Diagnosis of Diabetes from Tongue Image Using Versatile Tooth-Marked Region Classification

[1]G.Umadevi, [2]V.Malathy, [3]M.Anand

[1]Associate Professor, Department of Computer Science Engineering,
University of Engineering and Management, Jaipur

[2]Assistant Professor, Department of Electronics and Communication Engineering,
S R Engineering College, Warangal

[3]Professor, Department of Electronics and Communication Engineering,
Dr M.G.R Educational and Research Institute, Chennai

[1]guma_devi@rediffmail.com, [2]malathyvanniappan@gmail.com, [3]worldcommunication25@gmail.com

Abstract

Image analysis of the human tongue has been found to be useful in detecting various diseases in the body. The tongue image indicates the condition of different parts of the body and the changes in the tongue reflect the misbehavior of the internal parts of the body. So, the diagnosis of diseases is very much needed.The human tongue is detected and extracted effectively by implementing adaptive threshold segmentation. Gabor filter is used to identify the color and texture of the image. From the factors such as color, texture (coating), smoothness/cracks and size, the healthiness of the tongue is analyzed effectively. The threshold value taken from a healthy human tongue is used to classify a person's tongue whether it is normal or abnormal. If the tongue is abnormal, diseases, such as, thyroid, ulcer and diabetes can be diagnosed. In this paper, diagnosis of diabetes from tongue image using Versatile Tooth-Marked Region (VTMR) classification method is proposed. The proposed VTMR method is tested on the 96 BioHit tongue images collected from Sri Muthukumaran Medical College Hospital and Research Institute and 97 UV scanned tongue images captured from the patients by using IPhone with HD camera. The simulation work is carried out in MATLAB simulation environment utilizing the proposed VTMR method. The proposed VTMR method uses color and texture features for the diagnosis of diabetes. The color and texture features are beneficial to classify the tongue image and to diagnose diseases.

1. Introduction

Medical research deals with a variety of facts to explore automation technology for the diagnosis of disease conditions of a patient. The human organs for recognizing taste and smell are extremely intricate. The tongue is the most widely recognized site associated with the diagnosis of diabetes disease. The significant parameters of the tongue must be promptly pertinent to be used in routine laboratory administrations. Therefore, a need for the disease diagnosis by classifying the tongue image is necessary for further treatment.

The functioning of the tactile organs including the thyroid glands attracted the researchers and specialists from different branches not just to discover the complexities of the tool used, but also to motivate the researchers to build up the machine-learning methods.

Segmentation refers to the image-processing technique to distinguish or identify the diabetes affected organs/tissues using tongue images in this study (Malathy et al. 2019(1,2)). There are numerous segmentation or classification strategies for medicinal images to get an improved outcome on tongue images.

Bob et.al. (2014) recognized a tongue color range with 12 colors representative the tongue color features. Sujatha et al. (2015) captured the tongue image by a capture tool. The image was pre-processed by median filter. By utilizing three groups of features such as color, texture and geometry, diabetes was detected.Cibin et al. (2015) detected four groups of features extracted from tongue images-color, texture, geometry and energy. The tongue images were captured using the three-chip CCD camera with 8 bit resolution. Ramesh et al. (2015 proposed a technique to extract features from tongue images. The features were color, texture and geometry.

The tongue color characteristics were extracted with a Gabor filter mechanism.

This paper introduces Versatile Tooth-Marked Region (VTMR) classification for the diagnosis of diabetes from tongue images.

2. The Proposed VTMR Method

The detection of diabetes is described briefly using tongue image. The proposed VTMR method has four stages such as pre-processing, image segmentation, texture feature extraction and classification. The block diagram of the proposed VTMR method is shown in Figure 1.
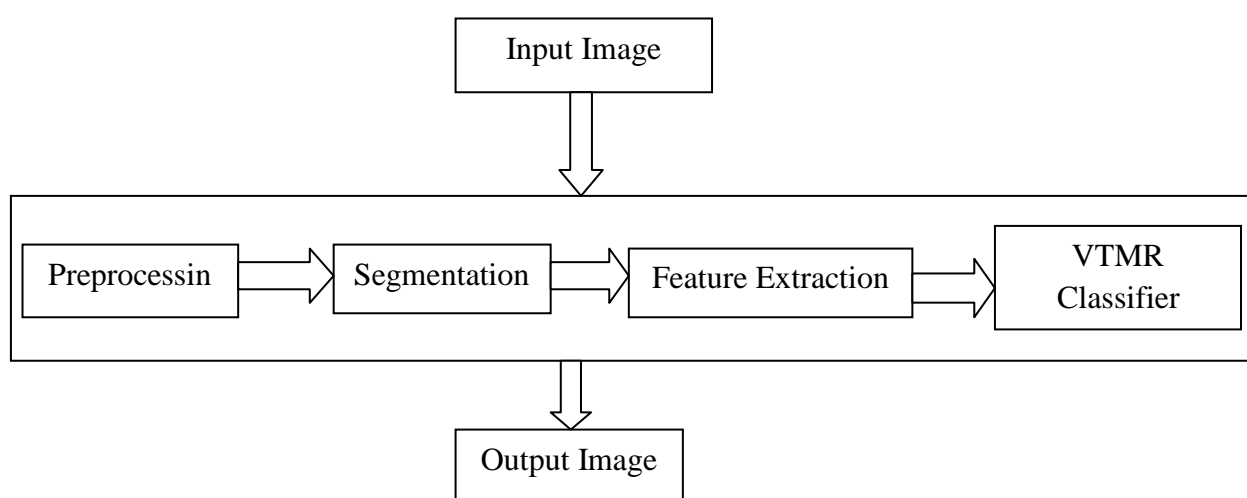

Figure 1: Block diagram of the proposed VTMR method

3. Pre-Processing

Pre-processing using temporal filtering

Denoising algorithm is a procedure in image reframing, which focuses on evacuating its commotion and consequently enhances the nature of the images (Sreedhar et al. 2018 (1),2019, 2018 (2), 2016). The proposed denoising algorithm is temporal filter (Sreedhar et al. 2018 (3)). This technique enhances the tongue images by removing the noise.

Temporal Filter Algorithm

Step 1: Load the input image.
Step 2: Set of images X depicted as $X_i = \{X_1, X_2, X_3...X_i\}$.
Step 3: Image set X portrayed as $Z_i = \{Z_1, Z_2, Z_3...Z_i\}$.
Step 4: Images to give output image $Y_i = \{Y_1, Y_2, Y_3...Y_i\}$.
Step 5: Contrast every X_i and the relating Y_i.

Step 6: for the rest of the qualities not thought about, where $Y_i(x, y) = 1$, look at Z at changed separation esteems j=1 to 10.
Step 7: for the rest of the qualities not analyzed, create the separation between each incentive in Y with each motivation in Z. Utilize any of the separation measurements.
Step 8: pre-processed image is obtained.
Step 9: Stop.

4. Image Segmentation

Clustering segmentation using heuristic algorithm

The Heuristic Algorithm (HA) procedure restricts the nearby esteems and does not consider the spatial relevant data. Henceforth, there is always an expanding requirement for the streamlining of the group, focusing on HA segmentation. For the segmentation of the tongue, cerebral tissues like gray matter, white matter and cerebral tissue locales are taken. At that point, the outcomes with different sorts of muscles are confined. HA has been acknowledged for the segmentation of the diabetes-affected tissues in tongue images.

Heuristic Algorithm

Step 1: Input image

Step 2: A feasible set can be divided into several subsets which are neighbors of the clusters. The feasible set can be taken as the union of the neighbor of these clusters.

Step 3: Collect uniform neighbor from every cluster.

Step 4: The best sample from each cluster.

Step 5: Determine the cluster K

Step 6: Construct the neighbor of S.

Step 7: Find each neighbor of S repeating from Step 2 to3, and choose the best sample among the neighbors of S.

Step 8: Output image.

Step 9: Stop.

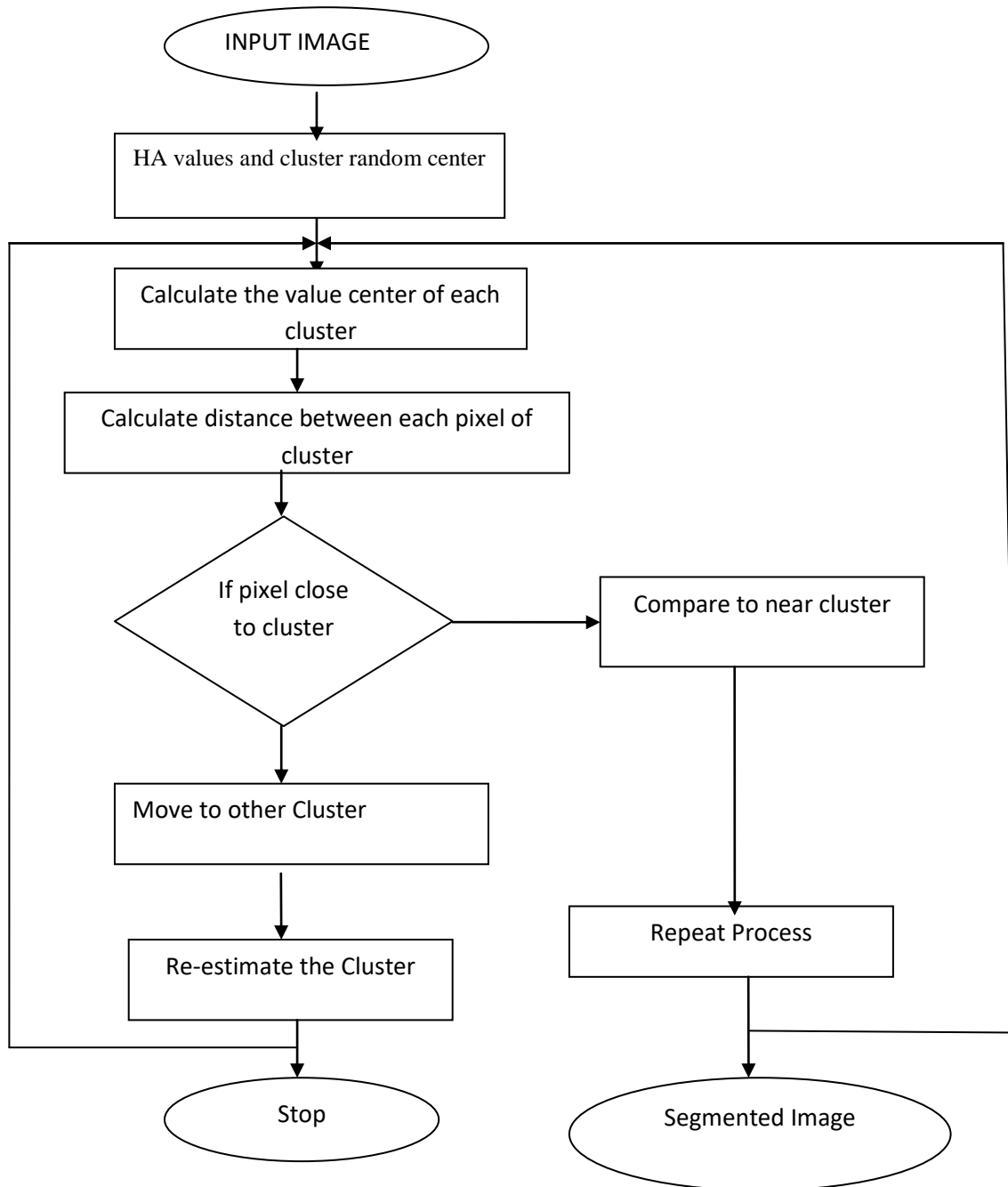Figure 2 shows the flow chart of clustering segmentation using HA algorithm.

Figure 2:Flow Chart of clustering segmentation using HA

Clustering is the process of finding the groups of objects such that the objects in a group would be similar to one another and different from the objects in other groups. High-quality clusters with high intra-cluster similarity

and low inter-cluster similarity are the useful clustering methods. In HA, initially the number of clusters is defined. Then the HA cluster centre is chosen randomly. The distance between each pixel and each cluster center is calculated.

5. Texture Feature Extraction and Classification

VTMR Classifier

In VTMR classification, the system takes the tongue images and divides each image into six areas. For each area, n numbers of color and texture features are extracted. From these n numbers of elements, VTMR identifies the four levels of severity of pancreas such as health, acute, chronic and regenerative.

VTMR Algorithm

Input: Tongue Images
Output: Tongue Diabetic or not
Step 1: Take Input Images I = {i_1, i_2… i_n}.
Step2: Image Pre-processing. In this step, 6 partitions from images
P = {p_1, p_2, p_3, p_4, p_5, p_6} are received.
Step 3: Segmentation and feature extraction. In this step, to get 2 sets of features, F = {f_1, f_2} Where f_1 = color features f_2 = texture features are received.
Step 4: Initialize weights W_{ij}
o   Set topological neighborhood parameters
o   Set learning rate parameters
o   While stopping condition is false
Step 5: For each input vector x, do step 6-8.
Step 6: For each j, compute:

$$D(j) = \sum iWij - Xi$$

Step 7: Find index J such that D (J) is a minimum.
Step 8: For all units j within a specified neighborhood of J and for all i

$$Wij(new) = wij(old) + \propto [Xi - wij(old)]$$

Color Analysis

The gabor functions include the process of color analysis and texture analysis. Let f(x,y) be the intensity at the coordinate g(x,y) in a gray scale image,

$$g_{f,\theta}(x, y) = f(x, y) \otimes \psi_{f,\theta}(x, y)$$

Texture Analysis

Texture will be extracted using Gabor function with g(x,y) is

$$g_{\lambda,\theta,\phi,\sigma,\gamma}(x, y) = \exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right]\cos\left(2\pi\frac{x}{\lambda} + \varphi\right)$$

Where   x=acosθ+bsinθ, y= -acosθ+bsinθ

6. Experiments and Discussion

Healthy Tongue Analysis

Figure 3 demonstrates the convolution of the input image with a Gabor filter in the tongue image pre-processing stage. As compared to input image, the pre-processed image shows a better view without noise.
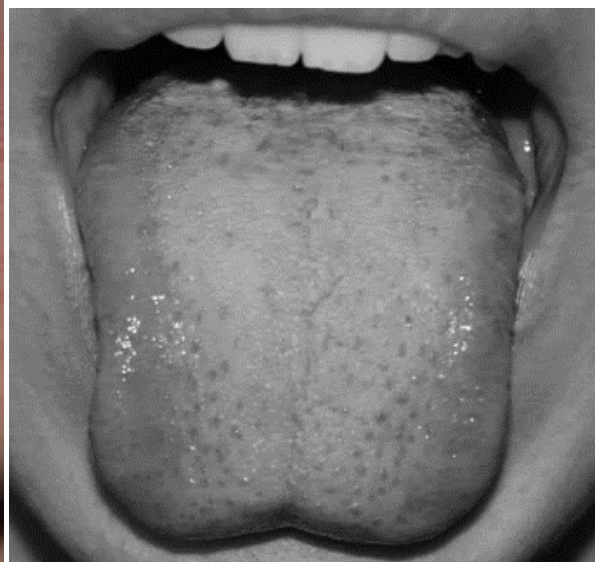
(a) Input Image   (b) Pre-processed image

Figure 3 Tongue image pre-processing

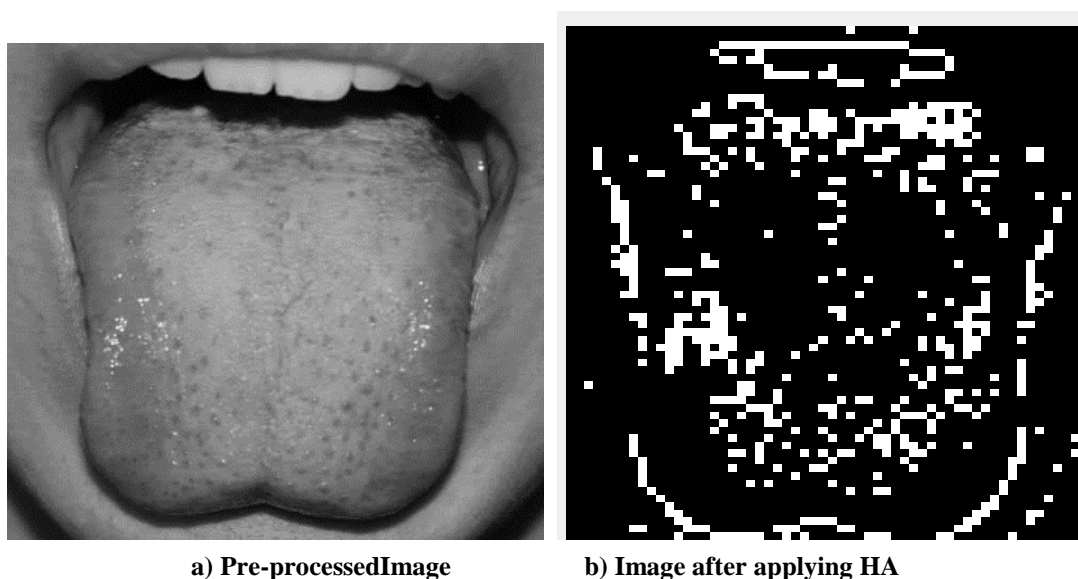Figure 4 shows the segmented tongue image after pre-processing.

a) Pre-processedImage   b) Image after applying HA

Figure 4Tongue image segmentation

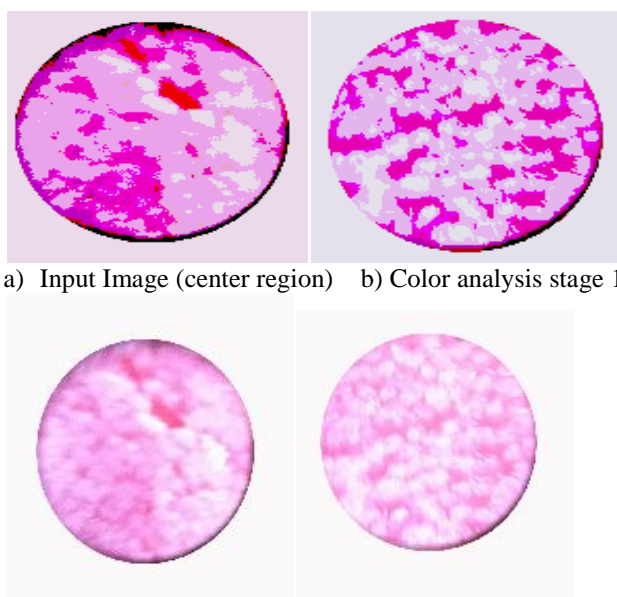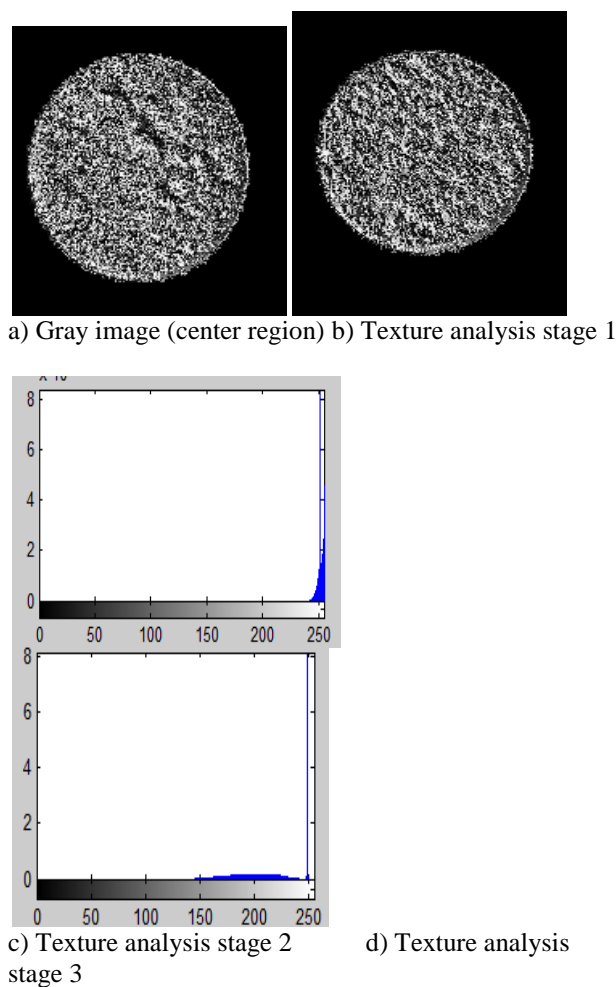Figure 5 shows the input tongue image and color analysis of tongue image.

a) Input Image (center region)   b) Color analysis stage 1

a) Gray image (center region) b) Texture analysis stage 1

a) Color analysis stage 2   b) Color analysis stage 3

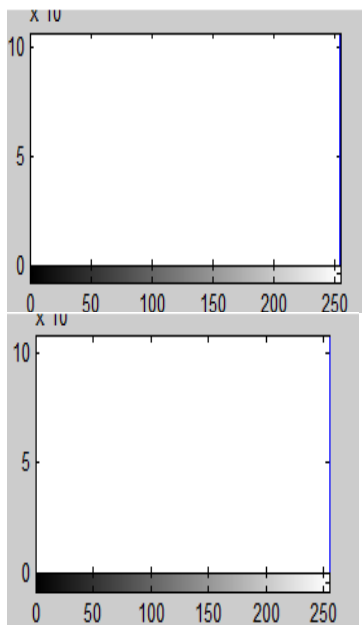Figure5: Tongue image color analysis

Figure 6 shows the gray image and texture analysis of tongue image.
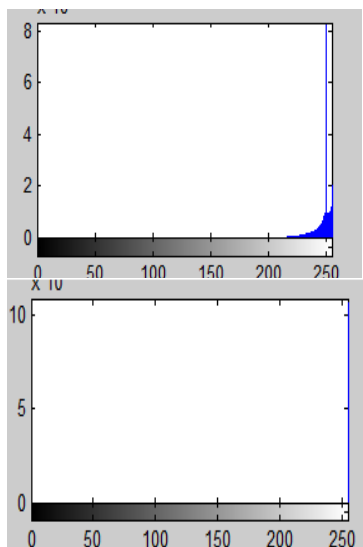
c) Texture analysis stage 2   d) Texture analysis stage 3

e) Texture analysis stage 4   f) Texture analysis stage 5

g) Texture analysis stage 6   h) Texture analysis stage 7

Figure 6:Tongue image texture analysis

Figure 7 confirms the result healthy tongue image.
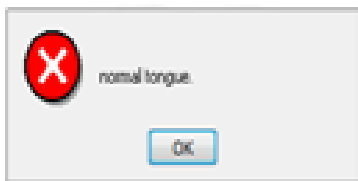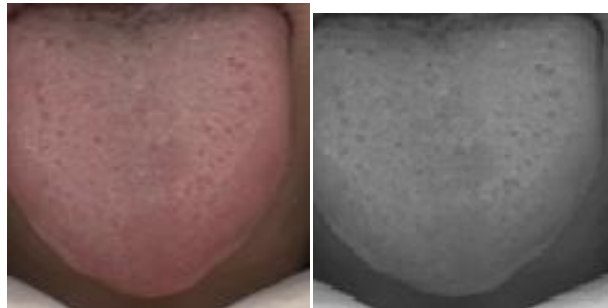

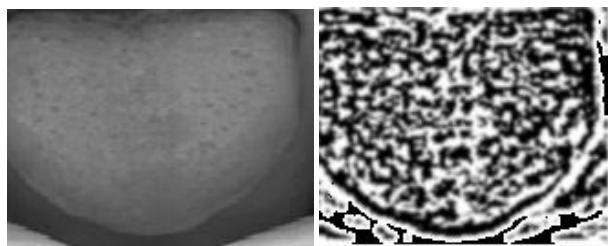Figure 7: Result of healthy tongue image

Diabetes Detection

Figure 8demonstrates the convolution of the input image with a Gabor filter in the pre-processing stage.

(a) Input Image   (b) Pre-processed Image

Figure 8: Tongue image pre-processing

Figure 9 shows the segmented tongue image after pre-processing.
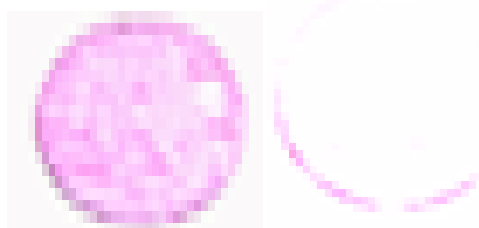

a) Pre-processed Image   b) Image after applying HA

Figure 9 Tongue image segmentation

Figure 10 shows the input tongue image and color analysis of tongue image.

a) Input Image(center region)   b) Color analysis stage 1

a) Color analysis Stage 2   b) Color analysis stage 3

Figure10: Tongue image color analysis

Figure 11 shows the gray image and texture analysis of tongue image.

a) Gray image (center region)   b) Texture analysis stage 1

c) Texture analysis stage 2   d) Texture analysis stage 3

e) Texture analysis stage 4   f) Texture analysis stage 5

g) Texture analysis stage 6   h) Texture analysis stage 7
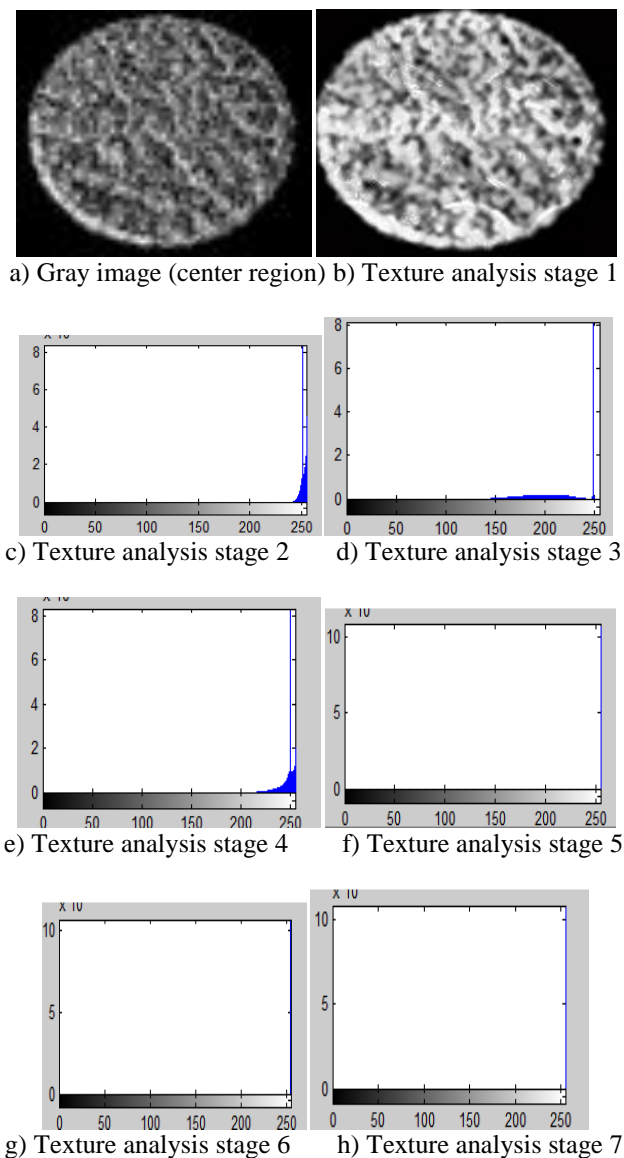
Figure 11: Tongue image texture analysis

The abnormality found in the surface of the tongue image reveals that the tongue is affected by diabetic disease. Figure 12 confirms the result of diabetic affected tongue.
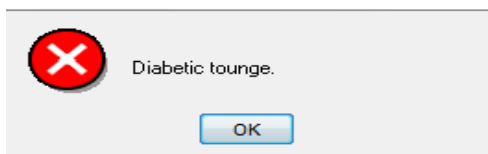

Figure 12: Result of diabetic affected tongue

To validate the performance of the proposed VTMR method, it is tested on 96 BioHit images. 50 images are classified as abnormal and 24 images are detected to be diabetes affected.

The proposed VTMR method is also applied on 97 UV scanned images. 48 images are classified as abnormal and 26 images are found to be affected by diabetes.

7. Results and Discussion

The tongue image for the diagnosis of diabetes using the proposed VTMR method is analyzed in this paper. The proposed VTMR method is implemented and evaluated for its performance. The proposed VTMR method has been simulated and tested using the popular tool MATLAB.

Table 1 shows the details of data set used to evaluate the performance of the proposed VTMR method.

Table 1: The Details of data set

Data Set Name	No. of Samples	Normal tongue image	Abnormal tongue image
BioHit	96	46	50
UV scanned image	97	49	48

Dataset Types are:
1. BioHit: Tongue images are collected from Sri Muthukumaran Medical College Hospital and Research Institute, Chennai, Tamil Nadu, India.
2. UV scanned image: UV scanned tongue images are captured from the patients by using IPhone with HD camera.

Table 2 shows the details of abnormal data set for the proposed VTMR method.

Table 2: Details of abnormal data set for the proposed VTMR method

Data Set Name	No. of abnormal tongue image	No. of diabetes affected tongue image	No. of other abnormal tongue image
BioHit	50	24	26
UV scanned image	48	26	22

8. Simulation Environment

MATLAB incorporates calculation, representation and programming to utilize the condition where the issues and arrangements are communicated in the original numerical documentation. Image processing tool compartment has large capacities for image rebuilding, improvement and

data extraction. With the assistance of MATLAB tool, tongue image analysis is performed for the diagnosis of diabetes using the proposed VTMR method.

9. Comparative Parameters

The proposed VTMR method is compared to prove the best. It is compared with the K-means with SVM and Fuzzy methods in terms of

- Etp
- SD
- CoV
- Classification Efficiency
- Classification Accuracy
- False Positive Ratio
- Time Complexity

For qualitative analysis, the parameters-Etp, SD and CoV are compared. For quantitative analysis, the parameters-Classification Efficiency Classification of Accuracy False Positive Ratio and Time complexity are compared for BioHit and UV scanned images.

Entropy (Etp)

The value of this metric is the mean absolute energy difference of the pixel components in the corresponding points of the image. The assessment of entropy is performed using condition 3.2.The entropy standard shows that expansive esteems have low precision rate while low esteems have high exactness rate.

$$e = entropy(I)$$

where e is a scalar value referring to the entropy of grayscale image I. Entropy is a factual measure of irregularity that can be utilized to describe the texture of the info image. Entropy is characterized as

$$entropy = sum(p,*\log 2(p))$$

Where p is contains the standardized histogram. Entropy utilizes two containers for coherent clusters and 256 canisters for uint8, uint16 or twofold exhibits.

Standard Deviation (SD)

The unbiased estimate of the standard deviation S_a, of the brightness within a region (r) with Λ pixels, is called the standard deviation. It is given by:

$$S_a = \sqrt{\frac{1}{\Lambda - 1} \sum_{m,n \in r} (a[m,n] - ma)^2}$$

Using the histogram formulation:

$$S_a = \sqrt{\frac{\left(\sum_a a^2 * h[a]\right) - \Lambda * m^2 a}{\Lambda - 1}}$$

The standard deviation S_a, isan estimate of a, the underlying brightness probability distribution.

Covariance (CoV)

The Covariance between two jointly distributed real-valued random variables X and Y with finite second moments is defined as the expected product of their deviations from their individual expected values:

$$\mathrm{cov}(X, Y) = E[(x - E[X])(Y - E[Y])]$$

where E[X] is the expected value of X, also known as the mean of X.

The covariance is sometimes denoted as σ_{XY} or $\sigma(X, Y)$, in analogy to variance. By using the linearity property of expectations, this can be simplified to the expected value of their product minus the product of their expected values.

Classification Efficiency

The efficiency mainly depends on the classification. Efficiency analysis is concerned with training time. At testing time, one has to test a quadratic number of classifiers in order to make the final prediction. Even though the constituent classifiers are quite likely to be simpler it can be expected that classification takes considerably longer for a round robin ensemble than for unordered binarization. The classification efficiency is defined as:

Classification of Efficiency= (No. of characterized images/No. of unclassified images) * 100

Classification of Accuracy

Image classification is a complex process that may be affected by many factors. The classification accuracy is defined as:

Accuracy = (correctly predicted class / total testing class) * 100   (6.2)

False Positive Ratio

When performing multiple comparisons, a false positive ratio (or false alarm ratio) is the probability of falsely rejecting the null hypothesis for a particular test. The false positive (FP) rate is calculated as the ratio between the numbers of negative events (N) wrongly categorized as positive (false positives) and the total number of actual negative events (regardless of classification).

$$\frac{FP}{N} = \frac{FP}{FP + TN}$$

Where FP is the number of false positives, TN is the number of true negatives and N=FP+TN is the total number of negatives.

Time complexity

It refers to the utilization of time required to finish the whole procedure to acquire an unmistakable arrangement of characterized images.
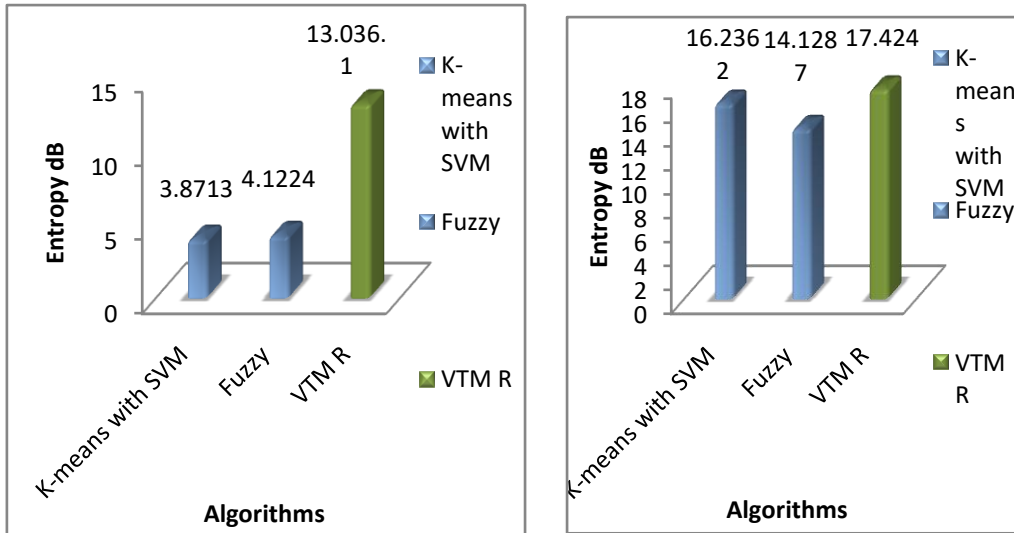
Running Time = [O (segmented images / Number of boundaries created)]

10. Comparative Analysis

Etp

Figure 13 shows the benefits of Etp that increased along with the performance of different methods such as K-Means with SVM, Fuzzy and VTMR. Figure 13shows that, the proposed VTMR method outperforms the K-means with SVM and Fuzzy for Etp.
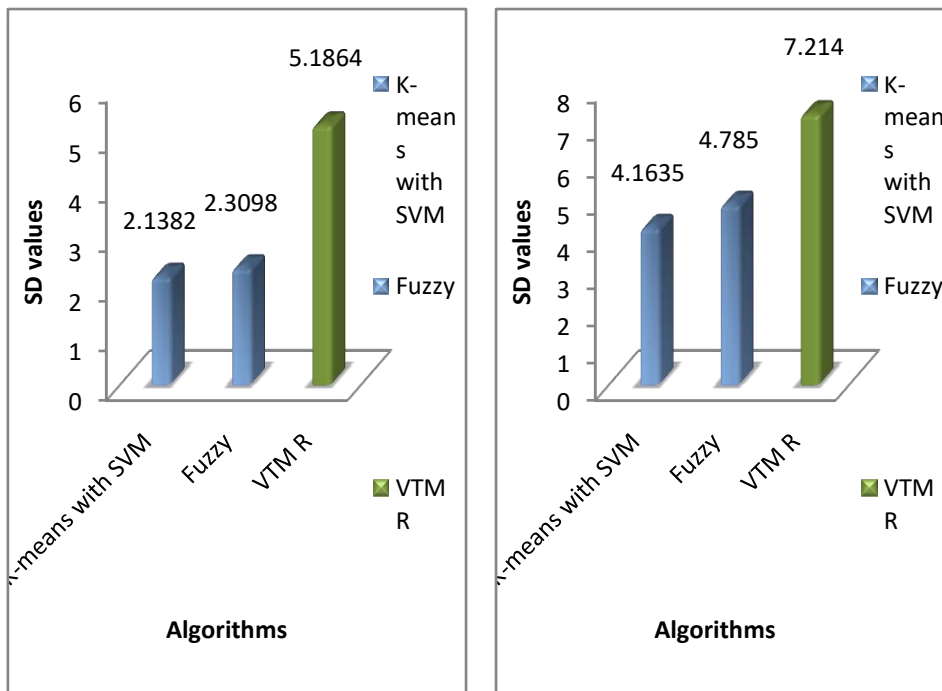

| Healthy image | Diabetes affected image |

Figure 13: Performance of Etp

SD

The SD metric depends on the maximum value of the image color component. Figure 14 shows the benefits of SD which does not make tongue image more distorted. The performance of different methods, K-means with SVM, Fuzzy and VTMR has been compared. Figure 14shows that the proposed VTMR method outperforms the standard K-means with SVM and Fuzzy in SD.
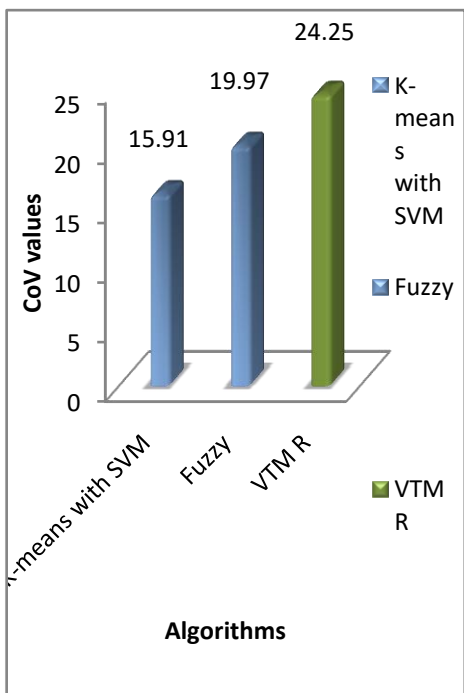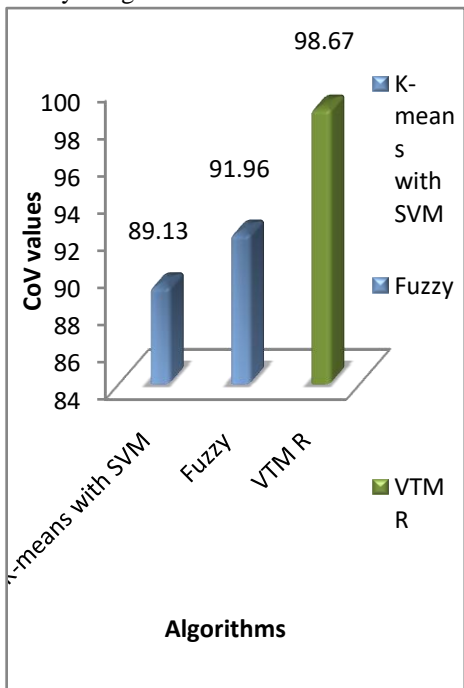

| Healthy image | Diabetes affected image |

Figure 14:Performance of SD

CoV

The datasets must have a significant influence on the performance of health quality metrics. Figure 15 shows the values of CoV increased by some amount in a tolerable limit. The performance of different methods K-Means with SVM, Fuzzy and VTMR is compared. Figure 15shows that the proposed VTMR method outperforms the K-means with SVM and Fuzzy in CoV.

Healthy image


Diabetes affected image

Figure 15: Performance of CoV

For qualitative analysis,Etp, SD and CoV are used to compare the three proposed methods with K-means with SVM and Fuzzy methods. Based on the above-tabulated values the comparative result charts are given. Figures 13, 14 and 15 present the comparison of Etp, SD and CoV respectively for different methods and it is proved that the proposed VTMR works well.

Classification Efficiency

Figure 16 presents the comparison of classification efficiency for different methods and it shows that, the proposed VTMR method has produced higher classification efficiency for two types of data set.
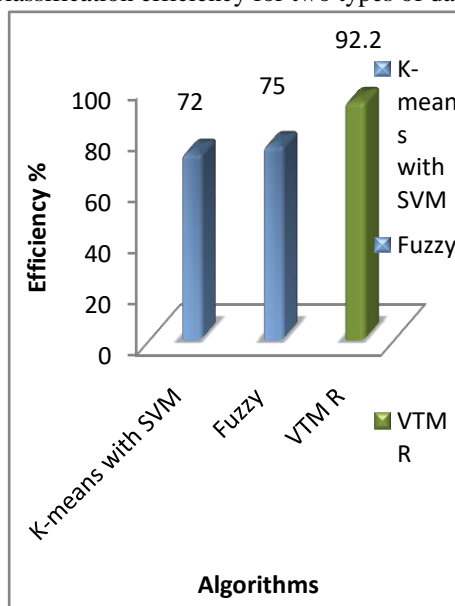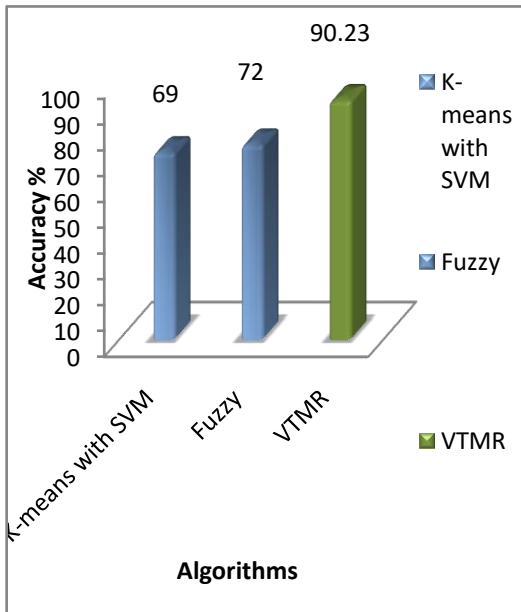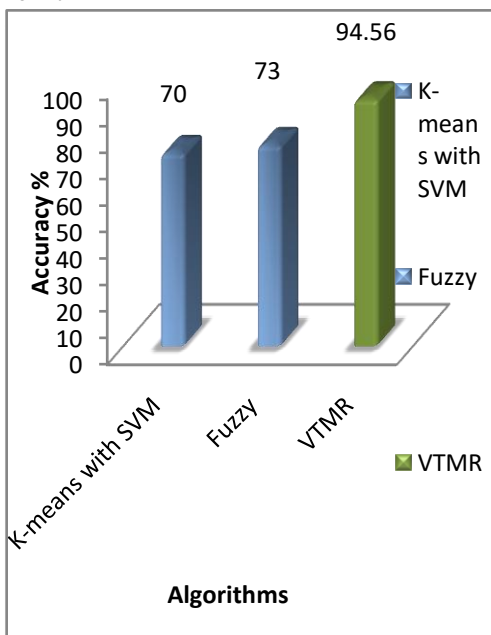

BioHit
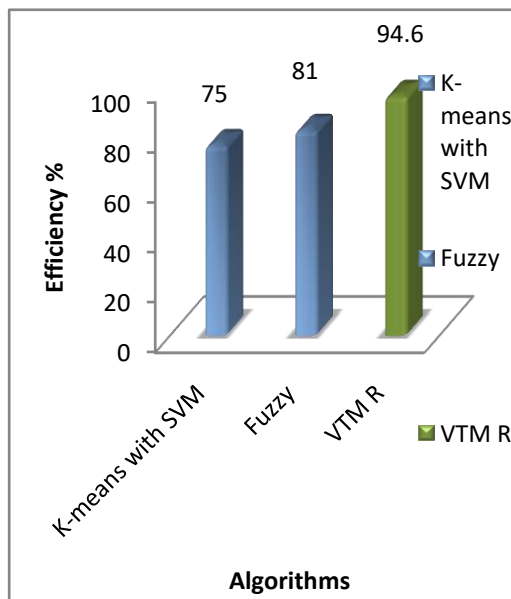
UV image

Figure 16: Comparison of Classification Efficiency

Classification of Accuracy

Figure 17 presents the comparison of classification accuracy for different methods and it shows that the proposed VTMR method has produced higher classification accuracy for two types of data set.
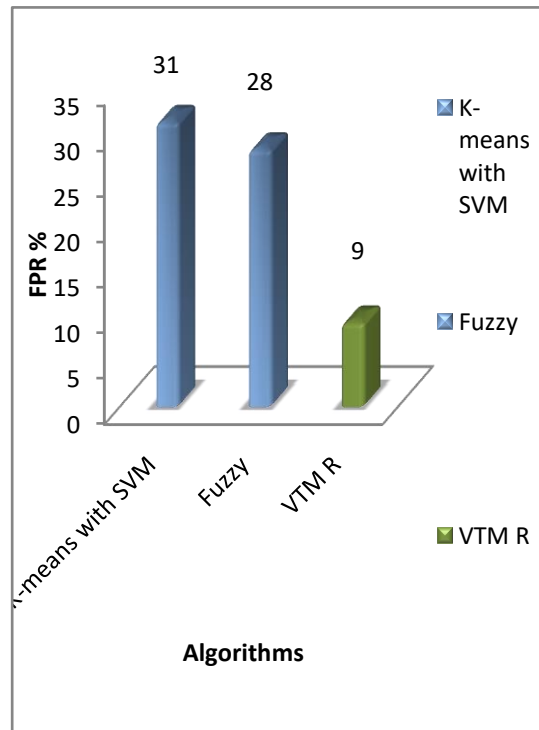

BioHit

UV image

Figure 17: Comparison of Classification Accuracy

False Positive Ratio

The level of significance that is used to test each hypothesis is set based on the image. When performing multiple comparisons in the false positive ratio, it usually refers to the probability of falsely rejecting the null hypothesis for a particular test. Figure 18 shows the comparison of results obtained from different ways of false classification. The proposed VTMR methodhas a less percentage for two types of data set.
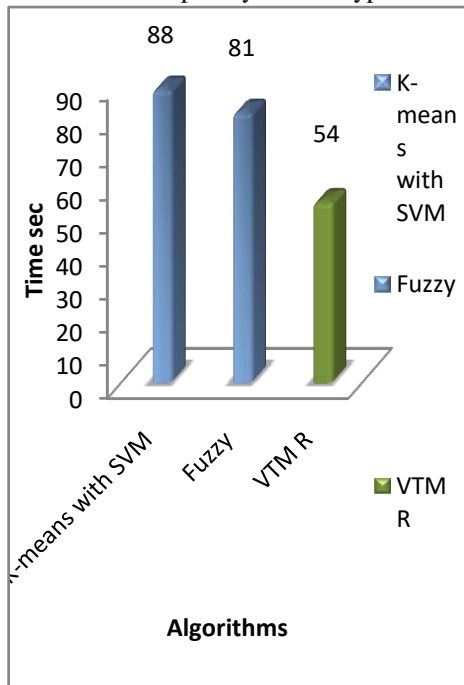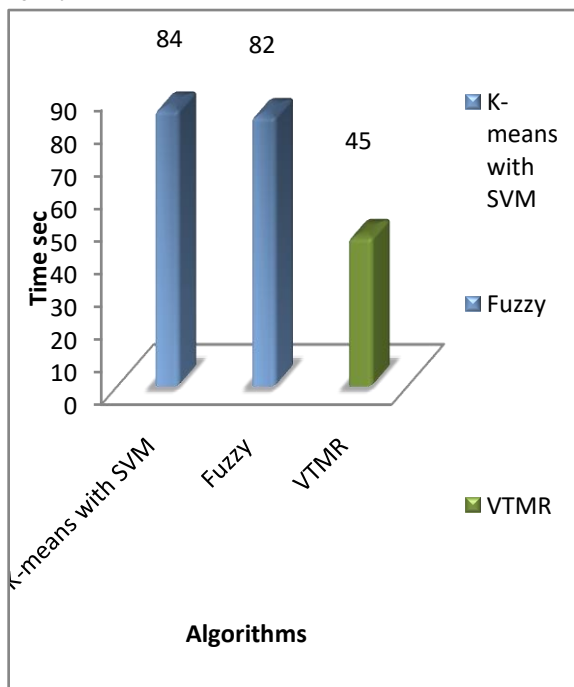

BioHit
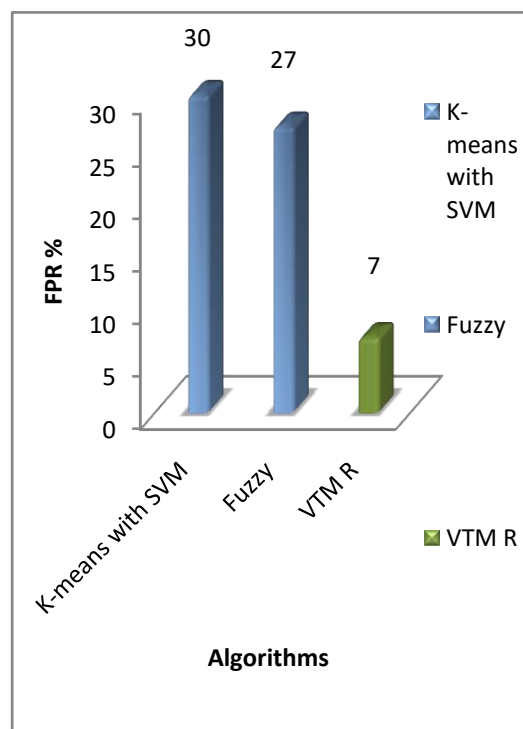
UV image

Figure 18: Comparison of different methods in the False Positive Ratio

Time Complexity

Figure 19 shows the comparison of different methods in time complexity and it depicts that the proposed VTMR has less time complexity for two types of data set.

BioHit

UV image

Figure 19: Comparison of Time Complexity of different methods

For quantitative analysis, Classification Efficiency, Classification Accuracy, False Positive Ratio and Time Complexity are used to compare the three proposed methods with K-means with SVM and Fuzzy methods for the BioHit and UV scanned images. Based on the above-tabulated values the comparative result charts are given. Figures 16, 17, 18 and 19 present the comparison of the parameters for different methods and it is proved that the proposed VTMR method works well.

11. Conclusion

From the results, it is inferred that, the proposed VTMR method performs well qualitatively and quantitatively in terms of Etp, SD, CoV, Classification Efficiency, Classification Accuracy, False Positive Ratio and Time Complexity. The proposed VTMR method is proved to be good when compared to K-means with SVM and Fuzzy methods.

Future Work

The present investigation can provide a broad experimental evaluation and further improvement on the algorithms.The Large Neighborhood Search algorithm is proposed for clinical data and it can be further redefined for genomic data.The stand-alone clinical decision support system (CDSS) can be further enhanced by the development of cloud-based CDSS or mobile-based CDSS applications.

References

[1]   Bob Zhang, B. V. K. Vijaya Kumar, David Zhang. Detecting diabetes mellitus and Non proliferative diabetic retinopathy using tongue color, Texture, and Geometry features. IEEE Transactions on Biomedical Engineering. 61(2), 491-501, 2014.

[2]   Sujatha S, Divya D. A novel approach for analyzing diabetes mellitus and non proliferative diabetic retinopathy using tongue. Indian Journal of Innovations and Developments. Vol. 4 (2), 104-110, June 2015.

[3]   Cibin N V, Wilfred Franklin S, Ajin N V. Diagnosis of diabetes mellitus and NPDR in diabetic patient from tongue images using LCA classifier. International Journal of Advanced Research Trends in Engineering and Technology (IJARTET) Vol. II, Special Issue XIX, February 2015.

[4]   Ramesh Kumar B, Sivapriya V. Diabetes Mellitus Discovery based on Tongue Texture Features using Log Gabor Filter Mechanism. International Journal of Innovative Research in Computer and Communication Engineering. Vol. 3, Issue 9, September 2015.

[5]   Malathy V &Kamali S M 2019 (1)," Integrating Fuzzy C-Means Algorithm with Level Set Methods for Segmentation of Injured Human spinal Cord in Computed Tomography Images",

Jour of Adv Research in Dynamical & Control Systems, Vol. 11, 02-Special Issue, pp.794-801.

[6] Malathy V &Kamali S M 2019 (2)," Brain Tumor Segmentation from Brain Magnetic Resonance Images using Clustering Algorithm", inInternational Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-8S, pp.625-629.

[7] Sreedhar K, Ramalinga Reddy, SreenivasaRao D," Image Denoising by using Modified SGHP Algorithm", in International Journal of Electrical and Computer Engineering, vol.8,No.2,April 2018 (1), pp.971-978.

[8] SreedharKollem, KattaRamalinga Reddy and DuggiralaSrinivasaRao(2018 (2)). Denoising and Segmentation of MR Images using Fourth Order Non-linear Adaptive PDE and New Convergent Clustering. International Journal of Imaging Systems and Technology, https://doi.org/10.1002/ima.22302, 1–15. (John-Wiley Publications)

[9] SreedharKollem, KattaRamalinga Reddy and DuggiralaSrinivasaRao(2019). A Review of Image Denoising and Segmentation Methods based on Medical Images .International Journal of Machine Learning and Computing, 9(3), 1–8.

[10] SreedharKollem (2018 (3)). Image Denoising by using Modified SGHP Algorithm. International Journal of Electrical and Computer Engineering, 8(2), 971–978.

[11] K. Sreedharand K. Shubhankar Reddy (2016). Image Retrieval Techniques: A Survey. International Journal of Electronics and Communication Engineering, 9(1), 19–27.

[12] S.M.Kamali and V.Malathy (2019), "Modern cane enabling obstacle detection and navigation based on human gesture using voice board" Test Engineering Management, November-December 2019 ISSN: 0193-4120 Page No. 3900 – 3906.