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 $\{T_1, T_2, \dots, T_{23}, F_1, F_2, \dots, F_{16}, TF_1, TF_2\}$ 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 $T_1, T_2, T_3, \dots, T_{23}$, 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 $F_1, F_2, F_3, \dots, F_{16}$, 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	$F_{13} = f_a$, $which\ \lvert X(a)\rvert = max\langle X(k)\rangle$	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	$TF_1 = \int \lvert X_k(w)\rvert^2 dw$	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.