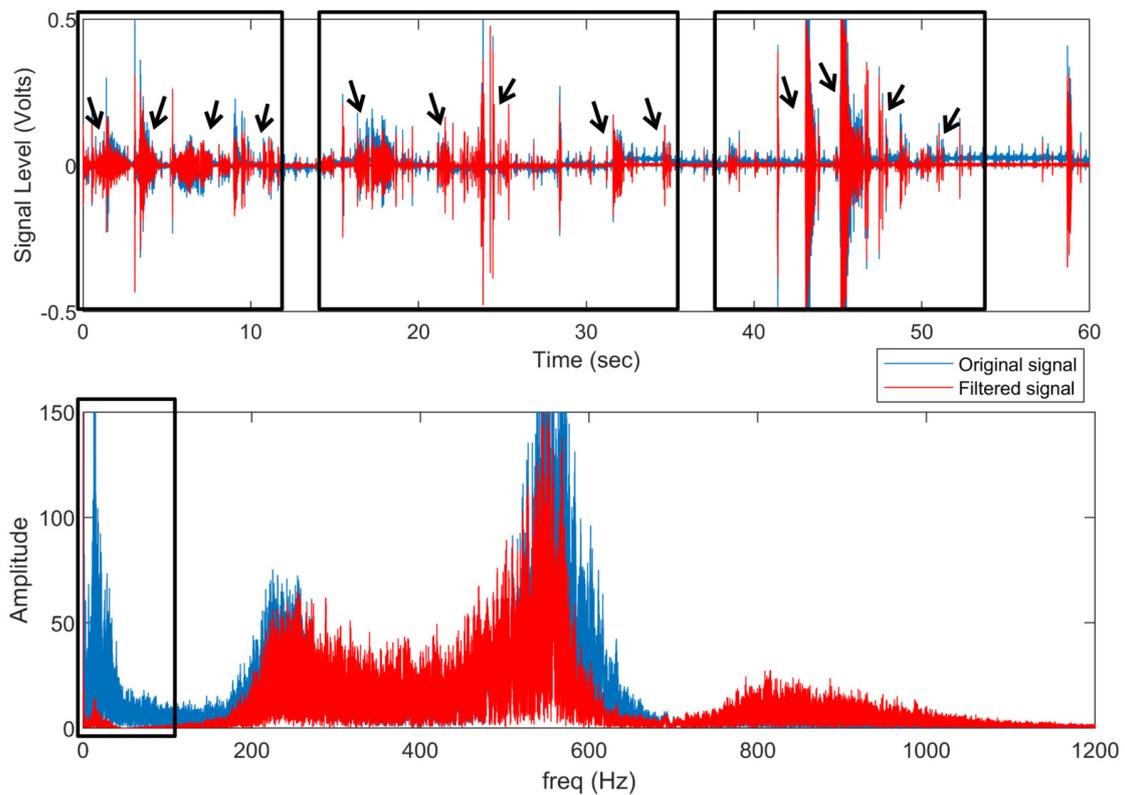

Figure 4. Comparison of before and after BS filtering.

Figure 5a,b are the time domain diagrams of typical BS signals after filtering with and without defecation intention. From the figure, we can see that the frequency and amplitude of BS are lower when there is no urge to defecate, while the frequency and amplitude of BS will increase when there is an urge to defecate, especially when there is a strong urge to defecate. This paper speculates that this phenomenon may be due to the movement of feces from the colon to the rectum, accompanied by more pronounced intestinal peristalsis.
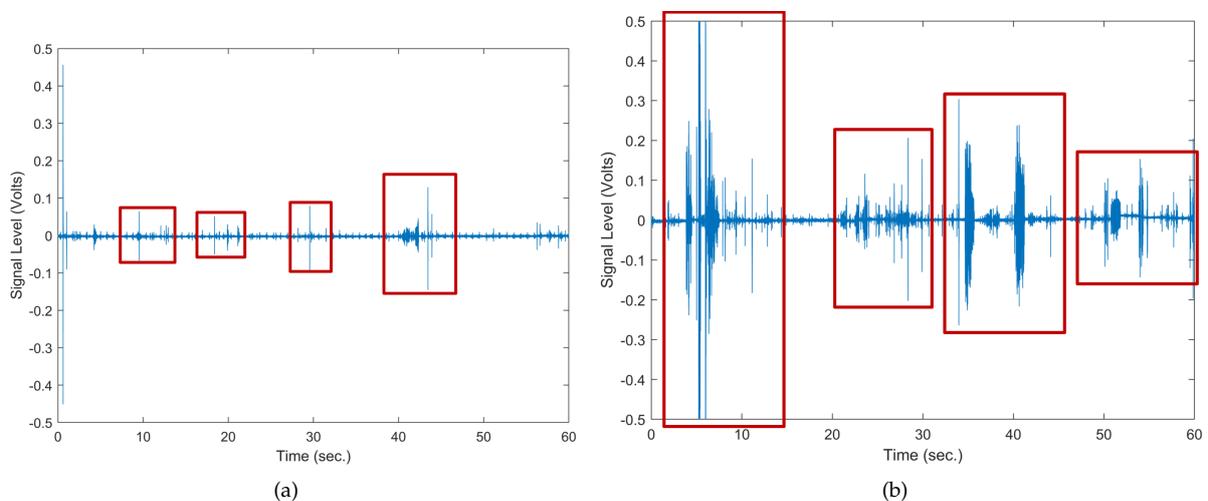

Figure 5. (**a**) BS signal without defecation intention. (**b**) BS signal with defecation intention.

3.2. Multi-Domain Features Eextraction

Due to the asymmetry, strong randomness, and wide dynamic range of BS signals, it cannot be directly used for defecation intention prediction, and it is difficult to fully extract suitable feature information only by linear correspondence or single-domain feature vector. Therefore, it is necessary to extract multi-domain features of BS signals.

The BS collected in this paper can be expressed as $x_n = \{x_1, x_2, \ldots, x_N\}$, where N is the number of data points of the signal x_n. Since the sampling frequency is 4000 HZ and the duration of each acquisition is 1 min, N is 240,000 here. In this paper, three widely used methods (statistical analysis, FFT, and wavelet packet transform) are used to extract the multi-domain features of the filtered BS signals.

3.2.1. Time Domain Features Extraction

Statistical analysis is a research method that uses statistical methods to analyze the research object from quantitative and qualitative. Using statistical analysis methods to extract the time domain features of the signal, we can obtain a distribution of the signal, which represents the waveform of the signal. When the human body has an obvious urge to defecate, the amplitude and distribution of the signal may be different from that when there is no urge to defecate. Therefore, this paper uses statistical methods to extract 16 time domain features of BS signals, and their expressions are listed in Table 1. Among them, there are 10 dimensional statistical parameters $T_1 \sim T_{10}$ such as mean, standard deviation, square root amplitude, and six dimensionless statistical parameters $T_{11} \sim T_{16}$ such as waveform index, peak index, and pulse index [33].

Table 1. Time domain feature expressions.

Feature Name	Feature Expression	Feature Name	Feature Expression
mean value	$T_1 = \frac{1}{N} \sum_{n=1}^{N} x_n$	Minimum Value	$T_9 = \min\|x_n\|$
standard deviation	$T_2 = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} [x_n - T_1]^2}$	peak-to-peak value	$T_{10} = T_8 - T_9$
square root amplitude	$T_3 = (\frac{1}{N} \sum_{n=1}^{N} \sqrt{\|x_n\|})^2$	waveform index	$T_{11} = \frac{T_2}{T_4}$
absolute mean value	$T_4 = \frac{1}{N} \sum_{n=1}^{N} \|x_n\|$	peak index	$T_{12} = \frac{T_8}{T_2}$
skewness	$T_5 = \frac{1}{N} \sum_{n=1}^{N} (x_n)^3$	pulse index	$T_{13} = \frac{T_8}{T_4}$
kurtosis	$T_6 = \frac{1}{N} \sum_{n=1}^{N} (x_n)^4$	margin index	$T_{14} = \frac{T_8}{T_3}$
variance	$T_7 = \frac{1}{N} \sum_{n=1}^{N} (x_n)^2$	skewness index	$T_{15} = \frac{T_5}{(\sqrt{T_7})^3}$
maximum value	$T_8 = \max\|x_n\|$	kurtosis index	$T_{16} = \frac{T_6}{(T_7)^2}$

3.2.2. Frequency Domain Features Extraction

FFT is a transform form that can transform a signal from the time domain to the frequency domain. It has a wide range of applications in the fields of acoustics and signal processing, because of its fast and efficient calculation algorithm, low computational cost, and can meet the premise that digital systems can process. Using FFT to extract the frequency domain features of the signal, the spectral information of the signal can be obtained, such as amplitude or phase, and it is more feasible to identify signal changes or patterns. The frequency domain analysis of the BS signal observes the characteristics of the signal according to the frequency, which makes the analysis of the signal more profound and convenient. When BS is frequent enough or occurs rarely, the collected signals will be relatively concentrated in the frequency spectrum, and due to the urge to defecate, there are more feces in the sigmoid colon, which may also affect the position of the main frequency band of the BS. The frequency domain feature expressions of the 13 BS signals $F_1 \sim F_{13}$ are shown in Table 2, and the definitions of the involved parameters are shown in Table 3. Among them, the feature value F_1 reflects the vibration energy of the BS signal in the frequency domain, and the feature values $F_2 \sim F_4$, F_6 and $F_{10} \sim F_{13}$ reflect the concentration and dispersion of the BS signal in the frequency spectrum. The feature values F_5 and $F_7 \sim F_9$ reflect the positional change of the BS signal in the main frequency band [33,34].

Table 2. Frequency domain feature expressions.

Number	Feature Expression	Number	Feature Expression
1	$F_1 = \frac{\sum_{k=1}^{K} y_k}{K}$	8	$F_8 = \sqrt{\frac{\sum_{k=1}^{K}(f_k^4 y_k)}{\sum_{k=1}^{K}(f_k^2 y_k)}}$
2	$F_2 = \frac{\sum_{k=1}^{K}[y_k - F_1]^2}{K-1}$	9	$F_9 = \frac{\sum_{k=1}^{K}(f_k^2 y_k)}{\sqrt{[\sum_{k=1}^{K}(f_k^4 y_k)][\sum_{k=1}^{K} y_k]}}$
3	$F_3 = \frac{\sum_{k=1}^{K}[y_k - F_1]^3}{K(\sqrt{F_2})^3}$	10	$F_{10} = \frac{F_6}{F_5}$
4	$F_4 = \frac{\sum_{k=1}^{K}[y_k - F_1]^4}{K(F_2)^2}$	11	$F_{11} = \frac{\sum_{k=1}^{K}[(f_k - F_5)^3 y_k]}{K(F_6)^3}$
5	$F_5 = \frac{\sum_{k=1}^{K}(f_k y_k)}{\sum_{k=1}^{K} y_k}$	12	$F_{12} = \frac{\sum_{k=1}^{K}[(f_k - F_5)^4 y_k]}{K(F_6)^4}$
6	$F_6 = \sqrt{\frac{\sum_{k=1}^{K}[(f_k - F_5)^2 y_k]}{K}}$	13	$F_{13} = \frac{\sum_{k=1}^{K}[\sqrt{|f_k - F_5|} y_k]}{K\sqrt{F_6}}$
7	$F_7 = \sqrt{\frac{\sum_{k=1}^{K}(f_k^2 y_k)}{\sum_{k=1}^{K} y_k}}$	—	—

Table 3. The definitions of the involved parameters in Table 2.

Parameter Name	Define
K	The number of spectral lines
y_k	Frequency spectrum obtained by using FFT
f_k	The frequency value of the k-th spectral line

3.2.3. Time-Frequency Domain Features Extraction

Both the time domain features and the frequency domain features describe the state information of the entire BS signal. It is not possible to analyze the BS signal locally, and it is difficult to observe the information of the frequency of the non-stationary component in the signal changing with time, that is, the time-frequency resolution is not high. With the help of time-frequency domain analysis, the frequency information of the BS signal can be observed in a small range, and the time-frequency resolution of the BS signal can be improved. Wavelet packet transform [35] is developed based on wavelet transform, and it is a modern time-frequency analysis and processing method that can effectively process all kinds of non-stationary random signals. It overcomes the shortcomings of wavelet transform with low resolution in the high-frequency part, and has better time-frequency resolution. It has been widely used in language, image, seismic, mechanical vibration and other fields.

In this paper, wavelet packet transform is used to extract the time-frequency domain features, and the collected BS signal can be decomposed into multiple two-dimensional parameters (time and frequency) to realize the feature decomposition in different frequency bands and different times. The time-frequency domain features include wavelet energy ratio $P_1 \sim P_{16}$, wavelet energy entropy EE, wavelet feature scale entropy $PE_1 \sim PE_{16}$ and wavelet singularity entropy SE. The specific steps are as follows:

(1) BS signal x_n is decomposed by a four-layer wavelet packet transform, and 16 sub-bands are obtained. The j-th layer wavelet packet decomposition of the signal x_n can be written as:

$$x_n = \sum_{i=1}^{2^j} x_j^i(n), \tag{1}$$

where $x_j^i(n)$ is the i-th sub-band signal decomposed by j-layer wavelet packet transform, and the sub-band signal is reconstructed to be the same length as the signal x_n.

(2) Extracting wavelet energy ratio $P_1 \sim P_{16}$ and wavelet energy entropy EE. After decomposition by wavelet packet, the total energy E_x of the signal x_n can be written as:

$$E_x = \sum_{i=1}^{2^j} E_j^i, \tag{2}$$

where $E_j^i = \int_{-\infty}^{+\infty} \left[x_j^i(n) \right]^2 \mathrm{d}n$ is the node energy of the *i*-th sub-band, so the wavelet energy ratio and wavelet energy entropy $P_1 \sim P_{16}$ can be defined as [36]:

$$P_i = \frac{E_j^i}{E_x}, \tag{3}$$

$$EE = -\sum_{i=1}^{2^j} P_i \log P_i, \tag{4}$$

(3) According to the wavelet packet coefficient sequence at each scale, extracting wavelet feature scale entropy $PE_1 \sim PE_{16}$. After the wavelet packet transform is performed on the signal x_n, the wavelet packet coefficient sequence at each scale can be obtained as: $C_j = \{ c_j(1), c_j(2), \cdots, c_j(N) \}$ $(j = 1, 2, 3, 4)$, where N is the length of the sub-band signal and c_j can be regarded as a division of the signal x_n. The measure of this division is defined as:

$$P_{jk} = \frac{c_{F(j)}(k)}{\sum_{k=1}^{N} c_{F(j)}(k)}, \tag{5}$$

where $c_{F(j)}(k)$ is the FFT of $c_j(k)$, so the wavelet feature scale entropy of the *j*-th scale of the signal x_n can be defined as [37]:

$$PE_j = -\sum_{k=1}^{N} P_{jk} \log P_{jk}, \tag{6}$$

(4) Extracting the wavelet singular entropy *SE*. Wavelet singular entropy [38] makes full use of the advantages of wavelet packet transform for adaptive time-frequency localization, the extraction function of singular value decomposition for time-frequency spatial feature patterns, and the statistical properties of information for signal uncertainty and complexity. It can be used to effectively identify BS signals in different states. The wavelet packet decomposition tree after the *j*-th layer wavelet packet decomposition is performed on the signal x_n is shown in Figure 6. The bottom *p* nodes of wavelet decomposition coefficients of length *q* can form a time-frequency distribution matrix $W_{p \times q}$, which reflects the time-frequency space energy distribution characteristics of the signal x_n. According to the singular value decomposition theory, $W_{p \times q}$ can be decomposed as:

$$W_{p \times q} = U_{p \times k} \Lambda_{k \times k} V_{k \times q}, \tag{7}$$

where $p = 2^j$, $q = N/2^j$, and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_k)$ is the singular value diagonal matrix, which satisfies the descending order: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$. So the wavelet singular entropy is defined as:

$$SE = \sum_{i=1}^{k} \nabla Q_i, \tag{8}$$

$$\nabla Q_i = -\left(\frac{\lambda_i}{\sum_{i=1}^{k} \lambda_i} \right) \ln \frac{\lambda_i}{\sum_{i=1}^{k} \lambda_i}, \tag{9}$$
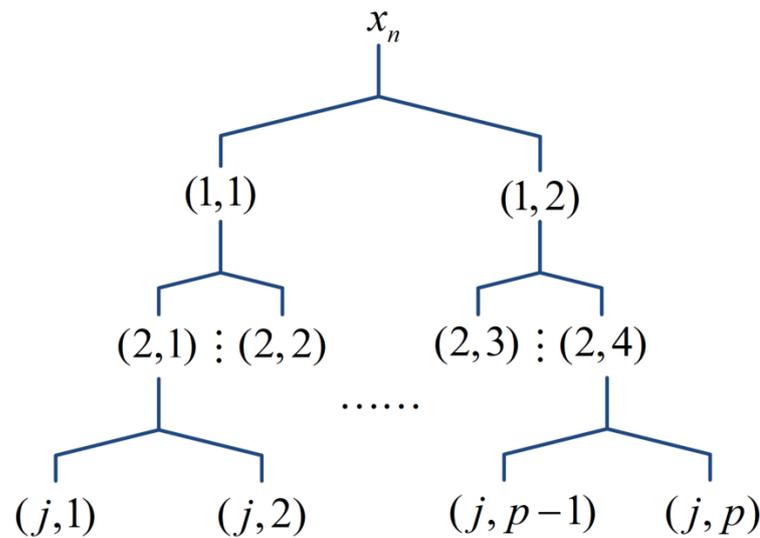
Figure 6. The *j*-layer wavelet packet decomposition tree.

Ultimately, by integrating the multi-domain features D_{data}^i of each data sample, high-dimensional feature sets H_{data} can be constructed:

$$
H_{\text{data}} = \begin{pmatrix} D_{\text{data}}^1 \\ \vdots \\ D_{\text{data}}^p \\ \vdots \\ D_{\text{data}}^q \end{pmatrix} = \begin{pmatrix} T_{1\times16} & F_{1\times13} & P_{1\times16} & EE_1 & PE_{1\times16} & SE_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ T_{p\times16} & F_{p\times13} & P_{p\times16} & EE_p & PE_{p\times16} & SE_p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ T_{q\times16} & F_{q\times13} & P_{q\times16} & EE_q & PE_{q\times16} & SE_q \end{pmatrix} \tag{10}
$$

where q represents the total number of data samples, and D_{data}^p represents the multi-domain features of the p-th sample. The dimension of the matrix H_{data} is 63, including 16-dimensional time domain features, 13-dimensional frequency domain features, 16-dimensional wavelet energy ratio features, one-dimensional wavelet energy entropy feature, 16-dimensional wavelet feature scale entropy features, and one-dimensional wavelet singular spectrum entropy feature.

3.3. Fisher Score Algorithm

We believe that the multi-domain features of normal human BS signals are distributed in centrosymmetric, which will change after adding the BS signals of having an urge to defecate. The multi-domain features of BS signals constitute a high-dimensional dataset H_{data}, which can reveal the state information and intrinsic properties of BS more broadly, but also brings some redundant and negative information. So before the model training, it is necessary to reduce the dimensionality of the high-dimensional dataset, which can select sensitive features with large changes in symmetry and meaningful for classification, and improve computational efficiency. FS algorithm is a typical filtering feature selection method [39]. It selects features that are effective for classification according to the score of candidate features, especially the features with the most distinguishing ability as candidate features. In the data space spanned by the selected features, the distance between data points in different classes is as large as possible, and the distance between data points in the same class is as small as possible. It can be considered that the higher the score of the feature, the greater the change in symmetry, and the more meaningful it is for classification. For a given high-dimensional dataset H_{data}, a widely used heuristic strategy is to calculate

the score of each feature independently according to the *F* criterion, then the FS value of the *j*-th feature is calculated as follows:

$$F(x^j) = \frac{\sum_{k=1}^c n_k (\mu_k^j - \mu^j)^2}{(\sigma^j)^2},$$

(11)

where $(\sigma^j)^2 = \sum_{k=1}^c n_k (\sigma_k^j)^2$, c is the number of classes, n_k is the number of the sample of the *k*-th class, μ_k^j and σ_k^j are the mean and standard deviation of the *j*-th feature in the data of the *k*-th class, μ^j and σ^j are the mean and standard deviation of the entire dataset of the *j*-th feature. After calculating the FS value of each feature in the multi-domain features of BS, according to the experimental test result, the top m features with the highest scores are selected as sensitive features, and a new low-dimensional feature set is established.

4. Support Vector Machine Optimized by Gray Wolf Optimization Algorithm

4.1. Support Vector Machine

There is a nonlinear mapping between the peristaltic mechanism of the intestine and the feature vectors of BS signal, and in general, the frequency of defecation in healthy people is about 1.2 times per 24 h [40]. Affected by this physiological characteristic, the number of BS data samples that can be collected during the experiment is limited, and it is necessary to consider the prediction of defecation intention based on small samples. SVM [41] is a machine learning method based on the principle of structural risk minimization, and its wide and successful application in the engineering and biomedical field in recent years has proved that it is an excellent classification model with strong generalization ability [42–44]. Therefore, this paper selects the SVM model to predict defecation intention.

Given a training set $M = \{(x_i, y_i) \mid x_i \in R^N, y_i \in \{-1, 1\}, i = 1, 2, \cdots, n\}$, where x_i are the sample data and y_i is the sample category, the basic principle of SVM is to find the optimal separation hyperplane, as shown in Figure 7. The solution of its optimal hyperplane can be transformed into the following constrained minimization problem:

$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 \\ s.t. y_i(\omega x_i + b) \geq 1 \end{cases},$$

(12)

where $\omega \in R^N$ is the coefficient or weight vector, and b is the bias term.

To improve the generalization ability of SVM, a slack variable $\theta_i \geq 0, i = 1, 2, \cdots, n$ is introduced, and Equation (12) is rewritten as: ...

$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^n \theta_i (\theta_i \geq 0) \\ s.t. \begin{cases} y_i(\omega x_i + b) \geq 1 - \theta_i \\ c \geq 0 (i = 1, 2, \cdots, n) \end{cases} \end{cases},$$

(13)

where c represents the penalty factor, and its value can weigh the empirical risk and the structural risk. By using the Lagrange multiplier method on the above equation, the Lagrange function of the optimization problem can be obtained:

$$L(\omega, b, \alpha, \theta, \mu) = \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^n \theta_i + \sum_{i=1}^n \alpha_i \left(1 - \theta_i - y_i \left(\omega^T x_i + b\right)\right) - \sum_{i=1}^n \mu_i \theta_i,$$

(14)

where $\alpha = (\alpha_1; \alpha_2; \cdots; \alpha_n)$ and $\mu = (\mu_1; \mu_2; \cdots; \mu_n)$ are Lagrange multipliers, among them, $\alpha_i \geq 0, \mu_i \geq 0$. Then the kernel function $G(x_i, x_j) = \langle \varphi^T(x_i) \varphi(x_i) \rangle$ is introduced, where $\langle \cdot \rangle$

represents the inner product operation. According to the KKT condition, Equation (14) can be described as the following dual optimization problem:

$$\begin{cases} \max L(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j G(x_i, x_j) \\ s.t. \sum_{i=1}^{n} \alpha_i y_i (0 \leq \alpha_i \leq c; i = 1, 2, \cdots, n) \end{cases}, \tag{15}$$

Common kernel functions in SVM include linear kernel function, polynomial kernel function, Gaussian kernel function, and radial basis function (RBF). Choosing different kernel functions, the classification accuracy may be very different. To better approximate any nonlinear function, the RBF is selected as the kernel function in this paper, which is defined as:

$$G(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2g^2}\right), \tag{16}$$

where g is the kernel parameter, which reflects the distribution complexity of the data samples in the high-dimensional space. Finally, the classification decision function of SVM can be defined as:

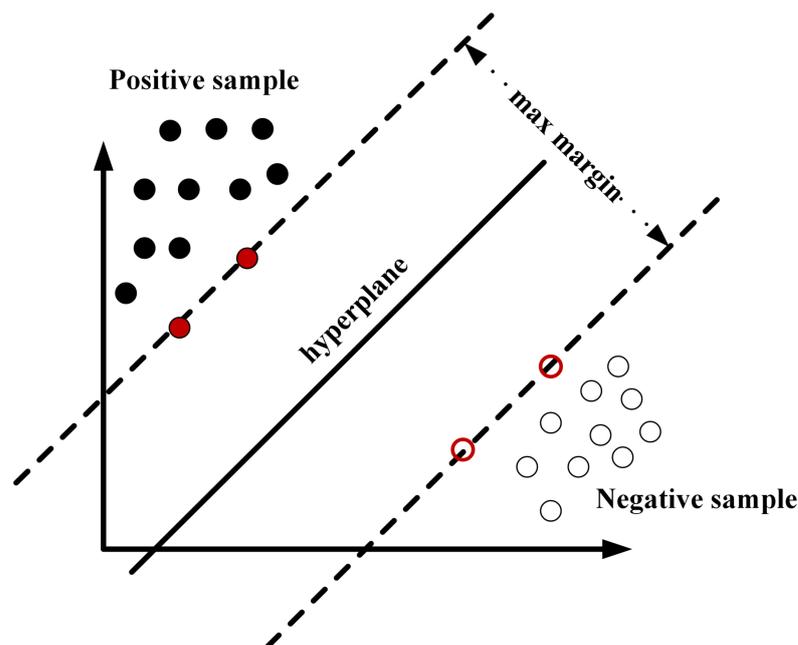$$f(x) = sign\left(\sum_{i=1}^{n} \alpha_i y_i G(x_i, x_j) + b\right), \tag{17}$$

Figure 7. The optimal hyperplane for binary classification.

4.2. Gray Wolf Optimization Algorithm

In the study, the choice of penalty factor c and kernel parameter g has a great influence on the classification accuracy of SVM, therefore, it is necessary to study the setting of SVM parameters [45]. Because GWO algorithm has a convergence factor that can be adaptively adjusted and an information feedback mechanism, it can achieve a balance between local optimization and global search. Besides, GWO has strong convergence and is easy to implement, so in this paper, GWO algorithm is used to optimize the penalty factor c and the kernel parameter g of the RBF kernel function of SVM.

The GWO algorithm was proposed by Mirjalili et al. [46] in 2014, which was inspired by the predation behavior of gray wolves and simulated the population structure and hunting behavior of wolves. The highest-ranking leader in the group becomes wolf A, who is mainly responsible for various decision-making matters in the group. Wolf B ranks

second only to wolf A in the wolf pack and is mainly responsible for assisting wolf A in decision-making. When wolf A's position becomes vacant, wolf B will become the new leader. Wolf C is the third level of the wolf pack and mainly follows the decisions of wolf A and wolf B, and wolf A and wolf B with poor fitness will also be downgraded to wolf C. Wolf Z is the lowest rank of wolves and is mainly responsible for the balance of relationships within the population. The GWO algorithm regards wolf A, wolf B and wolf C in the wolf pack as the three optimal solutions of the algorithm, and the rest of the wolves, including Z, revolve around wolf A, wolf B or wolf C to update the position.

The hunting process of gray wolves includes three parts: stalking and approaching prey, chasing and surrounding prey, and attacking prey. The behavior of gray wolves to update distance and position during stalking and approaching prey is defined as follows:

$$\vec{D} = \left| \vec{C} \times \vec{X}_p(t) - \vec{X}(t) \right|, \tag{18}$$

$$\vec{X}(t+1) = \vec{X}_p(t) + \vec{A} \times \vec{D}, \tag{19}$$

where $\vec{D}$ is the distance between the gray wolf and the prey, t is the number of iterations, $\vec{X}_p$ is the position of the prey, $\vec{X}$ is the position of the gray wolf, and its initial position coordinate is defined as (c, g). $\vec{A}$ and $\vec{C}$ represent coefficients, and the calculation formula is as follows:

$$\vec{A} = 2\vec{a} \times \vec{r}_1 - \vec{a}, \tag{20}$$

$$\vec{C} = 2\vec{r}_2, \tag{21}$$

where $\vec{a}$ is the convergence factor, which decreases linearly from 2 to 0 as the number of iterations increases. $\vec{r}_1$ and $\vec{r}_2$ are random values of $[0, 1]$ respectively. As shown in Figure 8, when $\left| \vec{A} \right| > 1$, it represents a global search, that is, the gray wolves expand the search range to find better prey, when $\left| \vec{A} \right| \leq 1$, it represents a local search, the gray wolves will narrow the surrounding area to search for nearby prey.
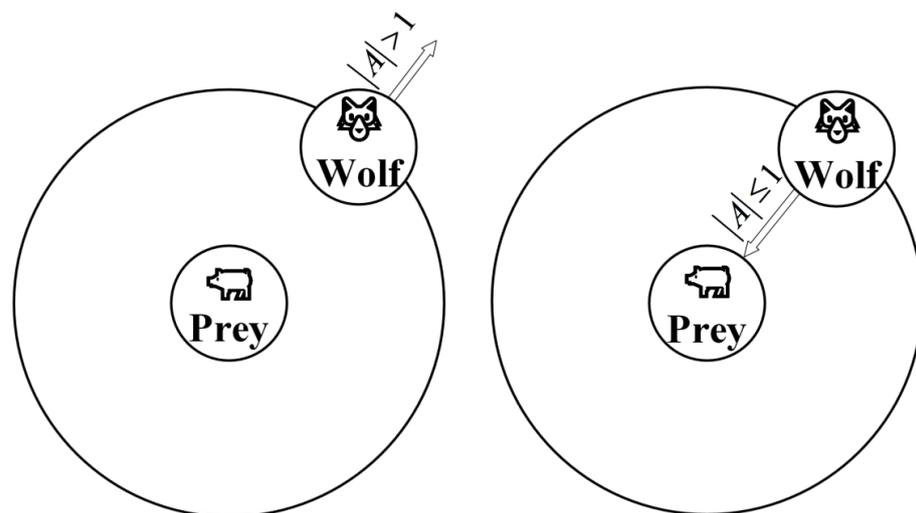

Figure 8. Gray wolf looking for and attacking prey.

When the gray wolves recognize the location of the prey, wolf B and wolf C surround the prey under the leadership of wolf A. Because wolves A, B, and C are closest to the prey, the positions of these three are used to determine the location of the prey. At the same time, the candidate wolves (including wolf Z) update their positions according to the positions of the three, and gradually approach the prey. The update mechanism of the individuals in the wolf group is shown in Figure 9.
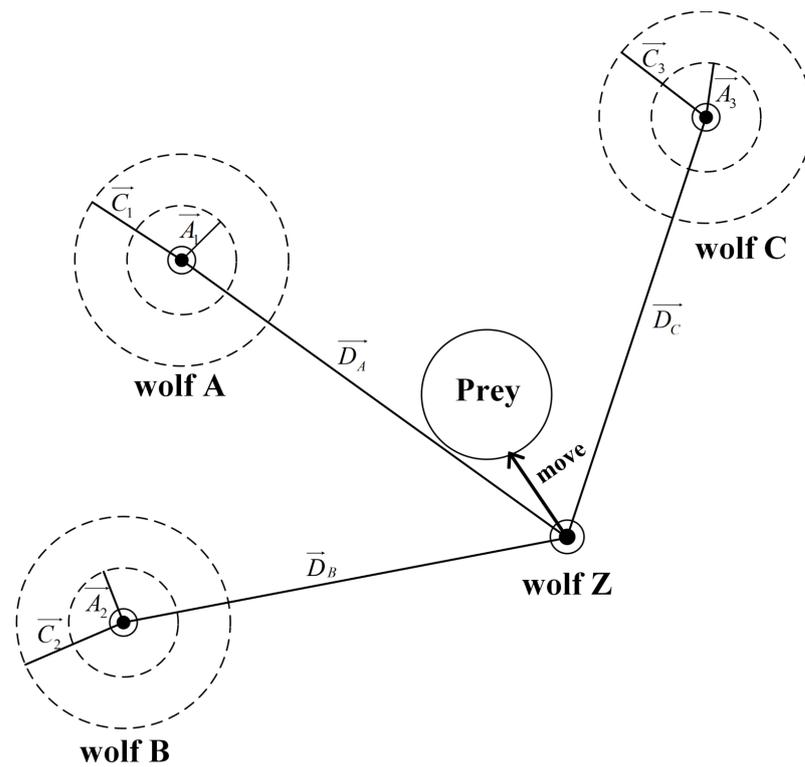
Figure 9. Gray wolf location update diagram.

Among them, the distance between wolf A, wolf B and wolf C and other individuals is expressed as:

$$
\begin{cases}
\vec{D}_a = \left| \vec{C}_1 \times \vec{X}_a(t) - \vec{X}(t) \right| \\
\vec{D}_b = \left| \vec{C}_2 \times \vec{X}_b(t) - \vec{X}(t) \right| \\
\vec{D}_c = \left| \vec{C}_3 \times \vec{X}_c(t) - \vec{X}(t) \right|
\end{cases},
\tag{22}
$$

where $\vec{X}_a$, $\vec{X}_b$ and $\vec{X}_c$ represent the current positions of wolf A, wolf B and wolf C respectively, $\vec{C}_1$, $\vec{C}_2$ and $\vec{C}_3$ are random variables, and $\vec{X}(t)$ represents the current position of the gray wolf. The step size and direction that wolf Z advances toward wolves A, B, and C are defined by Formula (23), and the final position of wolf Z is defined by Formula (24):

$$
\begin{cases}
\vec{X}_1 = \vec{X}_a - \vec{A}_1 \times \vec{D}_a \\
\vec{X}_2 = \vec{X}_b - \vec{A}_2 \times \vec{D}_b \\
\vec{X}_3 = \vec{X}_c - \vec{A}_3 \times \vec{D}_c
\end{cases},
\tag{23}
$$

$$
\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3},
\tag{24}
$$

Then, the position of the gray wolf is updated again until there is the best optimal solution. At that time, the prey stops moving, and the gray wolf completes the hunting process by attacking. The position coordinate value corresponding to the best optimal solution is defined as $(\text{best}c, \text{best}g)$. The calculation flowchart of GWO-SVM is shown in Figure 10. In the optimization process, this paper sets the population size M = 20, the maximum number of iterations $t_{\max} = 30$, and defines the fitness function as:

$$
f = \frac{y_f}{y_t + y_f},
\tag{25}
$$

where y_f and y_t represent the number of misclassified and correctly classified samples after 10-fold cross-validation, respectively. The smaller the fitness value, the higher the classification accuracy, that is to say, the parameter optimization process of SVM can be described as the solution to the fitness function minimization problem. GWO algorithm is used to optimize the penalty factor c and kernel function parameters g of SVM, which can effectively improve classification accuracy.
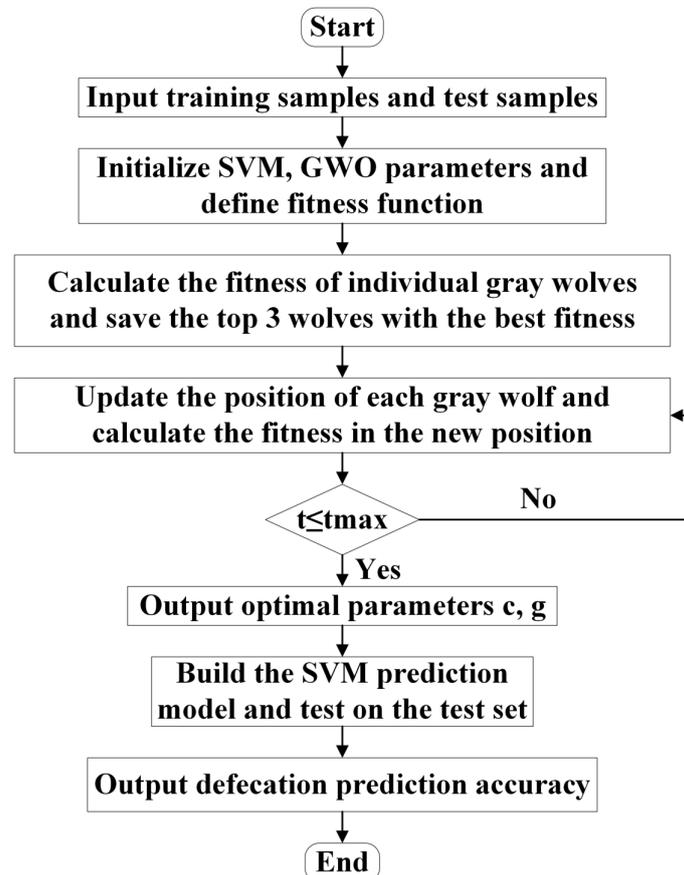

Figure 10. The flow chart of GWO-SVM.

5. Validation of the Proposed Method

All the data processing algorithms in this experiment are compiled in MATLAB R2019b, and the machine learning algorithm is implemented based on Scikit-Learn machine learning library. The operating environment is as follows: Intel(R) Xeon(R) CPU, 96 GB RAM and Windows 10 system.

5.1. Explanation of the Experimental Data

A total of 16 volunteers are recruited for the experiment. They have normal gastrointestinal function and did not take any medication that affected bowel motility in the recent period before data collection. And all the volunteers read the relevant precautions and signed the informed consent. According to the method in Section 2.2, we establish a dataset, and a total of 232 groups of BS data are collected. The duration of each group of BS data is 1 min. Among them, 117 groups of the data are marked as having the intention to defecate, and 115 groups of the data are marked as not having the intention to defecate. The data were collected from Beijing Bo'ai Hospital and South China University of Technology during November 2020 to February 2022. Among them, the data of bedridden elderly are all from Beijing Bo'ai Hospital, which is affiliated with the China Rehabilitation Research Center. The volunteers here are between 40 and 75 years old, both male and

female, and each volunteer's condition is somewhat different. For example, there are those who have had a stroke, those who could not speak, those who have a broken bone, and those who are not able to take care of themselves. In addition, in order to obtain more experimental data and increase the generalizability of the experimental data, we also conducted data collection at South China University of Technology. The volunteers here are between 22 and 35 years old. Therefore, the experimental data contains data from different age groups, different genders, and different physical conditions, which ensures the generalizability of the experimental data. Although the size of the BS dataset is limited by time and recruitment conditions, similar dataset sizes have been successfully used in proof-of-concept studies in the past [28,47]. Furthermore, we also plan to collect more BS data.

5.2. Experimental Results and Analysis

According to the method proposed in Section 3.2, multi-domain features extraction is performed on the filtered BS data. The 63 features extracted from each sample are combined to form a high-dimensional feature set with a size of 232×63, and then normalized, and the normalized interval is $[0, 1]$. After normalization, FS algorithm is used to select the candidate multi-domain features. According to the importance of different features, the multi-domain features are reordered as follows:

$FS_{31} > FS_{32} > FS_8 > FS_{10} > FS_{36} > FS_{33} > FS_{48} > FS_{25} > FS_{14} > FS_{13} > FS_{50} >$
$FS_{21} > FS_{12} > FS_{44} > FS_6 > FS_{23} > FS_{16} > FS_{53} > FS_{37} > FS_{11} > FS_{61} > FS_{40} >$
$FS_{45} > FS_{57} > FS_{19} > FS_{18} > FS_{20} > FS_{43} > FS_{47} > FS_{30} > FS_{62} > FS_2 > FS_{17} >$
$FS_{49} > FS_{22} > FS_{27} > FS_7 > FS_{29} > FS_9 > FS_{35} > FS_{15} > FS_{54} > FS_5 > FS_3 >$
$FS_{58} > FS_{56} > FS_{55} > FS_{34} > FS_4 > FS_{38} > FS_{42} > FS_{51} > FS_{60} > FS_{24} > FS_{59} >$
$FS_{28} > FS_{63} > FS_{46} > FS_{39} > FS_{41} > FS_{52} > FS_{26} > FS_1$

To study the contribution of FS algorithm, Figure 11 plots the classification accuracy of the proposed method under different feature dimensions selected by the FS algorithm. It can be seen from the figure that when the 19-dimensional feature set is selected by the FS algorithm, the highest classification accuracy rate is 92.86%. After continuing to increase the feature dimension, the classification accuracy rate does not increase but decreases. In addition, the higher the input feature dimensions, the longer the algorithm calculation time. Therefore, this paper adopts FS algorithm for feature selection, which can also balance the calculation efficiency and accuracy, and verifies the effectiveness of the FS algorithm.

The first 19 features with the highest scores after sorting by FS algorithm are used as sensitive features, and a new low-dimensional feature set with a size of 232×19 is constructed. Finally, the resulting low-dimensional feature set is fed into GWO-SVM classifier, and Figure 12 plots the relative curve between the number of iterations and the test accuracy. It can be seen from the figure that when the penalty factor c and the kernel parameter g are optimized to 306.6857 and 0.049902, respectively, the proposed method can achieve the best accuracy of 92.86% in predicting human defecation intention, which verifies the effectiveness of the proposed method.

To study the influence of parameters (c, g) on the results of the proposed method, the parameters *c* and *g* of SVM are adjusted according to the literature and some empirical values [34,48], and the corresponding test results are shown in Table 4. As can be seen from Table 4, using GWO algorithm to optimize the parameters (*c* = 306.6857 and *g* = 0.049902) achieves higher accuracy than using the six combined parameters listed in Table 3, that is to say, an inappropriate combination parameters (c, g) will reduce the classification performance of SVM, which illustrates the necessity and effectiveness of using GWO algorithm to optimize SVM.
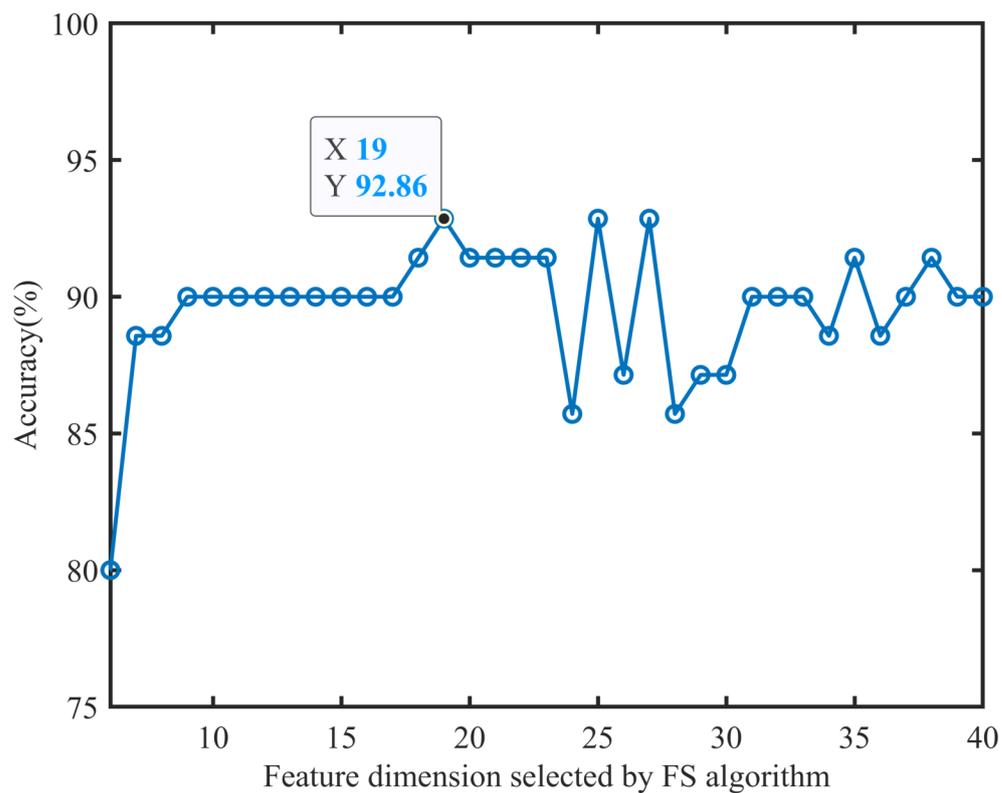
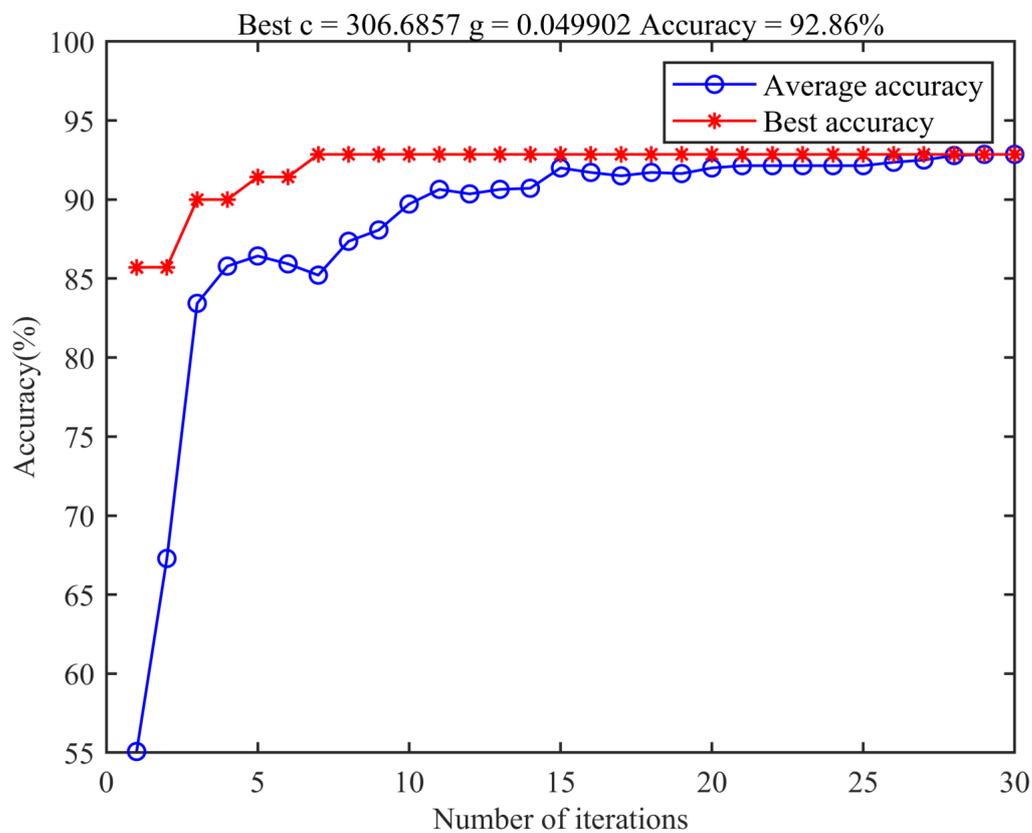**Figure 11.** Classification accuracy under different feature dimensions.

Figure 12. The optimization curve of GWO algorithm.

Table 4. Classification accuracy of SVM with different parameters.

Parameter Value	Accuracy
$c = 10 / g = 10$	87.14%
$c = 10 / g = 4$	90.00%
$c = 42.1 / g = 1.24$	87.14%
$c = 53.2 / g = 2.6$	82.86%
$c = 28.6 / g = 0.88$	90.00%
$c = 100 / g = 0.01$	91.43%
$c = 306.7 / g = 0.05$	92.86%

5.3. Comparison between Different Methods

In this subsection, a research comparison of different methods is carried out to verify the effectiveness and superiority of the proposed method. In this comparison, the proposed method is compared with methods in similar studies. These methods are KNN, NB, SVM, and LR. KNN is used by Saini et al. [14] to identify arrhythmia ECG signals, and it can ensure the simplicity and effectiveness of information in the classification process. NB is used by Ulusar [25] to evaluate the recovery of intestinal function of patients after surgery, and has a high accuracy rate. SVM is one of the classic machine learning methods with excellent generalization ability, and Yin et al. [27] used this model when recognizing BS. LR is the model used by Du et al. [28] to verify the feasibility of BS signals in the diagnosis of irritable bowel syndrome.

We take the extracted 16 time domain features, 13 frequency domain features, 34 time-frequency domain features, and 19 multi-domain features selected by FS algorithm as four different feature sets. The different input features are then combined with five classifiers (GWO-SVM, SVM, NB, KNN and LR). Table 5 shows the classification accuracy of different classifiers with different domain features. Table 6 shows the classification accuracy of different classifiers with different domain feature combinations. As can be seen from Tables 5 and 6, the method proposed in this paper (combination of multi-domain features and GWO-SVM) has the highest testing accuracy (92.86%) of all the combinations, which proves the superiority of the proposed method. It can also be seen in Tables 5 and 6 that combining GWO-SVM classifier with different types of features achieve an average test accuracy of 85.71% and 90.36%, respectively, which is higher than other classifiers. The effectiveness and superiority of GWO-SVM classifier is proved. In addition, we can also see in Table 5 and 6 that the combination of multi-domain features and different classifiers can achieve a higher average test accuracy (86.57%) than other input features. This shows that the extraction of multi-domain features can improve the classification accuracy, that is, multi-domain features have obvious advantages in classification. Further, we can see that the average test accuracy of the other input features in Table 6 is higher than that of the other input features in Table 5. This can be explained by the fact that richer feature information can be obtained with additional domain features extracted, which is helpful to obtain features sensitive to classification. Finally, we can also see from Table 6 that the average test accuracy decreases after removing any single-domain features from multi-domain features. This shows that the features of each domain are meaningful in multi-domain feature extraction. Further, we can see that the feature combination of time and frequency domains achieved an average test accuracy of 82.00%, which is lower than the combination of frequency and time-frequency domains (82.29%), and also lower than the combination of time and time-frequency domains (83.71%). This shows that among multi-domain features, time-frequency domain features provide more classification-sensitive information, which indicates the effectiveness of using wavelet packet transform to extract time-frequency features. This result is also consistent with the ranking result of the importance of different features by the FS algorithm in Section 5.2. Among the 19 sensitive features selected by us, there are 10 features in time-frequency domain, which are more than those in time domain (six features) and frequency domain (four features).

Table 5. Classification accuracy of different classifiers with different domain features.

Different Classifiers	The Testing Accuracy Obtained Using Classification Method with Different Features (%)				Average Accuracy (%)
	Multi-Domain Features	Time Domain Features	Frequency Domain Features	Time-Frequency Domain Features	
GWO-SVM	92.86%	82.86%	81.43%	85.71%	85.72%
SVM	87.14%	75.71%	72.86%	80.00%	78.93%
NB	82.86%	72.86%	68.57%	75.71%	75.00%
KNN	84.28%	71.43%	70.00%	72.86%	74.64%
LR	85.71%	72.86%	74.29%	78.57%	77.86%
Average accuracy (%)	86.57%	75.14%	73.43%	78.57%	—

Table 6. Classification accuracy of different classifiers with different domain feature combinations.

Different Classifiers	The Testing Accuracy Obtained Using Classification Method with Different Features Combinations (%)				Average Accuracy (%)
	Multi-Domain Features	Time and Frequency Domain Features	Time and Time-Frequency Domain Features	Frequency and Time-Frequency Domain Features	
GWO-SVM	92.86%	91.43%	87.14%	90.00%	90.36%
SVM	87.14%	88.57%	85.71%	82.86%	86.07%
NB	82.86%	75.71%	80.00%	81.43%	80.00%
KNN	84.28%	77.14%	81.42%	74.29%	79.28%
LR	85.71%	77.14%	84.28%	82.86%	82.50%
Average accuracy(%)	86.57%	82.00%	83.71%	82.29%	—

To demonstrate the generality of the experimental results, we input the multi-domain features into different classifiers for five different tests, and the results are shown in Figure 13. As can be seen from the figure, GWO-SVM used in this paper has higher classification accuracy than other classifiers (SVM, NB, KNN and LR) in five tests, which verifies the generalization and superiority of the proposed method. There are several reasons for this phenomenon. On the one hand, compared with standard SVM, GWO-SVM optimizes the parameters c and g, so the classification ability is better. On the other hand, compared with other classifiers, SVM can make more reliable decisions on a small number of data samples and has stronger generalization ability. It should be noted that the parameters of other comparison classifiers such as SVM, KNN, NB, and LR are chosen empirically, which may lead to a decrease in accuracy.

During the above experiment, we divided all data into training set and test set in a ratio of 7:3. Only the data in the training set is used for model training. After the model training is completed, the model is tested and verified on an unseen test set. Therefore, the obtained results are generalizable and can be applied to general situations. And compared with other methods, the accuracy of the proposed method reaches 92.86%, which is better than other methods.Therefore, we can conclude that the method proposed in this paper is more effective than traditional classification models. Since there is no related method proposed in the field of defecation prediction, this paper does not compare with other existing methods.
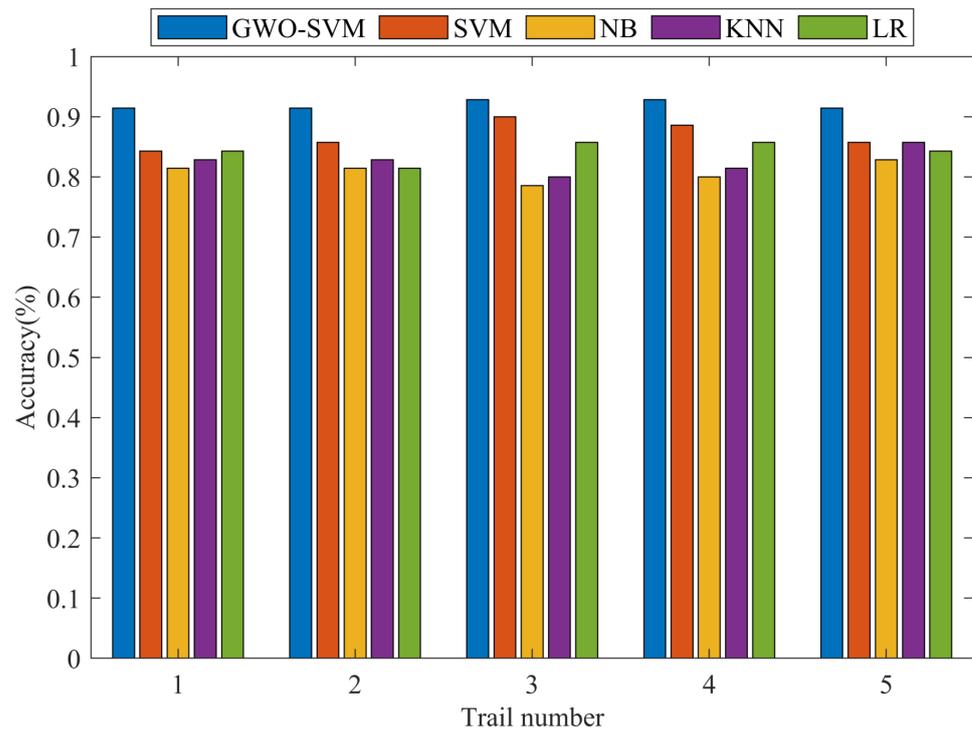
Figure 13. Five test results of using multi-domain features for different classifiers.

6. Conclusions, Limitations, and Future Research

This paper proposes a new defecation prediction method based on multi-domain features and GWO-SVM, so that nursing staff can know the defecation needs of the bedridden elderly in advance, better provide defecation care, reduce the dirty, messy, and smelly phenomenon in the defecation nursing process, and relieve the pain of the bedridden elderly in their later years.

To improve the quality of life of the bedridden elderly in their later years, a human defecation prediction method based on multi-domain features and GWO-SVM is proposed, so that nursing staff could know the defecation needs of the bedridden elderly in advance and provide better defecation care. In this method, the BS is used as the original signal, and three methods of statistical analysis, FFT and wavelet packet transform are applied to the filtered BS to extract the time domain, frequency domain and time-frequency domain features for building a high-dimensional feature set. In addition, feature selection is performed on high-dimensional feature sets, and FS algorithm is used to select meaningful and sensitive features according to the importance of each feature which can remove redundant information and construct low-dimensional feature sets. In this paper, the detailed process of defecation intention prediction based on GWO-SVM is designed, and the actual BS signals collected during the study are used as experimental data for experimental analysis, which verified the defecation intention prediction method proposed in this paper. The result shows that the proposed method can achieve a defecation prediction accuracy of 92.86%. Compared with the traditional classification method, the proposed method achieves better recognition results, and has more stable performance and higher reliability.

The innovations and contributions of this paper mainly include:

(1) The possibility of defecation prediction based on BS signals is proposed, and the correlation between BS signals and defecation intention is verified through experiments, which provide a new idea of defecation prediction;

(2) A BS monitoring system is established, and data were collected in Beijing Bo'ai Hospital affiliated to the China Rehabilitation Research Center, and a BS dataset for defecation prediction is established;

(3) Based on multi-domain features and GMO-SVM, we propose a new, cost-effective, and non-invasive method for human defecation prediction, which is an innovative application of machine learning in the field of healthcare.

Although the study in this paper has achieved good classification results, there are some aspects that need to be improved. On the one hand, the number of volunteers and the sample size are not sufficient, while more and higher quality data are necessary for future research. On the other hand, due to the noise in the data collection process, a further research of the noise reduction algorithm is needed.

To solve the above problems, the next research can start with the acquisition of more available BS data and data augmentation. At the same time, optimizing the noise reduction algorithm for the BS signals is also needed. Currently, research based on data augmentation is a popular direction.

References

1. Cudjoe, T.K.M.; Roth, D.L.; Szanton, S.L.; Wolff, J.L.; Thorpe, R.J. The Epidemiology of Social Isolation: National Health & Aging Trends Study. *J. Gerontol. Ser. B* **2020**, *75*, 107–113.
2. Musa, M.K.; Saga, S.; Blekken, L.E.; Harris, R.; Norton, C. The Prevalence, Incidence, and Correlates of Fecal Incontinence Among Older People Residing in Care Homes: A Systematic Review. *J. Am. Med. Dir. Assoc.* **2019**, *20*, 956–962. [CrossRef] [PubMed]
3. Mugita, Y.; Koudounas, S.; Nakagami, G.; Weller, C.; Sanada, H. 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* **2021**, *30*, 599–607. [CrossRef] [PubMed]
4. Zan, P.; Zhao, J.; Yang, L. Research on biomechanical compatibility for the artificial anal sphincter based on rectal perception function reconstruction. *IET Sci. Meas. Technol.* **2015**, *9*, 921–927. [CrossRef]
5. Devasahayam, S.R. Signals and Systems in Biomedical Engineering. *Topics in Biomedical Engineering International Book*; Springer: Berlin/Heidelberg, Germany, 2012.
6. Faust, O.; Bairy, M.G. Nonlinear analysis of physiological signals: A review. *J. Mech. Med. Biol.* **2012**, *12*, 1240015. [CrossRef]
7. Merletti, R.; Botter, A.; Cescon, C.; Minetto, M.A.; Vieira, T. Advances in surface EMG: Recent progress in clinical research applications. *Crit. Rev. Biomed. Eng.* **2010**, *38*, 347–379. [CrossRef]
8. Zhai, X.; Beth, J.; Chan, R.; Chung, T. Self-Recalibrating Surface EMG Pattern Recognition for Neuroprosthesis Control Based on Convolutional Neural Network. *Front. Neurosci.* **2017**, *11*, 379. [CrossRef]
9. Park, K.H.; Lee, S.W. Movement intention decoding based on deep learning for multiuser myoelectric interfaces. In Proceedings of the International Winter Conference on Brain-Computer Interface, Gangwon , Korea, 22–24 February 2016 ; pp. 1–2.
10. Fraiwan, L.; Lweesy, K. Neonatal sleep state identification using deep learning autoencoders. In Proceedings of the 2017 IEEE 13th International Colloquium on Signal Processing & Its Applications (CSPA), Penang, Malaysia, 10–12 March 2017.
11. Jang, S.W.; Lee, S.H. Detection of Epileptic Seizures Using Wavelet Transform, Peak Extraction and PSR from EEG Signals. *Symmetry* **2020**, *12*, 1239. [CrossRef]
12. Suwicha, J.; Setha, P.N.; Pasin, I. EEG-Based Emotion Recognition Using Deep Learning Network with Principal Component Based Covariate Shift Adaptation. *Sci. World J.* **2014**, *2014*, 627892.

13.   Li, D.; Wu, H.; Zhao, J.; Tao, Y.; Fu, J. Automatic Classification System of Arrhythmias Using 12-Lead ECGs with a Deep Neural Network Based on an Attention Mechanism. *Symmetry* **2020**, *12*, 1827. [CrossRef]

14.   Saini, I.; Singh, D.; Khosla, A. QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases. *J. Adv. Res.* **2013**, *4*, 331–344. [CrossRef]

15.   Barea, R.; Boquete, L.; Mazo, M.; Lopez, E. System for assisted mobility using eye movements based on electrooculography. *IEEE Trans. Nural Syst. Rehabil. Eng.* **2002**, *10*, 209–218. [CrossRef]

16.   Xia, B.; Li, Q.; Jie, J.; Wang, J.; Chaudhary, U.; Ramos-Murguialday, A.; Birbaumer, N. Electrooculogram based sleep stage classification using deep belief network. In Proceedings of the International Joint Conference on Neural Networks, Killarney, Ireland, 12–16 July 2015.

17.   Furness, J.B.; Callaghan, B.P.; Rivera, L.R.; Cho, H.J. The Enteric Nervous System and Gastrointestinal Innervatloe: Integrated Local, and Central Control. *Adv. Exp. Med. Biol.* **2014**, *817*, 39–71.

18.   Madsen, D.; Sebolt, T.; Cullen, L.; Folkedahl, B.; Mueller, T.; Richardson, C.; Titler, M. Listening to bowel sounds: An evidence-based practice project: Nurses find that a traditional practice isn't the best indicator of returning gastrointestinal motility in patients who've undergone abdominal surgery. *AJN Am. J. Nurs.* **2005**, *105*, 40–49. [CrossRef]

19.   Baid, H. A critical review of auscultating bowel sounds. *Br. J. Nurs.* **2009**, *18*, 1125–1129. [CrossRef]

20.   Yang, P.J.; LaMarca, M.; Kaminski, C.; Chu, D.I.; Hu, D.L. Hydrodynamics of defecation. *Soft Matter* **2017**, *13*, 4960–4970. [CrossRef]

21.   Acharya, U.R.; Faust, O.; Sree, S.V.; Ghista, D.N.; Dua, S.; Joseph, P.; Ahamed, V.I.T.; Janarthanan, N.; Tamura, T. An integrated diabetic index using heart rate variability signal features for diagnosis of diabetes. *Comput. Methods Biomech. Biomed. Eng.* **2013**, *16*, 222–234. [CrossRef]

22.   Rekanos, I.; Hadjileontiadis, L. An iterative kurtosis-based technique for the detection of nonstationary bioacoustic signals. *Signal Process.* **2006**, *86*, 3787–3795. [CrossRef]

23.   Dimoulas, C.; Kalliris, G.; Papanikolaou, G.; Petridis, V.; Kalampakas, A. Bowel-sound pattern analysis using wavelets and neural networks with application to long-term, unsupervised, gastrointestinal motility monitoring. *Expert Syst. Appl.* **2008**, *34*, 26–41. [CrossRef]

24.   Kim, K.S.; Seo, J.H.; Song, C.G. Non-invasive algorithm for bowel motility estimation using a back-propagation neural network model of bowel sounds. *Biomed. Eng. Online* **2011**, *10*, 69 . [CrossRef]

25.   Ulusar, U.D. Recovery of gastrointestinal tract motility detection using Naive Bayesian and minimum statistics. *Comput. Biol. Med.* **2014**, *51*, 223–228. [CrossRef]

26.   Liu, J.; Yue, Y.; Jiang, H.; Kan, H.; Wang, Z. Bowel Sound Detection Based on MFCC Feature and LSTM Neural Network. In Proceedings of the 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS), Cleveland, OH, USA, 17–19 October 2018.

27.   Yin, Y.; Jiang, H.; Feng, S.; Liu, J.; Chen, P.; Zhu, B.;Wang, Z. Bowel sound recognition using SVM classification in a wearable health monitoring system. *Sci. China Inf. Sci.* **2018**, *61*, 1–3. [CrossRef]

28.   Du, X.; Allwood, G.; Webberley, K.M.; Inderjeeth, A.J.; Osseiran, A.; Marshall, B.J. Noninvasive Diagnosis of Irritable Bowel Syndrome via Bowel Sound Features: Proof of Concept. *Clin. Transl. Gastroenterol.* **2019**, *10*, e00017. [CrossRef]

29.   Du, X.; Allwood, G.; Webberley, K.M.; Osseiran, A.; Wan, W.; Volikova, A.; Marshall, B.J. A mathematical model of bowel sound generation. *J. Acoust. Soc. Am.* **2018**, *144*, EL485–EL491. [CrossRef]

30.   Zhao, C.; Feng, Z. Application of multi-domain sparse features for fault identification of planetary gearbox. *Measurement* **2017**, *104*, 169–179. [CrossRef]

31.   Li, C.; Sanchez, R.V.; Zurita, G.; Cerrada, M.; Vásquez, R. Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. *Neurocomputing* **2015**, *168*, 119–127. [CrossRef]

32.   Klle, K.; Aftab, M.F.; Andersson, L.E.; Fougner, A.L.; Stavdahl, Y. Data driven filtering of bowel sounds using multivariate empirical mode decomposition. *BioMed. Eng. OnLine* **2019**, *18*, 28. [CrossRef]

33.   Lu, O.U. Rolling Bearing Fault Diagnosis Based on Supervised Laplaian Score and Principal Component Analysis. *J. Mech. Eng.* **2014**, *50*, 88.

34.   Liu, Z.; Cao, H.; Chen, X.; He, Z.; Shen, Z. Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings. *Neurocomputing* **2013**, *99*, 399–410. [CrossRef]

35.   Goksu, H. BCI oriented EEG analysis using log energy entropy of wavelet packets. *Biomed. Signal Process. Control* **2018**, *44*, 101–109. [CrossRef]

36.   Cao, Y.; Sun, Y.; Xie, G.; Wen, T. Fault Diagnosis of Train Plug Door Based on a Hybrid Criterion for IMFs Selection and Fractional Wavelet Package Energy Entropy. *IEEE Trans. Veh. Technol.* **2019**, *68*, 7544–7551. [CrossRef]

37.   Rodriguez, N.; Alvarez, P.; Barba, L.; Cabrera-Guerrero, G. Combining Multi-Scale Wavelet Entropy and Kernelized Classification for Bearing Multi-Fault Diagnosis. *Entropy* **2019**, *21*, 152. [CrossRef] [PubMed]

38.   Rodriguez, N.; Cabrera, G.; Lagos, C.; Cabrera, E. Stationary Wavelet Singular Entropy and Kernel Extreme Learning for Bearing Multi-Fault Diagnosis. *Entropy* **2017**, *19*, 541. [CrossRef]

39.   Bishop, C.M. Neural Networks for Pattern Recognition. *Agric. Eng. Int. Cigr J. Sci. Res. Dev. Manuscr. Pm* **1995**, *12*, 1235–1242.

40.   Rose, C.; Parker, A.; Jefferson, B.; Cartmell, E. The Characterization of Feces and Urine: A Review of the Literature to Inform Advanced Treatment Technology. *Crit. Rev. Environ. Sci. Technol.* **2015**, *45*, 1827–1879. [CrossRef] [PubMed]

41.   Burges, C. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [CrossRef]

42. Cai, J.; Chen, W.; Yin, Z. Multiple Transferable Recursive Feature Elimination Technique for Emotion Recognition Based on EEG Signals. *Symmetry* **2019**, *11*, 683. [CrossRef]

43. Wang, L.; Liu, Y.; Li, T.; Xie, X.; Chang, C. The Short-Term Forecasting of Asymmetry Photovoltaic Power Based on the Feature Extraction of PV Power and SVM Algorithm. *Symmetry* **2020**, *12*, 1777. [CrossRef]

44. Dong, Z.; Zheng, J.; Huang, S.; Pan, H.; Liu, Q. Time-Shift Multi-scale Weighted Permutation Entropy and GWO-SVM Based Fault Diagnosis Approach for Rolling Bearing. *Entropy* **2019**, *21*, 621. [CrossRef]

45. Bruzzone, L.; Chi, M.; Marconcini, M. A Novel Transductive SVM for Semisupervised Classification of Remote-Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3363–3373. [CrossRef]

46. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey Wolf Optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]

47. Markaki, M.; Germanakis, I.; Stylianou, Y. Automatic classification of systolic heart murmurs. In Proceedings of the IEEE International Conference on Acoustics, Vancouver, BC, Canada, 26–31 May 2013.

48. Wong, P.K.; Yang, Z.; Chi, M.V.; Zhong, J. Real-time fault diagnosis for gas turbine generator systems using extreme learning machine. *Neurocomputing* **2014**, *128*, 249–257. [CrossRef]