Evidential regression-based blood glucose detection using waveform features

Hongfeng Long, Chunping Yang, Wei Li, Tian Pu, and Zhenming Peng, *Member, IEEE*

Abstract—As one of the necessary diabetes control and treatment methods, the photoacoustic blood glucose detection technology has great potential due to its deep detection depth and low interference from stray light. Previous research mainly focused on improving the detection capabilities of hardware systems and ignored the exploration of the physical meaning of the signal itself. We analyzed the characteristics of the signal amplitude decay in the photoacoustic signal and employed the forced damping vibration equation to model the signal waveform. A new waveform feature was constructed to describe the amplitude attenuation rate. Moreover, facing low accuracy of blood glucose prediction in the case of small data, we proposed a stable and effective blood glucose detection combining time-frequency feature and waveform features with evidential regression. Finally, in human tissue and glucose solution experiments, the minimum error is achieved 1.02 ± 0.71 mg/dL and 13.28 ± 10.33 mg/dL, respectively.

Index Terms—waveform feature, photoacoustic signal of blood glucose, forced damping vibration equation, evidential regression

I. INTRODUCTION

DUE to the lack of a radical cure, diabetes has become one of the significant diseases threatening human health in the twenty-first century. Therefore, continuous monitoring and control of the patient's blood glucose level have become essential to combat diabetes [1]–[3]. Since frequent blood collection may lead to wound infection, non-invasive blood glucose detection technology has become a research hotspot in recent years [4]–[6]. Photoacoustic blood glucose detection technology has received much attention because of its higher signal-to-noise ratio and stronger skin penetration [7].

The main challenge of non-invasive blood glucose detection technology is, on the one hand, the sensitivity of the signal, which is easily affected by environmental factors; on the other hand, the relationship between the blood glucose signal and

physical properties is not clear, and the detection accuracy is difficult to guarantee [8]–[10]. In particular, in recent years, most studies on blood glucose prediction have focused on processing continuous glucose monitoring (CGM) data [11], [12]. Hence, there are few studies on the analysis of original sensor signals. In response to these two challenges, researchers have made some efforts [13]–[15]. In hardware systems, Tanaka [7] utilized amplitude modulation of dual wavelengths of light to detect small changes in glucose concentration and achieve the standard deviation of 48 mg/dL. Zhang [16] proposed the use of black tape (a strong absorption material) as a "guide star" to improve the sensitivity of photoacoustic signals. In the software, there are also relevant developments [17]. Jnawali K [18] mainly focused on the photoacoustic simulation of chirp excitation response, which helps design the signal acquisition system. Our previous research [19] focused on applying time-frequency analysis technology to blood glucose detection and proposed time-frequency features. However, the exact connection between the blood glucose signal and the physical model is still not found. Therefore, there is still much work to apply photoacoustic technology to actual blood glucose detection.

Inspired by [20], we found that the blood glucose signal is close to a damped sinusoidal signal in the waveform. Hence, the naive idea is to use a model that can generate damped sinusoidal signals to describe the photoacoustic signals, and a classic model is the forced damping vibration equation [21]. At the same time, introducing the forced vibration equation into the signal model helps establish the physical interpretation of the signal. Next, by establishing a suitable loss function and applying optimization techniques, we could calculate the equation parameters based on the known photoacoustic signals and utilized them as waveform features. Finally, we used the evidence regression to combine the waveform features and the previous time-frequency features [19] to establish a mapping relationship with the blood glucose concentration, thereby obtaining a new detection algorithm. In a large number of experiments, the minimum error can reach 1.02 ± 0.71 mg/dL.

In conclusion, our contribution can be mainly concluded as:

- 1) A new waveform feature is constructed to describe the amplitude attenuation characteristics of the photoacoustic signal using the forced damping vibration equation.
- 2) Combining time-frequency features and waveform features, a new blood glucose detection algorithm is proposed using evidence regression, and the prediction error is reduced to 1.02 ± 0.71 mg/dL.

Manuscript received Nov 28, 2021. This work was supported in part by National Natural Science Foundation of China (Grant No.61775030) and partially supported by the Key Laboratory Funds (Grant No.6142205190303). (Corresponding author: Zhenming Peng, Tian Pu)

The authors are with School of Information and Communication Engineering, University of Electronic Science and Technology of China, 611731, Chengdu, China (e-mail: lhfuestc@std.uestc.edu.cn, yangchunping@uestc.edu.cn, liwei@uestc.edu.cn, putian@uestc.edu.cn, zm-peng@uestc.edu.cn)

The authors are also with Laboratory of Imaging Detection and Intelligent Perception University of Electronic Science and Technology of China, 610054, Chengdu, China.

- 3) To a certain extent, the forced damping vibration equation can effectively explain the mechanism of blood glucose signal generation. Compared with the wave equation, this equation can better explain the decay of signal energy over time.

This paper is organized as follows: Section II briefly introduces the related basic theories involved. The waveform feature and the proposed blood glucose detection algorithm are presented in Section III. Section IV gives the necessary experiments to testify to our proposed method. Some discussions are in Section V, and the conclusion is in Section VI.

II. PRELIMINARIES

A. Forced damping vibration equation

Forced damping vibration is a classical mechanical vibration, which has been studied for many years [21]. In the photoacoustic signal generation, the glucose molecules that absorb laser energy also vibrate, and the mode of vibration can be modelled by forced damping vibration.

Different from the studies using wave equations [22], the forced damping vibration equation can better describe the tailing phenomenon of photoacoustic signals. The reason for trailing is that the energy of the laser pulse is gradually converted into the ultrasonic wave. The equation of the forced damping vibration is shown as follows:

$$\frac{d^2 x(t)}{dt^2} + 2\xi\omega \frac{dx(t)}{dt} + \omega^2 x(t) = F(t) \quad (1)$$

where $x(t)$ represents the photoacoustic signal in the time domain, and ξ and ω represent damping ratio and resonance frequency(or natural frequency), respectively. $F(t)$ is the excitation function, whose physical meaning is the input energy curve. When ξ, ω and $F(t)$ are known, the differential equation (1) can be solved by many numerical algorithms [23].

In general, we assume the excitation function $F(t)$ is a Gaussian function with small variance, and its peak is located at the maximum value of the photoacoustic signal. The formula is

$$F(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) \quad (2)$$

where μ is equal to the location of the maximum value of the photoacoustic signal, and σ is set as 1, generally. It is worth noting that when the model cannot output the damped vibration signal, σ needs to be adjusted, and the range is generally [1, 2].

B. Hilbert transform

In this paper, the Hilbert transform is utilized to extract the envelope of the photoacoustic signal, and its purpose is to match the solution results of (1) with the actual data. The details of the Hilbert transform are shown as follows.

$$\begin{aligned} H[x(t)] &= x(t) + i\hat{x}(t) \\ \hat{x}(t) &= x(t) * \frac{1}{\pi t} \end{aligned} \quad (3)$$

where $*$ represents the convolution operation, and $H[x(t)]$ is the Hilbert transform of the input signal $x(t)$. The complex signal $H[x(t)]$ is also called an analytical signal [24], [25], and its absolute value is the envelope of the original signal $x(t)$.

C. Evidential regression

Evidential regression [26] is a non-parametric regression method to analyze uncertain and imprecise data. Compared with some traditional regression methods, such as multiple linear regression(MLR) [27] and support vector regression(SVR) [28], this method can perform better in non-linear situation. Compared with the neural network regression(NNR) [29], in a small dataset, evidential regression can achieve a lower prediction error. The details of this regression are shown as follows.

Assume we have a dataset with N points $(\mathbf{x}_i, y_i)$ $i = 1, 2, \dots, N$ and $\mathbf{x}_i$ represents a vector. Firstly, We need to generate the corresponding basic belief assignment(BBA) of each point according to [30]. The frame of discernment(FOD) is $\Omega = \{y_1, y_2, \dots, y_k\}, k \leq N$, which means k different results in this dataset.

Then, the BBA of a point $(\mathbf{x}_i, y_i)$ can be computed by the following formula:

$$\begin{aligned} m_i(y_i) &= p \\ m_i(\Omega) &= 1 - p \end{aligned} \quad (4)$$

where p is the reliability of this point and $p \in [0, 1]$, generally set as 0.9. Each point can generate a corresponding BBA.

Next, a discounting [26] should be apply to each BBA, and the formula is

$$\hat{m}_i(\mathbf{x}, A) = \begin{cases} m_i(A)\phi(\|\mathbf{x} - \mathbf{x}_i\|) & A \in y_i \setminus \Omega \\ 1 - \phi(\|\mathbf{x} - \mathbf{x}_i\|) & A = \Omega \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where A represents a focal element which is a subset of the FOD, and ϕ and $\|\bullet\|$ are defined as:

$$\begin{aligned} \phi(d) &= \gamma \exp(-d^2) \\ \|\mathbf{x} - \mathbf{x}_i\| &= [(\mathbf{x} - \mathbf{x}_i)^T \Sigma (\mathbf{x} - \mathbf{x}_i)]^{1/2} \end{aligned} \quad (6)$$

where $\gamma \in [0, 1]$ is a real parameter, and Σ is a symmetric positive definite matrix. At the same time, γ is an optimization parameter, which needs to be determined in Algorithm 3.

After obtaining the corresponding BBA of each point, for an input vector $\mathbf{x}$, we should combine the information provided by each point of the training set Θ by utilizing the conjunctive rule of combination for BBAs. The final BBA is then:

$$\hat{m}_{\Theta}(\mathbf{x}, y_i) = \bigoplus_{i=1}^N \hat{m}_i(\mathbf{x}, y_i) \quad (7)$$

where $\bigoplus$ represents the combination rule of Dempster, and the relevant details are in [31].

The fitting result $\hat{y}$ of the input $\mathbf{x}$ can be computed by the following formula.

$$\hat{y} = \sum_{i=1}^k y_i \times \hat{m}_{\Theta}(\mathbf{x}, y_i) \quad (8)$$

III. THE PROPOSED METHOD

A. Waveform features

Although the common features can achieve acceptable performance [32], [33], they cannot directly describe the patterns of the blood-glucose photoacoustic signal. In 2020, Zhang [20] first proposed the time-domain photoacoustic waveform analysis to model this signal using a damped sinusoidal

signal, and build a concrete expression. However, a complex expression is hard to satisfy our expectations, and we wonder a more concise expression.

Inspired by the studies [21] in the forced damping vibration, we modelled the process of the photoacoustic signal generation using the equation of the forced damping vibration. The new model can achieve a more concise and accurate description with a differential equation. The details are shown as follows.

The waveform features are composed of the parameters or the derivative indexes of (1), such as resonance frequency ω , maximum overshoot, etc. Hence, the critical problem is how to use the known signal to solve these model parameters.

For convenience, we assume the solution of the (1) is $x(t)$ and the known photoacoustic signal is $\hat{x}(t)$. The plain idea is to construct an optimization problem to obtain these parameters, and the detail is

$$(\xi, \omega) = \arg \min_{\xi, \omega} \|x(t) - \hat{x}(t)\|^2 \quad (9)$$

where loss function is $L(x) = \|x(t) - \hat{x}(t)\|^2$.

However, due to the error between the measurement system and the model, the problem of directly solving (9) is unstable. Therefore, we designed a matching algorithm based on a signal envelope to solve this problem. The optimization problem can be transformed into the following form.

$$(\xi, \omega) = \arg \min_{\xi, \omega} |||H[x(t)] - |H[\hat{x}(t)]|||^2 \quad (10)$$

where $|H[\bullet]|$ represents the computation of the signal envelope using the Hilbert transform.

According to the principle of forced vibration, only when the damping ratio $\xi \leq 1$, a damped sinusoidal signal can be generated. However, through many experiments, we found that only when $\xi = 1$, the solution of (1) could get the desired result, which was discussed in detail in Section V-A. Therefore, we set $\xi = 1$.

The final optimization problem is

$$\omega_{opt} = \arg \min_{\omega} |||H[x(t)] - |H[\hat{x}(t)]|||^2 \quad \xi = 1 \quad (11)$$

where ω_{opt} represents the optimal matching frequency. Then, the other feature value can be calculated by

$$\begin{aligned} M_p &= \frac{|max(x(t)) - x(\infty)|}{x(\infty)} * 100\% \\ t_s &= \frac{3}{\omega_{opt}} \end{aligned} \quad (12)$$

where M_p, t_s represents the maximum overshoot and the adjustment time, respectively. $x(\infty)$ can be estimated by the mean value of $x(t)$.

Finally, the waveform feature is (ω_{opt}, M_p, t_s) . The pseudo-code of the algorithm flow is in Algorithm 1.

B. Blood glucose detection

Due to the expensive experimental equipment and the lack of public data sets, the problem of a small amount of data is inevitable in the processing of photoacoustic blood glucose signals. We proposed a new blood glucose detection algorithm combining the waveform features and evidential regression to face this problem. The pseudo-code is shown in Algorithm 2.

Algorithm 1 Waveform feature extraction

Require: $\mathbf{x}$: a data vector;

Ensure: $\mathbf{f}$: the waveform feature vector;

- 1: $\xi \leftarrow 1$
  - 2: $x_{max} \leftarrow \max(\mathbf{x})$
  - 3: $loc \leftarrow \text{find}(\mathbf{x}, x = x_{max})$
  - 4: $F(t) \leftarrow \text{solve (2) under } \mu = loc, \quad \sigma = 1$
  - 5: $x(t) \leftarrow \text{solve (1) and } \omega \text{ starts from 0}$
  - 6: $\omega_{opt} \leftarrow \text{solve (11)}$
  - 7: Following (12) to get M_p, t_s
  - 8: **return** $\mathbf{f} = (\omega_{opt}, M_p, t_s)$
-

Algorithm 2 Blood glucose detection algorithm

Require:

$\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)\}$: a set of data points for training(N points), where $\mathbf{x}_i$ and y_i represent i th data vector and the corresponding concentration of the blood glucose;

$\tilde{\mathbf{x}}$: a data point with unknown concentration.

Ensure: $\hat{y}$: an estimation of blood glucose concentration.

- 1: $\tilde{\mathbf{f}} \leftarrow \text{using Algorithm 1 to extract feature according to } \tilde{\mathbf{x}}$
  - 2: $\tilde{\mathbf{f}} \leftarrow [\tilde{\mathbf{f}}, \tilde{\mathbf{f}}_{tf}]$, $\tilde{\mathbf{f}}_{tf}$ represents the time-frequency features
  - 3: $\gamma \leftarrow \text{using Algorithm 3 to determine according to all training data}$
  - 4: **for** $i = 0 \rightarrow n$ **do**
  - 5: $\mathbf{f}_i \leftarrow \text{using Algorithm 1 and [19] to extract feature according to } \mathbf{x}_i$
  - 6: $\hat{m}_i(\tilde{\mathbf{f}}, y_i) \leftarrow \text{using (5) to generate according to } \mathbf{f}_i \text{ and } \gamma$
  - 7: **end for**
  - 8: $\hat{m}_{\Theta}(\tilde{\mathbf{f}}, y_i) \leftarrow \text{using (7);}$
  - 9:
  - 10: **return** $\hat{y} \leftarrow \text{using (8);}$
-

where the training process of the evidential regression has been shown in Algorithm 3, and the time-frequency features are extracted according to [19].

In Algorithm 2, we utilize two steps to ensure the accurate estimation of blood glucose under the condition of a small data set. Firstly, the waveform feature extracted from the input signal can help describe the signal's basic properties. Specifically, the resonance frequency ω can describe the energy attenuation rate of the photoacoustic signal. With the increase of resonance frequency, the energy decays faster, and this will be discussed in detail in Section V. In addition, compared with directly using the whole signal, the waveform feature is equivalent to reducing the dimension of the original data to facilitate the regression under a small data set.

Secondly, in small data sets, the evidential regression has a stronger fitting ability than classical methods [27] and can deal with strong non-linearity. For example, the multiple linear regression demand that the number of training data is greater than or equal to the dimension of a feature vector; otherwise, the regression accuracy cannot be guaranteed (rank defect problem). However, the more different features, the more information, so the better the performance. The evidential

regression can well reconcile the contradiction between the two, which is the multiple linear regression cannot handle. As for, the neural network regression [29] is difficult to converge under small data sets. Therefore, we comprehensively apply evidence regression to fit the nonlinear mapping function between waveform feature and blood glucose concentration.

Algorithm 3 The training process of the evidential regression

Require: $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)\}$: a set of data points for training(N points), where $\mathbf{x}_i$ and y_i represent i th data vector and the corresponding concentration of the blood glucose;

Ensure: γ : an optimization parameter.

```

1:  $\gamma \leftarrow 0.001$ 
2: while  $err$  tends to converge do
3:   for  $i = 0 \rightarrow n$  do
4:      $\mathbf{f}_i \leftarrow$  using Algorithm 1 to extract feature according to  $\mathbf{x}_i$ 
5:      $\mathbf{f}_i \leftarrow [\mathbf{f}_i, \tilde{\mathbf{f}}_{tf}]$ ,  $\tilde{\mathbf{f}}_{tf}$  represents the time-frequency feature [19]
6:   end for
7:   for  $i = 0 \rightarrow n$  do
8:     for  $j = 0 \rightarrow k$  do
9:        $\hat{m}_i(\mathbf{f}_i, y_j) \leftarrow$  using (5) to generate
10:    end for
11:     $\hat{m}_\Theta(\mathbf{f}_i, y_i) \leftarrow$  using (7);
12:     $\hat{y}_i \leftarrow$  using (8);
13:  end for
14:   $err \leftarrow \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$ 
15: end while

```

IV. EXPERIMENTS

A. Experimental setup

1) *Photoacoustic measurement*: The photoacoustic signal measurement was carried out by a self-built system. We used some glucose solutions as a single and uniform medium to verify whether our method could model the waveform of the photoacoustic signal. The schematic diagram and physical diagram of the experimental device are in Fig. 1.

In Fig. 1.(a), the wavelength of the pulsed laser is 1064nm, and the pulse width is 11.5ns. The repetition rate and the single pulse energy are 10Hz and 6.7mJ, respectively. The ultrasonic sensor(DAOBO, 8PΦ20) detected the photoacoustic signal and transformed it into electrical signals. The signal amplifier(Aigtek, ATA-5220) has a voltage gain of 60 dB and a bandwidth range of 1kHz to 5MHz. The real-time sampling rate of the equipment is 100Mpsps. Then, the details of this experiment steps are:

- 1) Firstly, we used a rubber-tipped dropper to add the measured glucose solution to the plastic pipe;
- 2) Then, we needed to turn on the pulse light source to irradiate the solution;
- 3) The generated photoacoustic signal could be received by the ultrasonic sensor and recorded;
- 4) Next, we used an invasive blood glucose meter(Sannuo, GA-3) to measure the concentration of the measured

Fig. 1. The schematic diagram and physical diagram of the experimental device. (a) The schematic diagram; (b) The physical diagram

glucose solution and took the average value of repeated measurements three times as a result;

- 5) Using our proposed algorithm with the measured signal, we could obtain the predicted value of the measured glucose concentration;
- 6) Finally, using the rubber-tipped dropper to suck out the measured solution and start the subsequent measurement.

One thing to note in the experiment is that the laser's power supply may produce serious interference signals. Therefore, it is necessary to ensure that the light source is physically separated from the sample under test when fixing the light source, and there is enough buffer in the middle to eliminate interference.

2) *Accuracy evaluation*: The prediction accuracy is evaluated by comparing the predicted glucose concentration $\hat{y}_i$ to the reference glucose value y_i using quantitative measures such as the root mean square error(RMSE) and the coefficient of determination r_s [34]. These evaluation indicators are calculated as follows:

$$\begin{aligned}
 RMSE &= \sqrt{\frac{1}{n} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \\
 r_s &= 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}
 \end{aligned} \tag{13}$$

where $\bar{y}$ represents the mean value of the reference glucose value y_i .

Furthermore, Clarke error grid(CEG) analysis and Bland-Altman(BA) plots are generally used as the "gold standard" to assess calibration accuracy by comparing the predicted and

reference glucose concentrations. In CEG, the points falling in Region A and B are considered clinically acceptable. As for the BA plot, the space between the two solid lines represents the error, and the smaller, the better. Moreover, the centered dotted line indicates the fixed error, and it is best to approach 0.

B. The glucose solution's signal

We obtained a total of 46 available samples using the experimental setup in Section IV-A, and the specific waveform of a sample(153 mg/dL) is shown in Fig. 2.

Fig. 2. The real signal of glucose solutions(153 mg/dL). The red curve is the original signal obtained from the experiment, and the blue one is the output of our proposed model.

Specifically, the blue curve is the solution of (1) under the most appropriate parameters. As can be seen directly from Fig. 2, our model mainly describes the gradual attenuation characteristics of the photoacoustic signals and models the attenuation rate of signals. However, there are still deficiencies in the rising part of the signal. For example, at the rising edge of the signal, the model is difficult to describe.

Next, due to the limited number of samples, we utilized 46 samples as the test set and randomly selected a part of them as the training set. Moreover, The extraction ratio gradually increased from about 20% to 100%. We compared our method with MLR [27], SVR [28] and NNR [29], combining the proposed waveform features and the time-frequency features in [19]. The details of the result are shown in Fig. 3 and Table. I.

From Fig. 3 and Table. I, our proposed method can achieve the best performance among these comparison methods, and the error can achieve 13.28 ± 10.33 mg/dL in the optimal condition. Without the influence of the skin and other tissues, the photoacoustic signal amplitude produced by glucose solution is significantly stronger than that of the blood glucose signal. Moreover, the blood glucose signal is significantly smoother than the glucose signal. The bias caused by this model may be a tricky part of our proposed model.

In addition, as can be seen from Fig. 3, the error curve gradually converges with the increase of training samples. Therefore, we employed the CEG and the BA plot to examine the proposed method. The CEG is shown in Fig. 4.

Fig. 4. The Clarke Error Grid of glucose solutions concentration

In Fig. 4, it is obvious that all data points are located in Region A and Region B, of which 93.48% (43/46) are located in Region A and the rest in Region B. It intuitively shows that our method is clinically acceptable in prediction performance. However, the data points do not fall evenly on the diagonal, so it still can be improved. In order to further quantitatively analyze the prediction error, we utilized the BA plot, as shown in Fig. 5.

Fig. 5. The Bland-Altman plot of glucose solutions concentration

From Fig. 5, each blue dot represents the corresponding sample; Most of the points are within the confidence interval(the range between two solid lines), and only one point slightly exceeds the interval. The center dotted line representing the mean difference is close to 0, indicating that the fixed bias of our method is small. However, $r_s = 0.62$ is low, which is a deficiency. That may be affected by the quality of the data itself, and the specific reasons need to be further studied.

Finally, the prediction error of our proposed method in this experiment achieved 13.28 ± 10.33 mg/dL and RMSE = 16.76 mg/dL.

C. The human finger's photoacoustic signal

In order to further explore the effect of our method in practical application, we used the actual photoacoustic data collected from human fingers for verification. Thirty samples collected were tested, and the signal acquisition equipment was introduced in [19]. It is worth noting that each data acquisition

Fig. 3. The prediction error of glucose solutions concentration. (a) The mean value of error(mg/dL); (b) The variance of error(mg/dL). The comparison methods are neural network regression(NNR), multiple linear regression(linear) and support vector regression(SVR).

TABLE I
THE PREDICTION ERROR OF GLUCOSE SOLUTIONS CONCENTRATION

Method	Training proportion(%)							
	20.0	37.8	55.6	64.4	73.3	82.2	91.1	100
MLR	70.73±54.70	55.73±50.07	44.85±33.06	40.32±31.54	41.02±31.46	40.65±30.82	39.99±30.43	39.87±30.56
SVR	80.57±54.91	71.81±53.80	69.00±55.08	67.81±50.84	65.16±50.20	64.58±49.49	65.21±50.50	64.15±50.25
NNR	287.97±239.46	147.25±166.80	73.94±100.41	74.30±94.19	52.41±77.46	44.73±65.27	42.59±59.32	39.72±51.98
Our method	72.82±55.86	55.27±57.71	32.79±43.30	31.71±39.70	23.56±36.92	19.49±24.25	16.11±16.47	13.28±10.33

The format of values in the table is the mean value (mg/dL) $\pm$ the standard deviation (mg/dL).

required the test to prick his finger and collect blood. The accurate value of blood glucose was obtained by the blood collected by the invasive blood glucose meter(Onetouch Verio Flex; LifeScan, Inc, Milpitas, CA, USA).

The waveform of the signal(marked red) and the output of our proposed model(marked blue) are shown in Fig. 6. Similarly, it can be seen from Fig. 6 that the model output

Fig. 6. The real signal of blood glucose(180 mg/dL). The red curve is the original signal, and the blue one is the model output.

results are well pressed into the real signal in the signal attenuation part. Therefore, the model can well model the amplitude attenuation rate of the photoacoustic signal.

If the fitting effect is carefully compared, it can be found that the fitting effect in the blood glucose signal(in Fig. 6) is better than that in glucose solution(in Fig. 2). The reason may be the smoothing effect of skin and other tissues on photoacoustic signal, which makes the final signal closer to the model in this paper.

In order to highlight the effectiveness of our proposed

algorithm, we compared our method with MLR, SVR and NNR, and the results are shown in Fig. 7.

According to Fig. 7, as more and more samples were used for training, the errors of various methods were gradually attenuated. Among them, our proposed method(marked blue) achieved the lowest error when the training data were sufficient, and it maintained good prediction performance when the training set was small. As for other methods, they cannot be achieved. For example, NNR cannot achieve the minimum error in the small-data case, and its error curve still does not converge. The specific details are shown in Table. II.

From Table. II, we can find that our proposed method can achieve the minimum prediction error and better stability. The marked red part represents the best performance in the corresponding situation. Moreover, the best performance can achieve 1.02 ± 0.71 mg/dL.

Fig. 8. The Clarke Error Grid of blood glucose

In Fig. 8, all data points are located in Region A and fit

Fig. 7. The prediction error of blood glucose concentration(human issue). (a) The mean value of error(mg/dL); (b) The variance of error(mg/dL). The comparison methods are neural network regression(NNR), multiple linear regression(linear) and support vector regression(SVR).

TABLE II
THE PREDICTION ERROR OF BLOOD GLUCOSE CONCENTRATION

Method	Training proportion(%)							
	33.3	48.1	55.6	63.0	77.8	85.2	92.6	100
MLR	37.35±69.76	22.99±24.26	22.69±15.82	18.24±16.89	19.29±16.54	18.26±17.41	17.58±14.59	17.66±14.34
SVR	22.11±18.80	19.12±14.96	18.36±14.24	18.85±14.35	18.50±14.06	18.52±14.73	18.11±14.22	18.27±14.09
NNR	60.39±58.03	56.95±60.11	37.12±41.27	25.96±34.14	30.23±35.91	28.15±29.57	18.75±27.01	18.92±23.97
Our method	18.11±21.52	12.75±14.49	8.03±9.68	8.40±13.86	6.44±12.93	5.38±13.16	1.56±2.54	1.02±0.71

The format of values in the table is the mean value (mg/dL) $\pm$ the standard deviation (mg/dL).

diagonally in CEG, which shows that our method has good prediction performance. Due to the small variation range of blood glucose in the actual human body, the data points are more concentrated, which is not as scattered as Fig. 4. Next, we employed the BA plot to quantitatively examine the prediction accuracy of our method, which is shown in Fig. 9.

Fig. 9. The Bland-Altman plot of blood glucose

According to Fig. 9, we can clearly find that the prediction errors are within the confidence interval, and the variance is small. At the same time, r_s achieves 1.0, which shows that our method has excellent prediction performance and can meet the needs of blood glucose detection. Finally, the prediction error of our proposed method in this experiment achieved 1.02 ± 0.71 mg/dL and $RMSE = 1.24$ mg/dL.

V. DISCUSSION

We want to discuss two questions in this section. The one is the influence of the damping ratio ξ in the forced damping vibration equation. The other one is a comparative experiment with the related works. The details are shown as follows.

A. The influence of the damping ratio ξ

The damping ratio ξ is an important parameter in the damped vibration equation, which is used to control the system's state. According to the specific value of ξ , the system status can be divided into three types:

- When $\xi < 1$, the system is in an under-damped state;
- When $\xi = 1$, the system is in a critical-damped state;
- When $\xi > 1$, the system is in an over-damped state.

In order to study the proposed model output in three different cases, we took $\xi = \{0.7, 1.0, 1.5\}$ and used a Gaussian function as excitation function $F(t)$. The results are shown in Fig. 10.

From Fig. 10, it is evident that only when $\xi = 1$, the model's output can meet expectations. Otherwise, the output result will have a large deviation. There is no clear explanation for the specific reasons for this phenomenon, but based on many experimental results, we set $\xi = 1$ in the subsequent experiments.

B. Comparative experiment

This experiment was mainly used to compare the performance of the method and related literature [16], [19], [33], [35]. We choose RMSE, r_s and CEG to measure the system's performance, and some implementation details are shown in Table. III.

In Table. III, since [33] only provided figures and lacked specific data, it is impossible to give the specific values of CEG, the coefficient of determination r_s and BA plot. The best performance has been marked(red).

According to Table. III, we can find that when only waveform features are used, a minor prediction error can

Fig. 10. The influence of different damping ratio. (a) The original signal (red) and the excitation function (blue), where $F(t)$ is used to solve (1). The blue curve in the (b), (c), and (d) represents the model output at different ξ ; (b) $\xi = 0.7$, (c) $\xi = 1.0$, (d) $\xi = 1.5$.

TABLE III
A COMPARISON WITH RELATED LITERATURE

Method	Implementation Details	RMSE(mg/dL)	CEG(%)				r_s	BA plot(mg/dL)
			A	B	C	D		
Zhang [16]	Human Blood Serum	17.30	100	0	0	0	0.9791	-0.15 ± 30.94
Zhang [20]	Human Blood Serum	9.84	100	0	0	0	0.9956	2.49 ± 29.19
Pai [33]	Whole Blood	15.61	-	-	-	-	-	-
Long [19]	Human issue	24.27	80	20	0	0	0.8412	0.00 ± 24.27
Only Waveform feature	Glucose solution	39.40	80	11	0	9	0.4200	26.38 ± 29.59
	Human issue	5.15	100	0	0	0	0.8400	3.11 ± 4.18
Our method	Glucose solution	16.76	93	7	0	0	0.6200	13.28 ± 10.33
	Human issue	1.24	100	0	0	0	1.0000	1.02 ± 0.71

be obtained under the same situation; for example, in the human issue, the error is only 5.15 mg/dL only using the waveform feature. At the same time, when we use multiple features (time-frequency features and waveform features) to express the photoacoustic signal together, better performance can be achieved. For example, the RMSE can achieve 1.24 mg/dL, and the BA plot reaches 1.02 ± 0.71 mg/dL. The result also shows that the new features we proposed are different from the existing ones and have better performance.

As for the reason that the test error in the glucose solution is higher than that of human tissue, the signal acquisition device itself may be unstable or human factors.

VI. CONCLUSION

In this paper, we mainly studied two issues: how to model the physical characteristics of the photoacoustic blood glucose signal and how to detect blood glucose concentration in a small dataset accurately. Given the differences in the characteristics of blood glucose signals of different concentrations in the

attenuation rate, signal amplitude, and tail length, we proposed a new waveform feature to model the photoacoustic blood glucose signals. Facing the problem of a small dataset, we employed the evidence regression combining multiple features to design a blood glucose detection algorithm to estimate the blood glucose concentration accurately.

The waveform feature is constructed based on the forced damping vibration equation because this equation can indicate that the attenuation of the signal is due to the gradual conversion of limited light energy into ultrasonic energy, rather than being attenuated by media such as skin. That is a more reasonable explanation for the apparent tailing phenomenon in the photoacoustic signal. ω_{opt} , M_p , t_s can express the attenuation rate, signal amplitude and tail length, respectively. We can calculate the corresponding model parameters using the optimization technology under the known photoacoustic signal to form the waveform features. Through the waveform features, we can establish a more reasonable signal model.

Next, we employed evidence regression combining the

proposed waveform features and time-frequency features to propose a new blood glucose detection algorithm. The significance of this work is to improve the accuracy of blood glucose concentration prediction under small-scale data and overcome the difficulty of the highly nonlinear fitting among multiple features. Through many experiments, we found that compared to MLR, SVR, and NNR, the evidence regression performs better in the case of small data sets. Furthermore, as the number of data increases, the performance can continue to improve.

The error analysis (such as CEG and BA plot) in real data is provided to illustrate our method's validity and rationality. In the case of human tissue and glucose solution, the minimum error is achieved 1.02 ± 0.71 mg/dL and 13.28 ± 10.33 mg/dL, respectively. However, our method still has shortcomings in describing the rising part of the signal, which will be studied in the future.

REFERENCES

- [1] C. Xu, S. Rassel, S. Zhang, A. Aloraynan, and D. Ban, "Single-wavelength water muted photoacoustic system for detecting physiological concentrations of endogenous molecules," *Biomedical optics express*, vol. 12, no. 1, pp. 666–675, 2021.
- [2] R. A. Karim, I. Vassányi, and I. Kósa, "After-meal blood glucose level prediction using an absorption model for neural network training," *Computers in Biology and Medicine*, vol. 125, p. 103956, 2020.
- [3] Y. Liu and Z. Peng, "Infrared small target detection based on resampling-guided image model," *IEEE Geoscience and Remote Sensing Letters*, 2021.
- [4] Y. Fu, Q. Zhou, Q. Liu, L. Tao, and X. Wei, "Photoacoustic signal classification for in vivo photoacoustic flow cytometry based on support vector machine," in *Optics in Health Care and Biomedical Optics IX*, vol. 11190. International Society for Optics and Photonics, 2019, p. 111902G.
- [5] A. Nawaz, P. Øhlckers, S. Sælid, M. Jacobsen, and M. Nadeem Akram, "Review: Non-invasive continuous blood glucose measurement techniques," *J. Bioinforma. Diabetes*, vol. 1, no. 3, pp. 1–27, 2016.
- [6] Y. Chen, Z. Peng, A. Gholami, J. Yan, and S. Li, "Seismic signal sparse time-frequency representation by lp-quasinorm constraint," *Digital Signal Processing*, vol. 87, pp. 43–59, 2019.
- [7] Y. Tanaka, T. Tajima, M. Seyama, and K. Waki, "Differential continuous wave photoacoustic spectroscopy for non-invasive glucose monitoring," *IEEE Sensors Journal*, vol. 20, no. 8, pp. 4453–4458, 2019.
- [8] S. Guin and M. Chattopadhyay, "Advancement and challenges for non-invasive monitoring of blood glucose: A review," *Advances in Medical Physics and Healthcare Engineering*, pp. 101–113, 2021.
- [9] B. Schuler, L. Kühner, M. Hentschel, H. Giessen, and C. Tarin, "Adaptive method for quantitative estimation of glucose and fructose concentrations in aqueous solutions based on infrared nanoantenna optics," *Sensors*, vol. 19, no. 14, p. 3053, 2019.
- [10] S. Li and Z. Peng, "Seismic acoustic impedance inversion with multi-parameter regularization," *Journal of Geophysics and Engineering*, vol. 14, no. 3, pp. 520–532, 2017.
- [11] J. Martinsson, A. Schliep, B. Eliasson, and O. Mogren, "Blood glucose prediction with variance estimation using recurrent neural networks," *Journal of Healthcare Informatics Research*, vol. 4, no. 1, pp. 1–18, 2020.
- [12] T. El Idrissi and A. Idri, "Deep learning for blood glucose prediction: Cnn vs lstm," in *International Conference on Computational Science and Its Applications*. Springer, 2020, pp. 379–393.
- [13] G. Chen, Z. Peng, and Y. Li, "A framework for automatically choosing the optimal parameters of finite-difference scheme in the acoustic wave modelling," *Computers & Geosciences*, p. 104948, 2021.
- [14] Z. Cao, X. Kong, Q. Zhu, S. Cao, and Z. Peng, "Infrared dim target detection via mode-klk2 extension tensor tubal rank under complex ocean environment," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 181, pp. 167–190, 2021.
- [15] X. Kong, C. Yang, S. Cao, C. Li, and Z. Peng, "Infrared small target detection via nonconvex tensor fibered rank approximation," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.
- [16] R. Zhang, F. Gao, X. Feng, H. Jin, S. Zhang, S. Liu, Y. Luo, B. Xing, and Y. Zheng, "'guide star' assisted noninvasive photoacoustic measurement of glucose," *ACS sensors*, vol. 3, no. 12, pp. 2550–2557, 2018.
- [17] Z. Ren, G. Liu, and Z. Huang, "Study on an improved wavelet threshold denoising for the time-resolved photoacoustic signals of the glucose solution," in *2015 International Conference on Optical Instruments and Technology: Optoelectronic Devices and Optical Signal Processing*, vol. 9619. International Society for Optics and Photonics, 2015, p. 96190L.
- [18] K. Jnawali, B. Chinni, V. Dogra, and N. Rao, "Photoacoustic simulation study of chirp excitation response from different size absorbers," in *Medical Imaging 2017: Ultrasonic Imaging and Tomography*, vol. 10139. International Society for Optics and Photonics, 2017, p. 101391L.
- [19] H. Long, B. Chen, W. Li, Y. Xian, and Z. Peng, "Blood glucose detection based on teager-kaiser main energy of photoacoustic signal," *Computers in Biology and Medicine*, p. 104552, 2021.
- [20] R. Zhang, Y. Luo, H. Jin, F. Gao, and Y. Zheng, "Time-domain photoacoustic waveform analysis for glucose measurement," *Analyst*, vol. 145, no. 24, pp. 7964–7972, 2020.
- [21] L. Bergman and J. Nicholson, "Forced vibration of a damped combined linear system," *Journal of Vibration, Acoustics, Stress, and Reliability in Design*, vol. 107, pp. 275–281, 1985.
- [22] T. Saratoon, T. Tarvainen, B. Cox, and S. Arridge, "A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation," *Inverse Problems*, vol. 29, no. 7, p. 075006, 2013.
- [23] L. F. Shampine, *Numerical solution of ordinary differential equations*. Routledge, 2018.
- [24] Y. Wang, R.-S. Wu, G. Chen, and Z. Peng, "Seismic modulation model and envelope inversion with smoothed apparent polarity," *Journal of Geophysics and Engineering*, vol. 15, no. 5, pp. 2278–2286, 2018.
- [25] X. Wang, Z. Peng, D. Kong, and Y. He, "Infrared dim and small target detection based on stable multisubspace learning in heterogeneous scene," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 10, pp. 5481–5493, 2017.
- [26] S. Petit-Renaud and T. Denœux, "Nonparametric regression analysis of uncertain and imprecise data using belief functions," *International Journal of Approximate Reasoning*, vol. 35, no. 1, pp. 1–28, 2004.
- [27] M. Tranmer and M. Elliot, "Multiple linear regression," *The Cathie Marsh Centre for Census and Survey Research (CCSR)*, vol. 5, no. 5, pp. 1–5, 2008.
- [28] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.
- [29] S. Cagan, M. Maci, M. Buldum, and C. Maci, "Artificial neural networks in mechanical surface enhancement technique for the prediction of surface roughness and microhardness of magnesium alloy," *Bulletin of the Polish Academy of Sciences. Technical Sciences*, vol. 67, no. 4, 2019.
- [30] L. M. Zouhal and T. Denœux, "An evidence-theoretic k-nn rule with parameter optimization," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 28, no. 2, pp. 263–271, 1998.
- [31] H. Long, Z. Peng, and Y. Deng, "A new structure of the focal element in object recognition," *International Journal of Intelligent Systems*, p. DOI: 10.1002/int.22675, 2021.
- [32] A. S. Alghamdi, K. Polat, A. Alghoson, A. A. Alshdadi, and A. A. Abd El-Latif, "Gaussian process regression (gpr) based non-invasive continuous blood pressure prediction method from cuff oscillometric signals," *Applied Acoustics*, vol. 164, p. 107256, 2020.
- [33] P. P. Pai, A. De, and S. Banerjee, "Accuracy enhancement for noninvasive glucose estimation using dual-wavelength photoacoustic measurements and kernel-based calibration," *IEEE transactions on instrumentation and measurement*, vol. 67, no. 1, pp. 126–136, 2017.
- [34] W. L. Clarke and B. Kovatchev, "Continuous glucose sensors: Continuing questions about clinical accuracy," *Journal of Diabetes Science and Technology*, vol. 1, no. 5, pp. 669–675, 2007.
- [35] A. K. Singh, M. V. Rebec, and A. Haidar, "Kalman-based calibration algorithm for agamatrix continuous glucose monitoring system," *IEEE Transactions on Control Systems Technology*, vol. 29, no. 3, pp. 1257–1267, 2020.