Application of Artificial Neural Network for Type 2 Diabetes Mellitus Detection Using Buccal Cell Images

Priyanka Kusuma Wardhani¹, Prihartini Widiyanti¹,²*, Franky Chandra Satria Arisgraha¹

1. Biomedical Engineering Study Program, Faculty of Science and Technology, Universitas Airlangga, Indonesia.
2. Institute of Tropical Disease, Universitas Airlangga, Indonesia.

Abstract

Diabetes mellitus (DM) is a metabolic disease causing hyperglycemia due to insulin action anomalies. DM can cause cellular changes, including buccal cell. Blood tests are used to diagnose diabetes, so non-invasive test is required for diagnosis of diabetes.

Accordingly, this research aims to design non-invasive system based on artificial neural network for type 2 DM detection using buccal cell images. Buccal cells smear were obtained from 30 subjects suffering from type 2 DM and 30 normal subjects. The smears were stained by using Papanicolaou method. Each slide were observed under digital microscope and were evaluated. The system was designed by using MATLAB with image processing and Probabilistic Neural Network (PNN) algorithm to classify features. Buccal cell images were segmented to get features. The features used in this study were nucleus area, nucleus perimeter and nucleus circularity.

Nucleus areas and perimeters in type 2 DM group were higher than those in control group with similar nucleus roundness in both groups. Forty nucleus feature datasets were used for training process, while 20 nucleus feature datasets were used for testing process.

The optimal PNN value was 0.4 g constant. The optimal accuracy of training was 92.5%, while the optimal accuracy of testing was 90%.

Keywords: Type 2 diabetes mellitus, buccal cell images, artificial neural network.

Introduction

Diabetes prevalence is high and tends to increase sharply. In 2013, International Diabetes Federation (IDF) noted that there were 382 million people with diabetes and predicted to increase 55% in 2035. According IDF (2013), Indonesia had 8.5 million people with diabetes. Diabetes Mellitus (DM) is an endocrine disorder in the metabolism of carbohydrates, lipids, and proteins due to deficiency of insulin in the body, the target tissue resistance to insulin, or both. Until now, there has been no therapy that can cure diabetes completely, but DM can still be controlled. DM is divided into two conditions, type 1 DM and type 2 DM. The number of patients with type 2 DM is growing faster as obesity rate is increasing, while human physical activity is decreasing.

Long-term high blood glucose causes diabetes patients prone to complications. A number of complications often occur in patients with DM, such as, microvascular abnormalities, coronary heart disease, and macrovascular disorders, such as stroke, heart disease systemic, and gangrene. Soft tissue abnormalities have been reported to be associated with diabetes mellitus in the oral cavity such periodontal diseases (periodontitis and gingivitis); salivary dysfunction leading to a reduction in salivary flow and changes in saliva composition, and taste dysfunction. Oral fungal and bacterial infections have also been reported in patients with diabetes. There are also reports of oral mucosa lesions in the form of stomatitis, geographic tongue, benign migratory glossitis, fissured tongue, traumatic ulcer, lichen planus, lichenoid reaction and angular cheilitis. In addition, delayed mucosal wound healing, mucosal neuro-sensory disorders, dental carries and tooth loss has been reported in patients with diabetes.
the chance of developing oral mucosal lesions were found to be higher in patients with diabetes compared to healthy controls. Diabetic patients have more severe periodontitis and Diabetes Mellitus negatively affects oral health as reflected by the higher decay, missing, filling, total score (DMFT) scores in diabetic patients.

Early diagnosis for DM is one of the important aspects to reduce the possibility of complications. Diagnosis of DM using blood tests is the most widely used and accurate. However, this method gives a psychological effect on patients, such as fear or discomfort due to blood sampling. In addition, this method also requires considerable time in blood analysis, so it will be useful to find a new tool that can replace, complement, or compare the results obtained, yet still provides patient with greater comfort.

Several soft tissue abnormalities have been reported to be associated with diabetes mellitus in the oral cavity. These complications include periodontal diseases (periodontitis and gingivitis); salivary dysfunction leading to a reduction in salivary flow and changes in saliva composition, and taste dysfunction. Oral fungal and bacterial infections have also been reported in patients with diabetes. There are also reports of oral mucosa lesions in the form of stomatitis, geographic tongue, benign migratory glossitis, fissured tongue, traumatic ulcer, lichen planus, lichenoid reaction and angular cheilitis.

In addition, delayed mucosal wound healing, mucosal neuro-sensory disorders, dental carries and tooth loss has been reported in patients with diabetes. The prevalence and the chance of developing oral mucosal lesions were found to be higher in patients with diabetes compared to healthy controls.

Several clinical and paraclinical techniques are available for evaluation of oral mucosa changes. Oral cytology is one of exfoliative cytology method, the study of cells taken or smeared on the surface epithelium of organ. Previously, exfoliative cytology method was very limited because of observer's subjectivity, so the possibility of wrong diagnosis was very large. But now, it can be overcome with quantitative methods, such as cytomorphometric analysis.

Cytomorphometric is a quantitative measurement of cell using graphic interface. Cytomorphological features include nucleus (nucleus size, shape, or membrane) and cytoplasmic qualities. Some characteristics are unique to certain disease processes, aiding greatly in diagnosis. Oral cytology has been used for early detection of premalignant and malignant oral lesion, microbial disease, Fe-deficiency, and cell proliferation in smoker.

Some of exfoliative cytology studies show there are changes in buccal mucosa cells and epithelial cells in mouth in the case of type 2 DM and it can be detected by cytomorphometric analysis.

According to Sonawane et al exfoliative cytology has potencies to be developed into a diagnostic method in type 2 DM. However, the previous studies by Sonawane et al, Rivera and Nunez-de-Mendoza, Suvarna et al, Ahmed and Garib and Jajarm et al provide quantitative value on nucleus, cytoplasm wide, and nucleus-cytoplasm ratio. Based on the previous studies, we conducted a study to classify buccal epithelial cells for non-invasive diagnosis of type 2 DM. This study is expected to build an application that is able to detect type 2 DM through buccal cell image and artificial neural network (ANN).

Materials and methods

Subjects were randomly selected from adult patients at the Department of Diabetes, Dr. Ramelan Marine Hospital. The subjects were divided into two groups, consisting of 30 subjects with type 2 DM and 30 non-diabetic subjects as the control group.

The subjects with smoking and tobacco chewing habits, habitual alcohol consumption, oral sepsis presence, other systemics disease presences, and non-diabetic medication consumption were excluded from this study.

Previous studies have demonstrated that cellular and nuclear size are influenced by these factors [Rivera].

The subject was rinsed with NaCL 0.9% to sterilize the buccal cavity. The surface of buccal mucosa was smeared with cytobrush and transferred to a clean and dry glass slide.

The slide was immediately fixed in 96% alcohol for 30 minutes and stained by using Papanicolaou technique. Each slide was observed under microscope (Olympus CX21) at 400× magnification.
Gaussian Low Pass Filter. The transfer function of GLPF is given as follow:\(^{19}\):

\[
H(u, v) = e^{-D^2(u, v)/2D_0^2}
\]

(2)

\(D_0\) is specified non-negative number. In this study, \(D_0\) is down to 0.7. Then, the image is segmented to get nucleus region.

Identification buccal cell of type 2 DM or is actually based on the nucleus’s geometric shape features. The features used in this study were nucleus area, perimeter, and circularity. The nucleus area can be calculated by adding the amount of the nucleus pixels, while the nucleus perimeter can be determined by adding the amount of the edge of nucleus pixels. Hence, the nucleus circularity can be defined as below:\(^{19}\):

\[
\text{circularity} (R) = \frac{4\pi A(R)}{p^2(R)}
\]

(3)

The grayscale image should be enhanced by using contrast stretching, brightness, and filtering process. The filter used in this study is Gaussian Low Pass Filter. The transfer function of GLPF is given as follow:\(^{19}\):

\[
H(u, v) = e^{-D^2(u, v)/2D_0^2}
\]

(2)

\(D_0\) is specified non-negative number. In this study, \(D_0\) is down to 0.7. Then, the image is segmented to get nucleus region.

Identification buccal cell of type 2 DM or is actually based on the nucleus’s geometric shape features. The features used in this study were nucleus area, perimeter, and circularity. The nucleus area can be calculated by adding the amount of the nucleus pixels, while the nucleus perimeter can be determined by adding the amount of the edge of nucleus pixels. Hence, the nucleus circularity can be defined as below:\(^{19}\):

\[
\text{circularity} (R) = \frac{4\pi A(R)}{p^2(R)}
\]

(3)

The grayscale image should be enhanced by using contrast stretching, brightness, and filtering process. The filter used in this study is Gaussian Low Pass Filter. The transfer function of GLPF is given as follow:\(^{19}\):

\[
H(u, v) = e^{-D^2(u, v)/2D_0^2}
\]

(2)

\(D_0\) is specified non-negative number. In this study, \(D_0\) is down to 0.7. Then, the image is segmented to get nucleus region.
After the nucleus features of all three cells on each slide were measured, the averages were derived from each individual, so there were three features in a feature dataset of each individual. The features were then used as an input of classification. The classification step used supervised classification algorithm, namely, Probabilistic Neural Network (PNN).

It was divided into two steps, namely training and testing. In this study, 20 feature datasets of each group were used for training process; 10 feature data sets of each group were used for testing process. A parameter that would be varied is $g$. The performance of the algorithm can be evaluated by computing the percentages of Accuracy, the respective definitions are as follow:

$$\text{Accuracy} = \frac{\Sigma(TP+TN)}{n} \times 100\%$$

Where $TP$ is the number of true positives, $TF$ is the number of true negatives, and $n$ is the number of data.

**Results**

**Image Processing**

Before buccal cells of patients were classified as with type 2 diabetes or normal, the buccal cell image had been conducted first by using digital image processing to obtain features. Next, the image of the buccal cell obtained from buccal cell slide were observed by using a digital microscope on the computer. Consequently, three cells on each slide were evaluated from three different points. Thus, according to Rivera and Nunez de Mendoza, it takes at least more than one cell to be able to make a diagnosis of a disease. Since the number of the subjects in this study was 60 subjects consisting of 30 DM and 30 normal subjects, and each subject was represented by three cells, the total image of the buccal cells used was 180.

Next, each image was cropped to the size of 180×180 pixels. Each image consisted of single buccal cell. There were three stages of digital image processing conducted to obtain features, such as grayscaling, image preprocessing, and segmentation. Figure 4. shows the stage depicting how a segmentation occurred in the nucleus. The first stage was the process of converting RGB images into grayscale image through grayscaling process. RGB images often need to be converted into a grayscale image because many image processing operations work in gray scale.

Moreover, there were several processes performed in image preprocessing, ie contrast stretching, brightening, and filtering. Contrast stretching and brightening aim to make images brighter and more contrast to facilitate the segmentation stage further. However, there were some visible noises on cells with certain intensity almost equal to the nucleus intensity. This could disrupt the segmentation process, so filtering process was needed to remove the noise. Filter used in this study was Gaussian Low Pass Filter (GLPF). GLPF has a constant variable as D0. Thus, the smaller the value D0 is, the more blurred image is obtained. In this study, the proper value of D0 was 0.07. At this value, noise found in the cell could be obscured, but did not affect the nucleus.

Furthermore, segmentation stage is the first step in the acquisition of features and image classification. Segmentation was performed to obtain nucleus region. The segmentation process
conducted was thresholding and nucleus edge detection. Thresholding was conducted with a limit of 140. The nucleus region was segmented into a white color. Edge detection was conducted to obtain nucleus edge.

![Image Processing](image1)

**Figure 4.** Image Processing.

The third image of the cells representing each subject was calculated and averaged. Each group consisted of nucleus features, perimeter, and circularity. It was used as training dataset and testing dataset for classification. Figure 5 is a distribution graph of subjects with type 2 DM and Normal for each nucleus feature area, perimeter, and circularity.

![Distribution Graph](image2)

**Figure 5.** Distribution graph of subjects with type 2 DM and Normal for each nucleus feature area, perimeter, and circularity.

### Discussion

#### Analysis Cytomorphometric

There were different nucleus features between the type 2 DM group and the normal group. Table 1 shows data of each group. The average nucleus area of the type 2 DM group was $252.23 \pm 41.03$, while the average nucleus area of the normal group was $165.54 \pm 24.96$. The average nucleus perimeter of the type 2 DM group was $68.59 \pm 5.84$, while the average nucleus perimeter of the normal group was $56.4 \pm 4.76$. The average nucleus circularity of the type 2 DM group was $0.67 \pm 0.01$, while the average nucleus circularity of normal group was $0.66 \pm 0.03$.

![Table 1](image3)

**Table 1.** Cytomorphometrical analysis of the type 2 DM group and the normal group.

Nucleus areas and perimeters in the type 2 DM group were higher than those in the control group. But, nucleus roundness in both group were similiar. According to Jajarmi\(^7\), one of factors causing the increasing of buccal cell...
nucleus of type 2 DM is inflammatory. High glucose levels creates the extracellular matrix in the body with a higher concentration. These conditions then can lead the extracellular liquid into the cell and provide a response in the nucleus, causing inflammation.

**Classification**

PNN algorithm was made in accordance with the flowchart in Figure 3. The advantage of PNN is that testing time needed is faster than Backpropagation ANN. PNN has various parameters of constant g value. The constant g will affect the value of smoothing parameter ($\sigma$). Smoothing parameter then affects the width of Gaussian curve and influences probability function at each sample point. If the value of $\sigma$ is too small, information on the classes will be lost, but if the value of $\sigma$ is too wide, the data can become blurry and detailed density will be biased. Selection of the value of $\sigma$ is usually conducted by trial and error of the variation of the constant g.\(^{13}\)

PNN testing process in this study used 10 feature datasets of the type 2 diabetes group and 10 feature datasets of the normal group that were not used in the training process. Those feature datasets would be classified as type 2 diabetes if the value of DM probability was greater than the value of normal probability and vice versa. The accuracy level of the classification results of network PNN then was determined by comparing to the results of the doctor's diagnosis.

Figure 6 is a graph showing the comparison of the training data to the test data. The best accuracy of the testing dataset was 90%, while the best accuracy of the training dataset was 92.5% with the constant g value from 0.36 to 0.45. Accuracy does not reach 100% because the value of $\sigma$ can not provide the appropriate grade estimation. Another possibility obtained by Gaussian curve was too wide so that the sample data information became blurred causing some sample points not classified correctly. In addition, an error in the training could also be caused due to the less number of the training data. This is based on research Anggraeny and Hatta\(^{18}\) that the number of training data affects the accuracy of the system.

![Figure 6](image_url)

**Conclusions**

1. Cell detection program of type 2 diabetes mellitus using buccal image was created based on digital image processing and Artificial Neural Network (ANN). Identification buccal cell of type 2 DM or normal was based on the nucleus’s geometric shape features. The features used in this study were nucleus area, perimeter, and circularity. A feature vector is a set of average features from buccal cell nucleus that was used as an input of classification. The algorithm used is JST Probabilistic Neural Network (PNN).

2. The classification used supervised classification algorithm, namely Probabilistic Neural Network (PNN). From the results of the training of PNN, the value of the constant g is located at the most optimal range up to 0.36 to 0.45. Konstana g, which is used for the classification is 0.40 because those values are in the middle range is optimal.
3. Optimal accuracy of the resulting program using 20 test data with the value of the constant \( g = 0.40 \) by 90%.

Acknowledgements

The author would like to thank Diabetic Clinic and Pathology Anatomy Laboratory Navy Hospital Dr Ramelan Surabaya for permission and technical assistance. We are also grateful to our patients for the participation in this study.

Declaration of Interest

The authors report no conflict of interest.

References


