



An IoT Based Predictive Modelling for Predicting Lung Cancer Using Fuzzy Cluster Based Segmentation and Classification

D. Palani¹ · K. Venkatalakshmi²

Received: 25 September 2018 / Accepted: 9 December 2018 / Published online: 18 December 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

In this paper, we propose a new Internet of Things (IoT) based predictive modelling by using fuzzy cluster based augmentation and classification for predicting the lung cancer disease through continuous monitoring and also to improve the healthcare by providing medical instructions. Here, the fuzzy clustering method is used and which is based on transition region extraction for effective image segmentation. Moreover, Fuzzy C-Means Clustering algorithm is used to categorize the transitional region features from the feature of lung cancer image. In this work, Otsu thresholding method is used for extracting the transition region from lung cancer image. Moreover, the right edge image and the morphological thinning operation are used for enhancing the performance of segmentation. In addition, the morphological cleaning and the image region filling operations are performed over an edge lung cancer image for getting the object regions. In addition, we also propose a new incremental classification algorithm which combines the existing Association Rule Mining (ARM), the standard Decision Tree (DT) with temporal features and the CNN. The experiments have been conducted by using the standard images that are collected from database and the current health data which are collected from patient through IoT devices. The results proved that the performance of the proposed prediction model which is able to achieve the better accuracy when it is compared with other existing prediction model.

Keywords Transitional image extraction · Edge detection · Segmentation · Fuzzy C-means clustering · Lung cancer image · Association rule mining (ARM) · Decision tree (DT) · Incremental classification · Convolutional neural network (CNN) · Internet of things (IoT)

Introduction

The rapid growth of Internet of Things (IoT) is very interesting for all people due to its various utilizations in different fields. It has enormous number of applications in various fields such as healthcare, smart city, etc. These applications

are used to reduce the time in emergency room waiting time, monitoring the patient health, monitoring the staff activities, inventories and managing the critical devices such as hearable which is useful for the people who affected hearing loss interact with the world. Next, the ingestible sensors are to be considered as a marvel in this modern-science world. These kinds of sensors are used to monitor the medication in our body and also to give a warning to us if it identifies any irregularities such as a symptom of diabetic, heart disease, etc. in human bodies. Moreover, computer technology along with Artificial Intelligence (AI) techniques that aimed to mimic the visual perception and hence decision making based on it. Specially, in drones like Skydio uses the computer technology for detecting obstacles and also used for navigating around them. In addition, healthcare charting is also can be done by using IoT devices including Audemix for reducing the manual work that has to do while preparing patient charting.

Predictive modelling is a process which applies the data mining and the conditional probability for predicting or forecasting the future outcomes. In this scenario, each of this

This article is part of the Topical Collection on *Image & Signal Processing*

✉ D. Palani
dpalani.phd2018@gmail.com

K. Venkatalakshmi
venkata_krish@yahoo.com

¹ Department of Electronics and Communication Engineering, University College of Engineering Villupuram, Villupuram, Tamilnadu 605103, India

² Department of Electronics and Communication Engineering, University College of Engineering Tindivanam, Tindivanam, Tamilnadu 604001, India

modelling can be made up of a set of predictors that are variables which are likely to be influenced in the future results. After collecting the necessary data by the predicting model, a new mathematical or statistics formula or soft computing technique based machine learning algorithm has been formulated and predicted for predicting the future effectively. The majority predictive models are available in the direction of e-business based on the past history. Other more complex predictive models include decision trees, k-means clustering and Bayesian inference and the complex predictive modelling is the neural network. Moreover, the neural network is working based on artificial intelligence including the image recognition, the smart assistants and the natural language generation. Hence, this paper proposes a new predictive model as application for predicting the lung cancer by using the image and smart data which has been collected from IoT device.

Cancer is a serious and kills disease in this world to all the human being life for the last few 10 years due to the human beings food habit. This disease affected peoples count is also gradually increasing every day. Moreover, it is also expecting to raise an essential health issue which is identified by the World Health Organization (WHO). In addition, the cause of death count is around 13% out of overall deaths in all over the world in [1]. Moreover, the WHO is also predicted that this count may increase reasonably in the year of 2030. Recently, the survival counts of cancer patients are reduced when it is compared to the earlier days due to the severity of the cancer disease. The cancer disease prediction is very important for reducing the severity of the disease and also for avoiding the death even the person affected by the cancer. The various types of cancers such as Bladder cancer, Lung cancer, Brain cancer, Melanoma Breast cancer, Non-Hodgkin lymphoma, Cervical cancer, Ovarian cancer, Colorectal cancer, Pancreatic cancer, Esophageal cancer, Prostate cancer, Kidney cancer, Skin cancer, Leukemia, Thyroid cancer, Liver cancer and Uterine cancer. From these, this research work focuses to detect the lung cancer. For this purpose, new segmentation algorithm, feature extraction, feature selection and classification algorithms have been proposed for predicting the lung cancer.

Lung cancer is a serious death disease and also carrying a miserable prognosis with a 5 years survival rate at only 18% [2]. The treatment therapy monitoring and the lung nodule analysis [3] by using the computed tomography (CT) medical images that are having useful strategies to diagnosis the lung cancer early and also to monitor the severity. These methods expressed that the need of lung nodule segmentation accuracy which is able to affect directly in the subsequent analysis results. Especially, the fact of growing the volumes of medical imaging that is given and developed a segmentation method that is data-driven of great medical needs for avoiding the tedious, manual processing and also reduces the variability of observer [4]. Image segmentation is a challenging task in

the medical image analysis that endeavours for dividing an image into many non-overlapping meaningful regions with homogeneous characteristics with respect to the image texture, image grey scale value, image position, etc. Sharma and Aggarwal [5] and also helps the radiologist for segmenting the part of tumour or unhealthy part of the medical images. Recently, the rapid growths of various methods have been proposed by many researchers for achieving better accuracy on image segmentation process recently [4, 6, 7]. For achieving better accuracy, continuous the segmentation performance which needs more attention due to the heterogeneity of the lung nodules on CT images.

Deep Learning is a major field of Machine Learning that applies neuron like mathematical formats to learn the tasks [8]. In the past, the Neural Networks (NN) are playing crucial role over the decision making process in many decades [8] and also to have gain and loose the favour of the researchers. The combination of deep learning and neural networks is called Deep Neural Networks (DNNs) that overcome the classification challenges on ImageNet. Since, DNNs have attracted the enormous amount of researchers and the multiple DNN structures including Convolutional neural networks (CNNs). Kunihiko [8] introduced the Convolution Neural Network (CNN) and it has three types of layers such as convolution, max-pooling and fully connected. Here, the convolution consists of neurons which are rectangular shape and each neuron take input from the previous neuron which is in same shape. The convolutional layers neurons have similar weightage. Max-pooling layer is the next layer of convolutional layer which takes rectangular grid from the convolutional layer and it produce a single output. Next layer is called fully connected which is occurs after the presence of many convolutional and max-pooling layers to select maximum of block. Generally, the CNN works in two ways such as forward propagation and back propagation. Both these ways works as per the regular NN functionality with convolutional layer and the Max-pooling layers.

In this paper, we propose a new IoT based Predictive Model for predicting the lung cancer disease by using the affected medical images which are given as input to the real-time application that has been developed as prediction system in this work. Moreover, the major contributions of this paper are as follows:

1. Propose a new image conversion methodology for performing effective pre-processing and the classification.
2. Introduce a new segmentation algorithm for effective segmentation over the lung cancer images.
3. Uses an effective clustering algorithm called Fuzzy C-Means clustering algorithm with temporal features for grouping the lung images.
4. Uses the Convolutional Neural Network for effective prediction.

5. Propose a new incremental classification algorithm with temporal features which is the combination of ARM and DT with temporal features for classifying the lung images.
6. Conducted the various experiments by using the medical dataset and the real time data that are collected from patient through IoT devices.
7. The experiments have been conducted for measuring the performance of the proposed model which is available in this paper.

The rest of this paper is organized as follows: Section 2 provides the detailed survey in the direction of diagnosing the cancer disease by using the report of image analysis on prediction systems. Section 3 explains the overall system architecture of the system. Section 4 described in detail about the proposed system. Section 5 shows the experimental results and discussion. Section 6 gives the conclusion and the future works.

Literature survey

Many works have been proposed by various researchers in the past in this direction. In addition, various IoT based healthcare applications have been developed by many researchers in the past [9–11]. Among them, Kannan et al. [12] developed an effective fuzzy c-means method for effective segmentation of the Magnetic Resonance Images that is seriously affected by intensity in homogeneities which are created by the radio-frequency coils. In their work, the initial cluster centers were assigned by using the newly proposed center initialization method for executing the fuzzy c-means clustering algorithm iteratively. Bing et al. [13] proposed a new algorithm called fuzzy level set algorithm for facilitating the medical image segmentation. Moreover, the proposed algorithm can evolve directly from the initial image segmentation process by using fuzzy clustering with spatial constraints. Pei et al. [14] proposed a new region based level set approach by using both the global image information and the local image information complementarily. They demonstrated that the efficiency and accuracy of their method in image segmentation over the medical image segmentation.

Manikandan et al. [15] real coded the existing genetic algorithm with the simulated binary crossover based multilevel thresholding which is used to segment the medical brain images. Nima et al. [16] introduced a new neural network based approach for medical image segmentation. They used the modified self-organizing map (SOM) network called moving average SOM (MA-SOM) that utilized for segmenting the medical images. After perform the initial segmentation, a new merging process was designed for connecting the image objects of a joint cluster together. Their approach segments the X-ray that is computerized tomography (CT) and the magnetic resonance (MR) head medical images is much better than the incremental supervised neural network (ISNN) and SOM-based methods.

Yangyang et al. [17] introduced a new particle swarm optimization algorithm which is working dynamic in nature with context cooperative and the quantum behaviour. Their algorithm incorporates new methodology for updating the context vector dynamically in each time and it completes a cooperation operation with other particles. They explained that how their method leads to enhance the search ability and the improved optimization over the previous methodologies and demonstrate that the empirical with experiment results using the standard benchmark test functions. Xiaoli et al. [18] proposed a new algorithm called bi-directional region growing segmentation algorithm for image segmentation. Their algorithm concentrates over the images noises and the pixel processing. Their method holds for the other segmentation applications in that image background region is simple but the target region of the image is complex. Moreover, they have proposed a neighbouring difference transform for selecting the right threshold.

Dwarikanath [19] developed a new method for combining the multiple expert annotations using graph cuts and the semi supervised learning. Sourav et al. [20] presented an application of a genetic algorithm (GA) which is based on segmentation algorithm for automatic grouping of unlabelled pixels of the MR images into different homogeneous clusters. Soumen et al. [21] incorporated the curve fitting method on an image for acquiring the segmented image thereby extracting the information from the medical images. Payel et al. [22] presented a new genetic algorithm for combining the representations of learned about the details such as shapes, the regional properties and the relative position of various objects into a single framework for automating the three-dimensional segmentation process. Their algorithm has been tested by conducting the various experiments for prostate the process of image segmentation over the medical images which are magnetic resonance.

Pratik et al. [23] designed a new tool called Medical Image Segmentation Tool that has two stages of algorithm. In the first stage, it generates an image which is binary marker of the region of interest using the morphological analysis using mathematical functions. The resulted image of the first stage is to use as mask image for the second stage that uses the method called Grab Cut for yielding an effective segmentation result. Their experimental results can be refined further by user interaction that was done by using the proposed Graphical User Interface (GUI). Their experimental results demonstrated that the proposed tool is accurate and also provides the satisfactory segmentation results with minimum user interaction over the medical and natural images. Sanping et al. [24] presented a new correntropy-based level set method (CLSM) for segmenting the medical images and also bias the fields of correction. They have considered two kinds of medical images such as synthetic and real medical images that facilitate to evaluate their method which shown the significant improvements over the efficiency and the accuracy of both of these images.

Yao-Tien Chen [25] proposed a new method for the medical image segmentation and also performs the visualization process with additional surface areas and the volume of measurements for the brain medical image analysis. Their method contains the edge detection and the Bayesian segmentation, surface and volume rendering, surface area and the volume of measurements for the 3D image objects of interest. Sheng-Chih Yang et al. [26] presented a new and robust segmentation method for medical image segmentation which is not only addressed the earlier challenges. It is also preserved the advantages of the existing methods for achieving the highest accuracy over the image segmentation. They have achieved best results delivered by the proposed progressive support-pixel correlation statistical method for the real time images such as medical images that are classified as simulated medical images, actual mammograms which is single-spectral and the multi-spectral breast magnetic resonance images. Finally, they compared their results with many existing and the similar kinds of segmentation algorithms for confirming the advantages and the contributions of their method.

Khanfir et al. [27] presented a new image segmentation algorithm for imitating the visual attention which is human focused in image interpretation using the possibilistic knowledge modelling concepts. Moreover, their pixel level approach consists over the Iterative Possibilistic Knowledge Diffusion over the immediate neighbourhood pixels. The major advantage of their model provides an iterative diffusion of per-pixel that has specific knowledge for surrounding the pixels to refine the segmentation process progressively. Moreover, their model achieved the lowest error rates and the possibilistic knowledge diffusion by using the existing filter called Nagao filter that is also adopted for the approach assessment.

Qiang et al. [28] introduced a new method which integrates the existing classifier called Support Vector Machine and the graph cuts for performing segmentation over the medical images. Moreover, they proposed a new training scheme which is localized for training a classification algorithm for each pixel that is working based on the target image information. In addition, a new graph cuts based segmentation method which is the combination of the constraint information of machine learning result, the edge information, the local information and the remote-local information is also proposed for post-processing. The experiments on synthetic and medical images demonstrate that the proposed method can achieve better performance than the state-of-the-art. Chunxiao et al. [29] developed a new segmentation algorithm which is weighted and variational selective approach. It contains two phases in which the first phase is for obtaining the related approximation that is smooth in nature using the existing method called Mumford-Shah method for achieving the target region in the input image. Here, the approximation method provided a bigger value for achieving the target region and the smaller values for other regions. The second of their method is to use the approximation and also to perform a thresholding task for obtaining the object of interest. The

approximation value was obtained by the multipliers and also established the convergence analysis over the proposed new approach.

Michal et al. [30] introduced a simple and yet powerful pipeline for medical image segmentation method which combines fully convolutional networks with fully convolutional residual networks. They also examined a design which takes the specific advantage of the recent advances in the understanding of both convolutional neural networks as well as ResNets. They focused over the most importance of a trainable pre-processing when using FC-ResNets and they show that a low-capacity FCN model can serve as a pre-processor to normalize medical input data. In their segmentation process, uses the fully convolutional neural networks for obtaining the normalized images that are refined iteratively to design a prediction model for segmentation.

Wencheng et al. [31] designed a new and improved version of active contour model for medical image segmentation that integrates the local and the global intensity information of the medical image with local fitting term and the global fitting term. They have achieved better performance on image segmentation process with less iteration while handling with medical images. Vardhana et al. [32] presented low-power architecture for edge detection to detect the biomedical images. The output of the edge detection process was given as output to their model that able to diagnose the various diseases by using convolutional neural network based image classification. They have used the existing algorithms called Sobel and Prewitt that were used for detecting the edges and implemented by using VLSI, and digital IC design of the architecture is presented. These algorithms used for edge detection which is co-simulated using MATLAB and Modelsim.

Jiaqing et al. [33] used the cosine function for expressing the data energy fitting of a traditional active contours model and also designed a new model that is based on the sectional image recovery according to the local cosine-fitting energy active contours that used for the medical and synthetic images. In addition, their model was designed for handling the images with noise, robust and the computational speed that were similar or faster when it compared with the existing models which are available in this direction. Moreover, they also described the model in a discrete form that is more convenient for adding a regular term for controlling the segmentation. Finally, they demonstrated that their segmentation results were accurate and also efficient when applied to the various kinds of medical images. Chandan and Anu [34] proposed a new approach by using the discrete cosine transform. In their approach is named the DCT-based local and nonlocal FCM (DCT-LNLFCM), the distance function of the FCM is represented as the sum of the local and nonlocal distances which themselves are the weighted values of the Euclidean distance used in the FCM.

Prashant and Rasika [10] proposed a new IoT based healthcare system for monitoring the patients. Jabaveersingh and Jebadurai [9] proposed a new IoT based healthcare application which combines the super-resolution of retinal images

consideration of multi-kernel SVR. They have achieved better accuracy when compared with the existing systems.

System architecture

The architecture of the proposed system is shown in Fig. 1. It consists of nine major components such as image database, IoT devices, user interface module, pre-processing module, clustering module, classification module, decision manager, rule manager and temporal manager. The selected lung images are stored in the image database. These images are to be collected by the user interface module which has the sufficient number of pixels in the input lung images. The user interface module has been collected the necessary medical inputs such as blood pressure, sugar level and heart beat details from IoT device. Moreover, all the lung images are forwarded into the pre-processing module for pre-processing the input lung images for classification. In this pre-processing module consists of six sub components such as image conversion from RGB to HSV, transition region identification, segmentation, feature extraction, thinning operation and cleaning operation. The pre-processing module provides the segmented image in the end of this phase. The pre-processed images can be sent to the classification module for classifying the

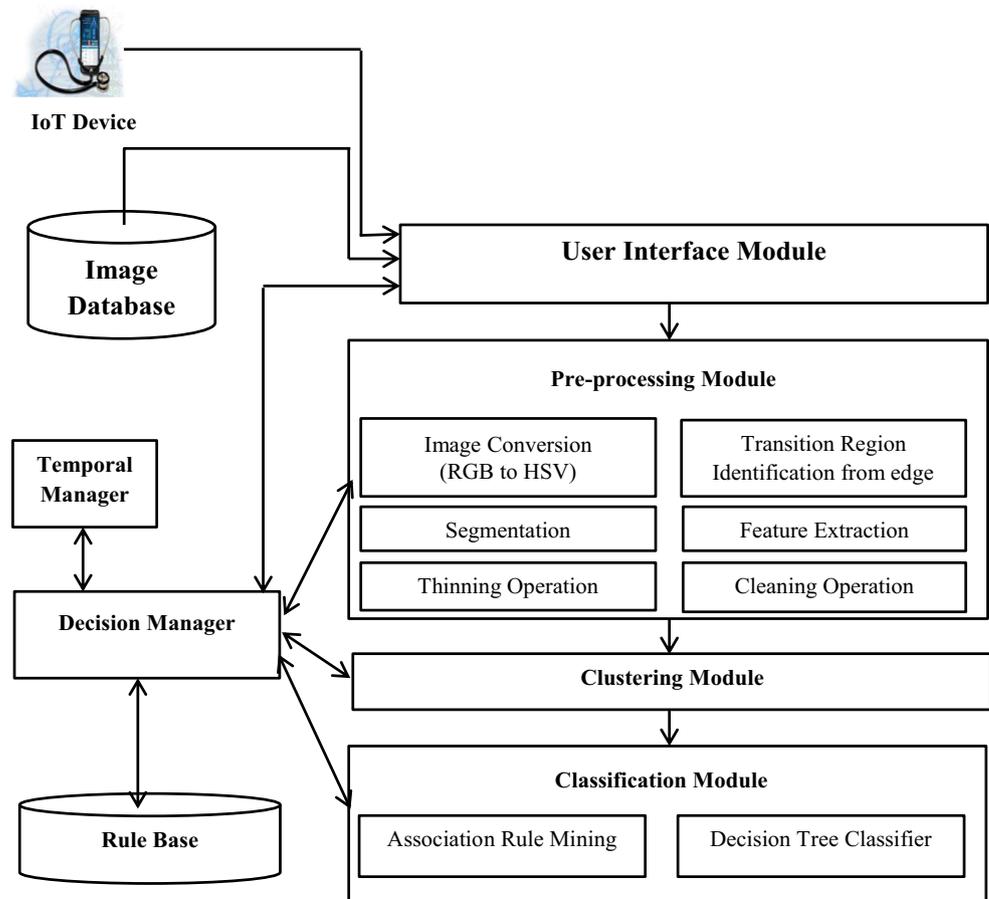
images. The clustering module is used an existing fuzzy c-means clustering algorithm with temporal feature for grouping the pre-processed lung images effectively. The classification module is also consists of two sub components namely Association Rule Mining and the Decision Tree Classifier. This module is responsible for making decision over the pre-processed lung images whether the given pre-processed image is normal or disease affected image.

The decision manager has the overall control over the architecture which is able to take decision on pre-processing module activities and the classification module. Decision manager can collect the necessary input images from the user interface module directly. It performs the classification process based on the rules which are available in the rule base. Moreover, it communicates with the temporal manager for making decision over the classification process in the proposed model.

Proposed system

This section is explained in detail about the proposed a new IoT based Predictive Modelling predicting the lung cancer images that are uploaded by the end users / public. Here, a new transition region extraction method based fuzzy c-means clustering

Fig. 1 System architecture



approach is proposed for effective image segmentation. In this work, the Otsu thresholding method is also used for extracting the transition region effectively from lung cancer images. In addition, the morphological cleaning process and the image region filling process are performed over edge lung cancer images. In addition, a new incremental classification algorithm which combines the existing Association Rule Mining (ARM) with Temporal features, the existing Decision Tree (DT) algorithm with temporal features and the standard CNN is also proposed according to [35] for making effective decision over the lung cancer images.

Pre-processing

This section described in detail about the local variance based image feature extraction process, Fuzzy C-Means Clustering based Transitional Feature Extraction in Medical Images, Transition Regions based on Extraction from Medical Image Features, Morphological thinning and cleaning operation, Morphological region filling to extract the object regions and Extraction of objects from object regions.

Local variance based image feature extraction process

In this work, the local variance features have been extracted from the original medical image for effective image classification. The extracting local variance features which are behind the region is that the image area having higher variance. Generally, it contains the edge of the image regions whereas the homogeneous regions with less variance when it is compared with other edge regions [36]. For a $n \times n$ local neighbourhood that has a pixel $x(k, l)$ which is central and the local variance features can be extracted by using the Eq. (1):

$$Lv(k, l) = \sigma^2 = \frac{1}{n^2-1} \sum_{k=1}^n \sum_{l=1}^n (f(k, l) - \bar{f})^2, \text{ for } \begin{matrix} k = 1, 2, \dots, M \\ l = 1, 2, \dots, N \end{matrix} \tag{1}$$

where (k, l) represent the coordinate of the neighbourhood of $n \times n$ which is local and the sub-image f and \bar{f} indicates the grey level that is also the mean value of neighbours. The equation is used throughout the medical image by sliding the window position ‘ n ’ from left part of the image to the right part of the image and also top of the image to bottom of the image for achieving the local variance of medical image feature. The width and the height of the given medical image that denoted by using M and the N . Moreover, the experiment has been considered the local neighbourhood as 4×4 for extracting the best feature. The local variance feature image indicates that the dominating features only like image edge of the image objects with background features that are also clustered further.

Fuzzy C-means clustering based transitional feature extraction in lung images

The local features that are extracted the transitional features from the existing clustering method. The existing clustering methods that are used to find the similarities between the different pixels or a cluster of pixels that are used to find the distinct structures in the image feature space. Moreover, it is also partitioning a group of pixels that are from input image feature space in a set of homogeneous pixels. In addition, the Fuzzy C-Means (FCM) clustering method is one of the unsupervised clustering methods. This is a most widely used method because of the simple functionality. Here, it also used for the feature variances that consist of various background image features, the transitional image features and the interior object of image features.

The existing FCM is used for finding an image partition that has ‘ p ’ with fuzzy clusters for a group of features $f_j \in R, j = 1, 2, \dots, C$ minimizing the cost function.

$$J(U, M) = \sum_{i=1}^p \sum_{j=1}^C (U_{i,j})^m d_{i,j} \tag{2}$$

Where $U = [U_{i,j}]$ is a fuzzy partition matrix, $U_{i,j} \in [1, \infty]$ is the membership coefficient of j^{th} image object in the i^{th} cluster. Here, the term $M = [m_1, \dots, m_p]$ is a cluster centre matrix. Moreover, the term $m \in [1, \infty]$ is also point out as a fuzzification parameter that uses the standard distance metric called Euclidean distance which is calculated between the regions x_j and m_i . The steps of the proposed fuzzy cluster based transition feature variance based segmentation algorithm works are as follows:

- a) Initialize the suitable values for the variables such as m, c and a small positive integer ϵ .
- b) Randomly choose the cluster center for the time interval t_1 and t_2 .
- c) Let assume the variable $t = 0$.
- d) Calculate the fuzzy value of partition matrix at $t = 0$ during the time interval t_1 and t_2 .
- e) Change the value $t > 0$ and the fuzzy partition matrix u as

$$u_{i,j}^{(t+1)}(t_1, t_2) = \frac{1}{\left(\sum_{l=1}^p \left(\frac{d_{l,j}}{d_{i,j}} \right)^{1/(1-m)} \right)} (t_1, t_2) \tag{3}$$

for $i = 1, \dots, p$ and $j = 1, \dots, C$

- f) Update the t value at $t = 0$ or update the condition is satisfied:

$$\|M^{(t+1)} - M^{(t)}\| (t_1, t_2) < \epsilon \tag{4}$$

In this work, ϵ value is 0.00001 according to [37]. Here, the three clusters are considered in the clustering process. Moreover, the image features are used as three different clusters in this work.

Transition regions based on extraction from medical image features

In this work, we have achieved three different disjoint sets of features that are classified as different classes such as class 1, 2, and class 3 features. From this classification, the background images are available in class 1, the inner image object regions and the background regions that are categorized as class transitional features and the boundary between the image foreground and the image background is also categorized as class 3 transitional features. In this work, the Otsu thresholding [38] is used for getting the image transition regions. End of this process, a thresholded image is also extracted where the value 1 represents the image transition regions that are available as separate image object from the image background.

Morphological thinning and cleaning operation

The transition regions have been extracted from the existing step which surrounds the image object regions appropriately. Here, all these image regions are several pixels width and also used for achieving the closed image object contour of a pixel width and the image transition regions subjected to the morphological thinning operation. Moreover, the thinning operation results in an edge image object with some isolated pixels near the edge. In addition, the morphological cleaning operation is carried out for getting these kinds of isolated pixels. Finally, we have extracted the contours of clean image objects that are continuous in nature.

Morphological region filling to extract the object regions

The image object contours are extracted in the previous thinning operation that is fully connected in this morphological region filling. The portion of the edge image is filled with a value of 1 that is remaining other values are as 0. Moreover, the region of image is filled with values 1 that can be referred as image object regions / segmentation mask.

Extraction of objects from object regions

The extraction of image objects from the previous step called morphological region filling for extracting the object regions where 1 denotes the image object masks whereas 0 indicates the background portion of the input image. The 1 values have been replaced with the original grey image values for

extracting the segmented image object. Moreover, the value 0 can be replaced with the intensity value 255 for making the background of the segmented image result as white.

Classification

This section is explained in detail about the proposed new image classification algorithm called Hybrid Temporal Association Rule with Decision Tree Classification (HTARDTC) algorithm for effective image classification dynamically. Here, the proposed algorithm has been developed according to the existing algorithm Hybrid Association Rule with Decision Tree Classification (HARC) algorithm [35]. The proposed hybrid classification algorithm uses the temporal features in both classifiers for making effective decisions dynamically. The decision tree based classifiers are used in decision making software applications which are available in this direction. In this prediction system, the decisions have been made by the physicians based on the maximum frequent item set which are found in the existing FP-Growth tree algorithm that has been compared with the maximum frequent item set of the input medical images that are considered as input [39, 40] for conducting the experiments in this work.

In this hybrid classification method, the decision tree classifier with association rule mining classifier which provides a better solution to the physicians for classifying the input lung images as benign images and the malignant images. Moreover, it is also can be performed by comparing the maximum frequent items that have been generated by using the temporal association rules in the training medical image that has been compared with the maximum frequent items of the test image. The Hybrid Temporal Association Rule and the Decision Tree Classification (HTARDTC) algorithm that are classified as the rules generated by the association rules in to normal, benign or malignant. The proposed HTARDTC developed new decision trees from a set of training image data by using entropy. The training image set is a set $SI = s_1, s_2, \dots, s_n$ of already classified samples. Each sample $s_i = x_{i1}, x_{i2}, \dots, x_{im}$ is a vector where $x_{i1}, x_{i2}, \dots, x_{im}$ indicates the features of the input images. The training image data is augmented with a vector $CI = c_1, c_2, \dots, c_m$ where c_1, c_2 indicate the class to which each sample belongs.

The proposed incremental classification algorithm selects any one rule of the image data which is able to split its set of samples into subsets as number of sets effectively in one class or the other class in this work. The information gain value is calculated for the process decision making by using all the input image features such as transition regions, range, morphological region, pixels as items sets. These all features with the highest normalized information gain value have been selected for making decision. The steps of the proposed incremental classification algorithm are as follows:

Incremental Classification algorithm.

Input: IR: a set of non-target image features, CI: the image feature, DI: a training image data set)

Output: Return a decision tree;

Step 1: Read the input image features and other features.

Step 2: If DI is empty then,
Returns 'Failure'

Step 3: If DI consists of records that are all with the same value with time constraints for the target feature then
Returns the value of leaf node

Step 4: If RI is empty then
Return most frequent values of the target image feature that are identified in records of SI;

Step 5: Let F be the image feature with largest Information Gain (F, SI) during the certain time interval $\langle t1, t2 \rangle$ among the features in RI;

Step 6: Let $\{a_j \mid j=1,2,\dots,m\}$ be the values of image feature F during the time interval $\langle t1, t2 \rangle$;

Step 7: Let $\{SI_j \mid j=1,2, \dots, m\}$ be the image subsets of SI consisting respectively of records with the value f_j for F in specific time interval $\langle t1, t2 \rangle$;

Step 8: Return a decision tree with root labelled D and arcs labelled d_1, d_2, \dots, d_m going respectively to the trees $(HTARC(RI-\{DI\}, CI, SI_1, \langle t1, t2 \rangle), HTARC(RI-\{DI\}, CI, SI_2, \langle t1, t2 \rangle), \dots, HTARC(RI-\{AI\}, CI, SI_m, \langle t1, t2 \rangle))$;

Step 9: Recursively apply HTARC to subsets $\{SI_j \mid j=1, 2, \dots, m\}$ until they are empty during the specific time interval between $t1$ and $t2$.

Step 10: Apply Convolutional Neural Network () with image data along with other features.

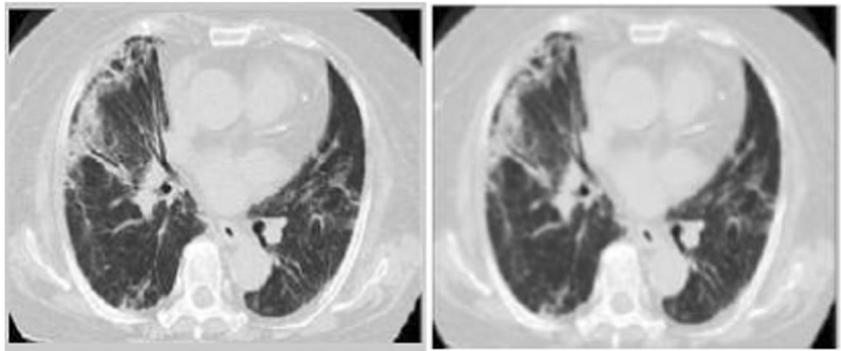
Step 11: Display the decisions over the uploaded images.

Step 12: End

In the traditional classification approach has been used like Association Rule Mining or Decision Tree Classifier whereas in this proposed method called temporal association rule and decision tree combined features have been used for the medical image classification. In this method

takes decision over the medical images by using the segmented input images and the transition features. Finally, it provides the result with the classification of the lung cancer images as normal image and disease affected image.

Fig. 2 Input images



Results and discussion

The proposed model has been implemented by using MATLAB. Here, the various experiments have been conducted for evaluating the proposed model. For this purpose, the well-known medical database of ILD in the lung image is utilized to evaluate the experimental results of the proposed method. Moreover, this work considered 113 lung images with high resolution CT images of dimension of 512 X 512 pixels in each slice. Generally, the left and right lungs have been segmented in this work from the chest CT scan lung images which are available in the database by using the standard Fuzzy C-Means Clustering method based segmentation process. In addition, other data such as blood pressure, heart beat, sugar level details have been collected through IoT device and also considered for making decision. This data will be given as input to the system in the form of text and numbers. Figure 2 shows the lung images which are taken as input in this work for carrying out the experiments.

Figure 3 shows the segmented image that uses the existing clustering algorithm called fuzzy c-means clustering algorithm for performing segmentation.

Figure 4 shows the segmented lung images by using the proposed fuzzy c-means cluster based segmentation algorithm that incorporates the transition features, morphological feature

extraction, transition based feature extraction and thinning and cleaning operation.

The performance evaluation of the proposed FCM based image segmentation method which incorporates the consideration of transition features and the feature extraction is done for making decision over the lung images using the hybrid temporal classification algorithm in this work. This evaluation process uses the various performance metrics such as Mean Square Error (MSE), the Peak Signal to Noise Ratio (PSNR), True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) ratios. These all performance metrics are employed to demonstrate the quality of lung images. The efficiency of the proposed model has been measured by using the Eqs. (5), (6) and (7).

$$\text{Prediction C Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (5)$$

$$\text{Sensitivity} = \frac{TP}{(TP + FN)} \quad (6)$$

$$\text{Specificity} = \frac{TN}{(TN + FP)} \quad (7)$$



Fig. 3 Segmented Image using FCM-CA



Fig. 4 Segmented Image using the proposed method

Table 1 Performance comparison between MSE and PSNR

Parameters	FCM	Enhanced FCM	Proposed FCM based Segmentation
MSE	0.3664	0.162	0.159
PSNR	52.49	56.09	57.02
Loss %	20	8	7
Accuracy (%)	97	99	99.54

Where TN represents the number of true negatives, FP holds the number of false positives, TP holds the number of true positives and FN indicates the number of false negatives. Moreover, the prediction accuracy is the direct proportional of the correctly predicted the disease affected lung images from the input images. The sensitivity is able to find the abnormal lung images that are available in the input image set. The specificity is used to identify the normal lung images from the given input images. In this proposed HTARC, this work has been considered the minimum confidence is 98%, the minimum support is 15% for making decision over the lung images, time interval and the threshold value as 0.002.

Table 1 shows the performance comparison between the performance of the FCM, Enhanced FCM and the proposed FCM based segmentation algorithm in terms of MSE and PSNR values over the input lung images.

From Table 1, it can be observed that the performance of the proposed FCM based Segmentation with transition features is better when it is compared with the other existing segmentation algorithms which are based on FCM and the Enhanced FCM. This is due to the fact that the use of transition features based extraction, thinning and cleaning operation over the transition features in the segmentation process along with FCM.

The five experiments have been conducted for evaluating the performance of the proposed prediction model and the existing models and classifiers which are available in this direction. Here, we have considered 300 images for conducting the various experiments. Figure 5 shows the

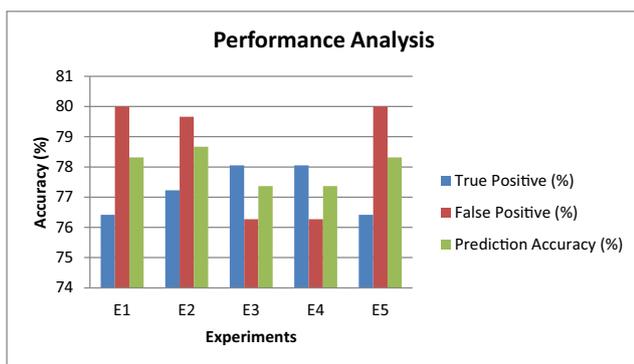


Fig. 5 Performance analysis

performance analysis in terms of true positive, false positive and the prediction accuracy.

From Fig. 5, it can be observed that the performance of the proposed model is uniform in all the five experiments in these performance metrics such as true positive, false positive and the prediction accuracy. This is due to the fact that the use of temporal features and the transition region based feature extraction and the FCM for effective segmentation which is used to predict the images effectively.

The performance of the proposed prediction model has been evaluated by using the metrics such as sensitivity, specificity and accuracy. The comparative analysis is made between the proposed prediction model and the existing methods that are available in this direction. Figure 6 shows the comparative analysis based on the sensitivity, specificity and accuracy. This comparative analysis is also considered the 300 lung images for conducting the experiments.

From Fig. 6, it can be observed that the proposed model is performed well when it is compared to all other models such as FCM based model and the Enhanced FCM based prediction model in terms of sensitivity, specificity and prediction accuracy. Here, the proposed model achieves 85% of prediction accuracy and the existing models accuracy is 81.68% and 82.67%. This is due to the fact that the use of FCM based segmentation, transition region based feature extraction, thinning and cleaning operation using transition features.

Figure 7 shows the comparative analysis between the IoT based healthcare applications. Here, we have considered the proposed IoT based healthcare system and the exiting healthcare system [9]. Here, we have considered two evaluation metrics such as MSE and PSNR accuracies.

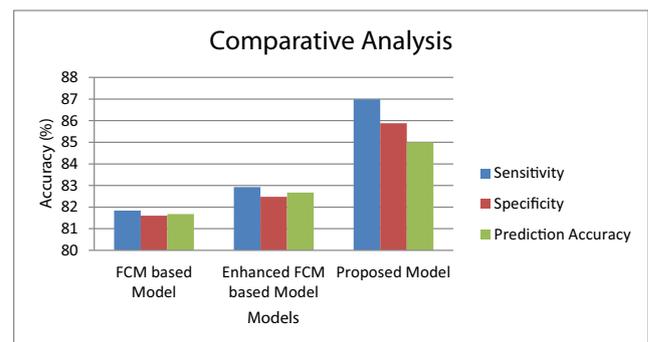


Fig. 6 Comparative analysis

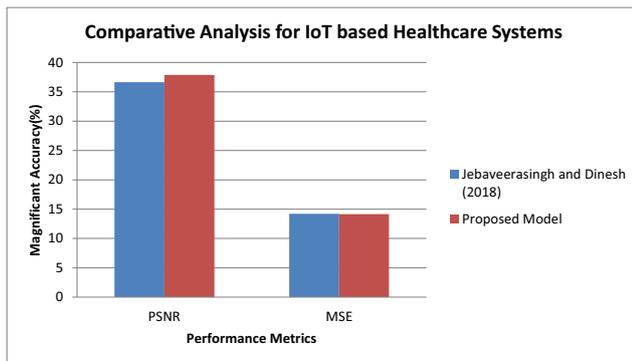


Fig. 7 Comparative analysis between the IoT based healthcare applications

From Fig. 7, it can be observed that the performance of the proposed IoT based healthcare application is better when it is compared with the existing IoT based healthcare application. This is due to the fact the use of effective segmentation and classification techniques for making effective decision over the standard lung cancer image data set and the IoT data that are collected from the lung cancer patients in remote.

Conclusion and future works

A new IoT based Predictive Modelling has been proposed and implemented in this work for predicting the lung cancer diseases as online application by using the transitional extraction, clustering and classification. This model uses the effective transition region extraction method for performing image segmentation effectively. Moreover, the existing Fuzzy C-Means Clustering algorithm has been used for segmentation which incorporates the transitional region based feature extraction, morphological cleaning and thinning operation from the feature of lung cancer image. The proposed segmentation process is used for improving the classification accuracy. In addition, a new incremental classifier is developed by using association rule mining with temporal features, the existing decision tree classifiers with temporal features and the CNN for making effective decision over the lung cancer image for predicting the lung cancer affected images. The experimental results proved that the efficiency of the proposed model in terms of prediction accuracy when it is compared with the existing models. The proposed model has been achieved 85% as prediction accuracy and it is more than 2% than the prediction accuracy of other existing models. Future works in this direction could be the introduction of fuzzy rough set theory for decision making over the lung cancer affected images.

Compliance with Ethical Standards

Conflict of Interest The authors have no conflict of interest.

(In Case Animals Were Involved) Ethical Approval Animals were not involved.

(And/or in Case Humans Were Involved) Ethical Approval This article does not contain any studies with human participants performed by any of the authors.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. World health organization cancer fact sheets. <<http://www.who.int/mediacentre/factsheets/fs297/en/index.html>>.
2. Siegel, R., Miller, K., and Jemal, A., Cancer statistics, 2016. *CA-Cancer J. Clin.* 66:7–30, 2016.
3. Aerts, H., Velazquez, E., Leijenaar, R., Parmar, C., Grossmann, P., Carvalho, S. et al., Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* 5:4006, 2014.
4. Kubota, T., Jerebko, A., Dewan, M., Salganicoff, M., and Krishnan, A., Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. *Med. Image Anal.* 15:133–154, 2011.
5. Sharma, N., and Aggarwal, L., Automated medical image segmentation techniques. *J. Med. Phys.* 35:3–14, 2010.
6. Farag, A. A., El Munim, H. E. A., Graham, J. H., and Farag, A. A., A novel approach for lung nodules segmentation in chest ct using level sets. *IEEE Trans. Image Process.* 22:5202–5213, 2013.
7. Lassen, B., Jacobs, C., Kuhnigk, J., van Ginneken, B., and van Rikxoort, E., Robust semi-automatic segmentation of pulmonary subsolid nodules in chest computed tomography scans. *Phys. Med. Biol.* 60:1307, 2015.
8. Fukushima, K., Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* 36(4):193–202, 1980.
9. Jebadurai, J., and Dinesh Peter, J., Super-resolution of retinal images using multi-kernel SVR for IoT healthcare applications. *Futur. Gener. Comput. Syst., Elsevier* 83:338–346, 2018.
10. Salunke, P., and Nerkar, R., IoT driven healthcare system for remote monitoring of patients. *J. Modern Trend Sci. Technol.* 3(06):100–103, 2017.
11. Gope, P., and Hwang, T., BSN-care: A secure IoT-based modern healthcare system using body sensor network. *IEEE Sens. J.* 16(5): 1368–1376, 2016.
12. Kannan, S. R., Sathya, A., Ramathilagam, S., and Devi, R., Novel segmentation algorithm in segmenting medical images. *J. Syst. Soft.* 83(12):2487–2495, 2010.
13. Nan, B., Che, L., Chui, K., Chang, S., and Ong, S. H., Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation. *Comput. Biol. Med.* 41(1):1–10, 2011.
14. Bai, P. R., Yi, Q., Lei, L., Sheng, L., Jing, H. T., Mao, L., and Cao, Y., A novel region-based level set method initialized with mean shift clustering for automated medical image segmentation. *Comput. Biol. Med.* 43(11):1827–1832, 2013.
15. Manikandan, S., Ramar, K., Willjuice Iruthayarajan, M., and Srinivasagan, K. G., Multilevel thresholding for segmentation of

- medical brain images using real coded genetic algorithm. *Measurement* 47:558–568, 2014.
16. Torbati, N., Ayatollahi, A., and Kermani, A., An efficient neural network based method for medical image segmentation. *Comput. Biol. Med.* 44:76–87, 2014.
 17. Li, Y., Jiao, L., Shang, R., and Stolkin, R., Dynamic-context cooperative quantum-behaved particle swarm optimization based on multilevel thresholding applied to medical image segmentation. *Inf. Sci.* 294:408–422, 2015.
 18. Zhang, X., Li, X. f., and Feng, Y., A medical image segmentation algorithm based on bi-directional region growing. *Optik Int. J. Light Electron Opt.* 126(20):2398–2404, 2015.
 19. Mahapatra, D., Combining multiple expert annotations using semi-supervised learning and graph cuts for medical image segmentation. *Comput. Vis. Image Underst.* 151:114–123, 2016.
 20. De, S., Bhattacharyya, S., and Dutta, P., Automatic magnetic resonance image segmentation by fuzzy intercluster hostility index based genetic algorithm: An application. *Appl. Soft Comput.* 47: 669–683, 2016.
 21. Biswas, S., Ghoshal, D., and Hazra, R., A new algorithm of image segmentation using curve fitting based higher order polynomial smoothing. *Optik* 127(20):8916–8925, 2016.
 22. Ghosh, P., Mitchell, M., Tanyi, J. A., and Hung, A. Y., Incorporating priors for medical image segmentation using a genetic algorithm. *Neurocomputing* 195(26):181–194, 2016.
 23. Kalshetti, P., Bundeale, M., Rahangdale, P., Jangra, D., Chattopadhyay, C., Harit, G., and Elhence, A., An interactive medical image segmentation framework using iterative refinement. *Comput. Biol. Med.* 83:22–33, 2017.
 24. Zhou, S., Wang, J., Zhang, M., Cai, Q., and Gong, Y., Correntropy-based level set method for medical image segmentation and bias correction. *Neurocomputing* 234(19):216–229, 2017.
 25. Chen, Y.-T., A novel approach to segmentation and measurement of medical image using level set methods. *Magn. Reson. Imaging* 39: 175–193, 2017.
 26. Yang, S.-C., A robust approach for subject segmentation of medical images: Illustration with mammograms and breast magnetic resonance images. *Comput. Electr. Eng.* 62:151–165, 2017.
 27. Khanfir Kallel, I., Almouahed, S., Solaiman, B., and Bossé, É., An iterative possibilistic knowledge diffusion approach for blind medical image segmentation. *Pattern Recogn.* 78:182–197, 2018.
 28. Zheng, Q., Li, H., Fan, B., Wu, S., and Xu, J., Integrating support vector machine and graph cuts for medical image segmentation. *J. Vis. Commun. Image Represent.* 55:157–165, 2018.
 29. Liu, C., Ng, M. K.-P., and Zeng, T., Weighted variational model for selective image segmentation with application to medical images. *Pattern Recogn.* 76:367–379, 2018.
 30. Drozdal, M., Chartrand, G., Vorontsov, E., Shakeri, M., Di Jorio, L., Adrian, A. T., Yoshu, R., Pal, B. C., and Kadoury, S., Learning normalized inputs for iterative estimation in medical image segmentation. *Med. Image Anal.* 44:1–13, 2018.
 31. Zhao, W., Xu, X., Zhu, Y., and Xu, F., Active contour model based on local and global Gaussian fitting energy for medical image segmentation. *Optik* 158:1160–1169, 2018.
 32. Vardhana, M., Arunkumar, N., Lasrado, S., Abdulhay, E., and Ramirez-Gonzalez, G., Convolutional neural network for biomedical image segmentation with hardware acceleration. *Cogn. Syst. Res.* 50:10–14, 2018.
 33. Miao, J., Huang, T.-Z., Zhou, X., Wang, Y., and Liu, J., Image segmentation based on an active contour model of partial image restoration with local cosine fitting energy. *Inf. Sci.* 447:52–71, 2018.
 34. Singh, C., and Bala, A., A DCT-based local and non-local fuzzy C-means algorithm for segmentation of brain magnetic resonance images. *Appl. Soft Comput.* 68:447–457, 2018.
 35. Rangaswamy, C., Raju, G. T., and Seshikala, G., Novel approach for lung image segmentation through enhanced fuzzy C-means algorithm. *Int. J. Pure Appl. Math.* 117(21):455–465, 2017.
 36. Parida, P., and Bhoi, N., Transition region based single and multiple object segmentation of gray scale images. *Eng. Sci. Technol. Int. J.* 19(3):1206–1215, 2016.
 37. Cai, W., Chen, S., and Zhang, D., Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation. *Pattern Recogn.* 40(3):825–838, 2007.
 38. Otsu, N., A threshold selection method from gray level histograms. *IEEE Trans. Syst. Man Cybern.* 9(1):62–66, 1979.
 39. Sun, S., Bauer, C., and Beichel, R., Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach. *IEEE Trans. Med. Imaging* 31(2):449–460, 2012.
 40. Taher, F., and Sammouda, R., Lung cancer detection by using artificial neural network and fuzzy clustering methods. *IEEE GCC Conf. Exhib.* 10:295–298, 2011.