International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Oral Cancer Detection: Modified KFCM Segmentation Clustering Algorithm

Shilpa Harnale[1], Dr. Dhananjay Maktedar[2]

Submitted: 28/04/2023 **Revised**: 29/06/2023 **Accepted**: 07/07/2023

Abstract: Neoplasm of the mouth is one of the world's life-threatening diseases. It grows in the oral cavity and nasopharynx as a malevolent neoplasm. It is part of a wider group of cancers called cancer of the head and neck. They mostly form in the mouth, tongue, cheeks, hard and soft palate, sinuses, lips and pharynx in squamous cells (throat). Because of the unregulated growth of abnormal cells, the mortality rate is on the increase. In order to minimize the probability of death, early detection and diagnosis of oral cancer are essential. MRI images Lesion is an innovative medical concept. MRI images reflect an inventive concept. An oral malignancy may typically be measured by the use of an MRI scan or a CT scan. The scale and extent of oral cancer spreading can be seen in MRIs. The distinction between normal and abnormal tissue can be seen. MRI scan is a pioneering tool for oral cancer diagnosis. The method for studying the advancement of oral malignancy is non-invasive, radiation-free imaging. The present article compares modified KFCM to the K-means and Fuzzy C-means (FCM) and analyzes the precision achieved with the segmentation. The K-Means algorithm is used as a clustering technique to eliminate the calibration time. The FCM algorithm is used to minimize the total iteration generated by the initialization of the exact cluster. The morphological operation is used to excerpt the appropriate area from the FCM cluster. The lesion area is finally calculated. The proposed approach focused on modified anisotropic diffused filter image pre-processing to remove artifacts from MRI images and segmentation techniques using KFCM clustering, segmentation, lesion removal from MRI and the exact region of lesion evaluation. The modified KFCM algorithm is used in this research to increase segmentation precision (accuracy).

Keywords: *Oral and maxillofacial surgery, Image Pre-processing, segmentation of KFCM, Operation of morphological.*

1. Introduction

Oral cancer is part of a wider community of cancers called head and neck cancers. The tumor is a rapid growth of abnormal tissue in the body. Oral cancer may grow in abnormal cells of the mouth or nasopharynx, such as the lip liner, oropharynx, gingiva, tongue, soft and hard palate, cheeks, and the roof of the mouth. Risk factors include smoking, alcohol use and Human Papilloma Virus (HPV).The death rate of oral malignancy is 90% due to heavy obesity, smoking and alcohol intake and 10% due to poor oral hygiene, denture pain, poor diet, vitamin deficiency and age factors. Oral precancerous tumors are leukoplakia, erythroplakia, lichen-planus, and so on. Early detection and diagnosis of oral cancer is critical for effective treatment. The MRI scan is a radiation-free procedure that is not detrimental to humans relative to the CT scan. It provides detailed information on the normal and abnormal tissue of the MRI image. Radiologists, oncologists and dentists are analyzing MRI photographs to identify and diagnose oral tumors. This proposed approach used anisotropic diffusion filter to minimize noise, undesired distortions and improve image quality. This approach is used to increase the precision of the segmentation.

Image segmentation is the method of partitioning a digital image into different subgroups called image objects, which may minimize the complexity of the image. Picture segmentation is used to locate artefacts and boundary in photographs. Similarity and discontinuity are the basic characteristics of the image. Segmentation algorithm used one of the characteristics of the segmentation image. The clustering approach is based on the similarity property in which the cluster is divided into homogeneous clusters[1-3].The clustering approaches are used to diagnose diseases. In this research, KFCM clustering is used as a segmentation method for tumor detection and tumor removal from MRI images. Minimizes the computation time of clustering in segmentation. To finish, the lesion is segmented from the MRI image and calculates the exact position of the lesion region.

[1]*Research Scholar, CSE Dept, GNDEC, Bidar,* [2]*Principal, GNDEC, Bidar*
[1]*shravi97@rediffmail.com ,* [2]*dhan_mak@yahoo.com*

Following approaches are typically refer to the segmentation techniques: thresholding techniques, techniques in the field, techniques based on the edge, cluster techniques, techniques based on the water shed, and techniques based on the artificial neural network.

1.1. Thresholding Techniques

The thresholding is the easiest way in which an image segment is compared with the threshold value in each pixel. If the value of the pixel is smaller than the threshold value, the value is 0, or the maximum value (generally 255). You may arbitrarily modify the threshold value. It is useful for distinguishing the context. Three models – local, global and adaptive – comprise thresholding strategies. In local threshold values, the value of the threshold is chosen by using local properties such as the default value or the local mean value of the various image areas. The single threshold value in the image as an entirety is used on the basis of an image histogram in global thresholds. For adaptive or dynamic threshold values, local threshold values for each pixel are chosen independently. The segmentation effect is determined by the threshold. The inacceptable threshold value is chosen to contribute to inappropriate segmentation. One method used to simplify the process of threshold selection is Otsu's thresholding.

1.2. Region based Techniques

The field-based segmentation is a technique for direct decision making in the region. It includes an algorithm that divides the image in several components with similar features (pixel set) (homogeneity, texture, intensity levels and sharpness). In the first place, this approach looks for smaller seed dots in the input image or larger parts. Subsequently these techniques are employed either by adding additional pixels to the seed points or by reducing the seed point to small areas and combining them with smaller seed points. Regional development is a clear method of regional segmentation. This approach is extremely simple and can divide image pixels properly with similar characteristics to wide areas. This approach is correctly used to segment spatially separate regions and areas with similar properties. Through this technology, related regions are also developed.

1.3. Edge Based Techniques

This method is used to indicate the boundary of the object and to identify those objects on the defined edges of the image. Edge-based techniques display the pixels of the picture that match the object's edges visible in the picture. The effect is the generation of a binary image with the placed edge pixels. Some examples of edge-based techniques are Sobel, Canny, Prewitt edge detection algorithms. Edge techniques are perfect for transparent and noise-free photographs. These methods may lead to further or incomplete corners of noisy pictures. These techniques are computer-fast and no previous information on the contents of the image is required. Fake edge detection, edge position, lack of real edges, noise problems, and high computing time pose many problems.

1.4. Clustering based Techniques

The picture is divided into clusters or disjoint groups of pixels with similar features by cluster dependent techniques. The data elements are grouped into clusters using data clustering properties such that they are more like one element in the same cluster than other clusters. Widely used clustering algorithms are K-means clustering and Fuzzy C Means (FCM).

1.5. Watershed based Techniques

This procedure uses morphology of the image. This method is based on the topological understanding theorem. The gradient magnitude of the topography and its slope is clearly measured by the grey values of the respective pixels. In regions known as catchment basins Watershed breaches the picture. The watershed approach is based on two easy approaches – the local minimum image gradient calculated in the first step is used as a marker and then combined. In the second phase, the watershed is transformed by markers and the same marker positions are used. The major drawback of this strategy is the problem of over-segmentation and under-segmentation.

1.6. Artificial neural network-based techniques

Image segmentation is also known as neural network image recognition. Used by AI, images such as objects, faces, text, handwritten text etc. can be processed and classified automated. The CNN architecture is used to classify and process high-definition image data, which specifically utilizes neural networks for this method[5-6]. The picture is used as a series of vectors or as a raster depending on the method employed. The vector or raster is converted into single components that represent a picture of the physical objects. Computer vision systems can logically analyze these structures and extract the most important elements and organize data using function extraction and algorithm of classification.

The paper showed the following: Part 2 describes the previous work and part 3 provides a summary of the methods of the proposed solution. Section 4 then discusses the outcomes and discussions. Chapter 5 evaluation of results, and Chapter 6 conclusion.

2. Previous Work

Different segmentation approaches for oral cancer detection in medical images are outlined in the literature:

A dental image segmentation process for cysts or tumors [7-8]. The colored images were converted into grey images, and the Gaussian filter is glued to these images. The active format models are then segmented (snake). This method has not the whole picture, and has been applied to the region of interest. Study of the operational characteristic receiver (ROC) as an efficiency measure for this method.

Adv: In opposition to conventional edge segmentation, snakes minimize the need to link the edge

The morphological watershed algorithm for the identification of oral cancer [9-10]. The watershed algorithm runs by the marker for lesion segmentation. This algorithm is then compared to the segmentation of watersheds. In preprocessing (x-ray), linear contrast stretching pictures are done, and then marked segmentation is utilized. The result was higher than the watershed algorithm.

An efficient approach focused on the proposed network of artificial neurons [11-13]. 120 sample images have been used in this process. 30 photos were used during preparation, testing and validation. The oral lesions were observed and used for four statistical features. This projected methodology is an operative method of dental and oral disease recognition and detection. The diagnosis and projection of planus, leukoplasmia and oral squamous cell carcinoma was deliberate with this method.

The system proposed for oral cancer detection [14-15]. Linear contrast stretching used for elimination of noise preprocessing. The Watershed segmentation algorithm is used in the current approach. Marker Segmentation Managed watershed is also used to minimize the watershed algorithm's sectional problem. An upgraded Marker Controlled Segmentation process is used to accomplishing highest segmentation accurateness. The over-segmentation problem has been solved. The time of the measurement is determined before and after comparison.

The suggested fusion approach to FCM and neutrosophic segmentation proposed in [16-17]. In this picture, the jaw is segmented to distinguish jaw lesion. Efficiency and consistency have been tested. The study was performed for precision, sensitivity and similarity to enhance the robust quality of the method. The work was developed in relation to the Firefly Algorithm, FCM, and the FCM fusion method. It produces the same area of lesion as the manual delimitation of the oral pathologist and assesses good outcomes.

Automatic symmetric axis analysis proposed segmentation in [18-20]. The CBCT image tumor automated division is used. For the acquirement of good image quality, diffusion filters are used. An effective approach with novel feature facts is this automated exposure. Multi-resolution FFD for multi-scale registration of stronghold and heftiness. This is a truer optimistic solution with less false positive outcomes. Dice, Jaccard and Hausdorff are efficiency indicators for this automated segmentation.

Adv: Hierarchic FFD with multi-resolutions for quick and robust multi-scale registration.

Disadv: Inverse FFD Calculated FFD Registration.

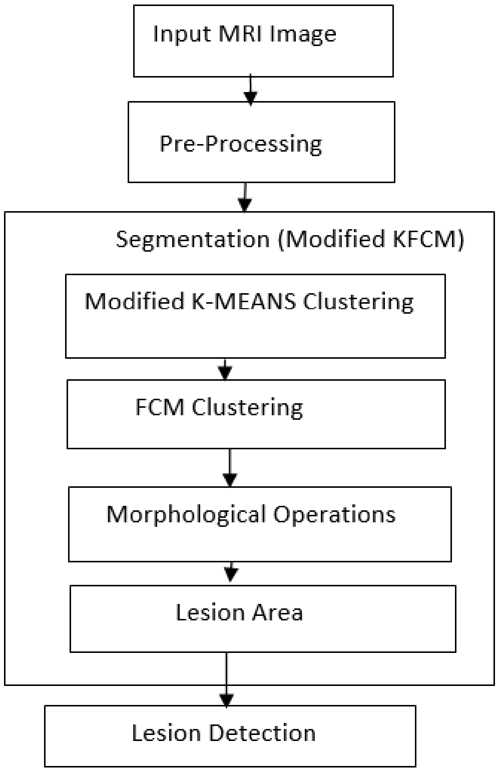
Limitations: This approach cannot handle tumors and lesions without unique limitations and cases with major asymmetric variables.

Fuzzy C-media for oral cancer detection Proposed clustering and filtering in [21]. This tactic focused on the anisotropic diffusion filter to segment the tumor in the pre-processing and clustering images to accomplish the quality of the CT image. Through the diffusion filter, they remove unwanted distortions in order to sharpen all boundaries and ensure the input picture is good enough. In the next step, the classification accuracy, properties and sensitivity in fuzzy c-measures were improved.

The proposed algorithm for morphological segmentation. CT images are used in this segmentation. Anisotropic diffusion filter used in pre-processing to reduce noise & improve input image configuration. Used segmentation based on segmentation In this region props used to extract features from the CT picture for the detection of oral cancer. Next used classifications like CNN, SVM & NAÍVE BAYS. The CNN, SVM & NAIVE BAYS grading accuracy is 96.15%, 93.20%, and 86.02% respectively. Proposed algorithm for adaptive threshold segmentation [23]. The images are used as CT images in this process. Anisotropic diffusion filter used in preprocessing to remove noise and improve image quality. Next segmentation of adaptive thresholds for segmentation. Features extracted from the GLCM and GLRLM intensity histograms. The accuracy of grading for the KNN, NB & SVM classifiers is 87.95 percent, 81.12 and 85.58 percent respectively.

3. Methodology

The approach established consists of various levels: 1) Pre-process Image 2) Segmentation of Image3) Detection of Lesion.



Input pictures (MRI) acquired are messy and originate from various sources. Diffusion filter named Modified Anisotropic is used in pre-processing to prevent undesirable falsification. Modified KFCM is a clustering strategy in segmentation, in which modified K-means accompanied by FCM. The morphology of the FCM cluster (predicting tumor), is then used to remove the lesion, for example dilation and erosion. The lesion region is eventually determined with the algorithm.

3.1 Image Preprocessing:

The MRI access is also susceptible to many forms of unintentional distortions. These distortions can diminish the correctness of segmentation. Smoothing image is an vital role to expand the superiority of the image in the processing of images.

3.1.1 Anisotropic Diffusion filter

It is a technique for minimizing noise without removing a significant part of the content of the image, typically edges, lines and further information that is necessary for picture interpretation. The diffusion is not a linear, but spatial transformation of the original image.

The anisotropic filter is commonly used to eliminate unwanted distortions, to improve grey image quality and to sharply sharpen the borders of other artifacts.

The overall anisotropic filter equation is:

$$\frac{\partial I}{\partial t} = div\big(c(x,y,t)\nabla I\big) = \nabla c.\nabla I + c(x,y,t)\Delta I$$

(1)

where: t is the time parameter

I is the original image

It is the filtered image

$\Delta I,\nabla I$ are the gradient and the laplacian of the image respectively

$c(x, y,t)$ is the diffusion constant

x, y, t: Represents the control rate of the diffusion coefficient.

It represents an image gradient which used for sharpening the boundaries in the image.

This Eq. 1 can be further reduced to:

$$It=c(x,y,t)\Delta I \quad (2)$$

The diffusion constant $c(x,y,t)$ is the primary edge stopping parameter which controls the filtering process and leads to the edge preservation. The value of the diffusion constant can be calculated by:

$c(x,y,t)=g(\nabla I)$ (3)

where: $g(\nabla I)$ is the conduction coefficient function and is represented by:

$g(\nabla I)=e(-(\|\nabla I\|/k)2)$ (4)

$$g(\nabla I) = \frac{1}{1+(\|\nabla I\|/k)^2} \quad (5)$$

where: k is the gradient modulus constant and $\|\nabla I\|$ is the parameterized image or the decomposed mask images. Eq (4) is used when high contrast is preferred over low contrast and Eq (5) is used when wider regions are preferred over smaller regions. It can be observed from Eq (3) that the value of the diffusion constant is dependent on the conduction coefficient. Further, the conduction coefficient depends on the parameterized images or the mask images. Therefore, these mask images play a crucial role in the estimation of the diffusion constant. In case of an MRI corrupted with the additive Gaussian noise, the decomposed mask images of that MRI will also contain biased pixel intensity values due to the noise. This will lead to an incorrect estimation of the diffusion constant, thereby, leading to degradation in the performance of the AD filtering. Therefore, in this, a modification in the AD approach has been proposed in order to remove this limitation.

3.1.2 Modified AD Filtering Algorithm

In the proposed algorithm, the Gaussian noise contaminated MRI is first decomposed into its respective set of mask images. In the decomposed mask images, the pixel intensities are corrupted due to the superimposition of the additive Gaussian noise leading to a bias in the intensities. Therefore, these mask images are subjected to the domain filtering function of the Bilateral Filter as a pre-processing step to the evaluation of the conduction coefficient. The domain function of the Bilateral Filter solely depends on the geometric closeness of the pixels and is therefore unaffected by the intensity bias produced by the noise. Thus, this function effectively reduces the noise present in the mask images. Mathematically, the domain filter function is represented as:

$$\mathrm{h}(x) = k_d^{-1} \sum_{y \epsilon \mho} f(y) c(y,x) \quad (6)$$

where: $h(x)$ is the output image, $c(y, x)$ is the parameter that measures the geometric closeness of the pixel x and its neighbour pixel y within the window Ω and d k is the normalization parameter given by:

$$k_d = \sum_{y \epsilon \mho} c(y,x) \quad (7)$$

The resulting mask images are then utilized for the calculation of the conduction coefficient using the Eq (4) or Eq (5) which is then further utilized by the AD algorithm for the noise suppression and structure preservation of the MRI.

3.2. Image Segmentation

The segmentation of the image is a well-known processing method which makes the foreground distinctive from the background. Segmentation is perfectly depending on the image pixel characteristics. The segmentation algorithm has two features, like discontinuity and similarity of the intensity values. Firstly, the image should be divided into different parts which meet the predetermined needs based on drastic changes in the levels of image strength, such as image limits. Secondly, an image should be divided into many related segments in accordance with predefined parameters.

3.2.1 Modified K-Means clustering

K-Means is most straightforward, iterative, unsupervised learning algorithms. The traditional kmeans process follows a simple and easy way to classify a given image through a certain number of clusters which are fixed apriori.The traditional K-means algorithm randomly selects k clustering centers, which has poor clustering.The complexity of modified K-means is lower and the method is easy to implement. Thus,modified K-means is adopted to initialize the cluster centroid in this approach. Modified K-means is based on K-means, which can initialize the centroids deterministically. The basic principle of the modified K-means algorithm for initialization of cluster centroids is to maximize the distance between the initial cluster centroids. This method allows deterministically initializing cluster centroids, overcoming the shortcomings of the Kmeans algorithm associated with its initialization instability.

. The initialization process of the modified K-means algorithm is as follows:

(1) Randomly select a sample point from the data set as the first initialized cluster centroid.

(2) Select the remaining cluster centroids:

(a) Calculate the distance between each sample point in the sample and the cluster centroid that has been initialized, and then select the shortest distance among them, denoted as di.

(b) Select the sample with the largest distance by probability as the new cluster centroid.

(c) Repeat the above process until k cluster centroids are determined.

(3) For the initial cluster centroids, the final cluster centroids are calculated using the K-means algorithm.

3.2.2 Fuzzy C-Means clustering

It is a unsupervised clustering algorithm. FCM makes data points in more than one cluster (pixels to be clustered). Some pixels can belong to more than one group or cluster, but can be related to various levels by group. The FCM algorithm is optimized by reducing it and the convergence of the algorithm.

There are several clusters using one data piece in this scheme. A partial value for each pixel of the image is allocated in the fuzzy logic. Fuzzy Set membership ranges from 0 to 1. By membership feature that represents the fuzzy role of this algorithm, data are linked to each cluster. The Fuzzy partition is achieved with the update of the member and cluster centre by iteratively optimizing Objective functions. A polyvalent logic is the clustering of fuzzy components. Clusters for each cluster are planned based on the distance between data points and cluster centers. Fuzzy C-Means (FCM) is a clustering technique, in which several clusters with different membership levels are identified by data object.

The following goal function is minimized by FCM:

$$J_m = \sum_{i=1}^{D} \sum_{j=1}^{N} u_{ij}^{m} \, ||x_i - c_j||^2$$

Where,

- D is the data points number.
- N is the total cluster number.
- M is an exponent of the fuzzy partition matrix for the control of the m>1 overlap.
- Fuzzy overlap shows how indistinct the boundaries between the groups are.
- Xi is the point of the information about ith.
- Cj is the centre of the j cluster.
- μij is the xi stage in combination with jth-cluster (group).The number of membership values for each cluster is one for a given data point, xi.

Algorithm:

The FCM clustering algorithmic steps are provided below:

1. Initialize the membership cluster randomly, μi.
2. The cluster centers should be determined:

$$C_j = \frac{\sum_{i=1}^{D} u_{ij}^{m} x_i}{\sum_{i=1}^{D} u_{ij}^{m}}$$

3. Apprise μ_{ij} as persucceeding:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{N} \left(\frac{||x_i - c_j||}{||x_i - c_k||}\right)^{\frac{2}{m-1}}}$$

4. Objective function must be calculated, Jm.

5. Step 2 to 4 repeat to boost Jm below the minimum threshold or to determine a maximum number of the iterations.

3.2.3. Morphological Operations

A series of non-linear operations relating to the form or morphology of image features is known as morphologic processing operations. These operations are suitable for binary image processing.

3.2.3.1. Dilation

The design element assigned to the image is used for the analysis and expansion of the forms found in the input image. Dilation affects erosion in opposite ways. It applies a pixel layer on both interior and external boundaries. That is, the maximum value of the output pixel in the vicinity is all pixels. A pixel is set to 1 when any of the next pixel is set to 1 in a binary image. Morphologic dilation enhances the visibility of the objects and covers them with small holes.

3.2.3.2. Erosion

The dilation is directly opposite. For testing and reducing the forms in the input picture, the assigned structuring factor acts like a local minimum filter in particular. Furthermore, the structuring elements reduce the image by removing a pixel layer both inside and outside regions. By means of erosion, we can remove the hole and distance between various areas. That is, the output pixel value for all neighborhood pixels is the minimum value. A pixel is fixed with a binary image to 0, when one of the next pixel value is 0. Morphological erosion is taking away islands and small objects, only substantial objects remained.

K-Means and FCM are used to collect the tumor predicted at the FCM prediction cluster (Fuzzy C-Means). Use morphological processes such as dilation and erosion to this output cluster section tumors.

3.2.4. Lesion area

In the estimated reasoning step, the lesion zone is calculated with the binarization process. Compared to a threshold, each transform coefficient. If the threshold value is lower, it is zero. It's one, otherwise. A picture has two values, whichever 0 or 1 for each pixel like black or white. Here is 256x256 jpeg determined image magnitude. To reflect the binary image, two colors are used.

Image,

I=

$$\sum_{W=0}^{255}\sum_{H=0}^{255}[F(0)+F(1)]$$

Pixels= Width (W) X Height (H) =256 X256

F (0) = white pixel (digit 0)

F (l) = black pixel (digit 1)

No_of_white_pixel

$$P=\sum_{W=0}^{255}\sum_{H=0}^{255}[F(0)]$$

Where,

P = number of white pixels (width*height)

1 Pixel = 0.264 mm

The area calculation formula is

Size_ of_ tumor, S = [(√p)* 0.264] mm2

Where,

P= no-of white pixels, W=width, H=height

3.3. Lesion Detection

Algorithm:

The oral lesion detection steps are described below.

1) Processes to initialize
2) Receive the JPEG input MRI image.
3) Check if the picture is in proper format you want and then go to step 4.
4) Convert the picture to grayscale if MRI image is in the rgb format else go to next step.
5) Detect the grayscale picture's edge.
6) Describe how many white pixels the image includes.
7) The lesion region using the formula is estimated.
8) Show the lesion location.
9) Compute the region area of interest which must be 5 mm2. The MRI image has an abnormal lesion if the region in the lesion is greater than 5mm2.
10) The procedure is over.

The RGB or grey image was first used by this algorithm. Next, the binarization method used to convert the grey to binary image. Afterward binarization, the edge of tumors pixels in the binary image was observed. Future, this algorithm calculates the extent of the tumor by binary white pixel representation (digit 0).

The projected area of the tumor is estimated in the last phase. Remember that, when the tumor region is greater than 5.00 mm^2, it will be detected as a tumor.

In the State flow procedure, 7 steps pseudo-code of the proposed work is stated. The steps can be designed to preprocess, clustering, segmentation, extraction, detection and calculation of lesion area from MRI images.

Pseudo code

1. Oral normal & abnormal lesion dataset inputs n MRI

2. Modified Anisotropic diffusion filter pre-processing data.

3. Apply modified K-means clustering method for a minimum calculation time cluster of data.

4. To minimize cumulative iterations, apply the FCM clustering process.

5. To remove oral tumors from the FCM cluster, use morphological operations like dilation and Erotic care.

6. Use an algorithm to calculate the region of lesion with exact position.

7. Detection of lesion (if lesion area is greater than 5mm2 then lesion present otherwise lesion absent)

4. Results And Discussion

In this analysis, modified KFCM is used to separate the tumor from MRI images in the segmentation technique. It reduces clustering or segmentation time. Lastly, the lesion is separated from the MRI image and the lesion region is calculated at the exact location.

.The MRI image is filtered to remove unnecessary falsification and enhance the high grade of the MRI image in image-preprocessing. The MRI image is subsequently divided into segmentation process. Initially, the clustering of modified K-Means separates the preprocessed image into clusters with or around the identical strength value. Then, FCM is used to minimize the total iteration to obtain an exact FCM cluster. FCM's output is a cluster that predicts tumor. Lesion from this cluster is observed. The tumor is extracted from predicting cluster using for morphological operations. Finally, use the algorithm to measure the region of the tumor.

The dataset of 100 MRI images are collected from JJM medical college, Davangere. For testing these images, MATLAB tool is used with version 2017. 10 MRI images which are addressed and selected randomly for the productivity of these images. The table units are: image name, lesion location, time spent and detection decisions i.e. lesion present and lesion absent.
Time delayed: the amount of time required for dividing an image input. Tic and Toc commands in Matlab measure the computer time of this algorithm.

Image name	**Area of the Lesion(mm2)**		
	Modified KFCM	**FCM**	**K-Means**
Sample 1	**13.58**	**13.78**	14.85
Sample 2	**0.64**	**0.69**	3.08
Sample 3	**0.95**	**0.95**	5.09
Sample 4	**12.4**	**12.9**	13.74
Sample 5	**1.73**	**1.77**	5.40
Sample 6	**14.5**	**14.4**	15.7
Sample 7	**7.67**	**7.49**	8.65
Sample 8	**8.51**	**8.43**	9.50
Sample 9	**1.69**	**1.69**	5.65
Sample 10	**1.37**	**1.49**	4.49

Table1: Table of Lesion Area

Image name	**Elapsed Time(Sec)**		
	FCM	**K-Means**	**Modified KFCM**
Sample 1	**15.3**	**1.92**	**2.64**
Sample 2	**15.1**	**1.83**	**2.67**
Sample 3	**15.1**	**1.90**	**2.62**
Sample 4	**15.0**	**1.70**	**2.76**
Sample 5	**15.3**	**1.60**	**2.70**

Sample 6	**15.7**	**1.52**	**2.56**
Sample 7	**15.7**	**2.28**	**2.55**
Sample 8	**15.2**	**2.15**	**2.54**
Sample 9	**14.4**	**1.82**	**2.62**
Sample 10	**14.8**	**2.16**	**3.01`**

Table2: Table of Elapsed Timex

Image Name	**Modified KFCM**		**Lesion De tection**
	Area of the Lesio n(mm2)	**Elapsed T ime(Sec)**	
Sample 1	**13.58**	**2.64**	**Present**
Sample 2	**0.64**	**2.67**	**Absent**
Sample 3	**0.95**	**2.62**	**Absent**
Sample 4	**12.4**	**2.76**	**Present**
Sample 5	**1.73**	**2.70**	**Absent**
Sample 6	**14.5**	**2.56**	**Present**
Sample 7	**7.67**	**2.55**	**Present**
Sample 8	**8.51**	**2.54**	**Present**
Sample 9	**1.69**	**2.62**	**Absent**
Sample 10	**1.37**	**3.01`**	**Absent**

Table3: Table of Lesion Area, Elapsed Time & Lesion Detection of Modified KFCM

Compared to three algorithms, Table 1 & Table 2 shows the lesion zone and the elapsed time comparison. Table 3. displays a message of lesion, whether present or missing, to reliably detect a regular, anomalous picture of the proposed.

5.Performance Evaluation

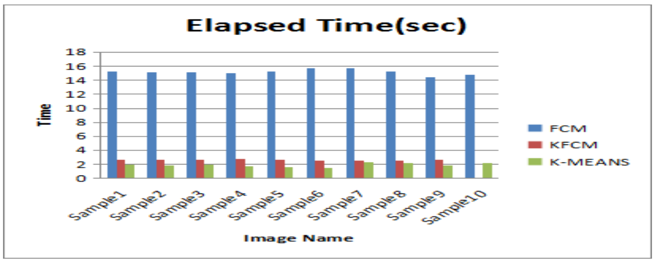


Fig. 1: A graph of Lesion Area Comparison

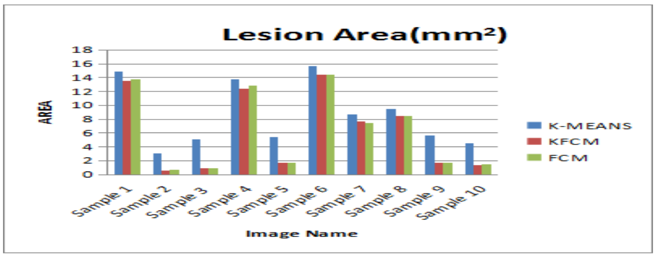


Fig. 2: A graph of Elapsed Time Comparison

Usage of algorithms for FCM and K means: 1) FCM Clustering is used to minimize the total iteration and accurate lesion area segmentation. 2) To reduce time for computation, clustering K-means is used.

This section clearly shows that, associated to the lesion part demonstrated in Fig.1, the lesion part of the modified KFCM system is nearer to the FCM lesion area. Delineation of FCM, modified KFCM segmentation is done under the guidance of radiologist manually. The calculation time of modified KFCM however, as shown in fig.2 is smaller than that of the FCM and close to the K-means. With modified KFCM segmentation, this procedure extracts accurate area of the tumor from the MRI image.

A dataset of 100 MRI images are collected with lesions and those without lesions in order to achieve the outcome of the methodology. Out of 100 sample pictures, the results are evaluated by 5 measures with 80 oral lesions and 20 without lesion pictures.

This is the Precision, Recall, Accuracy, Dice and Jaccard index for segmentation.

Segmentation Correctness: This means, the number of MRI images properly segmented lesions over the total number of MRI images.

$$\textbf{Accuracy} = \frac{(TP+TN)}{(TP+FP+TN+FN)}$$

Sensitivity: It denotes to the True Positive Rate of a number of properly segmented oral Lesion MRIs.

$$\textbf{Recall} = \textbf{Sensitivity} = TP / (TP+FN)$$

$$\textbf{Precision} = TP / (TP+FP)$$

Specificity: This refers towards true negative rate, which is the correctly segmented MRI without oral lesions.

$$\textbf{Specificity} = TN / (TN + FP)$$

$$\textbf{Dice} = 2TP / (2TP+FP+FN)$$

$$\textbf{Jaccard Index} = TP / (TP+FP+FN)$$

Where,

True Positive(TP) = Amount of correctly segmented oral lesion MRI images.

False Negatives(FN) = Amount of MRIs that are not orally damaged, incorrectly segmented.

True Negative(TN) = Amount of properly segmented, non-oral lesions MRI images.

False Positive(FP) = Amount of MRI frames with oral lesions wrongly segmented.

	Number Of MRI Images	**Correctly Segmented MRI Images**		
		Modified KFCM	**FCM**	**K-MEANS**
No Lesion	**20**	**18**	16	11
Lesion	**80**	**77**	74	69
Total	**100**	**95**	90	80
Accuracy		**95%**	90%	80%

Table 4. Segmentation Accuracy of the Proposed System

Method	Recall	Precision	Dice	Jaccard Index
K-MEANS	**0.884**	**0.862**	**0.873**	**0.775**
FCM	**0.948**	**0.925**	**0.936**	**0.880**
Modified KFCM	**0.974**	**0.962**	**0.968**	**0.939**

Table 5. Experimental Results

The experimental results of the proposed method show in Tables 4 & 5 that the Correctness of the segmentation obtained is 95 percent following testing of the proposed system. Recall, accuracy, dice and jaccard index are segmentation metrics used in this method. These measurements have been achieved by the values 0.974, 0.962, 0.968 and 0.939. In segmenting MRI image lesions, by means of modified KFCM segmentation method, the proposed approach shows an excellent performance.

6. Conclusion And Future Work

Many kinds of lesions exist. Lesion in the mouth or oral cavity may consist of a soft tissue or hard tissue. When a soft tissue remains, it is quickly isolated from the oral tissue by a K-Means algorithm. If the MRI image includes noise or some other falsification, it is deleted during pre-processing. The MRI image will then be given as a modified K-Means input. In this study, the initial clustering of the MRI images uses modified K-Means algorithm. Following an initial segmentation, an acceptable group was chosen according to the revised membership criteria. Then, a predictive cluster of tumor with FCM is obtained. Then, the picture is separated by using morphology and finally determined by the tumor region. The experimental results clarify the better results with less performance time and accurate lesion surfaces.

Modified KFCM's segmentation correctness, as shown in table 1 & 2, is compared to FCM & K-Means experiment. These findings demonstrated increased efficiency in the segmentation of MRI lesions by modified KFCM segmentation techniques with a high 95% accuracy. For more segmentation correctness, advanced algorithms can be used in future.

References

[1] Nurtanio I, Purnama IKE, Hariadi M, Purnomo MH. Cyst and tumor lesion segmentation on dental panoramic images using active contour models. IPTEK J Technol Sci 2011;22(3).

[2] Anuradha K, Sankaranarayanan K. Detection of oral tumor based on the marker-controlled watershed algorithm. Int J Comput Appl 2012;52(2).

[3] Tripathi, M.K. and Maktedar, D.D., 2020. A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: A survey. Information Processing in Agriculture, 7(2), pp.183-203.

[4] Maghsoudi R, Bagheri A, Maghsoudi MT. Diagnosis prediction of lichen planus, leukoplakia, and oral squamous cell carcinoma by using an intelligent system based on artificial neural networks. Dentomaxillofacial Radiol, Pathol Surg 2013;2(2):1–8.

[5] Anuradha K, Sankaranarayanan K. Oral cancer detection using an improved segmentation algorithm. Int J Adv Res Comput Sci SoftwEng 2015;5(1):451–6.

[6] Tripathi, M.K. and Maktedar, D.D., 2021. Detection of various categories of fruits and vegetables through various descriptors using machine learning techniques. International Journal of Computational Intelligence Studies, 10(1), pp.36-73.

[7] M. K. Alsmadi, “A hybrid Fuzzy C-Means and Neutrosophic for jaw lesions segmentation," [Press release], 2016. Retrieved from https://doi.org/10.1016/j.asej.2016.03.016.

[8] Tripathi, M.K. and Maktedar, D.D., 2016, August. Recent machine learning based approaches for disease detection and classification of agricultural products. In 2016 international conference on computing communication control and automation (ICCUBEA) (pp. 1-6). IEEE.

[9] F., Abdolali, R. A., Zoroofi, Y., Otake, & Y., Sato, " Automated classification of maxillofacial cysts in cone-beam CT images," Computers in biology & medicine, 72 108-119, 2016.

[10] Zhalong Hu, AbeerAlsadoon, Paul Manoranjan 1, P.W.C. Prasad, Salih Ali, A. Elchouemic, "Early Stage Oral Cavity Cancer Detection: Anisotropic Pre-Processing and Fuzzy C-Means Segmentation, "IEEE 978-1-5386-4649-6/18 2018.

[11] R. Prabhakaran,Dr. J. Mohana,”Detection of Oral Cancer using Machine Learning Classification Methods”,International Journal of Electrical Engineering and Technology (IJEET) Volume 11,

Issue 3, May 2020, pp. 384-393, Article ID: IJEET_11_03_041

[12] Shivendra, Kasa Chiranjeevi, and Mukesh Kumar Tripathi. "Detection of Fruits Image Applying Decision Tree Classifier Techniques." In Computational Intelligence and Data Analytics: Proceedings of ICCIDA 2022, pp. 127-139. Singapore: Springer Nature Singapore, 2022.

[13] Dr. S. Shajun Nisha1,M. Sakku Vidhya2,Dr. M. Mohamed Sathik3,’’A Comparative Analysis of Classification Algorithms for Oropharyngeal Cancer Detection’’, Journal of Composition Theory,Volume XIII, Issue IV, APRIL 2020

[14] Channapattana, Shylesha V., Srinidhi Campli, A. Madhusudhan, Srihari Notla, Rachayya Arkerimath, and Mukesh Kumar Tripathi. "Energy analysis of DI-CI engine with nickel oxide nanoparticle added azadirachta indica biofuel at different static injection timing based on exergy." Energy 267 (2023): 126622.

[15] Tripathi, M.K. and Maktedar, D.D., 2021. Optimized deep learning model for mango grading: Hybridizing lion plus firefly algorithm. IET Image Processing, 15(9), pp.1940-1956.

[16] NikhitaBiradar, Prakash H. Unki,’’Brain Tumor Detection Using Clustering Algorithms in MRI Images,’’International Research Journal of Engineering and Technology 1587-91, 2017;

[17] Tripathi, Mukesh Kumar, and Dhananjay D. Maktedar. "A framework with OTSU’S thresholding method for fruits and vegetables image segmentation." International Journal of Computer Applications 975 (2018): 8887.

[18] Eman Abdel Maksoud, Mohammed Elmoji, Rashid Al-Awadi, "Brain Tumor Segmentation Based On Hybrid Clustering Technique, " Elsevier B. V. on behalf of Faculty of Computers & Information, Cairo University.

[19] Chiranjeevi, K., Tripathi, M.K. and Maktedar, D.D., 2021, March. Block chain technology in agriculture product supply chain. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 1325-1329). IEEE.

[20] Eman A. Abdel Maksoud, Mohammed Elmogy, and Rashid Mokhtar Al-Awadi,” MRI Brain Tumor Segmentation System Based on Hybrid Clustering Techniques,” Springer International Publishing Switzerland 2014, CCIS 488, pp. 401–412, 2014.

[21] Tripathi, M.K. and Maktedar, D.D., 2022. Internal quality assessment of mango fruit: an automated grading system with ensemble classifier. The Imaging Science Journal, 70(4), pp.253-272.

[22] Mr. A. Kingsly Jabakumar. (2019). Enhanced QoS and QoE Support through Energy Efficient Handover Algorithm for UMTS Architectures. International Journal of New Practices in Management and Engineering, 8(01), 01 - 07. https://doi.org/10.17762/ijnpme.v8i01.73

[23] Krishna, K. S. ., Satish, T. ., & Mishra, J. . (2023). Machine Learning-Based IOT Air Quality and Pollution Detection. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2s), 132–145. https://doi.org/10.17762/ijritcc.v11i2s.6036

[24] Shukla, A., Juneja, V., Singh, S., Prajapati, U., Gupta, A., & Dhabliya, D. (2022). Role of hybrid optimization in improving performance of sentiment classification system. Paper presented at the PDGC 2022 - 2022 7th International Conference on Parallel, Distributed and Grid Computing, 541-546. doi:10.1109/PDGC56933.2022.10053333 Retrieved from www.scopus.com