# Electronic Nose for Detecting Multilevel Diabetes using Optimized Deep Neural Network

Riyanarto Sarno, *Member, IAENG,* Shoffi Izza Sabilla, *Member, IAENG,*
Dedy Rahman Wijaya, *Member, IAENG,* and Hariyanto

***Abstract*—Diabetes is a chronic disease which is still a major issue in the world. The common testing methods generally used to detect diabetes are urine dipstick, laboratory blood tests, and blood glucose monitors. However, those testing procedures are often perceived as painful and inconvenient for the patients. In this context, this study proposes an electronic nose (e-nose) for detecting three classes of diabetes (healthy, prediabetes, and diabetes) based on a patient breath. The proposed e-nose system is called DENS, which utilizes an optimized deep neural network for the classification. DENS also attempts to enhance accuracy and to reduce the error rate from previous studies. Therefore, this paper has three contributions: (i) the optimal gas sensors for capturing patient breaths; (ii) the optimal signal preprocessing; (iii) the fine-tuned parameters of deep neural network (DNN) for classifying multilevel diabetes. The proposed system successfully detected multilevel diabetes with an accuracy of 96.29% and showed a minimum classification error of 0.050.**

***Index Terms*—diabetes, deep neural network, electronic nose.**

## I. Introduction

DIABETES is a chronic disease that is still categorized as a major issue in the world. This is caused by insufficient production of insulin or the body not being able to effectively use the produced insulin[1]. Body cells cannot absorb and process glucose to energy without insulin. Thus, the body will not work normally as a result. According to the World Health Organization (WHO), in 2014 about 9% of the population in the world suffered from diabetes [1], with adults as the majority of the sufferers. On the other hand, based on the data of the American Diabetes Association, there are about 7 million people living with undiagnosed diabetes [2]. Diabetes can severely impact patients health if not treated and controlled properly. Diabetes may cause blindness, kidney failure, heart attacks, stroke, and lower limb amputation. Traditionally, diabetes can be detected by taking the blood sample of the patient as an invasive method to measure the blood glucose level (BGL), which are painful, unpractical, and inconvenient for patients [3]. Therefore, several non-invasive methods have been developed for detecting diabetes.

Recently, several non-invasive methods are available for detecting diabetes; such as urine testing [4], [5], foot-bacteria based detection [6], breath analysis [7], [8], and even a knowledge-based system using ontology [9]. Breath analysis is a common method to obtain information for clinical application by analyzing volatile organic compounds (VOC) from the breath [8]. In addition, breath analysis can also estimate the BGL by analyzing the correlation between concentrations of biomarkers in the breath and the BGL ranges which is generally used [10]. Diabetes can also be detected from the ketone inside the blood of the patient which is an acid compound resulted from fat burning as the source of the energy. Someone who suffers from diabetes usually has breath that smells like acetone because of high levels of ketone inside the blood. As a result, the differences in the breath of the diabetes patient can be easily distinguished. The electronic nose is able to capture the differences in a person breaths as mentioned in the previous section.

An electronic nose (e-nose) can identify components in odor and analyze the compounds to gain information [11]. E-noses have been utilized in many industries. For example, in the beverage industry e-noses are used to classify strawberry juice [12], detect odors under different concentration levels of tea or coffee [13], and evaluate the quality of water used in beer production [14]. Moreover, e-noses can also be used to monitor indoor air quality[15] and distinguish the breath of smoker and non-smoker [16]. E-nose has also attracted considerable interest because it is cheaper, faster, more portable, and easier to use than similar systems such as gas chromatography or mass spectroscopy (GC/MS) [17].

Regarding diabetes detection, a previous study proposed the use of an electronic nose for BGL monitoring [18]. The main component of the electronic nose is the gas sensor array. Gas sensors are combined to detect specific gas in the testing environment [19]. In the previous study, an electronic nose with 5 sensors was developed. Those sensors are able to detect Carbon Monoxide (CO), Carbon Dioxide ($CO_2$), Acetone, VOC, Humidity and Temperature, and DHT-22. DHT-22 can measure the humidity and the temperature of the testing environment.

There are two levels of diabetes based on BGL, level 1 and level 2. Patient on level 1 is called prediabetes who has BGL ranges between 120 milligrams per deciliter (mg/dl) to 150 mg/dl. The patient with diabetes level 2 has a BGL range above 150 mg/dl, which is considered as a severe condition. The result of prediabetes can be used as an early warning for the patients to prevent them from diabetes level 2.

This paper proposes an electronic nose for detecting multilevel diabetes using fine-tuned parameters of deep neural network (DENS). As a result, the patient is able to check the diabetes type without any pain and feel inconvenient. DENS uses human breaths to detect three classes, healthy, prediabetes, and diabetes. This system also attempts to en-

Manuscript received September 10, 2018; revised December 4, 2018. This work was supported by the Institut Teknologi Sepuluh Nopember Surabaya and Ministry of Research, Technology and Higher Education of Indonesia.
Riyanarto Sarno, Shoffi Izza Sabilla, Dedy RahmanWijaya, and Hariyanto are with Departement of Informatics, Institut Teknologi Sepuluh Nopember Surabaya, Kampus ITS Keputih, Sukolilo, Surabaya, 60111, Indonesia.
Tel: +6231 – 5939214 Fax: +6231 – 5913804.
Emails of the authors are riyanarto@if.its.ac.id, shoffi12@mhs.if.its.ac.id, dedrw@tass.telkomuniversity.ac.id respectively.

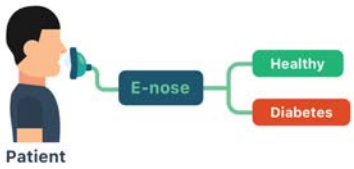
hance accuracy and reduce the error rate of a previous study. Therefore, this paper has three contributions; (i) the optimal gas sensors for capturing patient breaths; (ii) the optimal signal preprocessing; (iii) the fine-tuned parameters of deep neural networks (DNN) for classifying multilevel diabetes.

Recently, a lot of researches use the electronic nose with many sensors, but they do not look out whether the sensors are contributing significantly. DENS is able to solve the problem by observing the correlation between sensors and will exclude the sensors which are not contributing significantly. As a result, the gas sensors become more optimal for diabetes detection. The correlation value is calculated using principal component analysis (PCA). Many kinds of research use PCA to reduce the dimension of the independent variables [20], but in this study, PCA will be used to examine the correlation between sensors in this research.

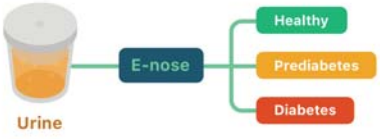
The output of the electronic nose is a signal. Those signals will always be affected by noise for about 20% [21]. Noise in this research was caused by the internal sensors, the changes of ambient conditions such as humidity and temperature, and the changes of electrical conditions such as the electrical voltage and current. Signal preprocessing is carried out to eliminate the noise in signals. Signal preprocessing is one of the important steps and support the development of accuracy and electronic nose sensitivity. Moreover, noise does affect the result, such as false detection. This is very crucial and unacceptable.

To improve accuracy, machine learning is needed to model the signal form into categories or classes. This algorithm has been used previously. However, most of the models used were very common, such as support vector machine (SVM), Naive Bayes, k-nearest neighbor (KNN), and artificial neural network (ANN) [22]–[24]. This research used deep-learning for the detection of multiple levels of diabetes. Deep-learning is a new model of machine learning that combines several models developed from supervised and unsupervised learning algorithms [25], resulting in better results. The purpose of using this approach, in addition to fast data processing, is that previously not labeled data will be labeled automatically [26]. Thus, the information can be optimally utilized and the performance will also be improved. In general, deep-learning or it is called a Deep Neural Network (DNN) is the development of an artificial neural network (ANN)[27] with more than one hidden layer. DNN has hyperparameters to increase the performance of the model that is formed [28]. In this study, a DNN model was developed to detect multiple levels of diabetes from the analysis of an e-nose response. The reliability of this model was assessed using two statistical parameters: confusion matrix and root mean squared error (RMSE).

The remainder of this study is organized into the following sections: Section 2 provides an explanation of related works and systems proposed in diabetes detection using an e-nose. The details of the proposed system are explained in Section 3. The results of the experiment and discussion are discussed in Section 4. In Section 5 the proposed system will be evaluated. A comparison between the proposed system and several other methods can be found in Section 5 too. Finally, Section 6 contains the conclusion of this work.




(a) Based on a patient breath




(b) Based on a patient urine odor

Fig. 1. Electronic nose for detecting diabetes

## II. Related Literature

The application of electronic nose in the diagnosis of diabetes mellitus is the technology with early diagnosis feature, non-invasive and convenient. Hence, it has been favored by doctors and patients. Diabetes complication can cause acidosis, which is directly related to the content of blood ketone acids body [18], [29]. Acetone is a nalidixic acid which is an end product of metabolism contained in the blood and can be excreted through the breath. According to this relationship, using the electronic nose technology for direct detection of acetone content in breath can indirectly evaluate blood glucose values and understand the blood glucose changes. There is a researcher [30] identified the exhaled gas of patients with diabetes mellitus acquired from the hospital by the electronic nose shown in Figure 1(a). Therefore, it demonstrates the feasibility of applying breath detection in the diabetes diagnosis.

The weakness of previous research [18] was that e-nose only divided patients into 2 classes, diabetes and healthy patient which has the accuracy of 95%.

The diagnosis of type 2 diabetes can also be conducted through urine odor detection. Moreover, there is a study [31] using collected urine samples from type 2 diabetic patients and normal healthy person, and detected urine odor molecules through the electronic nose technology shown in Figure 1(b). The result shows that the detection rate of patients with type 2 diabetes is 95.00%. Other paper [32] pointed out that the detection rate of inflammatory bowel disease or diabetes by the electronic nose was approximately 97% in the study on evaluation and application of electronic nose technology for human disease detection.

Although the research can identify diabetes type 2, the implementation of e-nose was still impractical, inefficient, and inconvenient for most patients because they have to urinate first. Another previous research which has been done by Dongmin Guo [10], the patient was divided into 4 levels to detect diabetes on e-nose, level 1 where BGL is between 81-100, Level 2 is 101-150, level 3 is 151-200, and level 4 is 201-421. The result showed that the accuracy of this research was 76.00%.

Therefore, this paper will modify the existing research by combining the previous methods, which is the detection of

TABLE I
EXHALED BREATH MOLECULES CONCENTRATIONS

| Concentration | Molecule |
|---|---|
| Percentage | Oxygen, Water, Carbon Dioxide |
| Parts-per-million (ppm) | Acetone, Carbon Monoxide, Methane, Hydrogen, Isoprene, benzene methanol |
| Parts-per-billion (ppb) | Formaldehyde, acetaldehyde, 1-pentane, ethane, ethylene, other hydrocarbons, nitric oxide, carbon disulfide, methanol, carbonyl sulfide, methanethiol, ammonia, methylamine, dimethyl sulfide, benzene, naphthalene, benzothiazole, ethane, acetic, aide |

multiclasses diabetes using e-nose on human breath analysis (DENS). As a result, diabetes test will be more practical, efficient, and convenient to the patients. This research is doing further review on detecting diabetes using electronic nose by using 3 classes according to the references of Blood Glucose Level in the world. Thus, the contribution of this research is significant because of the classification of 3 classes are more difficult than 2 classes.

The exhaled human breath is a mixture of $N_2$, $O_2$, $CO_2$, $H_2O$, inert gases, and thousands of other trace gases. These gases include inorganic molecules such as NO, $NH_3$, or CO, and volatile organic compounds (VOCs) such as acetone, methanol, or isoprene, with concentrations ranging from ppb or ppm [33] as shown in Table I.

## III. PROPOSED SYSTEM

### A. The Proposed Electronic Nose

An electronic nose (e-nose) is an instrument that is used to analyze odor or aroma. In this study, an e-nose was used with the aim to detect the content in some biomarkers, namely carbon monoxide (CO), carbon dioxide ($CO_2$), acetone, temperature, humidity, and volatile organic compounds (VOC). Those are the contents of breath that can differentiate which one is healthy, prediabetes, and diabetes patients based on the different concentrations of biomarkers. An Arduino ATM Mega 2560 microcontroller was used to generate the values from the sensor. It performed analog to digital conversion (ADC) and transferred the data from each gas sensor via a wireless network to the computer.

In the initial experiment, DENS used six gas sensors and one blood glucose level. The selection of gas sensors in the initial sensor array can be seen in Table II. The combination of sensors and their responses is based on our previous research [18]. Patients with diabetes have ketone in their breath and sweat which release acetone. This is caused by a side product of fat metabolism. In diabetes patients, insulin cannot be produced based on body need, so glucose cannot be delivered optimally to the cells of the body. In addition to the ketone content, which is a benchmark of diabetes patients, other compounds such as ethanol [34], carbon monoxide [35], alkane [36], and methyl nitrate [37] in the breath have also been shown to have different concentrations compared to healthy patients and have a correlation with the prediction of blood glucose level.

For this study, no comparable datasets of healthy, prediabetes, and diabetes patients based on e-nose sensor responses were available on the Internet. Therefore, the ground-truth data used were a combination of self-extracted data from the e-nose and from health institutions to determine whether a person was classified as 'healthy', 'prediabetes' or 'diabetes' based on blood sugar level [1]. The determination of ground-truth data class based on blood sugar level can be seen in Table III. 'Healthy' is defined by the BGL value of less than 120 mg/dl; while 'Prediabetes' is defined by the BGL value between 120-150 mg/dl; whereas 'Diabetes' is defined by the BGL value higher than 150 mg/dl.

TABLE II
INITIAL GAS SENSOR LIST IN THE SENSOR ARRAY

| No | Sensor | Selectivity |
|---|---|---|
| 1 | MQ 7 | Carbon monoxide (CO) |
| 2 | MQ 135 | $NH_3$ (ammonia), $NO_2$, alcohol, benzene, smoke, $CO_2$ |
| 3 | MQ 138 | Acetone |
| 4 | DHT-22 | Temperature & humidity |
| 5 | MiCS-5524 | Volatile organic compounds (VOC) |

TABLE III
BLOOD GLUCOSE LEVEL (BGL)

| No | BGL (mg/dl) | Class |
|---|---|---|
| 1 | $<$120 | Healthy |
| 2 | 120-150 | Prediabetes |
| 3 | $>$150 | Diabetes |

### B. The Optimal Signal Preprocessing

Data preprocessing is an important step in signal processing. Basically, the signal is very easy to get mixed with noise. Those noises are usually coming from the unstable voltage, or even from the changing temperature and air humidity [38]. Therefore, the signal must be reconstructed by using data preprocessing. It helps to improve the quality of the data, which will have an impact on the final result [39]. Data preprocessing in DENS is performed in two stages, namely signal de-noising and data normalization. In the experiment, DENS obtained the optimal parameters of the discrete wavelet transform (DWT) for fine-tuning the signal reconstruction. Coefficient from DWT in the formula is initialized as $dwt$. The DWT of the original signal $x(t)$ can be expressed with the following Equation 1 and 2:

$$dwt(i,k) = \langle x(t), \psi_{i,k}(t) \rangle \tag{1}$$

$$dwt(i,k) = \frac{1}{\sqrt{2^i}} \int_{\neg}^{-\neg} x(t)\psi^*(\frac{t-k2^i}{2^i})dt \tag{2}$$

where $i$,$k$, and $\psi$ are the scaling, the shifting parameter, and mother wavelets, respectively. The scaling parameter

establishes the time and frequency resolution of the scaled mother wavelet transform (MWT) [40]. Fix $i$ at a particular scale, so that $dwt(i,k)$ is a function of $k$ only. In light of the above equation, $dwt(i,k)$ can be viewed as a convolution of $x(t)$ with a dilated, reflected, and normalized version of the mother wavelet.

Prior to feature extraction, the signal response data are cleared from noise using discrete wavelet transform (DWT) and feature scaling using z-score normalization. De-noising the signals is an important stage for improving the performance of electronic nose (e-nose) accuracy [41] and sensitivity [42]. In this study, the coefficient taken was an approximation coefficient derived from a low-pass filter. The approximation value used depends on the decomposition level and the MWT [43]. The best-suited MWT is determined based on the information quality ratio (IQR) [44]. There are 38 Mother Wavelet Transforms (MWTs), including daubechies (db1-db10), symlet (sym2-sym8), coiflet (coif1-coif5), biorthogonal (bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, bior6.8), and dmey. The maximum value of IQR denotes the best-suited MWT for a particular signal. The IQR for the original signal $(x(t))$ and reconstructed signal $(y(t))$ can be computed with the following Equation 3:

$$IQR = \frac{\sum_{x_i \varepsilon x(t)} \sum_{y_j \varepsilon y(t)} p(x_i, yj) log_2(p(x_i)p(y_i))}{\sum_{x_i \varepsilon x(t)} \sum_{y_j \varepsilon y(t)} p(x_i, yj) log_2(p(x_i, yj))} - 1 \tag{3}$$

where $x_i$ and $y_j$, are particular values of $x(t)$ and $y(t)$ respectively. $P(x_i)$ and $P(y_j)$ are the marginal probability and $P(x_i, y_j)$ is the joint probability of $x_i$ and $y_j$.

Basically, the range of IQR is between 0 and 1. $IQR = 1$ indicates that the reconstructed signal can completely keep the essential information from the original signal and vice versa. The decomposition level also has to be determined carefully. If the level of decomposition is high, then the risk of damage to the signal is also big. If the decomposition level is low, then the signal will be sensitive to noise. The decomposition level can be determined using this rule:

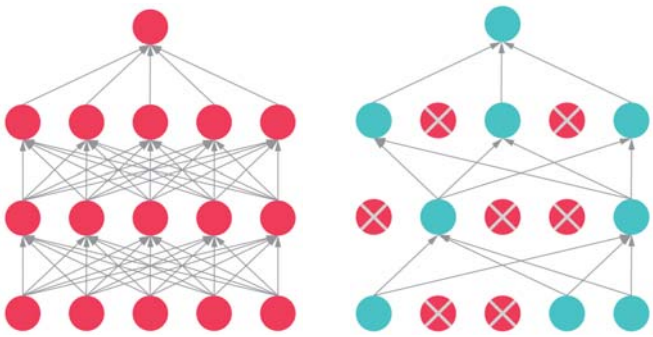
$$\frac{F_q}{2^{L+1}} \leq F_{char} \leq \frac{F_q}{2^L} \tag{4}$$

where $F_q$, $F_{char}$, $L$ denote the sampling frequency, dominant frequency, and decomposition level, respectively. After de-noising is done, the normalization process follows the approximation coefficient to standardize the scale of each attribute using the z-score.

Normalization or feature scaling is a technique to standardize or equate a range of data so that no attribute is too dominant over other attributes. In this study, normalization of the data was done once, using normalized standard scores (z-score normalization). The formula of z-score can be seen on the Equation 5. Scaling using normalized standard scores is done so that the signal data generated after noise removal have properties such as normal distribution.

$$z = \frac{x - \mu}{\sigma} \tag{5}$$

where $z$, $x$, $\mu$, and $\sigma$ are the default values, the original data (sensor response per time), the average of the original data, and the standard of the original data deviation, respectively.



(a) Traditional ANN (b) ANN with dropout technique

Fig. 2. The architecture of the Artificial Neural Network

### C. The Fine-tuned Parameters of Deep Neural Network

A fully connected layer is a layer in which all activation neurons of the previous layer are connected with the neurons of the next layer. This is the concept of traditional artificial neural network (ANN) [45]. Each connection has different weight and the weight is increased by bias in the next layer connection. The calculation of the weight and bias are $5x5$ $weight$ + 5 $bias$, $5x5$ $weight$ + 5 $bias$, and $5x1$ $weight$ + 1 $bias$. The total of 66 parameters will be updated to get the best result on the training process. Each layer has neuron and the activation function from the hidden layer and output layer will decide whether the neuron must be active or not. Some of the activation function that widely used are Sigmoid in Equation 6, Tanh, ReLu, Softmax.

$$\frac{f(y)}{f(x)} = \frac{1}{1+e^{-x}} x(1 - \frac{1}{1+e^{-x}}) \tag{6}$$

where $x$ is the weight dot bias. Sigmoid has ranged between 0 until 1 and Tanh -1 until 1 as shown in Equation 7.

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{7}$$

Sigmoid and Tanh has their own weakness that is covered by ReLu. ReLu performed the threshold from 0 until infinity can be shown in Equation 8. Another activation is $softmax$ as showed in Equation 9.

$$\frac{f(y)}{f(x)} = \begin{bmatrix} 1 & x > 0 \\ 0 & x \leq 0 \end{bmatrix} \tag{8}$$

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_j}} \tag{9}$$

where $\sigma$ is the probability of class member $j$ as the input, $z$ is the linear equation of the activation function, and $K$ is the class predicted with $j$ as part of that class. The formula for the linear equation of the activation function is as shown in Equation 10. The exponential value of weight $(w)$ is a coefficient summed with constant bias $(b)$, after which it is normalized.

$$softmax = normalize(exp(wx + b)) \tag{10}$$

The difference between a DNN and a traditional ANN is complexity, the more complexity on the feature, the more

cost and time on the ANN. This is happening because of its complex feature which affects the total layer. Thus, the neuron that was built is also complex. Another difference is that ANN has limited layer and neuron to get the optimal result. DNN has dropout which is not present in ANN [46]. Dropout is an algorithm to decide how many neurons are utilized. A number of randomly picked neurons are not used in the training process. In other words, these neurons are excluded based on the dropout parameter [47].

Figure 2(a) represents the architecture of a traditional ANN composed of 5 nodes in the input layer, 2 hidden layers, and 1 output layer, while Figure 2(b) is a neural network with dropout. The hidden layer consists of interconnected neurons. In Figure 2(a), all neurons inside the hidden layer are connected. Different from Figure 2(a), not all neurons in the hidden layer in Figure 2(b) are connected. This technique is used in DNN models to increase the performance of the training process and reduce the occurrence of overfitting [46]. Moreover, in DNN, there is a parameter called loss function, which is a function to show the loss value of each possibility that will be generated by the DNN model. Thus, the loss function estimates the quality of each given weight, bias, and parameter [48]. This function is executed when the learning model yields a high loss value. Thus, by utilizing this function errors are minimized. Loss function will compare the prediction of the output and the target with the formula on the Equation11. If the result is far from the target, then the performance of the model created is not optimal.

$$LossFunction = (Target - Prediction)^2 \tag{11}$$

Each error showed in the loss function, weight and bias must be readjusted. This process is called back-propagation. Before creating its network model, DNN has some parameter which can be used. Epoch, learning rate, batch_size , those all the hyperparameter which can be decided individually. Until now, in 2018, there is still no research on solving the ideal hyperparameter value.

To apply a deep-learning model, researchers use either handmade code or a deep-learning library. Libraries available today that can be used to solve DNN problems are Tensorflow, Keras, and Caffe [49]. The library most commonly used by researchers is Keras [39], [50]. The Keras library provides wrapper classes to allow the use of neural network models developed with Keras in scikit-learn [51]. The Keras library takes the name of a function as an argument. This function must return the constructed artificial neural network model when it is ready for the training process. Besides, this function also creates a baseline neural network for the diabetes classification problem. Argument required must be used on the Keras Library in order the create DNN model, as can be seen in Table IV.

To gather the parameters required, DENS ran some tests to compare the results of the available parameters. Some combination of Mother wavelet and level decomposition was made to find the best parameter and the results were used in the next step. Highest accuracy showed as the best parameter, where the accuracy can be calculated by using Equation 12.

$$Acc_i = \frac{cc_i}{N_i} \times 100\% \tag{12}$$

TABLE IV
ARGUMENT REQUIRED ON THE KERAS LIBRARY

| Required | Types |
|---|---|
| Optimizer | adagrad, adadelta, adam, stochastic gradient descent(sgd), adamax, nadam |
| Activation | softmax, linier, ReLu, sigmoid, Tanh |
| Kernel Initializer | Glorot, Uniform, RandomNormal, RandomUniform |

TABLE V
EXPERIMENTAL DATA

| ID | Amount of Data | Description |
|---|---|---|
| RH | 20 Data | RH = id for healthy patients |
| RA | 15 Data | RA = id for prediabetes patients |
| RC | 19 Data | RC = id for diabetes patients |

where $Acc_i$, $cc_i$, and $N_i$ are the accuracy of the classifier, the number of correctly classified sample, and the number of sample on $i$ testing data, respectively. The average of $Acc_i, i = 1, 2, .., k$ was calculated to obtain the final accuracy for the classifier.

## IV. RESULT AND DISCUSSION

### A. Data Acquisition

The patient gives a blood check result that contains information on the blood glucose level. The doctor gives the diagnosis (healthy or diabetes) to the patient based on the result of a random blood glucose check. Then, the patient is asked to breathe into the e-nose device for 150 seconds to retrieve the data to be used as the ground truth, using an oxygen mask connected to the sensors and breath container.

The microcontroller Arduino ATM MEGA 2580 sends the sensor response to the computer through a universal serial bus (USB). The data of the sensor response and the gas concentration in the breath of the patient were captured by the e-nose in the form of a text file (.txt) with comma separated value (csv) data format. The text file data of the sensor response and gas concentration along with some personal data and the ground-truth of the diagnose are stored, so that the ground-truth data are recorded in the database. The process of recording the ground-truth data can be seen in Figure 3.

Seven features are obtained from the data gathering process: CO, $CO_2$, ketones, humidity, temperature, VOC and the last reported blood glucose level. Each sample has around 2000 records of data that are gathered in a 150-second window. Then, to reduce the resources used in data processing, each feature value is averaged. An example of the data used is shown in Table V. The average of CO is initialized using avgCO, $CO_2$ is initialized using avgCO2, ketones are initialized using avgKetone, humidity is initialized using avgHumid, the temperature is initialized using avgTemp, and VOC is initialized using avgVOC. An example of the data used is shown in Table VI, where the ID column contains the identity of the patients. RH means the patient is healthy, RA represents prediabetes patients, and RC represents patients diagnosed with diabetes. Ground-truth data are required to

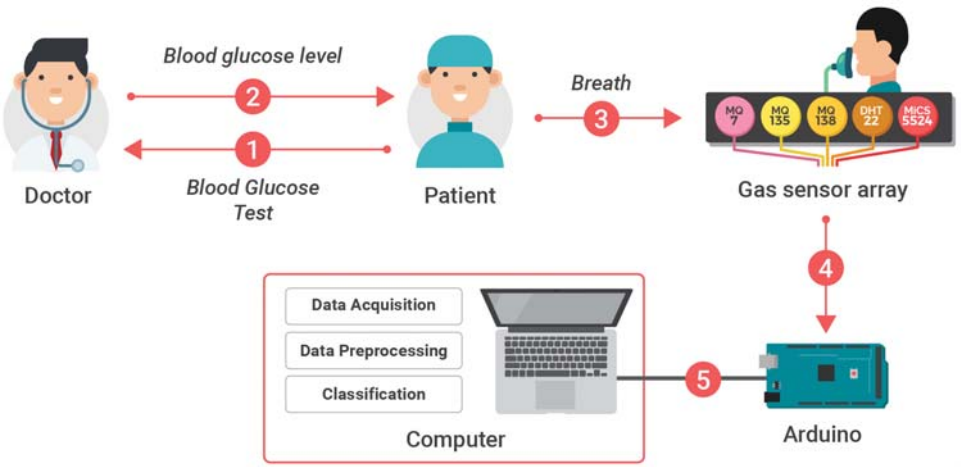



Fig. 3. DENS workflow for diabetes detection

TABLE VI
EXAMPLE OF ELECTRONIC NOSE RESULTS

| ID | avgCO | avgCO2 | avgKetone | avgHumidity | avgTemp | avgVOC | BGL |
|---|---|---|---|---|---|---|---|
| RH002 | 584.287 | 1.312 | 86.405 | 32.554 | 361.313 | 1.400 | 139 |
| RH004 | 521.110 | 1.232 | 801.695 | 33.947 | 370.709 | 1.318 | 133 |
| RH005 | 580.534 | 1.195 | 77.303 | 32.031 | 326.701 | 1.315 | 130 |
| RH006 | 693.127 | 1.529 | 916.682 | 33.639 | 320.177 | 1.460 | 123 |
| RH007 | 641.862 | 1.445 | 859.406 | 33.609 | 30.394 | 1.407 | 130 |
| RH008 | 1.008.369 | 1.547 | 779.528 | 2.633 | 362.038 | 1.365 | 127 |

get the training set, which is also used as the testing set. The ground-truth data were gathered in a clinic in Surabaya.

### *B. Data Preprocessing*

The ground-truth data contained high interference of noise, which needs Signal processing to avoid inaccurate data prediction. DWT was used in this research by finding the best parameters to reduce noise on the e-nose signal. Two best parameters were mother Wavelet and level decomposition. Some combinations were used from the parameters as be seen in Table VII, showed the optimal parameter obtained from mother wavelet db6 with level decomposition 1.

TABLE VII
FINE-TUNED PARAMETERS OF WAVELET TRANSFORM

| Level of Decomposition | Mother Wavelet | Accuracy |
|---|---|---|
| Level 1 | bior1.1 | 0.667 |
| Level 1 | bior1.3 | 0.667 |
| Level 1 | bior1.5 | 0.667 |
| Level 1 | bior2.2 | 0.667 |
| Level 1 | bior2.4 | 0.667 |
| Level 1 | dmey | 0.701 |
| Level 1 | coif1 | 0.736 |
| Level 1 | coif5 | 0.736 |
| Level 1 | db1 | 0.744 |
| Level 1 | db6 | 0.748 |

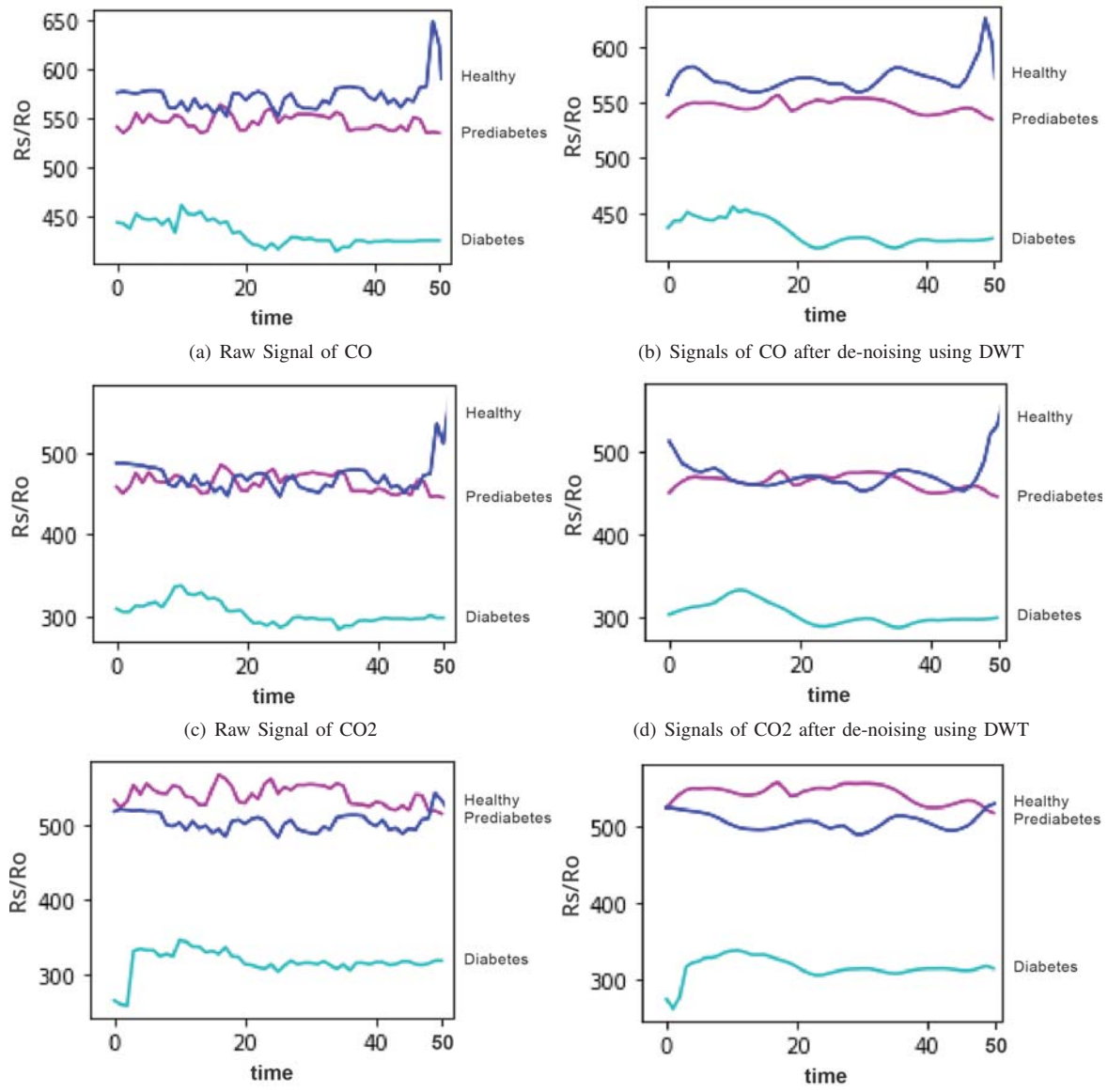
The experimental data obtained from the electronic nose are converted into graphical form. In Figure 4(a) on the left side is the result of signal for CO gas, which is still incompletely de-noised, and on the right side is the result of the signal after de-noising. $CO_2$ gas can be seen in Figure 4(d), where the noise from the signal has been reduced. The ketone signal looks smoother, as can be seen in Figure 4(f).

Not all of the features mentioned above may contribute to the detection of multilevel diabetes. A feature can be incremented down if it does not affect detection since the blood glucose level in none of the patients has a range that is too significant. After the signal is de-noised and normalized, the average value for each sample is calculated.

The feature extraction method used is principal component analysis (PCA). There are six features that are obtained from the sensors. CO gas, $CO_2$ gas, ketone gas, humidity, temperature, and VOC. In general, principal component analysis (PCA) is a technique for reducing the dimensions of data without eliminating the characteristics of the data by creating new variables from linear combinations of the original variables. This method extracts the features and removes features that are less influential, which means that there is a more significant difference between one class and the other classes.

TABLE VIII
REDUCE THE FEATURE USING PRINCIPAL COMPONENT ANALYSIS

| n-feature | Accuracy |
|---|---|
| 2 features | 0.667 |
| 3 features | 0.751 |
| 4 features | 0.736 |
| 5 features | 0.744 |
| 6 features | 0.748 |




(a) Raw Signal of CO

(b) Signals of CO after de-noising using DWT

(c) Raw Signal of CO2

(d) Signals of CO2 after de-noising using DWT

(e) Raw Signal of Ketone

(f) Signals of Ketone after de-noising using DWT

Fig. 4. Signals response of DENS

The input in this process is the data signal that has been normalized after preprocessing. Experiments were performed with feature reduction. The results produced satisfactory accuracy with only 3 features. Table VIII shows that some important features were missing when only 2 features were used, so three or more features obtained a better result. The result of using PCA with scaller was better than that only using PCA, as illustrated in Figure 5(a) and Figure 5(b), respectively.

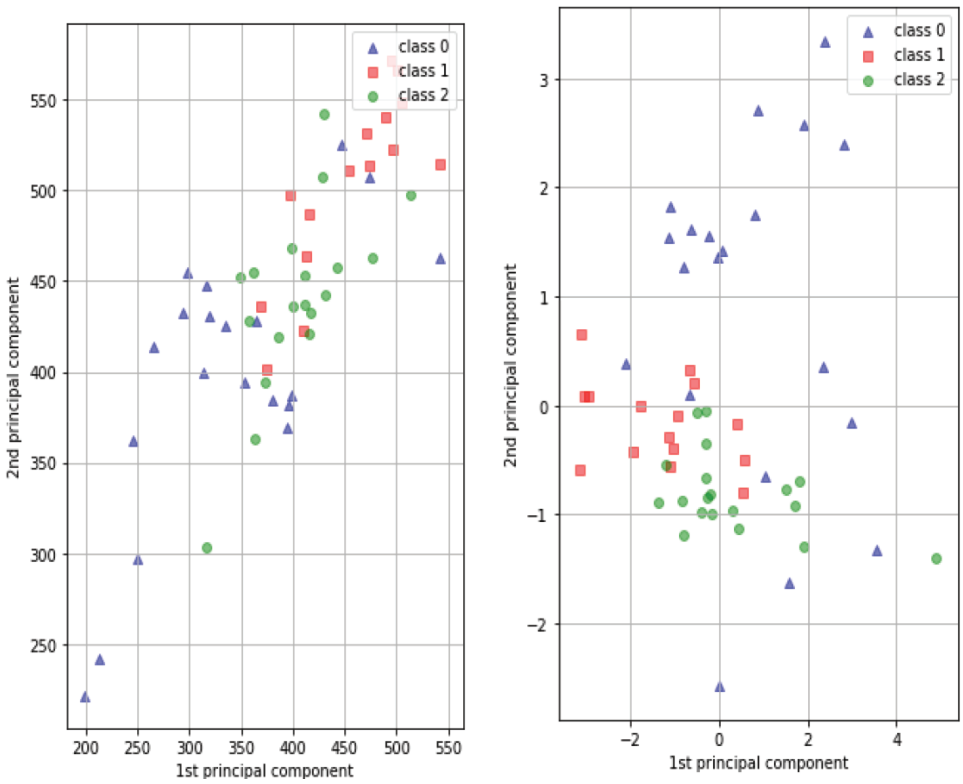
TABLE IX
OPTIMIZER PARAMETERS

| Optimizer | Accuracy |
|---|---|
| sgd | 0.759 |
| adagrad | 0.800 |
| adadelta | 0.926 |
| adam | 0.962 |

### C. Optimal Parameters of DNN

After getting 3 features from the extraction process, the next step is building the network model to set as reference. For this work, a fully connected layer model was created using a normal for the initializer kernel.

It creates a simple, fully connected network with 3 features or node in the input layer, and two hidden layers that contain 12 and 1000 neurons respectively. The hidden layer uses a rectifier activation function. Since DENS used one hot encoding for our dataset, the output layer must create 3 output values, one for each class. The output value with the largest value was taken as the class predicted by the model. DENS used $softmax$ activation function in the output layer. $Softmax$ activation was chosen because it is commonly used in classification with more than two classes. To gather the parameters required, DENS ran some tests to compare the results of available parameters. The best parameter from the test results was used in the next step. Optimizer parameter results are shown in Table IX. The accuracy value generated by Adam optimizer was better than




(a) Data signal after PCA (b) Data signal after PCA and StandarScaller

Fig. 5. The result of data signals after reduction feature using Principal Component Analysis with scaller

that of the other optimizers.

This study used a deep neural network (DNN) and the architecture of the DNN is shown in Figure 6. The parameter used was a normal initializer with 1000 neurons and 50 batches. ReLu was used to activate the movement of data from the input layer to the hidden layer, while $softmax$ activation was used to activate the movement from hidden layer to the output layer with 500 epochs and dropout value of 0.1.

This DENS found that the best parameters to be utilized are 1000 neurons and dropout value of 0.1. The dropout function is used to reduce unnecessary neurons. With a dropout value of 0.1 for 1000 neurons, the number of neurons that were used is 900 neurons as the 10% of the neurons is excluded. ReLu activation is used on this layer, which reduces the training time. After building the model, an output layer consisting of 3 nodes is created. These 3 nodes are the classes that represent the diabetes levels which are class 0 for 'healthy', class 1 for 'prediabetes', and class 2 for 'diabetes'.

The most important step of creating a neural network is the formation of the hidden layer. In this step, a fully connected layer needs to be created to which the set of nodes were connected. These nodes act as the input layer for the fully connected layer. Since this layer exists between the input layer and the output layer referred as a hidden layer. In the proposed method, the Adam optimizer is used on this layer, but the kernel used is a normal initializer. To activate the movement from the input layer to the hidden layer 1 ReLu activation was used. This activation is also used to activate the neuron in the hidden layer 1 to the hidden layer 2. To move the values generated by the hidden layer 2 to the output layer, the values were converted to non-linear values using $softmax$ activation. $Softmax$ was used to give more intuitive results which makes the classification process easier. A loss function was used with cross-entropy parameters to monitor errors generated during model building. Given 500 epochs, DENS was automatically searching for epochs to see how many training processes have reached the optimum value.

## V. Evaluation

The leave-one-out method, or n-fold cross validation, where n is the amount of data in the dataset, provides a more precise accuracy measurement but takes longer to execute. The k-fold cross validation method is considered to be less precise since it divides data into buckets randomly. Thus, when running the method for a second time it may have different accuracy due to random data placement. In the leave-one-out method, the spread of classes in various places is uneven, which makes the accuracy lower. Therefore, a stratified method was used here, which basically divides the data and sets the key of each class.

Stratified k-fold is a sampling method where the data are divided equally. In this study, there were 10 healthy patients, 10 prediabetes patients, and 10 diabetic patients. The data were divided equally into 2 places, with a key for each class such that each class had 5 healthy patients, 5 prediabetes patients, and 5 diabetic patients. This method guarantees that each place has the same type and the same properties.

### A. Confusion Matrix

The confusion matrix is widely used to test the performance of a classification method. It visualizes the table

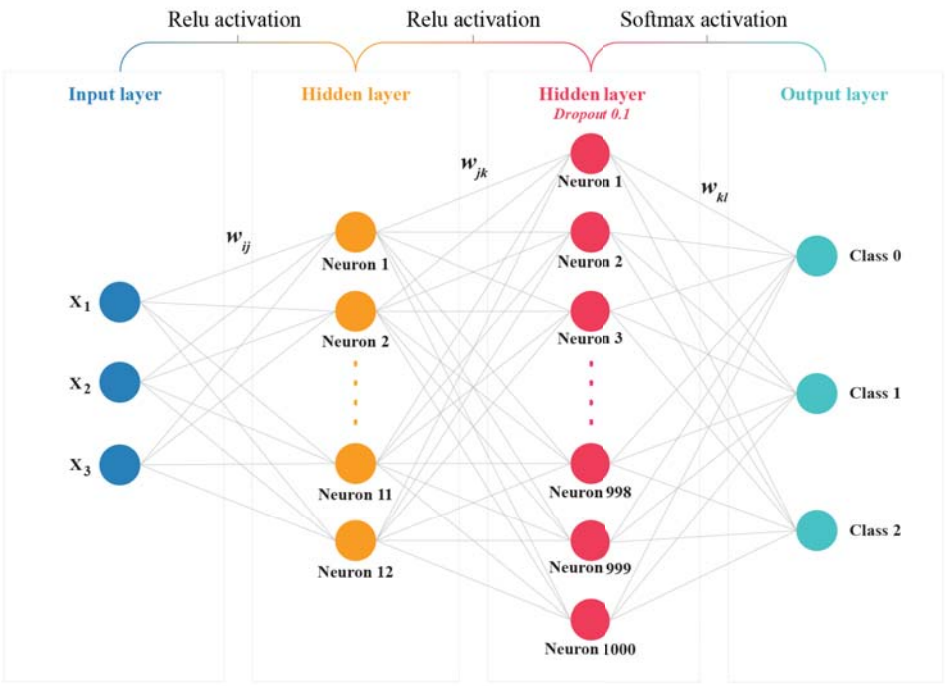



Fig. 6. Architecture of deep neural network for DENS

of comparison between the predicted class and the actual class. In this study, the confusion matrix stores classification information for healthy patients, prediabetes patients and diabetes patients by deep learning. In the DNN confusion matrix in Figure 7, there are 20 data of healthy patients. The actual data show that there were 19 patients that were predicted as healthy patients. Only 1 prediction said that the patient was prediabetes when it actually should have been healthy. From the prediabetes data, 14 out of 15 data were predicted correctly.

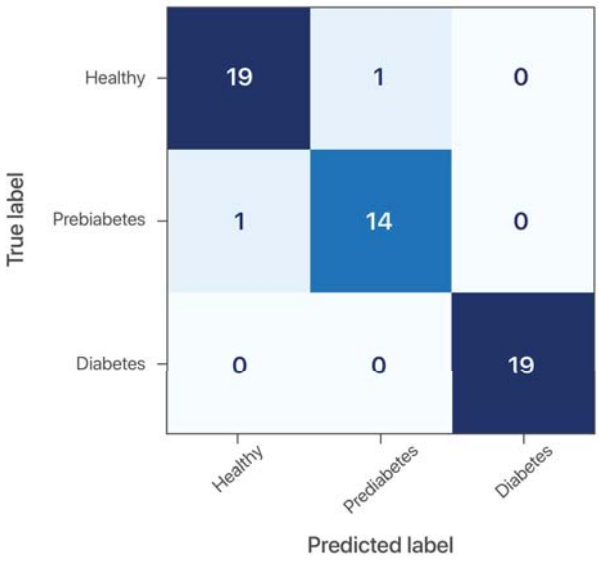



Fig. 7. Result of confussion matrix for classification multilevel of diabetes

Meanwhile, all of the diabetes data were predicted correctly. Hence, we can obtain various indicators of the performance of the classifier [52][53], such as accuracy, precision, and recall, as shown in Table X.

TABLE X
CLASSIFIER PERFORMANCE RESULT

| | Healthy | Prediabetes | Diabetes |
|---|---|---|---|
| Precision | 0.95 | 0.93 | 1.00 |
| Recall | 0.95 | 0.93 | 1.00 |
| F1-score | 0.95 | 0.93 | 1.00 |
| Accuracy | 0.96 | | |

TABLE XI
PERFORMANCE MEASURMENT OF CLASSIFIER

| | TP | FN | FN | TN | TPR | FNR | TNR | FPR |
|---|---|---|---|---|---|---|---|---|
| Class 0 | 18 | 2 | 0 | 34 | 0.9629 | 0.04 | 0.9815 | 0.02 |
| Class 1 | 15 | 0 | 1 | 38 | | | | |
| Class 2 | 19 | 0 | 1 | 34 | | | | |

For further analysis, the classification result of DNN in the experiment was divided into True Positive Rate (TPR), False Negative Rate (FNR), True Negative Rate (TNR), and False Positive Rate (FPR), as summarized in Table XI. As can be seen in Table XI, DNN produced TPR and TNR of 96.29% and 98.15% on average, respectively. It can be also seen in Table XI, on average DNN achieved low FNR and FPR of 0.04% and 0.02%, respectively. These results show that DNN could recognize the class of diabetes patients with a good performance.

Figure 8 show the ROC in each class is good by showing the result of each value is close to 1, class 0, class 1, and class 2 as 0.95, 0.99, and 0.99, respectively. Y-axis and x-axis are defined as TPR and FPR. TPR is equivalent to sensitivity where classifier marked as perfect if sensitivity moves toward coordinate (0,1) [54]. Macro average and micro average in the graph also indicate that the system performance is

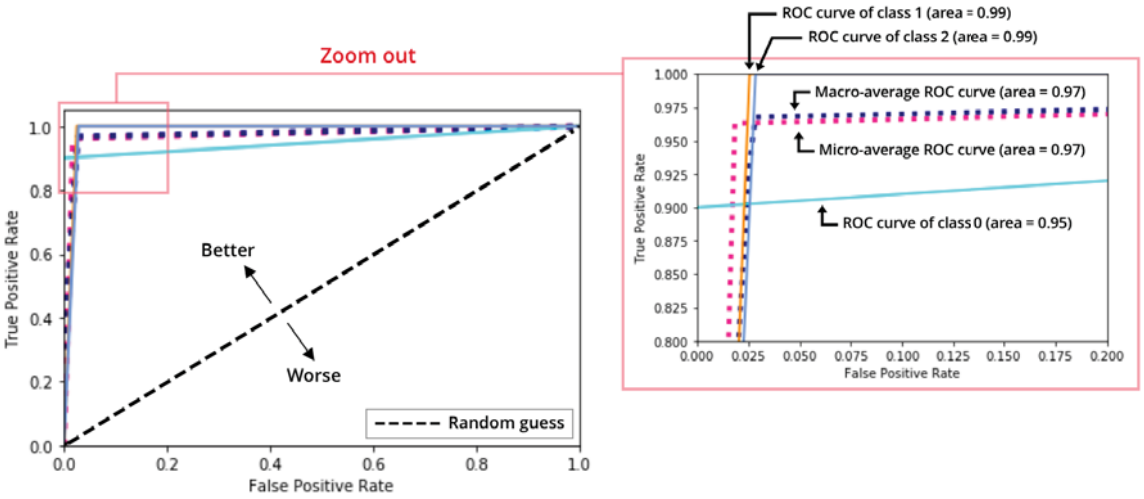



Fig. 8. Receiver Operating Characteristic (ROC) Curve of detecting multilevel diabetes

increased while using varied data, which are 0.97 in macro average and 0.97 in micro average.

### B. Root Mean Squared Error

Furthermore, several metrics can be used to measure the performance of a deep-learning method, such as root mean squared error (RMSE) and R-squared ($R^2$). RMSE measures the errors between actual and predicted values. A low RMSE value indicates a few prediction errors. The symbols $a$, $p$, $k$, $n$ in Equation 13 denote actual vector, predicted vector, number of predictors, and total samples, respectively. R-squared reflects the fitting between the actual and the predicted vectors, which means that a good fit is represented by a high $R^2$. Usually, the range of $R^2$ is between 0 and 1. A negative $R^2$ indicates that the chosen model does not follow the trend of the data, so it fits worse than a horizontal line. In this study, the actual vector reflects the observed glucose.

$$RMSE(a,p) = \sqrt{\frac{\sum_{i=1}^{n}(a-p)^2}{n}} \tag{13}$$

$$R^2(a,p) = 1 - \frac{\sum_{i=1}^{n}(a-p)^2}{\sum_{i=1}^{n}(a_i-\bar{p})^2} \tag{14}$$

The closer the value of $R^2$ is to one, the lower the error rate. It can be concluded that the DNN model could classify the data signal for diabetes very well, as can be seen from the value of $R^2$, which was 0.9496 for the DNN model.

### C. Performance Comparison

The result of the calculation of several classifier methods [55] are shown in Table XII, with the scaled and unscaled results for each classifier.

Table XII shows that, apart from the proposed method, the method with the best accuracy was SVM using rbf kernel. SVM uses the gamma value and the $C$ parameter as the main contributors. In this study, the gamma value ranged from $10^{-3}$ to $10^{1}$. The system finds the best gamma value based on the given range. C parameter received the same treatment as the gamma value. The range given for the C value was between $10^{-1}$ and $10^{3}$. This study tried to find the range closest to the optimal parameter.

TABLE XII
COMPARISON OF SEVERAL CLASSIFIERS

| Classifier Method | Unscaled Accuracy | Scaled Accuracy |
|---|---|---|
| Naive Bayes | 0.741 | 0.741 |
| Linier Discriminant Analysis | 0.741 | 0.741 |
| Decision Tree | 0.685 | 0.685 |
| LR | 0.704 | 0.741 |
| KNN | 0.685 | 0.815 |
| SVM | 0.759 | 0.833 |
| Proposed Method - DENS | 0.933 | 0.963 |

## VI. Conclusion

The DENS was implemented for non-invasive blood glucose level classification to categorize three levels of diabetes, i.e. 'healthy', 'prediabetes', and 'diabetes'. The DENS contained five gas sensors. The collected signals were reconstructed using DWT at db6 level 1. PCA reduced 2 out of 5 features available, as a result, DENS is more optimal by only using 3 sensors. Moreover, a deep-learning classification method to estimate the blood glucose level and detect multiple levels of diabetes was proposed. The classification done using a deep neural network has been proven to increase the performance of the classification compared to those of the other methods. The accuracy results of other classifiers, k-NN, SVM, Naive Bayes, and LDA, were 81.47%, 83.33%, 74.07%, 74.07%, respectively, while the proposed method achieved an accuracy of 96.29% and an error rate of 0.050. The future works can be (i) obtaining the optimal ambient temperature of the gas sensor, (ii) evaluating the effect of correlated gases for detecting diabetes, (iii) determining an optimal classification method.

## Acknowledgment

The authors would like to thank Institut Teknologi Sepuluh Nopember Surabaya and the Ministry of Research, Technology and Higher Education of Indonesia for supporting this research.

REFERENCES


[1] K. Ogurtsova, J. da Rocha Fernandes, Y. Huang, U. Linnenkamp, L. Guariguata, N. Cho, D. Cavan, J. Shaw, and L. Makaroff, "Idf diabetes atlas: Global estimates for the prevalence of diabetes for 2015 and 2040," Diabetes research and clinical practice, vol. 128, pp. 40– 50, 2017.

[2] K. M. Flegal, M. D. Carroll, C. L. Ogden, and L. R. Curtin, "Prevalence and trends in obesity among us adults, 1999-2008," Jama, vol. 303, no. 3, pp. 235–241, 2010.

[3] T. D. C. Minh, D. R. Blake, and P. R. Galassetti, "The clinical potential of exhaled breath analysis for diabetes mellitus," Diabetes research and clinical practice, vol. 97, no. 2, pp. 195–205, 2012.

[4] S. Siyang, C. Wongchoosuk, and T. Kerdcharoen, "Diabetes diagnosis by direct measurement from urine odor using electronic nose," in Biomedical Engineering International Conference (BMEiCON), 2012. IEEE, 2012, pp. 1–4.

[5] R. N. T. Shanty, R. H. Ginardi, and R. Sarno, "Interpretation of urine dipstick results based on color similarity using linear interpolation curve fitting," in Computational Intelligence and Cybernetics (CYBERNETICSCOM), 2013 IEEE International Conference on. IEEE, 2013, pp. 36–39.

[6] N. Yusuf, A. Zakaria, M. I. Omar, A. Y. M. Shakaff, M. J. Masnan, L. M. Kamarudin, N. A. Rahim, N. Z. I. Zakaria, A. A. Abdullah, A. Othman et al., "In-vitro diagnosis of single and poly microbial species targeted for diabetic foot infection using e-nose technology," BMC bioinformatics, vol. 16, no. 1, p. 158, 2015.

[7] S. Lekha and M. Suchetha, "Non-invasive diabetes detection and classification using breath analysis," in Communications and Signal Processing (ICCSP), 2015 International Conference on. IEEE, 2015, pp. 0955–0958.

[8] K. Yan, D. Zhang, D. Wu, H. Wei, and G. Lu, "Design of a breath analysis system for diabetes screening and blood glucose level prediction," IEEE transactions on biomedical engineering, vol. 61, no. 11, pp. 2787–2795, 2014.

[9] A. Arwan, M. Sidiq, B. Priyambadha, H. Kristianto, and R. Sarno, "Ontology and semantic matching for diabetic food recommendations," in 2013 International Conference on Information Technology and Electrical Engineering (ICITEE). IEEE, 2013, pp. 170–175.

[10] D. Guo, D. Zhang, and N. Li, "Monitor blood glucose levels via breath analysis system and sparse representation approach," in Sensors, 2010 IEEE. IEEE, 2010, pp. 1238–1241.

[11] A. Rashed and H. Santos, Odour user interface for authentication: Possibility and acceptance: Case study, Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2010, 17-19 March, 2010, Hong Kong, pp. 219–222.

[12] S. Qiu, L. Gao, and J. Wang, "Classification and regression of elm, lvq and svm for e-nose data of strawberry juice," Journal of Food Engineering, vol. 144, pp. 77–85, 2015.

[13] S. Omatu and M. Yano, "E-nose system by using neural networks," Neurocomputing, vol. 172, pp. 394–398, 2016.

[14] A. Quarto, D. Soldo, F. Di Lecce, A. Giove, V. Di Lecce, and A. Castronovo, "Electronic nose for evaluating water use in beer production," in Olfaction and Electronic Nose (ISOEN), 2017 ISOCS/IEEE International Symposium on. IEEE, 2017, pp. 1–3.

[15] J. He, L. Xu, P. Wang, and Q. Wang, "A high precise e-nose for daily indoor air quality monitoring in living environment," Integration, the VLSI Journal, vol. 58, pp. 286–294, 2017.

[16] T. Saidi, T. G. Welearegay, O. Zaim, O. G. Le´on, R. Ionescu, N. El Bari, and B. Bouchikhi, "Ability of discrimination of breath from smoker and non-smoker volunteers by using an electronic nose based on wo 3 nanowires and sno 2 sensors," in Olfaction and Electronic Nose (ISOEN), 2017 ISOCS/IEEE International Symposium on. IEEE, 2017, pp. 1–3.

[17] K. Yan and D. Zhang, "A novel breath analysis system for diabetes diagnosis," in Computerized Healthcare (ICCH), 2012 International Conference on. IEEE, 2012, pp. 166–170.

[18] R. Sarno, D. R. Wijaya et al., "Detection of diabetes from gas analysis of human breath using e-nose," in Information & Communication Technology and System (ICTS), 2017 11th International Conference on. IEEE, 2017, pp. 241–246.

[19] D. R. Wijaya, R. Sarno, and E. Zulaika, "Gas concentration analysis of resistive gas sensor array," in 2016 International Symposium on Electronics and Smart Devices (ISESD). IEEE, 2016, pp. 337–342.

[20] D. Patra, M. K. Das, and S. Pradhan, "Integration of fcm, pca and neural networks for classification of ecg arrhythmias." IAENG International Journal of Computer Science, vol. 36, no. 3, pp. 229–233, 2009.

[21] S. Adhikari, R. Amini, L. A. Stolz, and M. Blaivas, "Impact of point-of-care ultrasound on quality of care in clinical practice," Reports in Medical Imaging, vol. 7, pp. 81–93, 2014.

[22] O. Herman-Saffar, Z. Boger, S. Libson, D. Lieberman, R. Gonen, and Y. Zeiri, "Early non-invasive detection of breast cancer using exhaled breath and urine analysis," Computers in biology and medicine, vol. 96, pp. 227–232, 2018.

[23] Q. Li, F. Yang, L.-S. Liu, Z.-Z. Zheng, X.-J. Lin, and Q.-H. Wu, "Classification of diabetes disease using tcm electronic nose signals and ensemble learning," in Computer Science & Education (ICCSE), 2014 9th International Conference on. IEEE, 2014, pp. 507–511.

[24] S. I. Sabilla, R. Sarno, and J. Siswantoro, "Estimating gas concentration using artificial neural network for electronic nose," Procedia Computer Science, vol. 124, pp. 181–188, 2017.

[25] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural networks, vol. 61, pp. 85–117, 2015.

[26] A. Kamilaris and F. X. Prenafeta-Bold´u, "Deep learning in agriculture: A survey," Computers and Electronics in Agriculture, vol. 147, pp. 70–90, 2018.

[27] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2016.

[28] A. Koutsoukas, K. J. Monaghan, X. Li, and J. Huan, "Deep-learning: investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data," Journal of cheminformatics, vol. 9, no. 1, p. 42, 2017.

[29] J. W. Gardner and P. N. Bartlett, "A brief history of electronic noses," Sensors and Actuators B: Chemical, vol. 18, no. 1-3, pp. 210–211, 1994.

[30] W. Ping, T. Yi, X. Haibao, and S. Farong, "A novel method for diabetes diagnosis based on electronic nose1," Biosensors and Bioelectronics, vol. 12, no. 9-10, pp. 1031–1036, 1997.

[31] E. I. Mohamed, R. Linder, G. Perriello, N. D. Di, S. P¨oppl, and A. L. De, "Predicting type 2 diabetes using an electronic nose-based artificial neural network analysis." Diabetes, nutrition & metabolism, vol. 15, no. 4, pp. 215–221, 2002.

[32] R. Arasaradnam, N. Quraishi, I. Kyrou, C. Nwokolo, M. Joseph, S. Kumar, K. Bardhan, and J. Covington, "Insights into 'fermentonomics': evaluation of volatile organic compounds (vocs) in human disease using an electronic 'e-nose'," Journal of medical engineering & technology, vol. 35, no. 2, pp. 87–91, 2011.

[33] A. Kumar, I. Singh, and S. Sud, Indoor Air Quality Estimation by Using Smart Sensing System, Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2009, 18-20 March, 2009, Hong Kong, pp1218-1220.

[34] P. R. Galassetti, B. Novak, D. Nemet, C. Rose-Gottron, D. M. Cooper, S. Meinardi, R. Newcomb, F. Zaldivar, and D. R. Blake, "Breath ethanol and acetone as indicators of serum glucose levels: an initial report," Diabetes technology & therapeutics, vol. 7, no. 1, pp. 115–123, 2005.

[35] P. Paredi, W. Biernacki, G. Invernizzi, S. A. Kharitonov, and P. J. Barnes, "Exhaled carbon monoxide levels elevated in diabetes and correlated with glucose concentration in blood: a new test for monitoring the disease?" Chest, vol. 116, no. 4, pp. 1007–1011, 1999.

[36] M. Phillips, R. N. Cataneo, T. Cheema, and J. Greenberg, "Increased breath biomarkers of oxidative stress in diabetes mellitus," Clinica Chimica Acta, vol. 344, no. 1-2, pp. 189–194, 2004.

[37] B. Novak, D. Blake, S. Meinardi, F. Rowland, A. Pontello, D. Cooper, and P. Galassetti, "Exhaled methyl nitrate as a noninvasive marker of hyperglycemia in type 1 diabetes," Proceedings of the National Academy of Sciences, vol. 104, no. 40, pp. 15 613–15 618, 2007.

[38] D. R. Wijaya, R. Sarno, and E. Zulaika, "Sensor array optimization for mobile electronic nose: Wavelet transform and filter based feature selection approach," International Review on Computers and Software, vol. 11, no. 8, pp. 659–671, 2016.

[39] L. Sharma and N. Mehta, "Data mining techniques: A tool for knowledge management system in agriculture," international journal of scientific & technology research, vol. 1, no. 5, pp. 67–73, 2012.

[40] J.-s. Wang, Y.-l. Ruan, B.-w. Zheng, and S.-z. Gao, "Face recognition method based on improved gabor wavelet transform algorithm." IAENG International Journal of Computer Science, vol. 46, no. 1, pp. 12–24, 2019.

[41] S. I. Sabilla and R. Sarno, "Development of wavelet transforms to predict methane in chili using the electronic nose," in 2017 International Conference on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation (ICAMIMIA). IEEE, 2017, pp. 271–276.

[42] H. Kim, B. Konnanath, P. Sattigeri, J. Wang, A. Mulchandani, N. Myung, M. A. Deshusses, A. Spanias, and B. Bakkaloglu, "Electronic-nose for detecting environmental pollutants: Signal processing and analog front-end design," Analog Integrated Circuits and Signal Processing, vol. 70, no. 1, pp. 15–32, 2012.

[43] K. Kunadumrongrath, and A. Ngaopitakkul, Discrete Wavelet Transform and Support Vector Machines Algorithm for Classification of Fault Types on Transmission Line, Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2012, 14-16 March, 2012, Hong Kong, pp1042-1046.

[44] D. R. Wijaya, R. Sarno, and E. Zsulaika, "Information quality ratio as a novel metric for mother wavelet selection," Chemometrics and Intelligent Laboratory Systems, vol. 160, pp. 59–71, 2017.

[45] J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu, "Deep learning for smart manufacturing: Methods and applications," Journal of Manufacturing Systems, vol. 48, pp. 144–156, 2018.

[46] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," The Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929–1958, 2014.

[47] K. Hara, D. Saitoh, and H. Shouno, "Analysis of dropout learning regarded as ensemble learning," in International Conference on Artificial Neural Networks. Springer, 2016, pp. 72–79.

[48] K. Janocha and W. M. Czarnecki, "On loss functions for deep neural networks in classification," arXiv preprint arXiv:1702.05659, 2017.

[49] B. J. Erickson, P. Korfiatis, Z. Akkus, T. Kline, and K. Philbrick, "Toolkits and libraries for deep learning," Journal of digital imaging, vol. 30, no. 4, pp. 400–405, 2017.

[50] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," Computers and Electronics in Agriculture, vol. 161, pp. 272–279, 2019.

[51] F. Chollet et al., "Keras," https://keras.io, 2015.

[52] A. S. Patil and B. Pawar, Automated Classification of Web Sites using Naive Bayesian Algorithm, Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2012, 14-16 March, 2012, Hong Kong, pp519-523.

[53] M. Lamy, R. Pereira, J. C. Ferreira, F. Melo, and I. Velez, "Extracting clinical knowledge from electronic medical records," IAENG International Journal of Computer Science, vol. 45, no. 3, pp. 488–493, 2018.

[54] N. M. Saad, N. Noor, A. Abdullah, S. Muda, A. Muda, and H. Musa, "Segmentation and classification analysis techniques for stroke based on diffusion-weighted images." IAENG International Journal of Computer Science, vol. 44, no. 3, pp. 388–395, 2017.

[55] T. N. Phyu, Survey of Classification Techniques in Data Mining, Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2009, 18-20 March, 2009, Hong Kong, pp727-731.