Diabetes Detection Using Carbon Nanomaterial Coated QCM Gas Sensors and a Convolutional Neural Network through Urine Sample

Misbah[1]   **Muhammad Rivai[1]***   **Fredy Kurniawan[2]**

[1]Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia
[2]Department of Chemistry, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia
* Corresponding author's Email: muhammad_rivai@ee.its.ac.id

Abstract: Diabetes is a disease that the entire world is afraid of. Early detection measures are required in this instance. Diabetes diagnosis by blood samples has the disadvantage of being uncomfortable. This study uses a urine sample to estimate a person's risk of developing diabetes. The electronic nose is a system that has the potential to recognize diabetes in this way. This method consists of quartz crystal microbalance gas sensors coated with carbon nanomaterials, including single-walled carbon nanotubes, double-walled carbon nanotubes, multi-walled carbon nanotubes, and graphene oxide. A reciprocal counter implemented in a field programmable gate array (FPGA) device is used to measure the frequency shift on the sensor. The convolutional neural network technique is used to detect diabetes. The results of the experiments suggest that this system can distinguish between healthy and diabetic people with an accuracy of 91%.

Keywords: Carbon nanotubes, Convolutional neural network, Diabetes, Diseases, Gas sensor.

NOMENCLATURE

QCM side

Δf	Resonant frequency shift
f	Base frequency
ρ	Crystal density
μ	Shear modulus of quartz crystal
Δm	Change in mass per unit area

CNN side

$g(\tau)$	Original signal
$h(t-\tau)$	A function shifted to the original signal
x	Input value
y	The output, which has a value between 0 to 1
w_t	Weight value in the current iteration
w_{t-1}	Weight value in the previous iteration
η	Learning rate
η'	Updated learning rate
α	Summation of gradient square
ε	A tiny number used to avoid division zero value
$\partial L/\partial w$	The gradient of the loss function

1. Introduction

Recently, diabetes has become a severe health problem. This disease can cause various complications such as cardiovascular, kidney disease, and others [1]. The cause of this disease is a decrease in insulin production. Therefore, the glucose in the blood becomes high. Kidneys can also filter glucose in the blood, but only in specific amounts. If it exceeds its capacity, the kidneys will excrete it through the urine, as shown in Fig. 1.

Currently, clinical diagnosis relies on the examination of glucose in the blood. This diagnosis is considered fast and accurate. However, the possibility of infection and pain during blood sampling is problematic. Many attempts have been made to improve the sensitivity of glucose detection.

Biosensors have attracted much attention because of their low cost, simplicity, and practicality [3]. However, natural enzymes still have some intrinsic disadvantages, such as easy to change properties,
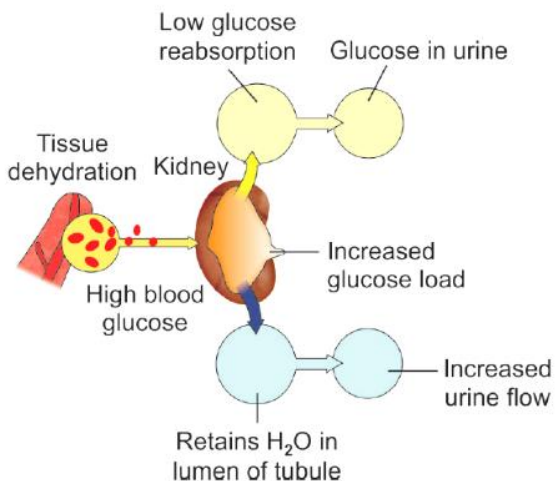
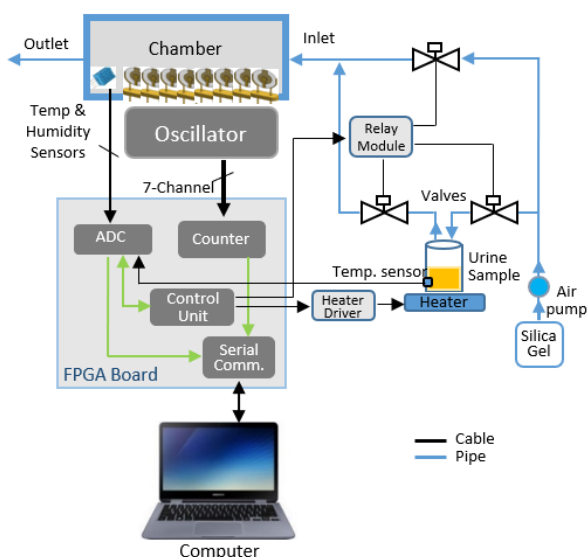Figure. 1 Process of producing glucose in the urine [2]
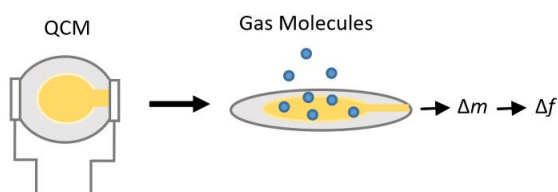

Figure. 2 Electronic nose system block diagram

Figure. 3 The basic principle of the QCM sensor [17]

high purification costs, and difficulties in preparation and storage. Several sensor materials have unique peroxidase-like catalytic activity and are applied to fabricate hydrogen peroxide and glucose sensors. These materials are based on metal oxide nanomaterials of Fe_3O_4 nanospheres [4], Fe-Pd/rGO [5], $NiCo_2O_4$ [6], graphene [7], and commercial metal oxide [8]. Metal oxide materials have low sensitivity, selectivity, and energy consumption disadvantages. Furthermore, nanomaterials for manufacturing glucose sensors are based on precious metals such as

PtAu and carbon-based nanomaterials such as carbon nanotubes [9]. All nanomaterials are used to improve sensor capabilities. Carbon nanomaterials have the advantages of being soluble in water, having a fast response, very light, large surface area, and high sensitivity. In addition, glucose can be detected with the lowest limit using localized surface plasmon resonance (LSPR) optical fiber [10, 11] and spectrometer [12]. LSPR sensors and spectrometers are optical sensors with expensive equipment, complex systems and are not portable. Electrode sensors are used in detecting glucose [13]. This sensor has a large size and limited sensitivity. Other sensors, such as quartz crystal microbalance (QCM), are also used in glucose detection [14]. This sensor has the advantage of being sensitive to changes in gas mass and is stable.

Diabetes identification through urine has been carried out in previous studies with several machine learning methods, including k-nearest neighbors (k-NN), support vector machine (SVM), and PCA-logistic regression (PCA-LR), ANN, and convolutional neural network (CNN) [8, 15, 16].

The use of QCM sensors and non-enzyme materials to increase sensor sensitivity needs to be developed. In addition, the method of identifying diabetes with pattern recognition algorithms still needs to be improved. Therefore, this study developed an electronic nose system using a QCM gas sensor array coated with carbon nanomaterial to diagnose diabetes from urine samples. Several methods were used to identify the disease, including Naïve Bayes, k-NN, SVM, PCA-LR, and CNN. This study contributes to applying electronic nose systems to analyze conditions through urine samples.

This study has several sections, including section 2, which describes the research method used. Section 3 discusses the results of the system evaluation. Section 4 presents the conclusions of the research that has been done.

2. Methods

The system design in this study consists of a clean air pump including silica gel, a urine sample container, three valves, a sensor chamber in which there are seven QCM sensors, an oscillator circuit, a heater driver, a field programmable gate array (FPGA) board, and a computer as shown in Fig. 2. On the FPGA board there is a frequency counter, control unit, analog to digital converter (ADC), and serial communication. The on/off control method adjusts the temperature in the sample container.

2.1 Quartz crystal microbalance (QCM)

The QCM sensor is highly sensitive to gas molecules attached to the electrode, as shown in Fig. 3. This sensor will change the oscillation frequency of the quartz resonator (Δf), which is defined as:

$$\Delta f = -\left(\frac{2f^2}{\sqrt{\rho\,\mu}}\right)\Delta m \qquad (1)$$

where f is the base frequency (Hz), ρ is the density of quartz crystal (g cm^{-3}), μ is the shear modulus of quartz crystal (g cm^{-1} s^{-2}), Δm is the change in mass per unit area (ng cm^{-2}) [18].

2.2 Carbon nanotubes (CNT)

CNTs can be described as sheets of graphite rolled into a tube. This graphite sheet consists of hexagonal sp2 carbon atoms. This structure is formed from hexagonal carbon chains with one or many layers. The diameter of the tube is approximately 1 nm, and the length is on the micro-scale. Single-walled carbon nanotubes (SWCNT) CNTs have a tensile strength of 50-100 GPa and Young's modulus of 1-2 TPa. CNTs of the multi-walled carbon nanotubes (MWCNT) have a tensile strength of 11-63 GPa and Young's modulus of 270-950 GPa. From these characteristics, several studies make CNTs a supercapacitor material [19], electrochemical sensors [20], and gas sensors [21]. This study used six CNT and graphene oxide (GO) materials along with their types and sizes, as shown in Table 1.

The sensor coating solution was obtained by mixing carbon nanomaterials, polyvinylpyrrolidone (PVP), and demineralized water. PVP is a chemical substance that functions to dissolve carbon material into water. After that, ultrasonication was carried out for 1 hour so that the solution was mixed entirely. Then, the coating process was carried out on the sensor using an ultrasonic atomizer with an operating frequency of 110 kHz, as shown in Fig. 4. By noting a drop in its working frequency; the coating was applied to both sides of the QCM surface. The decrease in frequency on QCM is monitored via a computer. The operating frequency was reduced not to exceed 10,000 Hz, allowing QCM to operate reliably.

2.3 Oscillator circuit

An oscillator circuit is needed to generate a vibration signal on the QCM sensor. The oscillator circuit often used in digital applications is the Pierce oscillator type, as shown in Fig. 5. The QCM sensor

Table 1. Types of carbon nanomaterials

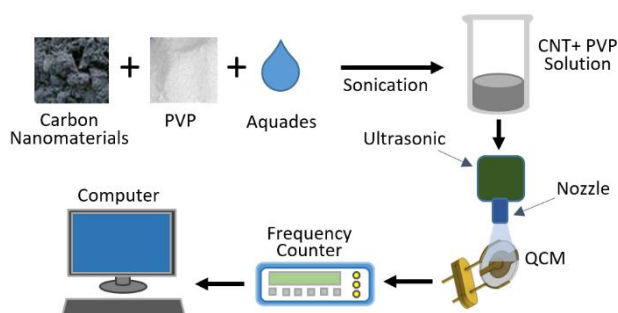| Nanomaterials	Diameter (nm)	Length (μm)	Purity (%)
SWCNT (SW)	1 - 2	5 - 30	95
DWCNT (DW)	2 - 4	50	60
MWCNT (M0406)	4 - 6	0.5 - 2	98
MWCNT (M0515)	5 - 15	0.5 - 2	95
MWCNT (M1030)	10 - 30	0.5 - 2	98
MWCNT (M3050)	30 - 50	0.5 - 2	98
GO	-	8 - 15	98

Figure. 4 Coating process on QCM
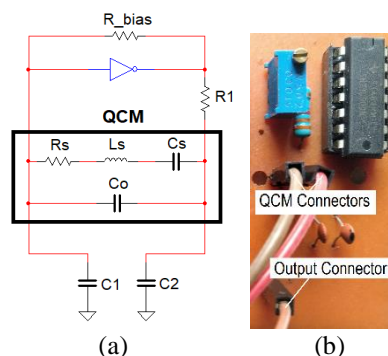

(a)   (b)

Figure. 5 Oscillator circuit: (a) Schematic diagram, and (b) Printed circuit board

used in this study has a vibration frequency of around 10 MHz. The component values in the oscillator circuit include Rbias=2kΩ, R1=3kΩ, and C1=C2= 25pF.

2.4 Field programmable gate array (FPGA)

FPGA offers high resolution and precision in counting pulses and delay times [22]. Therefore, FPGA is suitable for frequency measurement tools utilizing traditional [23] and reciprocal [24] methods. In this study, the oscillator circuit requires a frequency counter system with channels according to the number of QCM sensors. With many frequency
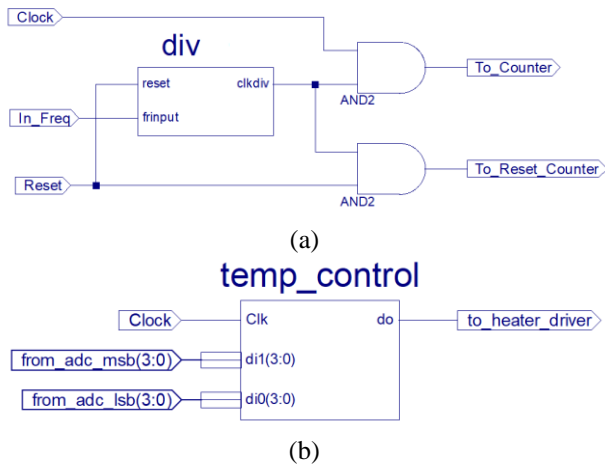
420

(a)

(b)

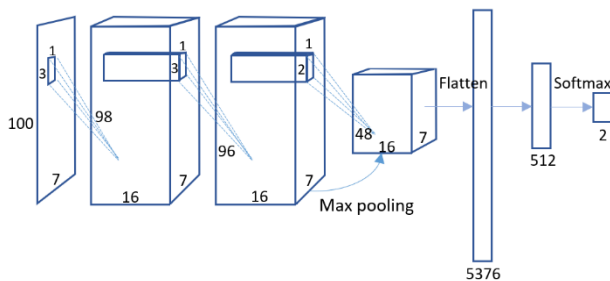Figure. 6 Design on FPGA: (a) Reciprocal counter, and (b) Temperature control

Figure. 7 The CNN architecture

counter systems, using FPGAs is quite effective in implementing frequency counters using the reciprocal method. This system requires a 24-bit counter that corresponds to the sensor frequency. The FPGA module used in this study is Spartan 6 XC6SLX9 made by Xilinx Corp. This module has 9,152 logic cells, 1,430 slices, 5,720 6-input LUTs, 90 distributed RAM, 45 shift registers and 11,440 flip-flops.

The reciprocal counter design consists of div, count, latch, encode, and serial blocks, as shown in Fig. 6 (a). Div serves to divide the input frequency, count to count the number of pulses, latch to lock the counting results, encode to set the data to be sent, and serial to send data to the computer. Meanwhile, Fig. 6 (b) shows the temperature control scheme, with input from the ADC and output to the heater driver.

2.5 Convolutional neural network

CNN is capable of studying features and classifications for raw input data at the same time. The convolution kernel is a filter that can extract features from raw data [25]. The CNN architectural model in this study consists of two layers of convolution, dropout, max-pooling, flattening, and fully connected. The convolution operation using one

dimension for discrete signals is expressed as follows:

$$(g * h)[t] = \sum_{-\infty}^{\infty} g[\tau]h[t - \tau]d\tau \qquad (2)$$

where g is the original signal, and h is a function shifted to the original signal. The convolution process on the data aims to extract features from the input. Convolution produces a linear transformation of the input data based on the spatial information contained in the data. Dropout functions so that the model does not overfit. Layer pooling is a non-linear downsampling technique. Max-pooling is one of the non-linear operations performed during pooling. The principle of max-pooling is to find the highest value of several input values. Rectified linear unit (ReLU) effectively removes negative values from the activation map by replacing zeros. The ReLu function is defined as follows:

$$y = ReLu(x) = \max(x, 0) \qquad (3)$$

where x is the input value, *and y is* the output with a value between 0 and 1. The adaptive gradient optimizer (ADAGRAD) is used during the learning process, has the following formula:

$$w_t = w_{t-1} - \eta'_t \frac{\partial L}{\partial w_{t-1}} \qquad (4)$$

$$\eta'_t = \frac{\eta}{\sqrt{\alpha_t + \varepsilon}} \qquad (5)$$

$$\alpha_t = \sum_{i=1}^{t} \left(\frac{\partial L}{\partial w_{t-1}} \right)^2 \qquad (6)$$

where w_t is the weight value in the current iteration, w_{t-1} is the weight value in the previous iteration, η is the learning rate, η' is the updated learning rate, α is the summation of gradient square, and ε is a tiny number used to avoid division zero value, $\partial L/\partial w$ is the gradient of the loss function. Flatten is used to convert semi-structured data into a relational representation. The cross-entropy loss function is used to determine the model's accuracy. The fully connected layer is a layer that is often used in the use of multi-layer perceptron to modify data dimensions that can be classified linearly. Its activation function is Softmax.

The CNN architecture used in this study consists of 1-D convolution, dropout, max-pooling, flattening, and fully connected, as shown in Fig. 7. This architecture has seven inputs from sensors, each consisting of 100 points. The 100 × 7 data was
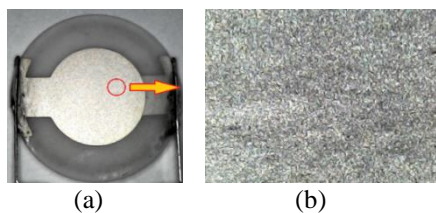
(a)   (b)

Figure. 8 Sensor coating: (a) QCM surface coated with carbon nanomaterial, and (b) The degree of homogeneity
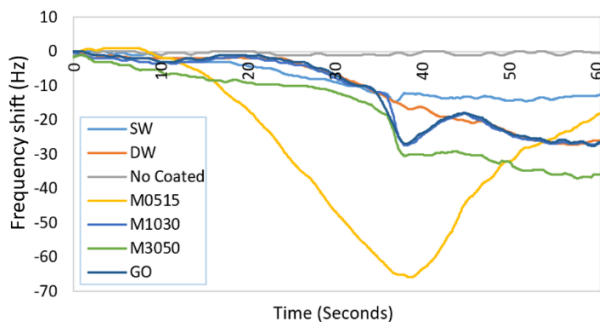

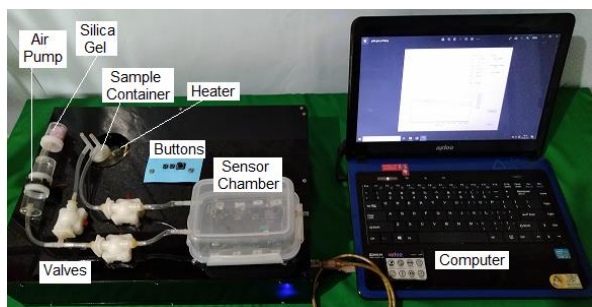Figure. 9 The response of the QCM sensors
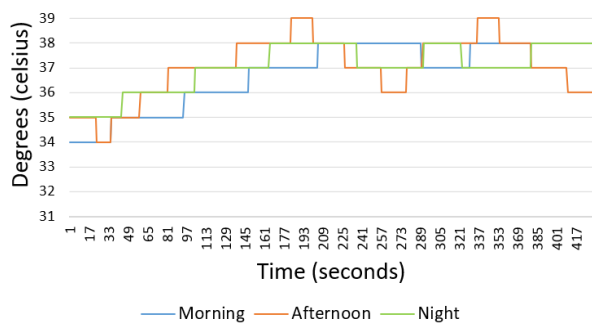

Figure. 10 The experimental equipment

Figure. 11 Temperature control with a setpoint of 38 °C in the sample container
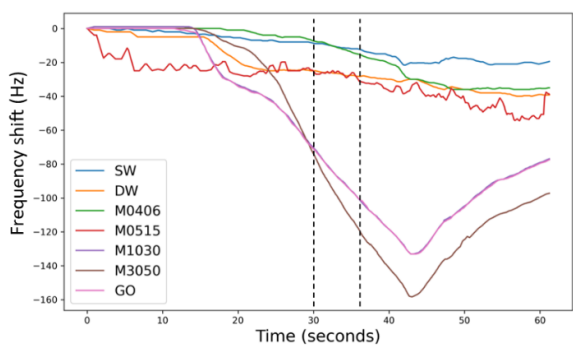

Figure. 12 Sensor responses when exposed to urine vapor

processed twice in convolution with 16 kernels and a filter size of 3×1 to $98 \times 16 \times 7$ and $96 \times 16 \times 7$. The data were reduced using max-pooling 2×1 to $48 \times 16 \times 7$. Then the data was made into one column using flatten to 5376, and from flatten entered the fully connected layer, which consists of 512 neurons in the hidden and two neurons in the output as the number of identification classes.

2.6 Urine samples

Urine samples were taken from 38 subjects at Muhammadiyah Gresik Hospital. There were 21 healthy subjects and 17 diabetic subjects. Subjects consisted of 22 women and 16 men with an age range of 30-60 years. The collection process involves fasting for 8 hours, after which about 10 ml of urine is collected.

3. Results and discussion

The coating of the QCM surface with carbon nanomaterials using an ultrasonic atomizer is shown in Fig. 8.

Fig. 9 shows the difference in response between uncoated and coated sensors. When the urine vapor enters the sensor chamber, the uncoated QCM does not change significantly. Meanwhile, the coated QCM shows a decrease in frequency response.

The experiment was carried out three times, namely in the morning, afternoon, and evening, so there were 114 data with details of 63 healthy and 51 diabetic subjects. The experimental equipment is shown in Fig. 10. The sensor chamber has a volume of 650 ml. The sample was flowed at approximately 50 ml/min for 40 seconds and allowed to stand for up to 30 seconds, for a total of 70 seconds. The sampling frequency on the system is 16 Hz. Therefore, 16 x 70 = 1120 points are obtained for a sensor in one take. The system has seven sensors, so 7 x 1120 = 7840 points are obtained. The temperature in the sample container is controlled with a set point of 38ºC, as shown in Fig. 11.

Fig. 12 shows the sensor response when given urine vapor. The dotted line is the boundary of the points taken from the 30th to the 36.25th second. From these limits, 100 points were collected by finding the average value and normalized, as shown in Table 2. From this normalization, the response patterns between healthy and diabetic subjects were used, as shown in Fig. 13.

Besides that, each data attribute was searched for the mean, standard deviation, variance, and median, as shown in Table 3 and Table 4. Visualization of the attribute values of healthy and diabetic subjects helps

Table 2. The normalized value of the sensor response

Carbon Nanomaterials	Samples	
	Healthy	Diabetes
SW	0.05	0.11
DW	0.08	0.26
M0406	0.01	0.12
M0515	0.35	0.26
M1030	0.73	0.87
M3050	1.00	1.00
GO	0.73	0.87

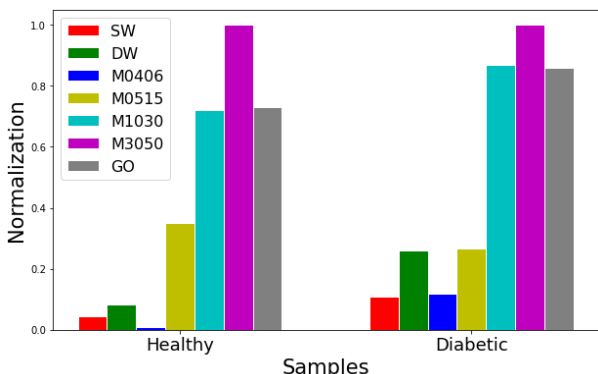
Figure. 13 Diabetic and healthy sample data patterns

Table 3. Attribute statistic description for healthy

Attribute	Mean	Std. Dev.	Variance	Median
SW	10.3	4.34	18.9	10.0
DW	14.3	6.59	43.54	15.0
M0406	12.34	7.6	57.78	11.0
M0515	21.23	8.96	80.3	22.0
M1030	69.66	23.06	531.96	71.0
M3050	94.3	27.7	767.32	97.0
GO	69.88	23.43	549.08	71.0

Table 4. Attribute statistic description for diabetic

Attribute	Mean	Std. Dev.	Variance	Median
SW	12.92	5.3	28.11	13.0
DW	13.43	7.96	63.5	12.0
M0406	12.09	6.34	40.20	10.0
M0515	23.19	10.96	120.19	22.0
M1030	62.52	25.03	626.68	62.0
M3050	92.03	26.46	700.27	99.0
GO	63.0	24.64	549.08	62.0

in providing information, making it easy to identify patterns, as shown in Fig. 14.

In data processing, two treatments were carried out. First, by finding the average value of 100 points for a sensor. Because there are seven sensors, seven values are obtained and used as data in the Naïve Bayes, k-NN, SVM, and PCA-LR methods. Meanwhile, the second treatment comprises 100 points for a sensor, with 700 points for seven sensors.
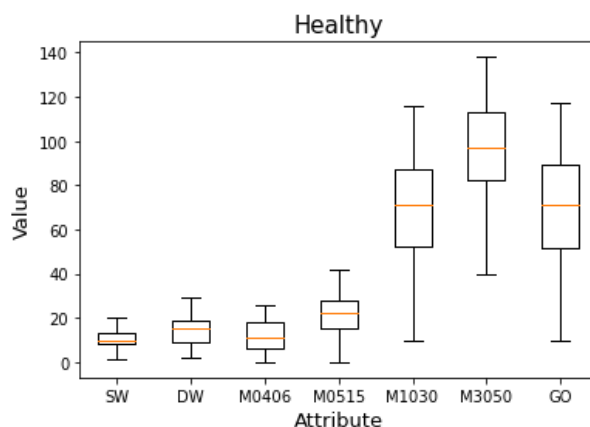
(a)
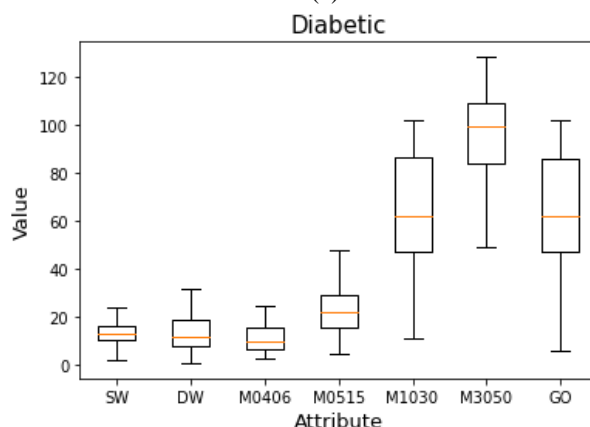
(b)

Figure. 14 Attribute statistics for: (a) healthy, and (b) diabetic subject
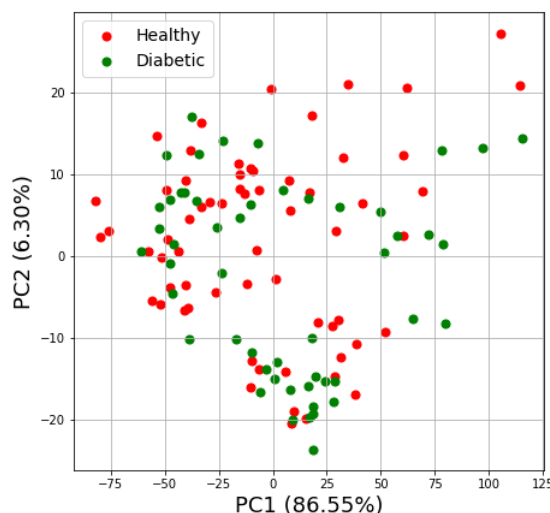

Figure. 15 Distribution of data using PCA

This data is used for the CNN method. The training and test data ratio is 70:30, with details of 80 training data and 34 test data.

Before evaluating several identification methods, it is necessary to review the data distribution so that the data's position for each category is known. The

data totaled 114 samples consisting of healthy and diabetic classes. The principal component analysis (PCA) method was used to identify data distribution in these two categories. This method can reduce the number of variables and visualize the data distribution. Fig. 15 shows that the initial data processed by PCA looks spread out with a variance of 86.55% for PC1 and 6.3% for PC2, so the total variance is 92.85%.

The next step is to evaluate several identification methods, as listed in Table 5. The first method, Naïve Bayes with the Gaussian distribution, has a higher accuracy than the Bernoulli distribution, which is 76.5%. The second method, k-NN, produces the best accuracy at k = 5, which is 71%. The third method, SVM with a radial basis function (RBF) kernel, obtained better accuracy than sigmoid and polynomial kernels, which is 76.5%. The fourth method, PCA-LR, obtained an accuracy of 64.7%. The fifth method, CNN with eight kernels and a size of 5×1, obtained an accuracy of 79.4%. The CNN method is the best candidate as a classifier that can be further optimized.

It is necessary to know the distribution of the data to identify how effective the convolution process is on CNN as the initial data processor in this method. After the data has gone through the convolution process, they start to collect and group together with a 95% variance for PC1 and 2.5% for PC2, as shown in Fig. 16, so that the total variance of the two components is 97.5%. There are two CNN architectural models evaluated in this study. The first model consists of two convolutional layers, two max-pooling layers, dropout, flatten, and fully connected, as shown in Table 6. The second model consists of two convolutional layers, dropout, max-pooling, flatten, and fully connected, as shown in Table 7.

A total of 80 x 700 points is carried out in the training process. To determine the success of the process is measured by training loss. Fig. 17 shows each training loss curve that decreases with increasing epoch. The best weight and bias values will be used to identify the 34 test data. The number of neurons in a fully connected layer during the training process is 512.

Training losses from the second CNN model involving 8 kernels with a size of 5×1 is 0.6, 8 kernels with 9×1 is 0.34, and 16 kernels with 3×1 is 0.33. From the three curves, the training process fluctuates in achieving the smallest losses.

The CNN model is evaluated using two parameters, including the number of kernels and filter size. In these two parameters, namely 8, 16, 32, and 3×1, 5×1, 7×1, 9×1. The results of the first model

Table 5. Evaluation results of several methods

Methods	Accuracy (%)	Precision (%)	Recall (%)
Naïve Bayes (Gaussian)	76	72	94
Naïve Bayes (Bernoulli)	67.6	72	68.4
k-NN (k=3)	61.7	62	79
k-NN (k=5)	71	70	84
k-NN (k=7)	64	67	74
SVM (RBF)	76.5	76.2	84
SVM (Sigmoid)	61.7	62	79
SVM (Polynomial)	64.7	62	95
PCA-LR	64.7	64	84
CNN	79.4	78.9	83
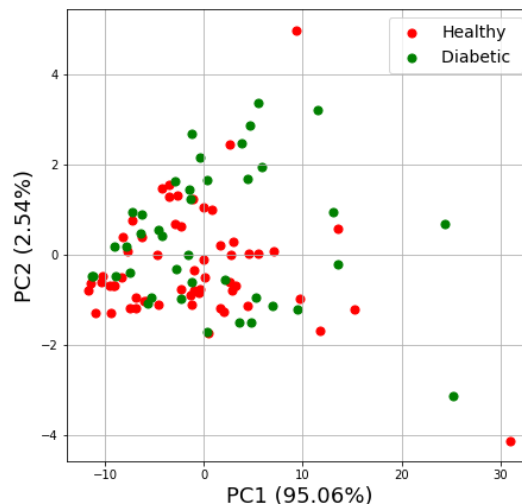
Figure. 16 Distribution of data after convolution

Table 6. Summary of the architecture of the first CNN model

Layer	Type	Output Shape	Filter size	Stride
0-1	*1D Convolution*	98, 32	3 × 1	1 × 1
1-2	*1D Max-pooling*	49, 32	2 × 1	2 × 1
2-3	*1D Convolution*	47, 32	3 × 1	1 × 1
3-4	*1D Max-pooling*	23, 32	2 × 1	2 × 1
4-5	*Dropout*	23, 32	-	-
5-6	*Flatten*	736	-	-
6-7	*Fully connected*	512	-	-
7-8	*Fully connected*	2	-	-

Table 7. Summary of the architecture of the second CNN model

Layer	Type	Output Shape	Filter size	Stride
0-1	*1D Convolution*	98, 16	3 × 1	1 × 1
1-2	*1D Convolution*	96, 16	3 × 1	1 × 1
2-3	*Dropout*	96, 16	-	-
3-4	*1D Max-pooling*	48, 16	2 × 1	2 × 1
4-5	*Flatten*	768	-	-
5-6	*Fully-connected*	512	-	-
6-7	*Fully-connected*	2	-	-

Table 8. CNN architecture evaluation

Number of Kernels	Filter size	Model CNN	
		First	Second
		Accuracy (%)	
8	3×1	73.5	70.5
8	5×1	73.5	79.4
8	7×1	76.4	76.4
8	9×1	70.5	82.3
16	3×1	76.4	91.1
16	5×1	70.5	76.4
16	7×1	70.5	82.3
16	9×1	79.4	73.5
32	3×1	82.3	73.5
32	5×1	67.6	73.5
32	7×1	79.4	76.4
32	9×1	61.7	82.3
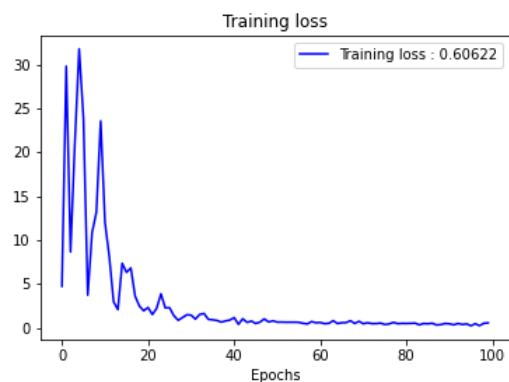
(a)
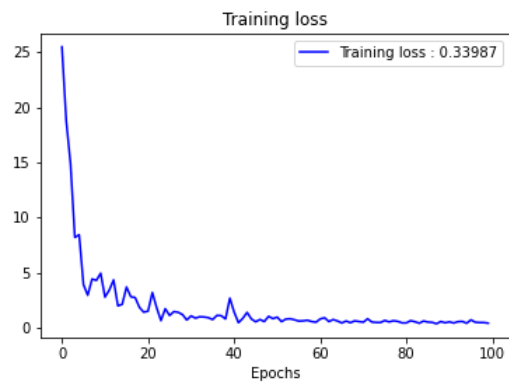
(b)
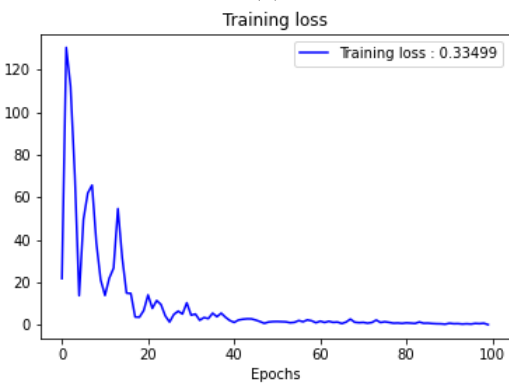
(c)

Figure. 17 The second model training losses curve with the kernel number and size of: (a) 8 (5×1), (b) 8 (9×1), and (c)16 (3×1)

evaluation obtained the lowest accuracy of 61.7% on 32 kernels with a size of 9×1. Meanwhile, the highest accuracy was 82.3% on 32 kernels with a size of 3×1.

In the second model, the evaluation results obtained the lowest accuracy obtained by 8 kernels with a size of 3×1 of 70.5%. The highest accuracy was obtained for 16 kernels with a size of 3×1 of 91.1%, as presented in Table 8. From the evaluation of each model, it can be concluded that the second CNN architectural model is better than the first model. Table 9 shows the results identified by this system compared with the subject's clinical tests. The clinical test measures blood sugar using the ACCU-CHEK tool in units of mg/dL.

The confusion matrix can be used for extra information in evaluating the identification method. This method is usually described as a comparison between actual conditions and predictions. Fig. 18 shows 17 healthy subjects according to actual conditions, and two healthy subjects are in the diabetes category. Then, there are 14 diabetic subjects according to actual conditions and one diabetic subject in the healthy category. For further analysis, the results of CNN identification are divided into true positive rate (TPR), false positive rate (FPR), true negative rate (TNR), and false negative rate (FNR). TPR measures the proportion of correctly identified positives, also called sensitivity.

The TPR value of 0.944 indicates that the evaluation of the CNN method in identifying is good, as shown in Table 10. From evaluating several identification methods, including Naïve Bayes, k-NN, SVM, PCA-LR, and CNN, the best accuracy results were 76%, 71%, 76%, 65%, and 91%, respectively, as shown in Table 11.

Table 9. Diabetic identification results

No	Subject (mg/dL)	Indication	Identification
1	77	Healthy	Healthy
2	97	Healthy	Healthy
3	98	Healthy	Healthy
4	108	Healthy	Healthy
5	111	Healthy	Diabetic
6	117	Healthy	Healthy
7	118	Healthy	Healthy
8	121	Healthy	Healthy
9	121	Healthy	Healthy
10	123	Healthy	Healthy
11	126	Healthy	Healthy
12	129	Healthy	Healthy
13	132	Healthy	Healthy
14	136	Healthy	Healthy
15	141	Healthy	Healthy
16	142	Healthy	Healthy
17	147	Healthy	Diabetic
18	148	Healthy	Healthy
19	156	Healthy	Healthy
20	173	Diabetic	Diabetic
21	193	Diabetic	Diabetic
22	197	Diabetic	Diabetic
23	199	Diabetic	Diabetic
24	208	Diabetic	Diabetic
25	212	Diabetic	Healthy
26	231	Diabetic	Diabetic
27	217	Diabetic	Diabetic
28	252	Diabetic	Diabetic
29	220	Diabetic	Diabetic
30	223	Diabetic	Diabetic
31	234	Diabetic	Diabetic
32	241	Diabetic	Diabetic
33	255	Diabetic	Diabetic
34	342	Diabetic	Diabetic

Figure. 18 Confusion matrix

Table 10. Performance measurement of the CNN model

TP	FP	TN	FN	TPR	FPR	TNR	FNR
17	14	2	1	0,944	0,125	0,875	0,055

Table 11. Comparison of the five pattern recognition methods

Classifier	Accuracy (%)	Precision (%)	Recall (%)
Naïve Bayes	76	72	94
k-NN	71	70	84
SVM	76	76	84
PCA-LR	65	64	84
CNN	91	89	94

Table 12. Comparison of several methods for the classification of diabetes

Ref.	Sensor	Classifier	Accuracy (%)
[26]	Graphene based sensor	Naïve bayes	70
[27]	Blood Glucose, Blood Pressure, Load cell	RB-Bayes	72.9
[28]	Loadcell, blood glucose	MLFF-NN	84.17
[29]	Blood Glucose, Blood Pressure, Load cell	ANN	85.1
[30]	Blood Glucose, Blood Pressure, Load cell	LSTM	87.26
[31]	Metal oxide	GBT	90.4
[32]	ECG	SVM	90.5
This work	QCM + Carbon nanomaterial	CNN	91.1

Several previous studies related to the identification of diabetes have been carried out. The methods included the Naïve Bayes classifier with graphene-based sensors [26], RB-Bayes with blood glucose sensors, blood pressure, and load cells [27], multi-layer feed-forward neural networks (MLFF-NN) that use blood glucose sensors and body weight [28], artificial neural networks (ANN) with blood glucose sensors, blood pressure, and load cells [29], long short-term memory (LSTM) with blood glucose sensors, blood pressure, and load cells [30], gradient boosting tree (GBT) using metal oxide sensors [31],
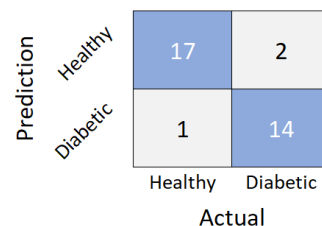
SVM uses a heart rate sensor [32], and CNN uses a carbon nanomaterial coated QCM sensor. The results of the accuracy of these studies are listed in Table 12. From the table, it can be seen that the CNN method has the highest accuracy in detecting diabetes.

4. Conclusion

This study developed a system of carbon nanomaterial-coated QCM sensors coupled with a pattern recognition algorithm to identify diabetes through urine. The carbon nanotubes involved in

experiments include SW-CNT, DW-CNT, MW-CNT, and GO. Some methods are deployed to become the classifier model, including Naïve Bayes, k-NN, SVM, PCA-LR, and CNN. The CNN architecture, which involves 16 kernels with a size of 3×1, has the highest accuracy rate of 91.1% in identifying diabetes. Further studies will develop the system to be more portable with lower power consumption.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

In this research article author contribution are as follows: "Conceptualization and Methodology, Misbah, Muhammad Rivai, and Fredy Kurniawan; Software, Misbah; Data curation, Misbah; Writing-original draft preparation, Misbah; Writing-review and editing by Misbah and Muhammad Rivai; Supervision, Muhammad Rivai and Fredy Kurniawan."

Acknowledgments

References

[1] N. Tofte, M. Lindhardt, K. Adamova, S. J. L. Bakker, J. Beige, J. W. J. Beulens, A. L. Birkenfeld, G. Currie, C. Delles, I. Dimos, L. Francova, M. F. Moller, P. Girman, R. Goke, T. Havrdova, H. J. L. Heerspink, A. Kooy, G. D. Laverman, and H. Mischak, "Early detection of diabetic kidney disease by urinary proteomics and subsequent intervention with spironolactone to delay progression (PRIORITY): a prospective observational study and embedded randomized placebo-controlled trial", *The Lancet*, Vol. 8. Issue 4, pp. 301–312, 2020.

[2] R. Gundamaraju and R. Vemuri, "Pathophysiology of Greedy Colon and Diabetes: Role of Atropine in Worsening of Diabetes", *Euroasian Journal of Hepato-Gastroenterology*, Vol. 4, No. 1, pp. 51–54, 2014.

[3] Y. Pang, Z. Huang, Y. Yang, Y. Long, and H. Zheng, "Colorimetric Detection of Glucose Based on Ficin with Peroxidase-like Activity", *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, Vol. 189, pp. 510–515, 2018.

[4] J. Xu, Y. Sun, and J. Zhang, "Solvothermal synthesis of Fe3O4 nanospheres for high-performance electrochemical non enzymatic glucose sensor", *Scientific Reports*, Vol. 10. No. 1, pp. 1–8, 2020.

[5] C. Yang, W. Feng, Y. Li, X. Tian, Z. Zhou, L. Lu, and Y. Nie, "A Promising Method for Diabetes Early Diagnosis via Sensitive Detection of Urine Glucose by Fe–Pd/rGO", *Dyes and Pigments*, Vol. 164, pp. 20–26, 2019.

[6] H. J. Jo, A. Shit, H. S. Jhon, and S. Y. Park, "Highly sensitive non-enzymatic wireless glucose sensor based on Ni-Co oxide nanoneedle-anchored polymer dots", *Journal of Industrial and Engineering Chemistry*, Vol. 89, pp. 485–493, 2020.

[7] J. Zhu, S. Liu, Z. Hu, X. Zhang, N. Yi, K. Tang, M. G. Dexheimer, X. Lian, Q. Wang, J. Yang, J. Gray, and H. Cheng, "Laser-induced graphene non-enzymatic glucose sensors for on-body measurements", *Biosensors and Bioelectronics*, Vol. 193, pp. 1–12, 2021.

[8] S. F. Gumelar, H. M. Budiyanto, M. F. Mayda, and B. A. A. Sumbodo, "Exploration of elcetronic-nose potenstial as diabetes urine detection using machine learning algorithms", *International Journal of Advanced Research in Science, Engineering and Technology*, Vol. 6, No. 7, pp. 1–10, 2019.

[9] L. Lin, S. Weng, Y. Zheng, X. Liu, S. Ying, F. Chen, and D. You, "Bimetallic PtAu alloy nanomaterials for nonenzymatic selective glucose sensing at low potential", *Journal of Electroanalytical Chemistry*, Vol. 865, pp. 1–8, 2020.

[10] W. Zheng, B. Han, E. Siyu, Y. Sun, X. Li, Y. Cai, and Y. Zhang, "Highly-sensitive and reflective glucose sensor based on optical fiber surface plasmon resonance", *Microchemical Journal*, Vol. 157, pp. 1–7, 2020.

[11] L. Singh, R. Singh, B. Zhang, S. Cheng, B. K. Kaushik, and S. Kumar, "LSPR Based Uric Acid Sensor Using Graphene Oxide and Gold Nanoparticles Functionalized Tapered Fiber", *Optical Fiber Technology*, Vol. 53, pp. 1–8, 2019.

[12] Y. Yu, J. P. Huang, J. Zhu, and S. L. Liang, "An accurate noninvasive blood glucose measurement system using portable near-infrared spectrometer and transfer learning framwork", *IEEE Sensors Journal*, Vol. 21, No. 3, pp. 3506–3519, 2020.

[13] N. Singer, R. G. Pillai, A. I. D. Johnson, K. D. Harris, and A. B. Jemere, "Nanostuctured nickel oxide electrodes for non-enzymatic electrochemical glucose sensing", *Microchimica Acta*, Vol. 187, pp. 1–10, 2020.

[14] S. Wang, G. Liu, B. Yang, Z. Zhang, D. Hu, C. Wu, Y. Qin, Q. Dou, Q. Dai and W. Hu, "Low-fouling CNT-PEG-hydrogel coated quartz crystal microbalance sensor for saliva glucose detection", *Royal Society of Chemistry Adv*, Vol. 11, No. 37, pp. 22556–22564, 2021.

[15] L. Zhang, Y. Wang, M. Niu, C. Wang, and Z. Wang, "Machine learning for characterizing risk of type 2 diabetes mellitus in a rural Chinese population: the Henan rural cohort study", *Scientific Reports*, Vol. 10, No. 1, pp. 1–10, 2020.

[16] Misbah, M. Rivai, F. Kurniawan, Z. Muchidin, and D. Aulia, "Identification of Diabetes through Urine Using Gas Sensor and Convolutional Neural Network", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 1, pp. 520–529, 2022.

[17] M. Rivai, A. Arifin and E. I. Agustin, "Mixed vapour identification using partition column-QCMs and Artificial Neural Network", *International Conference on Information & Communication Technology and Systems*, Surabaya, Indonesia, pp. 172–177, 2016.

[18] J. Kim, P. Urchaga, S. Baranton, C. Coutanceau and G. Jerkiewicz, "Interfacial structure of atomically flat polycrystalline Pt electrodes and modified Sauerbrey equation", *Physical Chemistry Chemical Physics*, Vol. 19, No. 33, pp. 21955–21963, 2017.

[19] P. Xie, W. Yuan, X. Liu, Y. Peng, Y. Yin, Y. Li, and Z. Wu, "Advanced carbon nanomaterials for state-of-the-art flexible supercapacitors", *Energy Storage Materials*, Vol. 36, pp. 56–76, 2021.

[20] A. Moghaddam, H. A. Zamani, H. K. Maleh, "A new electrochemical platform for dasatinib anticancer drug sensing using Fe_3O_4-SWCNTs/ionic liquid paste sensor", *Micromachines*, Vol. 12, pp. 1–12, 2021.

[21] S. J. Young, Y. H. Liu, Z. D. Lin, K. Ahmed, M. D. N. I. Shiblee, S. Romanuik, P. K. Sekhar, T. Thundat, L. Nagahara, and S. Arya, "Multi-walled Carbon Nanotubes decorated with silver nanoparticles for acetone gas sensing at room temperature", *Journal of The Electrochemical Society*, Vol. 167, No. 16, pp. 1–7, 2020.

[22] P. Kwiatkowski, D. Sondej, and R. Szplet, "Subpicosecond resolution time interval counter with multisampling wave union type B TDCs in 28 nm FPGA device", *Measurement*, Vol. 209, pp. 1–11, 2023.

[23] Misbah, M. Rivai, and F. Kurniawan, "Quartz crystal microbalance based electronic nose system implemented on Field Programmable Gate Array", *Telkomnika*, Vol. 17, No. 1, pp. 370–376, 2019.

[24] M. Todorokihara, "A 9 ps DNL/INL/RMS FPGA-Based Sigma Accumulation TDC with Unlimited Dynamic Range for Time-Based Analog Front End Applications", *IEEE Sensors*, Sydney, Australia, pp. 1–4, 2021.

[25] P. Peng, X. Zhao, X. Pan, and W. Ye, "Gas Classification using Deep Convolutional Neural Networks", *Sensors*, Vol. 18, No. 1, pp. 1–11, 2018.

[26] R. Kalidoss, S. Umapathy, R. Kothalam, and U. Sakthivelu, "Adsorption kinetics feature extraction from breathprint obtained by Graphene based sensors for diabetes diagnosis", *Journal of Breath Research*, Vol. 15, No. 1, pp. 1–11, 2020.

[27] R. Rajni and A. Amandeep, "RB-Bayes algorithm for the prediction of diabetic in Pima Indian dataset", *International Journal of Electrical and Computer Engineering*, Vol. 9, No. 6, pp. 4866–4872, 2019.

[28] S. Kumar, B. Bhusan, D. Singh, and D. K. Choubey, "Classification of Diabetes using Deep Learning", *International Conference on Communication and Signal Processing*, Chennai, India, pp. 0651–0655, 2020.

[29] N. Pradhan, G. Rani, V. S. Dhaka, and R. C. Poonia, "Diabetes prediction using artificial neural network", *Deep Learning Techniques for Biomedical and Health Informatics*, Academic Press, pp. 327–339, 2020.

[30] U. M. Butt, S. Letchmunan, M. Ali, F. H. Hassan, A. Baqir, and H. H. R. Sherazi, "Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications", *Journal of Healthcare Engineering*, Vol. 2021, pp. 1–17, 2021.

[31] Z. Ye, J. Wang, H. Hua, X. Zhou, and Q. Li, "Precise Detection and Quantitative Prediction of Blood Glucose Level with an Electronic Nose System", *IEEE Sensors Journal*, Vol. 22, No. 13, pp. 12452–12459, 2022.

[32] Y. Aggarwal, J. Das, P. M. Mazumder, R. Kumar, and R. K. Sinha, "Heart rate variability features from nonlinear cardiac dynamics in identification of diabetes using artificial neural network and support vector machine", *Biocybernetics and Biomedical Engineering*, Vol. 40, No. 3, pp. 1002–1009, 2020.