Self-Monitoring and Detection of Diabetes with Smart Toilet based on Image Processing and K-Means Technique

Lailis Syafaah*
Department of Electrical Engineering,
Faculty of Engineering,
Universitas Muhammadiyah Malang
Malang, Indonesia
lailis@umm.ac.id

Desy Fatkhi Azizah
Department of Electrical Engineering,
Faculty of Engineering,
Universitas Muhammadiyah Malang
Malang, Indonesia
desiazizah17@gmail.com

Inda Rusdia Sofiani
Department of Electrical Engineering,
Faculty of Engineering,
Universitas Muhammadiyah Malang
Malang, Indonesia
indarusdias05@umm.ac.id

Merinda Lestandy
Department of D3 Electrical Technology,
Facuty of Engineering,
Universitas Muhammadiyah Malang
Malang, Indonesia
merindalestandy@umm.ac.id

Amrul Faruq
Universitas Muhammadiyah Malang, Indonesia
Graduate Student, Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia,
Kuala Lumpur, Malaysia
faruq@umm.ac.id

***Abstract*— Diabetes is a disorder due to insulin cannot be produced by the pancreas. Furthermore, it may also occur because the body can not effectively use the insulin produced. Sometimes, blood samples are checked by piercing with a needle on the arm or finger three to four times a day. Taking blood to measure the level of sugar in diabetic patients can lead to infection because diabetics did not produce insulin. The authors therefore developed a smart toilet for the diagnosis of diabetes by the colour of urine. The aim of this tool is to make it easier for someone to monitor their health, particularly diabetes. This tool is use the camera as a colour sensor and use the raspberry pi as the main control tool. The smart toilet is installed with the camera and is designed to take an image to produce a red green blue (RGB) colour model. This RGB model is then clustered using the K-means method. Patients are also able to monitor their health by integrating the Internet of Things (IoT) platform. The experimental work is carried out by 20 sample test objects, the results obtained are of high quality and the tool can be used correctly. Whereas in terms of quantity, the urinary colour detection error is about 5%. It can be observed that developed smart toilet is an option for monitoring diabetic activities and controlling the patients glucose levels.**

Keywords—diabetes, urine, internet of things, K-means, smart toilet

I. Introduction

The Internet of Things (IoT) sometimes referred to as the Internet of Objects, will make a huge difference including ourselves. Nowadays, the IoT concept has been widely developed to assess human activities in order to improve their life and health conditions [1],[2]. IoT is one of the major communication advances technology that linked the internet with sensors, working devices, and capturing the data for scientific purposes. Smart devices, an IoT based including healthcare application is a promising tool for redesigning technological aspects, economic, and social prospects.

A number of healthcare application utilizing IoT technology has been proposed and developed over these years [3]. Smart devices implementing the IoT concept for diabetic detection and management are one of the most common technologies. This paper introduced an IoT based smart toilet health care application to detect diabetics through the colour of urine. This tool is IoT based urine detection to capture the RGB value of urine in order to mintor and assessment. Urinary assessment and culture therefore very important to track infection for each patients [4].

Diabetes Mellitus (DM) is a metabolic issue described as hyperglycaemia (high glucose) that is progressively affecting the world population [5]. Diabetes is a disorder because the pancreas is unable to produce insulin. It may also occur because the body can not effectively use the produced insulin. It is a hormone whose job is to balance blood sugar levels. If insulin can not be produced properly, it can cause hyperglycemia, i.e. an increase in the blood glucose concentration [6].

Four non-communicable diseases can cause the highest death, one of them is diabetes mellitus. Three other diseases, namely cancer, blood vessels, and heart. Several hereditary diseases such as hypertension and heart disease may also be caused by diabetes.

Diabetes can be identified by symptoms such as blurred vision, leg pain, and numbness, changes in the skin, red and swollen gums, range of disease, healing in slow wounds. Diabetes prevention can be done to regulate a healthy lifestyle, exercise on a regular basis, reduce sugar consumption, consume a lot of fiber, reduce smoking and change other bad habits.

Funding: Under the 2020-Development of Scientific Research Work for Doctoral degree, research scheme of Universitas Muhammadiyah Malang

Currently, the number of diabetics in the world is around 285 million and is expected to increase to 366 million in 2030 [3]. In Indonesia, more than 7 million people had diabetes in 2012, and 84% did not know that they had diabetes [7].

For control of sugar levels, blood samples are usually perforated three or four times a day with a needle on the arm or finger. Blood can cause infection as the diabetic do not produce insulin by measuring the levels of sugar in diabetic patients. Another way to measure blood sugar levels is by using saliva and urine. In this study, the authors used urine as a result of a high correlation of 0.99 glucose levels in the urine compared to blood sugar levels [8].

In general, normal urine has a value that tends to be clear and has a higher RGB value compared to urine with diabetes. Previous studies found that the RGB value in normal human urine was R = 227-233, G = 258-2279, B = 273-307. Whereas in the urine of diabetics, urine tends to be more turbid and has a lower RGB value than normal urine. Previous studies have the RGB value of R = 205-222, G = 242 = 261, B = 243-261 [9].

A toilet is a place that everybody goes to. Many parents are abandoned today by their children because of work and other things. Therefore, the health conditions of the parents could difficult for a child to monitor. Older people are highly vulnerable to certain conditions, such as diabetes. The cause of diabetes is a number of factors, such as heredity, lack of exercise, smoking, poor eating patterns, and so on. The authors therefore want to create a smart toilet that can make it easier for children to monitor parents' health, especially in the case of diabetes.

This smart toilet uses colors to detect diabetes-affected person's urine. This urine test is based on particular specifications for urine colour. In previous studies using the TCS230 sensor to detect the colour of urine [9], while this study uses a camera to get more accurate results. In this study, the RGB approach is to detect red, green, and blue colours using the RGB method. This study uses the K-means method for clustering the colour of urine as an indicator of diabetes.

II. METHODS

A. Structural Design of Hardware and Software

This smart toilet is designed with a prototype that uses a white background to get maximum exposure to urine from the camera. This system generates a reading output that is displayed on the web.

The process performed to make this smart toilet can be seen from the block diagram in Figure 1. First, there is a camera to detect the difference in urine colour which is accommodated in the toilet, urine as a medium that digital image processing is identified. In addition, a microcontroller that uses raspberry pi to process data and system controllers. After that, the data then provided to the user, which can be viewed on the remote interface web monitoring.

B. Data Collection and Processing using K-means Technique

In this study, using the k-means clustering method, several stages, such as pre-processing data, must first be done before the data is clustered. These are the following stages:

Data Collection. The data collected in this study is to use the data obtained in the urine of a person with diabetes and in the urine of a normal person. Sugar content values have also been taken as a correlation in this study to match the data taken from the urine RGB value. Twelve data samples has been carried out in this study, advancement from the previous study [9]. The data contains 10 sample with diabetic patients, and 10 samples with a normal patients. The result of RGB data samples is shown in Table 2.

Data Pre-processing. Pre-processing data is a step to separate data based on the type of data. In this study, data are available in the form of numbers obtained from the RGB value of the urine. Data is processed for clustering based on the predetermined value of 2 clusters, i.e. diabetes urine and normal urine. Pre-processing data is an advanced stage after data collection. At this stage, the data shape is changed so that it becomes binary data. In this study, the data collected did not experience missing data, or there were blank data, so no missing data was performed. In this case, the data was divided into 80% training data and 20% as testing data.

Data Processing. A data processing with clustering method is done using the k-means. To do clustering using k-means, the following steps are required. First, determine the desired number of clusters. In this study, the cluster is divided into two, the urine of normal people and the urine of diabetics. Next, determine the initial centre point of the cluster or the centroid. Choose from random data to determine the centroid cluster. Each data is measured by the distance to the centroid. The grouping is based on the minimum distance between the data and the centroid cluster.

The centroid recalculation is performed by finding the middle value of each cluster once it has been obtained. After obtaining the new centroid, the distance between the object and the centroid is recalculated. Next, an iteration is performed to determine whether the k-means calculation is appropriate. The iteration will stop if the members of the cluster remain or remain unchanged. Clustering is a method for grouping data with similarities or predetermined characteristics. Clusters are groups of data that are similar in the same group and superimposed on different cluster objects. The objects are clustered into one or more objects so that they have a high similarity between each other in one cluster [5].

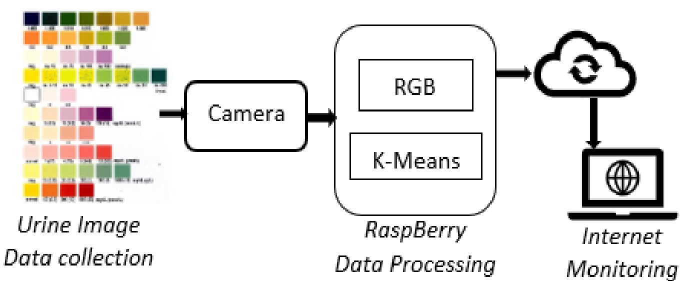

Fig. 1. Illustrated hardware and software design

K-Means is a data clustering method that divides existing data into one or more clusters so that data with the same characteristics are grouped into one cluster and data with different characteristics are grouped into other groups [10].

K-Means is a distance-based clustering method that divides data into a number of clusters and algorithms that only work on numerical data. This method separates data into clusters so that data with the same characteristics are grouped into the same cluster [11]. The clustering process begins by recognizing the data to be clustered, Xij (i = 1, ..., n; j = 1, ..., m) where n is the amount of data to be clustered while m is the number of variables. The centre of each cluster is randomly assigned at the beginning of the iteration, Ckj (k = 1, ..., k; j = 1, ..., m). The distance between each data is then calculated with each centre of the cluster. Calculation of the distance of the is data (xi) at the center of the $k-th$ cluster named (dik), can be written in the formula, as in Equation (1):

$$d_{ik} = \sqrt{\sum_{j=1}^{m}(x_{ij} - c_{ij})^2} \quad (1)$$

The data is entered in the cluster if the distance of the data to the k-th cluster centre has the smallest value compared to the distance to the other cluster centre. Calculation can be made using Equation (2).

$$\min\sum_{k=1}^{k} d_{ik} = \sqrt{\sum_{j}^{m}\left(x_j - c_{ij}\right)^2} \quad (2)$$

III. Result and Discussion

Testing is done in several stages, the testing phase consists of hardware testing and software testing. Testing the camera sensor is done several times with a white background.

This hardware is tested to measure the sensors supports this system, such as a camera, as a colour sensor. The camera used for this study is a USB camera module. This camera takes urine images from the toilet prototype that has been made and the image has an RGB image value is processed on raspberry. To take pictures manually, the user must press the button. This prototype also has a buzzer to find out if the camera has been able to take pictures.

The camera as can be seen in Figure 2 on this prototype is on the side of the toilet hole where it holds the urine. The purpose of the camera is to make it easy to shoot. In order to dispose of urine, there is a 5v dc motor works to drain the urine. It is pressed the motor button until the urine is drained.

Fig. 2. Illustrated toilet with smart camera

A. Clustering with K-means

The first data collection process that needs to be done is to take samples of urine from diabetics and healthy people. The urine sample is placed in the toilet prototype that has been produced, the urine sample used in this study is 10 urine with diabetes and 10 urine from healthy people. There are no missing data or blank data in data collection, therefore the data is divided into two parts, namely test data and training data. In this case, the data is divided into 80% training data and 20% as testing data.

Using Equation (1) we can determine the cluster distance according to the data RGB value as collected. For instance, according to Equation (1), we have calculated the cluster minimum distance for data #1 and #7 as shown in Table I. And Table II shown the completed data with the clustering grub value according to the minimum distance value.

TABLE I. Clustering Calculated Data

Cluster number	Centroid		Cluster grub
	X	Y	
K1 (117,125,28)	0	71,93	1
K2 (110,110,98)	71,93	0	0

TABLE II. Clustering and RGB Urine Data

R	G	B	C1	C2	Min Distance	Cluster
117	125	28	0,00	71,93	0,00	1
116	118	37	11,45	61,81	11,45	1
124	123	58	30,87	44,33	30,87	1
117	111	15	19,10	83,30	19,10	1
135	128	48	27,07	58,73	27,07	1
126	122	49	23,04	52,92	23,04	1
110	110	98	71,93	0,00	0,00	0
110	118	103	75,65	9,43	9,43	0
102	108	98	73,58	8,25	8,25	0
118	112	17	17,06	81,42	17,06	1
110	118	117	89,55	20,62	20,62	0
116	123	78	50,05	24,60	24,60	0
108	100	99	75,81	10,25	10,25	0
105	110	102	76,45	6,40	6,40	0
128	118	16	17,72	84,33	17,72	1
132	120	14	21,12	87,41	21,12	1
105	100	101	78,09	11,58	11,58	0
108	101	96	72,67	9,43	9,43	0
112	100	101	77,32	10,63	10,63	0
114	100	117	92,49	21,84	21,84	0

In Figure 3, the detailed patient data contains required information, shows individual patient data, and there are generated RGB values generated. The accuracy of the system can be measured by comparing with the actual result from k-means and the output from the IoT based toilet system. Table 3 shows these results.

Sample of data obtained 80% of the total data used as training data and 20% of the data used as test data. The results of the clustering diagram obtained as shown in Figure 4. The red color indicates normal urine, while the blue color indicates diabetes urine and yellow color as centroid or middle value.

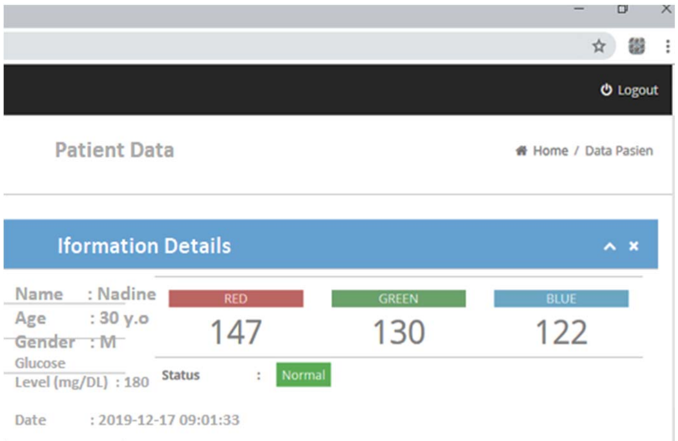

Fig. 3. Internet monitoring of diabetics recorded by the color of urine image processing

TABLE III. RESULT ON IOT CLUSTERING SYSTEM

Sample number	*Clustering*		Remark
	Calculation	IoT measured	
1	1	1	TRUE
2	1	1	TRUE
3	1	1	TRUE
4	1	1	TRUE
5	1	1	TRUE
6	1	1	TRUE
7	0	0	TRUE
8	0	0	TRUE
9	0	0	TRUE
10	1	1	TRUE
11	0	0	TRUE
12	0	1	FALSE
13	0	0	TRUE
14	0	0	TRUE
15	1	1	TRUE
16	1	1	TRUE
17	0	0	TRUE
18	0	0	TRUE
19	0	0	TRUE
20	0	0	TRUE

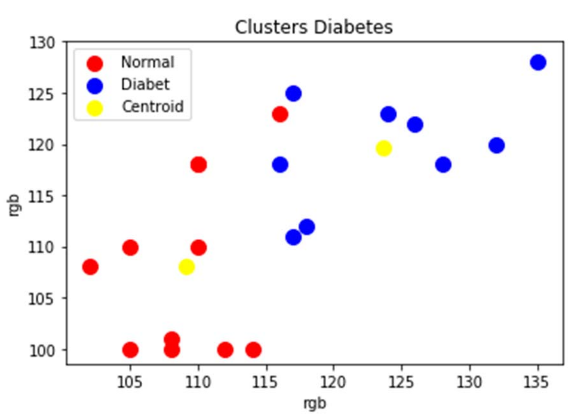

Fig. 4. Clustering chart result

From the results in Table 3, error performance of the IoT based toilet can be measured with total number of error devided with te total number of data respectively are 1 and 20. The error is about 5%, and the accuracy of the system is about 95%.

IV. Conclusion

This chapter conluded exerimentaly study, namely the detection of diabetics using urine with digital image processing and k-means algorithm to detect the urine-color level. This chapter also explains the suggestions made on the system.

Several conclusions are drawn from the experiments that have been carried out in this thesis, namely smart toilets: self-monitoring and diabetes detection, including: (1) Smart toilets to detect diabetes that have been successfully performed using a prototype toilet and certain components such as a high-end raspberry version along with the program. (2) In this study, clustering using the k-means method was carried out successfully, namely the clustering of diabetic urine and normal urine. (3) Use the encoder command in the source code to convert binary data to data labels.

In advance, a smart toilet should detect the urine of diabetics needed several parameters so that the results obtained are more accurate. If the data obtained can only be less accurate using the camera as a colour sensor, then it should be added to the ammonia sensor so that the resulting data has multiple parameters and the data obtained is more accurate. The toilet colour should be white to prevent interruption of the sensor reading. Toilet lighting must be stable so that the sensor readings can be accurate. In addition to the detection of diabetes by the use of urine, patients should also be subject to further tests by the doctor or may also be monitored by the doctor.

Acknowledgment

The authors wish to express their gratitude to the Faculty of Engineering, Universitas Muhammadiyah Malang (UMM) for supporting and facilitating this study. The first author acknowledges the Directorate of Research and Community Services UMM under the Development of Scientific Research Work for Doctoral scheme to undertaking this work.

References

[1] S. Deshkar and V. G. Menon, “A Review on IoT based m-Health Systems for Diabetes,” *Int. J. Comput. Sci. Telecommun.*, vol. 8, no. 1, pp. 13–18, 2016.

[2] R. Vasanthakumar, K. D. Darsini, S. Subbaiah, and K. Lakshmi, “IoT for monitoring diabetic patients,” *Int. J. Adv. Res.*, vol. 4, no. 2, pp. 2149–2157, 2018.

[3] A. J. Jara, M. A. Zamora, and A. F. G. Skarmeta, “An Internet of things-based personal device for diabetes therapy management in ambient assisted living (AAL),” *Pers. Ubiquitous Comput.*, vol. 15, no. 4, pp. 431–440, 2011.

[4] Z. Y. Yildirim *et al.*, “Urine levels of matrix metalloproteinases and tissue inhibitor of metalloproteinases in children with type 1 diabetes mellitus,” *JCRPE J. Clin. Res. Pediatr. Endocrinol.*, vol. 11, no. 2, pp. 157–163, 2019.

[5] W. L. Yun and M. R. K. Mookiah, “Detection of diabetic retinopathy using K-means clustering and self-organizing map,” *J. Med. Imaging Heal. Informatics*, vol. 3, no. 4, pp. 575–581, 2013.

[6] N. Sneha and T. Gangil, “Analysis of diabetes mellitus for early prediction using optimal features selection,” *J. Big Data*, vol. 6, no. 1, 2019.

[7] Z. P. Shaw JE, Sicree RA, “Global estimates of the prevalence of

diabetes for 2010 and 2030," *Diabetes Res Clin Pr.*, vol. 87, pp. 4–14, 2010.

[8] D. J. Newman *et al.*, "Systematic review on urine albumin testing for early detection of diabetic complications," *Health Technol. Assess. (Rockv).*, vol. 9, no. 30, 2005.

[9] R. Budianto, "Prototype Urine Analyzer Telemetry Menggunakan Sensor Warna Untuk Mendeteksi Penyakit Diabetes Pada Seseorang," *J. Elektron. Pendidik. Tek. Elektron.*, vol. 7, no. 4, pp. 79–85, 2018.

[10] M. Nair and D. Mishra, "Classification of diabetic retinopathy severity levels of transformed images using K-means and thresholding method," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 4, pp. 51–59, 2019.

[11] S. B. Manojkumar and H. S. Sheshadri, "Classification and detection of diabetic retinopathy using K-means algorithm," *Int. Conf. Electr. Electron. Optim. Tech. ICEEOT 2016*, pp. 326–331, 2016.