International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-7 Issue-4s, November 2018

Non-Invasive Diabetes Mellitus Detection Using Facial Block Color

S. Sathyavathi, K. R. Baskaran, S. Kavitha

Abstract: ***Diabetes Mellitus (DM) is a condition in which glucose level in the body is much higher than the normal. The traditional way to diagnosis DM is Fasting Plasma Glucose (FPG) test. As this method is slightly painful and uncomfortable several another method which are more comfortable and non-invasive are found. In this paper, we propose a new non-invasive method to detect DM based on facial block color features using various classification algorithms. Facial images are first captured using a specially designed non-invasive device, and calibrated to ensure consistency in feature extraction and analysis. Four facial blocks are extracted automatically from face image and used to represent a face features. A facial color gamut is constructed with six color centroids (red, yellow, light yellow, gloss, deep red, and black) to compute a facial color feature vector, characterizing each facial block. Finally, the features are classified using J48. ForJ48, two sub dictionaries, a Healthy facial color features sub dictionary and DM facial color features sub dictionary, are employed in the classification process. Apart from this we also use ZeroR, Support vector machine (SVM)[8] ,J48 to determine the accuracy, precision and recall using the data set that comprises of healthy and DM samples. Finally, we compare all these algorithms and choose the efficient one using its accuracy level.***

Keywords: ***Non-Invasive, Algorithm Efficiency, Health Enhancement***

I.INTRODUCTION

Diabetes mellitus is a group of metabolic diseases characterized by hyperglycemia resulting from defects in insulin secretion, insulin action, or both. The chronic hyperglycemia of diabetes will lead to long-term damage, dysfunction, and failure of various organs, especially the eyes, kidneys, nerves, heart, and blood vessels. Diabetes will lead to pathogenic processes. Two main types of DM: insulin dependent diabetes and non-insulin dependent diabetes. People with insulin dependent diabetes fail to produce insulin, and therefore require injections of it. Non-insulin dependent diabetes can be categorized by insulin resistance and is the most common type. Currently, there is no cure for insulin dependent diabetes or non-insulin dependent diabetes. The method practised commonly for diagnosing DM is plasma glucose (FPG) test. This method of diagnosing is accurate, it can be considered invasive, and slightly painful. Hence there is a need to develop a non-invasive yet accurate detection method.

Manuscript published on 30 November 2018.
*****Correspondence Author(s)
S. Sathyavathi, Department of Information Technology, Kumaraguru College of Technology, Coimbatore, Tamil Nadu, India.
Dr. K. R. Baskaran, Department of Computer science and Engineering, Kumaraguru College of Technology, Coimbatore, Tamil Nadu
S. Kavitha, Department of Information Technology, Kumaraguru College of Technology, Coimbatore, Tamil Nadu

We propose a non-invasive method to detect DM by distinguishing Healthy and DM samples (using facial block color features) via a J48 algorithm. A non-invasive facial image is first captured using a non-invasive device. The facial image is divided into four blocks. The blocks like, one on the forehead and nose, and two below the left and right eyes are extracted automatically and used to represent a face. Next a facial color gamut is developed with six color centroids to compute a facial color feature vector, characterizing each facial block. Then a comparison was made to determine the Healthy and DM samples of the facial blocks. The comparison is done with following algorithms like J48, ZeroR and Support vector machine (SVM). [8]

II. FACIAL IMAGES AND DATASET

The description for conversion of facial images into color centroids value for the blocks are been specified.

A. Facial Image Acquisition Device:

Facial images from various health grades must be captured and maintained at a standardized setting in order to ensure unbiased feature extraction and analysis. An image size of 640 × 480 pixels is maintained for all the images that are been handled.

B. Facial Block Definition:

Thehuman face can be partitioned into four regions where the regions can reflect the health status of internal organs. Four facial blocks are named as A,B,C and D. Block A is located on the forehead. Blocks B and D are symmetrical and found below the left and right eyes, respectively. Finally, Block C is situated on the nose, B and D's midpoint.

If a patient is positioned further away from the capture device, both the camera and block size will require recalibration. Binarization is first applied to each image, where morphological operations are used to locate the pupils. From there, the distance between the pupils is used to map out the blocks

C. Facial Image Dataset*:*

The facial image database comprises of images split into healthy samples and DM samples.

III.FACIAL BLOCK COLOR FEATURE EXTRACTION

Color feature extraction is represented using six centroid values. Six centroid values are used to calculate a color feature vector for each block. Facial color gamut and their color centroids are important in facial block analysis. Facial color gamut shows the range of facial block colours. Each pixel is extracted from facial block and combined to form the facial block color distribution.

The six centroids characterize the most commonly found colors in the facial block (since it is within the black boundary) and are spread out to ensure that two or more colors do not overlap. These are then used to calculate the vector. Further we depict them into six centroids using the facial color gamut as a solid color square with its label on top and correspondingly RGB value below. These six centroids are red, yellow, light yellow, gloss, deep red, and black. The mean color feature

vectors of four blocks (A, B, C and D) for Healthy and DM samples are further obtained.

IV.VARIOUS CLASSIFICATION ALGORITHMS

A. healthy versus dm classification with j48

Once the facial color feature vectors [3] are extracted from the facial blocks, they are classified using the J48. J48 is a powerful decision tree technique and it has the additional features of accounting for missing values, decision trees pruning, continuous attribute value ranges, derivation of rules, etc. The WEKA tool provides many options associated with tree pruning. In other algorithms the classification is performed recursively till every single leaf is pure. But in J48, it generates the rules from which the identity of the data is generated. Using J48, the potential information is calculated for every attribute. Then the gain in information is calculated that would result from a test on the attribute. Then the best attribute will be found out using present selection criterion. Then that attribute will be selected for branching. Pruning: Because of the outliers this is a significant step to the result. Some instances are present in all data sets which are not well defined and differ from the other instances its neighborhood. J48 algorithm is performed on the training set and a tree will be formed. It is performed for decreasing classification errors which are being produced by specialization in the training set. Pruning is performed for the generalization of the tree.

J48 is mainly performed because this algorithm can handle both continuous and discrete values. J48 algorithm can fix a threshold values for continuous values. It can split the entire data set into two categories namely values below threshold and values above threshold.

B. Healthy Versus Dm Classification With Zeror

In pattern recognition, ZeroR algorithm is simple method used for classification. Input consists of the training examples in the feature space. The output depends mainly relies on target and ZeroR algorithm will ignore all predictors. Although there is no predictability power in ZeroR, it is useful for determining a baseline performance as a benchmark for other classification methods. It usually constructs a frequency table for the target value and select the most frequent value for the target.

C. Healthy Versus Dm Classification With Svm

In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked for belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. When data are not labeled, supervised learning is not possible, and an unsupervised learning approach is required, which attempts to find natural clustering of the data to groups, and then map new data to these formed groups. The clustering algorithm which provides an improvement to the support vector machines is called support vector clustering and is often used in industrial applications either when data is not labeled or when only some data is labeled as a preprocessing for a classification pass.

Computing the (soft-margin) SVM classifier amounts to minimizing an expression of the form

$$\left[\frac{1}{n}\sum_{i=1}^{n}\max\left(0,1-y_i(w.x_i+b)\right)\right]+\lambda\parallel w\parallel^2$$

We focus on the soft-margin classifier since, as noted above, choosing a sufficiently small value for λ yields the hard-margin classifier for linearly classifiable input data.

V.PROBLEM DESCRIPTION

The main idea of implementing such noninvasive method is to determine the accuracy of various classification and clustering algorithms using a combination of various facial block color features to know in which particular block the level of diabetes seem to be high and distinguish between the diabetes and healthy samples.

Initially, average accuracy of the samples will be calculated using J48. Average accuracy is used to evaluate the classification result of the healthy versus DM samples. To achieve the optimal result using facial blocks A, B, and C, all combinations were tested. Altogether there were seven combinations (without repeating): A, B, C, AB,AC,BC and ABC. If any combinations have the same accuracy, then only one will be selected on the computational cost. Sensitivity and specificity values have been calculated for the highest average accuracies.The J48 was compared with traditional classifiers such as ZeroR and SVM. In these algorithms, the results using kernel function will be mapped with training data into the kernel space. Then the average accuracy of these three algorithms will be compared and algorithm with high.

VI.MODEL IMPLEMENTATION

A. Facial blocks extraction

In this module, binarization is first applied to face image where morphological operations are used to locate the pupils. From there, the distance between the pupils is used to map out the blocks.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Block A is located on the forehead. Blocks B and D are symmetrical and found below the left and right eyes, respectively. Finally, Block C is situated on the nose, B and D's midpoint.

B. Feature extraction

Given a facial block, its RGB pixels values are first extracted and converted to CIELAB by translating RGB to CIEXYZ using

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

followed by CIEXYZ to CIELAB by

$$\begin{cases} L^* = 166 f\left(\frac{Y}{Y_0}\right) - 16 \\ a^* = 500\left[f\left(\frac{X}{X_0}\right) - f\left(\frac{Y}{Y_0}\right)\right] \\ b^* = 200\left[f\left(\frac{Y}{Y_0}\right) - f\left(\frac{Z}{Z_0}\right)\right] \end{cases}$$

Where $f(x) = x^{\frac{1}{3}}$ if $x > 0.008856$ or $f(x) = 7.787x + 16/116$ if $x \leq 0.008856$. Then, $X0$, $Y0$, and $Z0$ are the CXYZ tristimulus values of the reference white point. The CIELAB values are then compared to six centroids from the facial color gamut and assigned the color which is closest to it. After evaluating all the pixels of a facial block, the total of each color is summed and divided by the total number of pixels. This ratio forms the facial color feature vector v, where $v = [c1, c2, c3, c4, c5, c6]$ and ci represents the sequence of the six centroids.

C. The input files

The input file consists of the pre- processed data of the both healthy and diabetes mellitus samples. The attributes that are considered are facial color samples, facial color feature vector, etc. The sample training set for the model implementation is provided below. Created a predictive model based on the use of historical data (1128 records) collected from a well renowned hospital). In the data EXCEL sheets the observations are situated on rows and columns represent the variables. That data includes facial color gamut with its six centroids such as red(R), yellow(Y), light yellow(LY), gloss(GL), deep red(DR), black(BK), block and class.

VII.FIGURES

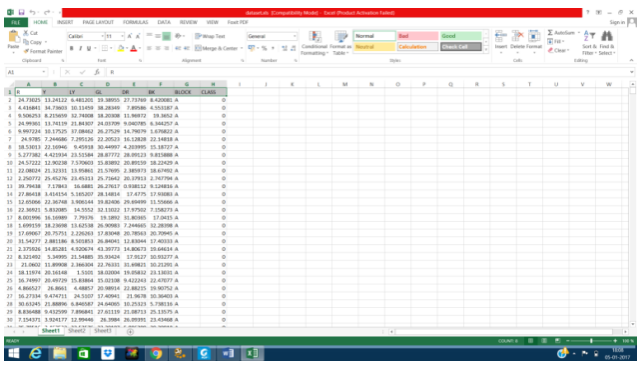

Fig. Data Set

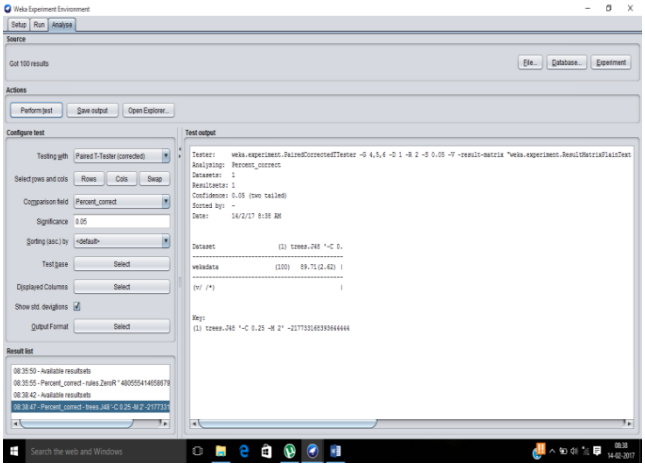

Fig Accuracy of J48

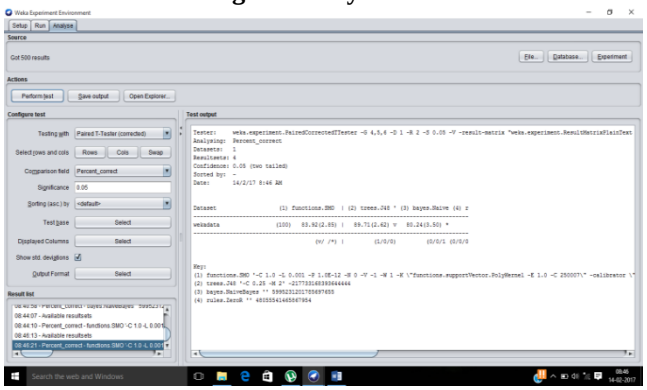

Fig Accuracy Comparison of Various Algorithms

VIII.CONCLUSION

This system is used to provide a comparative study of different noninvasive approaches to classify Healthy and DM samples using facial color features extracted from facial blocks. While comparing the algorithms like J48, ZeroR and Support Vector Machine(SVM), J48 has the highest accuracy in classifying the diabetes and healthy samples than the other algorithms.

REFERENCES

1. A. Ganesh , A. Yang, J. Wright, S. Sastry and Y. Ma "Robust face recognition via sparse representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 2, pp. 210–227, Feb. 2009.
2. D. Zhang and X. Wang, "An optimized tongue image color correction scheme," IEEE Trans. Inf. Technol. Biomed., vol. 14, no. 6, pp. 1355–1364, Nov. 2010.
3. B. Pang, D. Zhang and K. Q. Wang, "Tongue image analysis for appendicitis diagnosis," Inf. Sci., vol. 175, no. 3, pp. 160–176, 2005.
4. I. Lemahieu , J. M. A. D. Naeyaert, W. Philips and Y. V. Haeghen , "An imaging system with calibrated color image acquisition for use in dermatology," IEEE Trans. Med. Imaging, vol. 19, no. 7, pp. 722–730, Jul. 2000.
5. R. Kenet et al., "Clinical diagnosis of pigmented lesions using digital epiluminescence microscopy," Arch. Dermatol., vol. 129, pp. 157–174, 1993.
6. T. Schindewolf et al., "Evaluation of different image acquisition techniques for a computer vision system in the diagnosis of malignant melanoma," J. Amer. Acad. Dermatol., vol. 31, no. 1, pp. 33–41, 1994.
7. B. Pang, D. Zhang , N. M. Li, K. Q. Wang, and Z. Zhang, "Computerized diagnosis from tongue appearance using quantitative feature classification," Amer. J. Chin. Med., vol. 33, no. 6, pp. 859–866, 2005.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)