TRANSFER AND COLLABORATIVE LEARNING METHOD FOR PERSONALIZED NONINVASIVE BLOOD GLUCOSE MEASUREMENT MODELING

Weijie Liu[1,4]*, Anpeng Huang*[2,*]*, Ping Wang*[3,1,4,*] *and Hebin Yao*[5]

[1]School of Software and Microelectronics, Peking University
[2]National Institute of Health Data Science, Peking University
[3]National Engineering Research Center for Software Engineering, Peking University
[4]Key Laboratory of High Confidence Software Technologies (Peking University), Ministry of Education
[5]Endocrinology Department in Navy General Hospital PLA China

ABSTRACT

Non-invasive Glucose Measurement (NGM) technology is promising and desired for patients with hyperglycemia or hypoglycemia. In various kinds of NGM technologies, a prediction algorithm model plays a special role that is to map a group of physical signals to a glucose level of a person at a given time. Unfortunately, there is no practical solution available to the public, under the circumstances of different skin color, skin thickness, physiological differences, etc. In this paper, a Transfer and Collaborative Learning (TCL) method for personalized NGM modeling is proposed, and an Artificial Neural Network Model with TCL (ANN-TCL) is established for predicting glucose concentration, in which model parameters can be tuned according to individual physiological conditions. To verify performance of the proposal, it is embedded into our developed NGM system, and compared with alternate solutions. Clinical trials in the PLA NAVY General Hospital demonstrate that this proposal can reduce the impact of individual discrepancies (IDs), and achieve expected results (R^2=0.82) for all test subjects. Obviously, it is helpful to each patient with our NGM system for the purpose of glucose self-monitoring. The unique gain benefits from a special design of the self-learning strategy which can fully take advantage of both universal and personal information.

Index Terms— Non-invasive, Blood Glucose, Transfer Learning, Artificial Neural Network, Clinical Trials

1. INTRODUCTION

Blood glucose monitoring technology is necessary for people with diabetes mellitus to control their Blood Glucose Concentration (BGC). The glucose monitoring devices currently available in markets are either invasive or mini-invasive, which requires pricking the skin to collect blood samples for tests. These measurement methods not only bring pain to the patients but also put users at infection risks due to skin-cuts. Both invasive and mini-invasive glucose measurement methods are not suitable for daily routine monitoring. For these reasons, the research of Non-invasive Glucose Monitoring (NGM) has become one of the hottest topics in healthcare.

Although many kinds of technologies have been proposed, there is still no methods that meet clinical requirement [1]. The most commonly used measurement principle is based on optics, e.g. absorption spectroscopy [2, 3, 4, 5], Raman spectroscopy[6], optical coherence tomography [7, 8, 9] and fluorescence [10, 11]. Regardless of which theories based on, it is necessary to establish a blood glucose prediction model, which maps original optical signals to BGC.

In NGM modeling, the most common models are configured based on Partial Least Squares Regression (PLSR) [12, 3, 2], Support Vector Regression (SVR) [13, 14], and/or Artificial Neural Networks (ANN) [15, 16]. E.g., Ashok et al. [16] presented a supervised three-layer ANN for estimating BGC based on reflected laser beam from the index finger. In most studies, the blood glucose prediction model is trained on a full dataset, which contains data onto all subjects. This type of model does not consider the impact of Individual Discrepancies (IDs) on signals, resulting in a lower accuracy for some individuals. There are many IDs in NGM, e.g., thickness and color of skin tissue, oxygen, body temperature, etc. To overcome IDs, researchers have tried a variety of methods [2, 17]. However, different measurement methods need different ways to overcome IDs. Therefore it is complicated to propose a universal solution, which needs further research.

This study makes the following contributions. A system that can gather light absorbance on human earlobes is developed, and clinical trials obtain a clinical dataset that contains 4012 samples. We proposed a "Transfer and Collaborative Learning method for personalized Noninvasive Blood Glucose Measurement Modeling" that can overcome IDs. Compared with alternate solutions, our proposal shows the best performance in both accuracy and universality.

*Correspondence to *Ping Wang (pwang@pku.edu.cn)* and *Anpeng Huang (hapku@pku.edu.cn)*. This work is supported in part by National Key R&D Program of China No.2016YFB0800603, No.2017YFB1200700, and National Natural Science Foundation of China No.61701007.

2. SIGNALS AND DATASET

In order to collect data for modeling, a device, called Earlight, that can gather earlobe transmitted light intensity at four different NIR wavelengths (940nm, 1450nm, 1550nm and 1710nm) is developed, as shown in Fig. 1 (**a**) and (**b**).

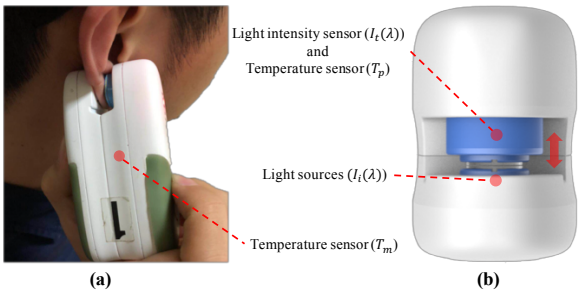

Fig. 1. Our self-developed Earlight system: (a) Using Earligt to measure an earlobe; (b) Probe of the Earlight.

Earlight's probe acts like a clip, which can apply different pressure on the earlobe to change the thickness of it, and simultaneously measure transmitted light intensity signals $I_t(\lambda, t)$, earlobe temperature $T_p(t)$ and machine temperature $T_m(t)$. As the thickness of the earlobe changes, the measured $I_t(\lambda)$ also changes. An example of the original $I_t(\lambda, t)$, $T_p(t)$ and $T_m(t)$ signals measured during 30s is illustrated with Fig. 2 (**a**) and (**b**). The sample interval is 40ms.

We conducted clinical trials at the PLA Navy General Hospital from December 17, 2017, to February 3, 2018. While nurses are measuring BGC for a patient by an invasive glucometer, we use the Earlight to record the optical and temperature signals and save these signals in a database. All of the clinical trials were done under the guidance and supervision of the PLA Navy General Hospital and following the principles of the Declaration of Helsinki [18]. Each patient measured his/her BGC seven times a day. The dataset obtained at this hospital is called the NAVY.

There are a total of 4012 records from 89 patients in the NAVY dataset, and these 89 patients are indexed by PID (#1 $\sim$ #89). To conduct universal and personal training, validating and testing for NGM models, we split the NAVY dataset into four sub-datasets based on patient's PID, as shown in Table 1. The personal training dataset is composed of the first half (by acquisition time) of the data from #51 $\sim$ #89 patient, while the personal testing dataset contains the other half. The two sub-datasets have no intersection.

Table 1. Statistics of different subdatasets in the NAVY.

Dataset	PID	count	min	max	mean
All	#1∼#89	4012	3.1	27.8	10.85
Universal Training	#1∼#40	2253	3.1	27.8	10.68
Universal Validating	#41∼#50	302	3.1	27.6	11.09
Personal Training	#51∼#89	727	3.1	27.4	10.96
Personal Testing	#51∼#89	730	3.1	27.2	10.80

3. MODEL

In this section, a model for estimating BGC from Earlight signals is developed. The overall framework of the non-invasive BGC prediction model is shown in Fig. 2.

3.1. Features Construction

In this paper, the features are selected from the Earlight original signals according to the Beer-Lambert law [19], which is adopted for measuring the concentrations of components in a medium by optical methods. First, the transmitted light intensity signals are divided into Direct Current (DC) signals ($f < 0.5Hz$) and Alternating Current (AC) signals ($1Hz < f < 2Hz$) by Fast Fourier Transform (FFT) [20] and inverse Fourier Transform (iFFT)[20]. Signals with other frequency are filtered as noise. In this paper, the DC signal is regarded as a Pressure Difference Signals (PDS), and the AC signal corresponding to the low phase of the DC signals is concerned as a Pulse Wave Signals (PWS), which is caused by the heart rate. PDS and PWS are shown in Fig. 2(**c**) and (**d**).

3.1.1. *Features from PDS*

In PDS, four types of absorbance-related features (mean Low pressure Absorbance $mLpA$, mean High pressure Absorbance $mHpA$, Middle pressure Absorbance MpA and mean Relative pressure Absorbance $mRpA$) are defined as (1) $\sim$ (4).

$$mLpA(\lambda) = \frac{1}{N_l} \sum_{t \in T_l} log(\frac{I_i(\lambda)}{I_t^{pds}(\lambda, t)}), \tag{1}$$

$$mHpA(\lambda) = \frac{1}{N_h} \sum_{t \in T_h} log(\frac{I_i(\lambda)}{I_t^{pds}(\lambda, t)}), \tag{2}$$

$$MpA(\lambda) = log(\frac{I_i(\lambda)}{I_t^{pds}(\lambda, t_m)}), \tag{3}$$

$$mRpA(\lambda) = \frac{1}{N_l \times N_h} \sum_{t_l \in T_l, t_h \in T_h} log(\frac{I_t^{pds}(\lambda, t_h)}{I_t^{pds}(\lambda, t_l)}). \tag{4}$$

In (1) $\sim$ (4), $I_i(\lambda)$ and $I_t^{pds}(\lambda, t)$ indicate incident and transmitted light intensity in PDS. T_l, T_h and t_m, represent the time sets of low and high pressure phases, and intermediate time of the rising phase, respectively. N_l and N_h represent the number of samples in T_l and T_h phases.

3.1.2. *Features from PWS*

In PWS, three types of absorbance-related features (mean Expend vessel Absorbance $mEvA$, mean Constrict vessel Absorbance $mCvA$ and mean Relative vessel Absorbance $mRvA$), are defined as (5), (6) and (7).

$$mEvA(\lambda) = \frac{1}{N_e} \sum_{t \in T_e} log[\frac{I_i(\lambda)}{I_t^{pws}(\lambda, t)}], \tag{5}$$

$$mCvA(\lambda) = \frac{1}{N_c} \sum_{t \in T_c} log[\frac{I_i(\lambda)}{I_t^{pws}(\lambda, t)}], \tag{6}$$

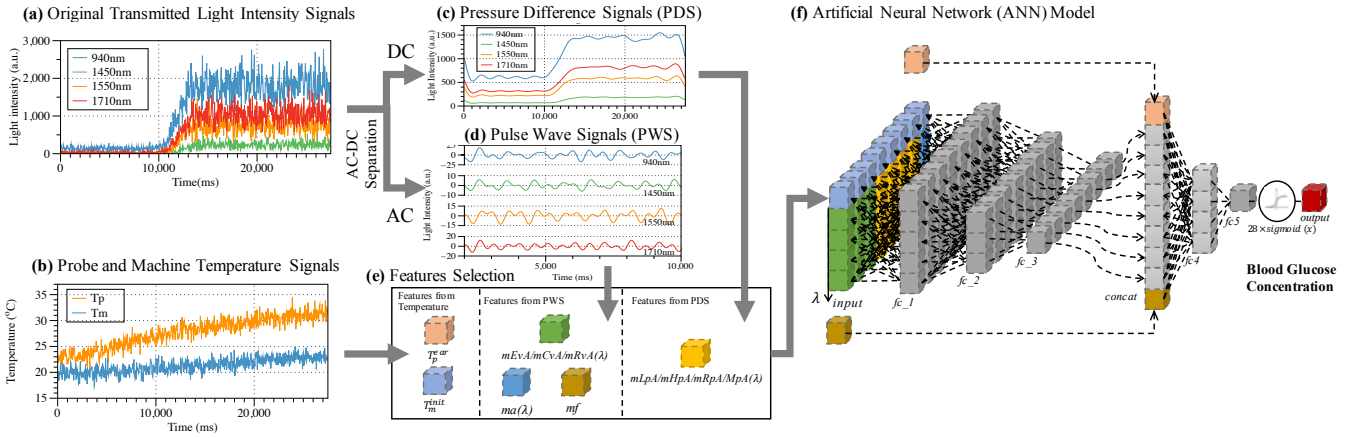

Fig. 2. The overall framework of non-invasive BGC prediction model: (**a**) Original transmitted light intensity signals ($I_t(\lambda, t)$); (**b**) Original temperature signals ($T_p(t)$ and $T_m(t)$); (**c**) Pressure Difference Signals (PDS) taken from the DC signals; (**d**) Pulse Wave Signals taken from the AC signals; (**e**) Features selected from PDS, PWS and temperature signals; (**f**) The ANN model structure we designed for BGC estimation.

$$mRvA(\lambda) = \frac{1}{N_{ec}} \sum_{j=1}^{N_{ec}} log[\frac{I_t^{pws}(\lambda, t_{c_j})}{I_t^{pws}(\lambda, t_{e_j})}]. \quad (7)$$

In (5) $\sim$ (7), $I_t^{pws}(\lambda, t)$ indicates transmitted light intensity in PWS. $T_e = \{t_{e_i} | i = 1, 2, ..., N_e\}$ and $T_c = \{t_{c_i} | i = 1, 2, ..., N_c\}$, which represent the time points of blood vessels expending and constricting. N_e and N_c represent the number of samples in T_e and T_c. $N_{ec} = min(N_e, N_c)$.

In addition to absorbance-related features, some features related to physiology can also be deduced from PWS. Here, the mean amplitude (ma) and mean frequency (mf) of PWS are extracted, which represent the average of amplitude in PWS and approximate value to human heart rate.

3.1.3. Features from Temperature Signals

In order to reduce the influence of temperature on BGC estimation, some features related to the earlobe and the circuit temperature should be extracted as a input of the prediction model. T_p and T_m signals are obtained by the Earlight, however, the 30s measurement time is not enough for converging T_p to the earlobe temperature, as shown in Fig. 2(b). We use $T_p(t) = k_p e^{-\tau_p t} + \beta_p$ and $T_m(t) = k_m t + \beta_m$ to fit $T_p(t)$ and $T_m(t)$. The temperature corresponding to the first intersection of $T_p(t)$ and $T_m(t)$ can be considered as the initial circuit temperature T_m^{init}, and the second intersection can be taken as earlobe temperature T_p^{ear}.

3.2. ANN Model Structure

To estimate BGC value from these features, we have tried many models including PLSR [12], SVR [13] and ANN [15], and found that ANN works best. After experimenting with a variety of ANN structures, the structures shown in Fig. 2(**f**) presents the best BGC prediction performance.

There are one input layer, six hidden layers and one output layer in the ANN. T_p^{ear} and mf are feed directly to $concat$ layer as physiological background information, because there is a clear high-level relationship among body temperature, heart rate, and BGC. In the first three hidden layers (fc_1, fc_2 and fc_3), features (neurons) with the same type are united through fully-connected networks, and there is no connection between different types of features (neurons). The activate function used by fully-connected layers is ReLU [21]. Since the blood glucose concentration (BGC) ranges from 0 to 28 mmol/L, a $28 \times sigmoid(x)$ activate function is applied to limit the output.

Table 2. Testing result (MSE / R^2) of different models on the NAVY personal testing dataset.

PID	Universal Model			Personalized Model	
	PLS[12]	SVR[13]	ANN	ANN-TL	ANN-TCL
ALL	25.9/0.48	21.8/0.54	13.9/0.66	14.5/0.77	**9.7/0.82**
#54	52.8/0.48	30.7/0.63	14.0/0.71	21.3/0.79	**15.4/0.83**
#57	17.6/0.39	10.9/0.79	12.2/0.79	**11.1/0.91**	9.7/0.89
#58	153.4/0.50	108.4/0.17	36.7/0.46	28.4/0.61	**20.5/0.74**
#59	40.7/0.48	26.3/0.38	10.7/0.64	**15.8/0.86**	10.2/0.79
#61	16.7/0.35	9.5/0.69	9.9/0.61	8.0/0.87	**7.5/0.88**
#64	8.7/0.84	7.6/0.86	7.6/0.85	**4.3/0.95**	5.6/0.92
#65	8.1/0.67	6.7/0.83	6.6/0.71	**5.0/0.93**	7.4/0.85
#72	25.9/0.43	71.6/0.19	33.5/0.54	21.4/0.66	**18.4/0.79**

3.3. Collaborative Learning Method

Before we propose the Collaborative Learning (CL) method, we have tried to train the ANN model directly using the universal training dataset by mean squared error (MSE) loss function and ADAM algorithm [22], and test it on the personal testing set. However, although our ANN model is superior to the PLS and SVR models in both MSE and R^2, as shown in Table. 2, it shows a lower accuracy in some

individuals, e.g., #58 and #72. The reason for these poor performances in some individuals is that the IDs between these patients and others. If the patients in the testing dataset differ significantly from those in the training set, the model is not suitable for him/her.

To solve this problem, we put forward an idea of personalized training, which means, for different individuals, each has a set of individual-customized model parameters w_i (i indicates the ith patient). The information learned by a model from the data should be divided into two types. The first type is Common Information (CI) that characterizes physical laws, e.g., the absorption coefficient of glucose molecules, the refractive index of water, etc. The second type is Personal Information (PI) that characterizes the individual's properties, e.g., the skin thickness and colors, etc.

Based on the above ponder, we decided to modify the MSE loss function so that it can learn both CI and PI at the same time during model training. Assuming there are M patients, each patient has his/her own model parameters $w_i (i = 1, 2, ...M)$. Each w_i is an L-dimensional vector. We define $PD(w_1, w_2, ..., w_M)$ to describe the difference among these these personal model parameters, as (8),

$$PD(w_1, w_2, ..., w_M) = \frac{\sum_{i=1}^{L} \sum_{j=1}^{M} (w_j(i) - \overline{w}(i))^2}{M \times L}, \tag{8}$$

where $\overline{w}(i)$ ($\overline{w}(i) = \frac{1}{M} \sum_{M}^{j=1} w_j(i)$) represent the mean of the ith model parameters across all patients. The greater the difference among w_i, the greater the PD. With PD, we propose the TCL loss function as (9),

$$\begin{aligned} Loss(w_1, w_2, ..., w_M) = & \frac{1}{M} \sum_{i=1}^{M} MSE(ANN(f(i), w_i), bgc(i)) \\ & + \alpha ||w||^2 \\ & + \beta PD(w_1, w_2, ..., w_M), \end{aligned} \tag{9}$$

where β is a coefficient as α. Compared with the MSE loss function, there is a PD item, which makes the difference among everyone's personal model parameters w_i not be too large. This means the model must take CI into account while learning PI. Since the new loss function in (9) requires multiple peoples's models to be trained at the same time, we call this training method Collaborative Learning (CL).

4. EXPERIMENT AND CONCLUSION

We use CL to train our ANN model on the personal training dataset but found that there is too little data to make the model converge. To speed up the model training, we use the method of transfer learning (TL) [23, 24], which means that we first use the classic training method to get an initial model parameter value w_{init} on the universal training dataset, and then use CL to carry personalized model training from w_{init} for those patient in the personal training dataset. The training method combined with TL and CL is called Transfer and Collaborative Learning (TCL) in this paper. These experiments are implemented on a Python platform with Tensorflow.

The performance of our ANN model trained with TCL (ANN-TCL) is tested on the personal testing dataset, compared with other universal models (PLS, SVR, and ANN without TCL) and a personalized model (ANN-TL, only trained with TL). Testing results are shown in Table. 2 and Fig. 3, which demonstrated that our ANN-TCL is superior to other models in both R^2 and EGA [1] . It could be noted that, for the ANN-TCL model, there is no case where some individuals perform poorly, which indicates that TCL can overcome IDs to some extent.

Overall, this paper proposed a transfer and collaborative learning method for noninvasive blood glucose measurement, and experiments proved that it could reduce the impact of IDs.

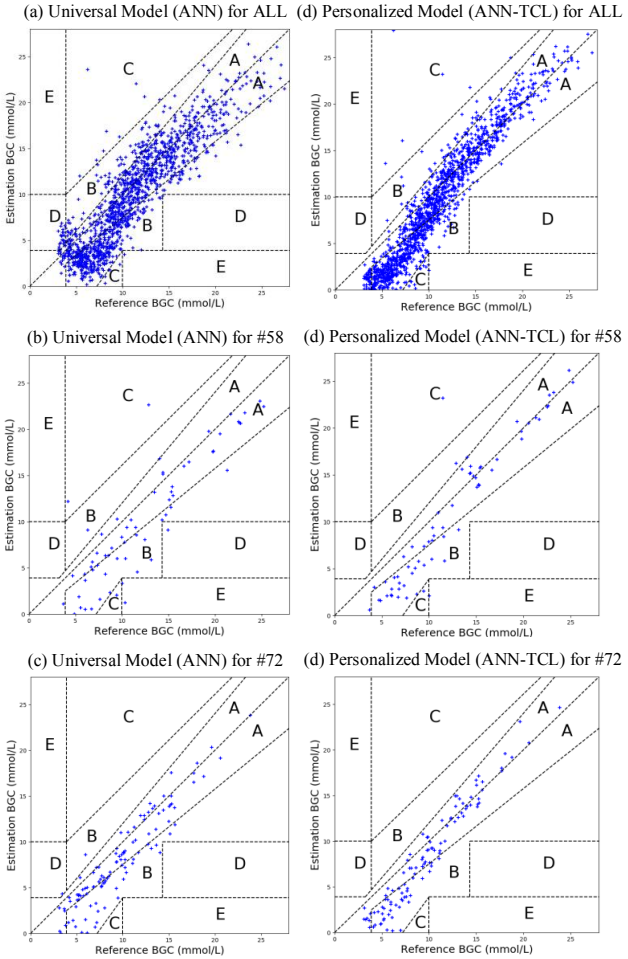

Fig. 3. Clarke Error Grid Analysis (EGA) [25] of universal model (ANN) and personalized model (ANN-TCL). For the #58 and #72 patients, the performance of glucose level prediction is significantly improved.

[1] EGA is the "gold standards" for determining the accuracy of glucose meters, and the grid breaks down a scatterplot of a reference and an estimated BGC value into five regions: A and B are medically acceptable. C, D, and E would confuse treatment. The more points that fall in the region A, the better.

5. REFERENCES

[1] Yulia Mayzel Keren Horman Tamar Lin, Avner Gal and Karnit Bahartan., "Non-invasive glucose monitoring: A review of challenges and recent advances," *Engineering and Biosciences*, vol. 6, no. 5, 2017.

[2] Jingying Jiang, Xiaolin Min, Da Zou, and Kexin Xu, "Mathematical modeling on experimental protocol of glucose adjustment for non-invasive blood glucose sensing," pp. 413–423, 2012.

[3] H. M. Heise, R Marbach, T Koschinsky, and F. A. Gries, "Noninvasive blood glucose sensors based on near-infrared spectroscopy.," *Artificial Organs*, vol. 18, no. 6, pp. 439–447, 2010.

[4] David D. Cunningham and Julie A. Stenken, *Near-Infrared Spectroscopy for Noninvasive Glucose Sensing*, John Wiley and Sons, Inc., 2009.

[5] Md. Koushik Chowdhury, Anuj Srivastava, Neeraj Sharma, and Shiru Sharma, "Noninvasive blood glucose measurement utilizing a newly designed system based on modulated ultrasound and infrared light," *International Journal of Diabetes in Developing Countries*, vol. 36, no. 4, pp. 439–448, 2016.

[6] L. I. Yan, Ling Jing Meng, Jun Xian Liu, Yu Zheng, Chun Sheng Song, and Hui Lu Yao, "Gelatinization process of single rice starch particle using laser tweezers raman spectroscopy system," *Acta Photonica Sinica*, vol. 39, no. 7, pp. 1324–1329, 2010.

[7] Hafeez Ullah, Fayyaz Hussain, and Masroor Ikram, "Optical coherence tomography for glucose monitoring in blood," *Applied Physics B*, vol. 120, no. 2, pp. 355–366, 2015.

[8] Kirill V. Larin, Irina V. Larina, Irina Y. Petrova, Massoud Motamedi, and Rinat O. Esenaliev, "Optical coherence tomography and noninvasive blood glucose monitoring: a review," *Proc Spie*, vol. 5, no. 29, pp. 285–290, 2015.

[9] R. A. Gabbay and S Sivarajah, "Optical coherence tomography-based continuous non-invasive glucose monitoring in patients with diabetes," *Diabetes Technology and Therapeutics*, vol. 10, no. 3, pp. 188–93, 2008.

[10] F Hussain, "Fluorescence-based glucose sensing.," *Biosensors and Bioelectronics*, vol. 20, no. 12, pp. 2555–2565, 2005.

[11] University of Leeds, "Noninvasive device could end daily finger pricking for people with diabetes," *Medical press*, July 15, 2015.

[12] Paul Geladi and Bruce R. Kowalski, "Partial least-squares regression: a tutorial," *Analytica Chimica Acta*, vol. 185, no. 86, pp. 1–17, 1985.

[13] Mariette Awad and Rahul Khanna, "Support vector regression," *Neural Information Processing Letters and Reviews*, vol. 11, no. 10, pp. 203–224, 2007.

[14] Sarul Malik, Rajesh Khadgawat, Sneh Anand, and Shalini Gupta, "Non-invasive detection of fasting blood glucose level via electrochemical measurement of saliva," *Springerplus*, vol. 5, no. 1, pp. 701, 2016.

[15] Simon Haykin, "Neural networks: A comprehensive foundation," *Neural Networks A Comprehensive Foundation*, pp. 71–80, 1994.

[16] V. Ashok, S Rajan Singh, and A. Nirmalkumar, "Determination of blood glucose concentration by back propagation neural network," *Indian Journal of Science and Technology*, vol. 3, no. 8, 2010.

[17] L. I. Gang, Wang Yan, Li Qiu Xia, Lin Ling, Li Xiao Xia, and Liu Yu Liang, "Theoretic study on improving noninvasive measurement accuracy of blood component by dynamic spectrum method," *Journal of Infrared and Millimeter Waves*, vol. 25, no. 5, pp. 345–348, 2006.

[18] Word Medical Association, "Declaration of helsinki: Ethical principles for medical research involving human subjects," *JAMA*, vol. 310, no. 20, pp. 2191–2194, 2013.

[19] Dean Calloway, "Beer-lambert law," *Journal of Chemical Education*, vol. 74, no. 7, pp. 744, 1997.

[20] Sen M Kuo and Bob H Lee, *Fast Fourier Transform and Its Applications*, Prentice Hall, 1988.

[21] Vinod Nair and Geoffrey E Hinton, "Rectified linear units improve restricted boltzmann machines," *International Conference on Machine Learning (ICML)*, pp. 807–814, 2010.

[22] Diederik P. Kingma and Jimmy Lei Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations (ICLR)*, 2015.

[23] Sinno Jialin Pan and Qiang Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[24] Karl Weiss, Taghi M. Khoshgoftaar, and Ding Ding Wang, "A survey of transfer learning," *Journal of Big Data*, vol. 3, no. 1, pp. 9, 2016.

[25] W. L. Clarke, D Cox, L. A. Gonder-Frederick, W Carter, and S. L. Pohl, "Evaluating clinical accuracy of systems for self-monitoring of blood glucose," *Diabetes Care*, vol. 23, no. 11, pp. 1711–1712, 2000.