Intuitionistic Fuzzy C-means Clustering Method for Automatic Liver Tumor Segmentation from CT Scans

Dissertation submitted to Jawaharlal Nehru University in partial fulfilment of the requirements for the award of the degree of

Master of Technology

In

Computer Science and Technology

By

ASHUTOSH KUMAR SINGH

Under the supervision of

Prof. R. K. Agrawal

SCHOOL OF COMPUTER AND SYSTEMS SCIENCES

JAWAHARLAL NEHRU UNIVERSITY

NEW DELHI-110067

DECEMBER 2022

जवाहरलाल नेहरु विश्वविद्यालय

SCHOOL OF COMPUTER & SYSTEMS SCIENCES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI-110067
INDIA

Certificate

This is to certify that the dissertation entitled "*Intuitionistic Fuzzy C-means Clustering Method for Automatic Liver Tumor Segmentation from CT Scans*" is being submitted by **Mr. Ashutosh Kumar Singh** to **School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi-110067, India** in the partial fulfilment of the requirements for the award of the degree of **Master of Technology in Computer Science and Technology**. This is entirely his own work, carried out in the School of Computer and Systems Sciences under the supervision of **Prof. R. K. Agrawal**. The matter personified in this dissertation has not been submitted for the award of any degree of this or any other university.

Supervisor 27.12.2022

Prof. R. K. Agrawal

School of Computer & Systems Sciences

Jawaharlal Nehru University

New Delhi -110067

Dean 28.12.2022

School of Computer & Systems Sciences

Jawaharlal Nehru University

New Delhi -11006

जवाहरलाल नेहरु विश्वविद्यालय

SCHOOL OF COMPUTER & SYSTEMS SCIENCES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI-110067
INDIA

Declaration

I hereby declare that the dissertation entitled **"***Intuitionistic Fuzzy C-means Clustering Method for Automatic Liver Tumor Segmentation from CT Scans***"** in partial fulfilment of the requirements for the degree of **Master of Technology in Computer Science and Technology** submitted to **School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi-110067, India** is an authentic record of my own work carried out under the supervision of **Prof. R. K. Agrawal**. The matter personified in this dissertation has not been submitted for the award of any degree of this or any other university.

Ashutosh Kumar Singh

Ashutosh Kumar Singh

M.Tech. (2020-2022)

School of Computer & Systems Sciences

Jawaharlal Nehru University New Delhi -110067

Acknowledgement

I am extremely thankful to my thesis supervisor, Prof. R.K. Agrawal for his invaluable assistance, continuous support during my work. His immense knowledge and plentiful experience have encouraged me all the time. His motivation and continuous insights have helped to make this research work possible in these harsh times of pandemic. His involvement in this work has immensely helped to successfully complete my thesis. My lab mates have also been very helpful and supportive. I would like to express gratitude to my family, friends for their patience, understanding and their infinite support. Last but not the least, I am very grateful to my school and this wonderful university for nurturing the infrastructure and research environment.

Ashutosh

Abstract

Accurate Liver segmentation from computed tomography (CT) scan images is an essential and crucial step in computer-aided diagnosis of liver tumors. The manual method of liver tumor diagnosis done by a radiologist is time consuming, hard, and tedious work. To find tumors in the liver, radiologists must segment the liver from CT scans first. The liver tumor is the cause of death of many people around the world. So, if we detect liver tumors early then we can save the lives of many people. Therefore, there is a need for an automated method of liver tumor detection. To solve the problem, we have investigated three methods Fuzzy C-Means (FCM) clustering method, Fuzzy C-Means with Spatial constraints (FCM_S) and Spatial Intuitionistic Fuzzy C-Means (SIFCM) method for liver segmentation. All the three methods are unsupervised clustering methods for segmentation. All the three methods are a two-step process. In the first step, CT Windowing is done on the input CT image for better clarity of organs. In the second step, FCM and its variants are applied on the preprocessed image to segment the liver.3D-IRCADb-01 dataset which is publicly available and is used with added gaussian noise for the comparison of liver segmentation using the three existing methods. The performance of these existing methods is compared against the available ground truth of the liver. We calculated the Jaccard similarity score (JSC) and Dice similarity coefficient (DSC) for the three methods on the 3D-IRCADb-01 dataset. We found that SIFCM is giving better results than FCM and FCM_S. SIFCM uses spatial features of image data for noise handling and Intuitionistic Fuzzy sets (IFS), which deals with uncertainty. For future work, we are going to segment the tumor from the segmented liver.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview and Motivation

The liver is considered as one of the main body parts in the human being. It has lots of important roles in blood refining, blood agglomeration, production of protein and detoxification process of medications. Abnormal growth of liver cells in or within the liver is called Liver tumors (also known as hepatic tumors). Liver is made up of different types of cells. So, there is a probability of developing different types of tumors. Liver tumors are mainly of two types (i) benign (non-cancerous) and (ii) malignant (cancerous). Cancer types like lung, stomach, liver, prostate and colorectal are common in men and some cancer types like breast, cervical, colorectal, lung and thyroid are common in women. Cancer is the second most cause of death across the world after heart related disease according to the report of the World Health Organization (WHO). According to data from WHO, cancer was responsible for 9.9 million deaths in 2020 out of which 830 180 (8.3 %) deaths were caused by liver cancer worldwide. Liver cancer cases occur often in less developed countries, which is about 83 % of the total cases. The highest cases occur in Asia and Africa. The liver has two lobes: Right and Left and is on contact with various organs of the body such as gallbladder, pancreas, and intestines. As there are many organs attached to the liver, liver cancer can have two types. (i) primary (originating from various cells that build the liver) and (ii). Secondary (originating from cancerous cells from other organs). Therefore, the detection of liver tumors is very essential in early stages. If detected early, then it can be cured, and the life is saved. There are various imaging techniques for the finding of liver disease like magnetic resonance imaging (MRI), ultrasonography, Percutaneous Transhepatic Cholangiography, radionuclide scanning, computed Tomography (CT)etc. Amid these techniques, CT scanning is extensively used to identify liver diseases. CT scanning is non-operative and less expensive than other methods.

Figure 1-1: CT scanner

In CT images, not only the liver is present, but nearby organs are also present. So, from CT images, segmentation of liver and tumor is required for early detection, treatment arrangement, and observation of liver cancer. Currently, mostly manual segmentation of liver tumors is done by radiologists. Manual segmentation of CT images is lengthy and tedious because CT images are huge amounts of data and there are various organs attached in abdominal CT images with almost equal intensity. Manual segmentation of CT images also requires expertise. Therefore, there is a need for computer-aided accurate and efficient tumor segmentation methods. Many methods have been proposed and developed to automate the liver tumor segmentation task. There are mainly two types of segmentation algorithms. (i) supervised learning methods and (ii.) unsupervised learning methods. Unsupervised methods do not require labelled data. The most commonly unsupervised learning methods used for liver segmentation are Graph Cut, watershed transform (WT), fuzzy entropy, clustering method, etc. In supervised learning training data along with labelled data is required for segmentation. The most used supervised learning methods for liver segmentation is artificial neural networks, convolutional neural networks, and their variants. In unsupervised learning, clustering is most widely used for segmentation. Clustering methods can be either hard or soft. Among different clustering approaches, Fuzzy C-Means (FCM) based clustering approach are extensively used. FCM is a soft-clustering method. For liver segmentation, many FCM-based methods are proposed by researchers around the world. We have investigated some of the FCM based methods like FCM, FCM_S and SIFCM for liver segmentation in CT images with noise and provided comparative results of this. We found that SIFCM performs better for segmentation as it can handle noise and uncertainty in CT images.

In the next section, we present brief details of liver tumor segmentation methods based on FCM and its variants.

1.2 Objective

The Objective of the thesis work is as follows:

- To investigate and implement the existing method Fuzzy C-means (FCM), Fuzzy c-means with local spatial information (FCM_S) and spatial intuitionistic fuzzy c-means (SIFCM) method for liver segmentation in CT images.
- To compare the performance of these three methods on the publicly available 3d-ircadb-01 dataset with the added gaussian noise.

1.3 Contents of thesis

Chapter 2 describes existing methods for liver segmentation in CT images. Chapter 3 discusses the three FCM based methods for liver segmentation in CT images, which we have investigated. Chapter 4 discusses experimental results and Discussion. Finally, Chapter 5 discusses the conclusion and future direction.

Chapter 2

Related Work

Li et al. (2009) [1] suggested tumor segmentation in liver based on integrating fuzzy c-means (FCM) and level sets. In the first part, unsupervised clustering by FCM is used. It initially segments and detects tumors. In the output of the first part, a series of morphological operations are carried out for refinement. In the last stage, an enhanced level set method is used for clear extraction of tumors. This approach segments tumors more clearly and is more reliable as compared to FCM. However, this approach only focuses on segmenting tumors from the liver. Liver segmentation from abdominal CT scans is also a major issue not considered in this work.

Kumar et al. (2013) [2] suggested a possibilistic alternative fuzzy C-means (PAFCM) clustering approach for liver tumor segmentation. This algorithm works in two stages. In the first stage, liver from abdominal CT images is segmented and in the last stage tumor is segmented from segmented liver. In the first stage, a series of steps like pre-processing, intensity analysis, region growing, and post-processing are carried out. In the final stage, PAFCM is used to divide liver tumors from the divided liver regions. PAFCM differs from the alternative fuzzy c-means (AFCM). In AFCM cluster membership sums to one but in PAFCM a possibilistic partition is used. This algorithm is rapid, fast, and reliable. It handles noisy images efficiently.

Obayya and Rabaie (2015) [3] proposed automated segmentation of liver tumors in CT images using fuzzy c-means (FCM). This method works in two stages. This method works in two steps. In the first stage, several preprocessing steps are applied to the original CT image to enhance the image quality. Firstly, morphological filters are used for de-noising CT images as well as separating attached organs. Then the thresholding process is used to convert it into binary image. Further noise reduction is carried out using morphological processes. The connected component is then obtained from the obtained binary image using a connected component algorithm. Then the largest boundary is taken out because the liver is the largest organ in the input CT images. Finally, a mask is formed using a region-filling algorithm. Only used for packing within max bounds. After this element-wise multiplication is done between the original input and developed mask which results output image with the original values of intensities of the liver and rest are black.

The first stage output serves as the second stage input. The FCM clustering algorithm is used to cluster the intensity levels of the image into three clusters: background pixels (black pixels), liver region (light pixels) and suspicious region (dark pixels). The method is computationally fast. But the method provides poor performance in the presence of complex textures and noise.

Das and Sabut (2016) [4] proposed liver tumor segmentation using kernelized fuzzy c-means (KFCM) clustering with adaptive thresholding. This algorithm also works in two stages. In the first stage the selected image is pre-processed with 3D-guassian filter. It helps to minimize the consequence of the noise. On filtered CT image adaptive threshold is applied to segment the liver from it.

The segmented liver is given as input for the next stage. At this stage, the KFCM method is used to segment the tumor from the liver. KFCM replaces the Euclidean distance originally used in FCM with a kernel-derived distance metric. The performance of this approach is found to be better as compared to FCM for handling noisy CT images.

Yugandar and Reddy (2017) [5] suggested distance regularized level set evolution (DRLSE) based on Fuzzy C-Means Clustering for liver tumor detection. This approach is proposed basically for noisy CT images. In this approach firstly median filter is used which takes CT images with noise as input. A median filter is a spatial nonlinear filter used for denoising noisy images. The denoised image from the median filter is given as input to the FCM for image segmentation. In the last segmented image is given as input to the DRLSE model. This model identifies liver tumors. This approach is efficient for processing noisy CT images.

Rela et al. (2020) [6] proposed liver super pixel based Fast fuzzy C means (SFFCM) clustering algorithm for liver tumor detection. In this approach first multiscale morphological gradient reconstruction-wavelet transform (MMGR-WT) is used to convert CT images to superpixel image. Then FCM is used to segment the image. After that Histogram of the segmented image is formed. It is then used for selecting the grey level of tumor region. Finally, grey level thresholding is used to obtain tumor region. This method requires less time and less interaction from human for tumor segmentation. It is also very fast for color image segmentation. But it has limited use in real life applications, because in this method we need to set number of clusters initially.

Khan and R (2020) [7] suggested automated liver tumor segmentation in CT Images using alternative fuzzy c-means (AFCM). This method consists of two stages. In the first part liver segmentation from abdominal CT scans is done using threshold-based slope difference differentiation (SDD) technique. The output of the first stage is taken as input to the next stage where tumor detection is done by AFCM. In AFCM the Euclidean distance used in FCM is replaced by an alternate distance function. This method is also able to handle noisy image efficiently. The results are highly accurate like manual segmentation. This method detects tumor but doesn't predict the liver cancer.

Al-Saeed et al. (2020) [8] proposed Liver Segmentation from CT images using Fast-Generalized Fuzzy C-Means (FG-FCM). This method more accurately segments the liver as compared to Fuzzy C-Means. This method is fast and reliable for noisy CT scans. This method is of has two stages. In the first step preprocessing is done. CT images are converted to greyscale images. Then Adaptive histogram equalization (AHE) is used for contrast enhancement and Adaptive median filtering (AMF) is used for noise reduction. In the second stage liver segmentation is done. FG-FCM performs liver segmentation after taking input taken from first stage . But the limitation of this work is that it is only segmenting the liver not the tumor present in the liver.

Pohle and Toennies (2001) [9] proposed an adaptive region growing method for liver segmentation. This method automatically learns uniformity criteria from the properties of the region to be segmented. This method cannot handle scenarios where the segmented regions are heterogeneous.

Zhao et al. (2010) [10] proposed Fuzzy C-means Clustering-based multilayer perceptron neural network for Liver CT Images Automatic Segmentation. In this method first threshold method is used on the initial input CT image to remove ribs and spines. Then segment CT images using FCM and morphological filters. After that training of multi-layer perceptron neural network is done using the first CT image as the sample image and the first segmented image as the expected image. This trained multi-layer perceptron neural network is used to segment adjacent slices. This process is iterated until all the slices are segmented. But this method is only for liver segmentation, it is not for the tumor segmentation.

Song et al. (2014) [11] proposed Liver Segmentation Based on Spatial kernelized Fuzzy C-means (SKFCM) and Improved GrowCut for CT Images. It consists of four steps. The first step is pre-processing step in which median filter is used to remove noise from CT image. In the second step SKFCM is used for rough segmentation. In the third step refined segmentation is done using GrowCut algorithm. In the fourth step post-preprocessing is done to obtain a smoother contour of liver using morphological operations. This method only segments the liver and not segmenting the tumor from liver.

2.1 Dataset

The 3D IRCADb-01 database consists of 3D CT scans of 10 women and 10 men with liver tumors in 75% of cases. This dataset has 20 folders. Each folder corresponds to each different patient. This dataset is freely available. This data is publicly available. Each folder consists two nifty files. One file is of CT scans of person liver which includes abdomen part also. This file is named like “ircad_e01_orig.nii” for first person similarly for second person “ircad_e02_orig.nii” and so on. The second file contains the liver mask of the respective person. This file is named like “ircad_e01_liver.nii” for first person similarly for second person “ircad_e02_liver.nii” and so on. For our experiment we have used central slice of each patient. In the central slice of each patient, we have added gaussian noise with mean=128 and standard deviation=20.

Chapter 3

Methods

In this chapter in the first section, we briefly discuss the three existing methods Fuzzy clustering means (FCM), Fuzzy clustering means with local spatial information (FCM_S) and spatial Intuitionistic Fuzzy Clustering means (SIFCM). In the second section systematic workflow of the Liver Segmentation from CT Scans is discussed which consists of three steps (i.) Pre-processing (ii.) FCM based segmentation (iii.) post-processing.

3.1 Fuzzy c-means clustering (FCM) and its Deviations

3.1.1 Fuzzy c-means (FCM)

Fuzzy c-means (FCM) [12] is an unsupervised soft clustering approach. It comes from the traditional k-means clustering algorithm. In K-means clustering each data point is assigned to only one cluster. The FCM uses a fuzzy membership which assigns membership value for each class.

The objective function of FCM is defined as below:

$$min\ \ J_{FCM} = \sum_{j=1}^{n} \sum_{i=1}^{c} \left(u_{ij}\right)^{q} d^2\left(x_j, v_i\right)$$

subject to $\sum_{i=1}^{c} \mu_{ij} = 1, 0 \leq \mu_{ij} \leq 1$ and $\sum_{j=1}^{N} \mu_{ij} > 0$

here $X = \{x_1, x_2, \cdots, x_n\} \subseteq R^z$ is the n number of data points in the z-dimension, c is the total number of clusters with $2 \leq c < n, u_{ij}$ is the degree of membership of x_j in the i^{th} cluster, $q(q > 1)$ is a constant controlling fuzziness, v_i is the i[th] cluster center , $d^2\left(x_j, v_i\right)$ is a Euclidian distance between data x_j and cluster center v_i.

the degree of membership μ_{ij} and the centroids v_i are updated in each cycle by equation (1) and (2)

$$\mu_{ij} = \frac{\|x_j - v_i\|^{-2/(m-1)}}{\sum_{k=1}^{c} \|x_j - v_k\|^{-2/(m-1)}} \qquad (1)$$

$$v_i = \frac{\sum_{j=1}^{N} \mu_{ij}^m x_j}{\sum_{j=1}^{N} \mu_{ij}^m} \qquad (2)$$

The algorithm of FCM is as follows:

1 Initialize values for c, q ,ϵ $(threshold\ value)$ and max_iter

2 Randomly initialize fuzzy partition matrix $U = [u_{ik}]$.

3 initialize counter variable $y = 0$.

4 Calculate the c cluster centers $\left\{v_i^{(y)}\right\}$ using equation (2) where 1≤i≤c

5 Update the membership matrix using equation (1)

6 If $\|U^{(y)} - U^{(y+1)}\| < \epsilon$ or y=max_iter then stop; else, set $y = y + 1$ then go to step 4.

7 De-fuzzify the obtained membership values.

FCM provides better clustering as compared to K-means clustering and gives improved outcome for overlapped dataset. It requires more number of iterations to give the results and it does not handle noise present in the image.

3.1.2 Fuzzy C-means with local spatial information (FCM_S)

The Fuzzy C-means with local spatial information (FCM_S) is proposed by Ahmed et al. [13]. In FCM_S spatial term is added with FCM objective function. The spatial term smooths a pixel intensity value by considering its specified neighborhood pixels. The FCM_S handles noisy images better as compare to FCM. The objective function of FCM_S is given as:

$$min\, J_m(\mathbf{U},\mathbf{V}:\mathbf{X}) = \sum_{i=1}^{c}\sum_{j=1}^{N} \left(u_{ij}\right)^m \left\|x_j - v_i\right\|^2 + \frac{\alpha}{N_R}\sum_{i=1}^{c}\sum_{j=1}^{N} \left(u_{ij}\right)^m \sum_{r\in N_j} \left\|x_r - v_i\right\|^2$$

$$\text{s.t} \sum_{i=1}^{c} u_{ij} = 1, 1 \le j \le N \qquad (3)$$

where $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$ are N image pixels , m is the constant controlling fuzziness with constraints $(1 < m < \infty)$, c is the total count of clusters which is constant and with constraints $(1 \le c < N)$, $\mathbf{V} = (v_1, v_2, \dots, v_c)$ indicates the centers of clusters, $u_{ij}\left(0 \le u_{ij} \le 1\right)$ is the degree of membership of jth pixel in ith cluster , $\mathbf{U} = \left(u_{ij}\right)_{c\times N}$ is fuzzy partition matrix, α is the controlling parameter. The fuzzy partition matrix $\mathbf{U}$ has constraints $0 < \sum_{j=1}^{N} u_{ij} < N, \forall i$. After solving equation (3) using Lagrange method of undetermined multiplier, centroids and degree of membership is given by equation (4) and equation (5) respectively as below:

$$v_i = \frac{\sum_{j=1}^{N} \mu_{ij}^m \left(x_j + \frac{\alpha}{N_R}\sum_{r\in N_j} x_r\right)}{(1+\alpha)\sum_{j=1}^{N} \mu_{ij}^m}, 1 \le i \le c \qquad (4)$$

$$\mu_{ij} = \left(\frac{\left(\left\|x_j - v_i\right\|^2 + \frac{\alpha}{N_R}\sum_{r\in N_j} \left\|x_r - v_i\right\|^2\right)}{\sum_{k=1}^{c} \left(\left\|x_j - v_p\right\|^2 + \frac{\alpha}{N_R}\sum_{r\in N_j} \left\|x_r - v_k\right\|^2\right)}\right)^{\frac{-1}{(m-1)}}, 1 \le j \le N, 1 \le i \le c \quad (5)$$

Although FCM_S handles noise in the image, but its execution time is high. It also does not handle uncertainty in the data.

The algorithm of FCM_S is as below:

1. Set No_of_cluster c, fuzzifier constant m, Number of maximum iteration F, $\alpha(regularizing\ parameter), stopping\ criteria\ \epsilon$

2. Perform random initialization of degree of membership matrix U_0 and loop counter z=0

3.Reiterate

4.for every i= 1, … , c

4.1 Calculate the cluster centers V_z by equation (4)

5.Compute the membership matrix U_z by equation (5)

6. z$\leftarrow$ z +1;

7. If $\|U_z - U_{z-1}\| < \epsilon$ or the F= max_iter then stop else update $U_z = U_{z-1}$ then move to step 4

3.1.3 Spatial Intuitionistic Fuzzy C-means (SIFCM)

Spatial Intuitionistic Fuzzy C-means (SIFCM) [14] uses intuitionistic fuzzy sets with spatial information. The spatial neighborhood helps to remove noise and intuitionistic fuzzy sets helps to handle uncertainty and vagueness in the image. The intuitionistic Fuzzy Set (IFS) was proposed by Atanassov [15]. Intuitionistic fuzzy sets guard the uncertainty that may arise owing to the insufficient information in telling the membership degree. An Intuitionistic fuzzy set is a 3-tuple set which includes degrees of membership, non-membership, and hesitation.

An image I in IFS notation is denoted as:

$$I = \{(s_{ij}, \mu_I(s_{ij}), v_I(s_{ij}), \pi_I(s_{ij})\}$$

where $\mu_I(s_{ij})$ is degree of membership, $v_I(s_{ij})$ is degree of non-membership and $\pi_I(s_{ij})$ is degree of hesitation of the image pixel s_{ij}. With the following condition,

$\mu_{\tilde{I}}(s)\colon \mathrm{S} \rightarrow [0,1], v_{\tilde{I}}(s)\colon \mathrm{S} \rightarrow [0,1]$ and $0 \leq \mu_{\tilde{I}}(s) + v_{\tilde{I}}(s) \leq 1; \forall \mathrm{S} \in \mathrm{S}$

And hesitation degree satisfies,

$$\pi_I(s_{ij}) = 1 - \mu_I(s_{ij}) - v_I(s_{ij}), 0 \leq \pi_I(s_{ij}) \leq 1$$

For building of Intuitionistic fuzzy sets Sugeno proposed a negation function. The negation function proposed by Sugeno is as follows:

$v_I(s_{ij})$=$Negation(\mu_I(s_{ij})) = (1 - \mu_I(s_{ij}))/(1 + \lambda . \mu_I(s_{ij})), \lambda > 0, N(1) = 0, N(0) = 1$

In IFS the membership function $\mu_I(s_{ij})$ is calculated as the addition of its membership function $\mu_{FI}(s_{ij})$ and its hesitation degree $\pi_I(s_{ij})$.

$$\mu_I(s_{ij}) = \mu_{FI}(s_{ij}) + \pi_I(s_{ij}) \text{ ------ } (6)$$

The spatial function is defined as:

$$h_{ij} = \sum_{k \in NB} (x_j) u_{ik}, \text{------} \qquad (7)$$

where $\mathrm{NB}(x_j)$ indicates the neighboring pixels of $x_{j.}$. The spatial function h_{ij} characterizes the notch of likelihood that x_j is in the i^{th} cluster. Spatial term is adjusted in the membership function as follows:

$$u'_{ij} = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^{c} u_{kj}^p h_{kj}^q}$$

where p and q are the controlling parameter which control weightage of fuzzy membership calculated using equation (6) and spatial function calculated using equation (7).

The algorithm of SIFCM is as follows:

1 Initialize centers v_i for $i = 1, \ldots, c$

2 Calculate the membership as:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\|x_j - v_i\|}{\|x_j - v_k\|} \right)^{2/(m-1)}},$$

where i ϵ[1, c]; j ϵ[1, N]

3 Compute the hesitation degree as follows:

$$\pi_{ij}(x) = 1 - u_{ij}(x) - \left(1 - u_{ij}(x)\right)/\left(1 + \lambda \cdot u_{ij}(x)\right),$$

where i ϵ[1, c]; j ϵ[1, N]

4 Calculate the fuzzy membership function incorporating hesitation degree:

$$u'_{ik} = u_{ik} + \pi_{ik}$$

where i ϵ[1, c]; j ϵ[1, N]

5 Compute the spatial function as follows:

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik}$$

where i ϵ[1, c]; j ϵ[1, N]

6 Calculate the new membership function which includes the intuitionistic and spatial information as:

$$u''_{ij} = \frac{u'^{p}_{ij} h^{q}_{ij}}{\sum_{k=1}^{c} u'^{p}_{kj} h^{q}_{kj}}$$

where i ϵ[1, c]; j ϵ[1, N]

Update $u_{ij} = u''_{ij}$, where i ϵ[1, c]; j ϵ[1, N]

7 Compute the new centers as follows:

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}^{m} x_j}{\sum_{j=1}^{N} u_{ij}^{m}}$$

where i ϵ[1, c]; j ϵ[1, N]

9 If | u_{ij}(new) $-u_{ij}$(old) |< ϵ then stop, else go to step 2.

10 Defuzzify the membership value

SIFCM performs better for noisy image and the images with uncertainty and vagueness.

3.2 Systematic Workflow of liver Segmentation

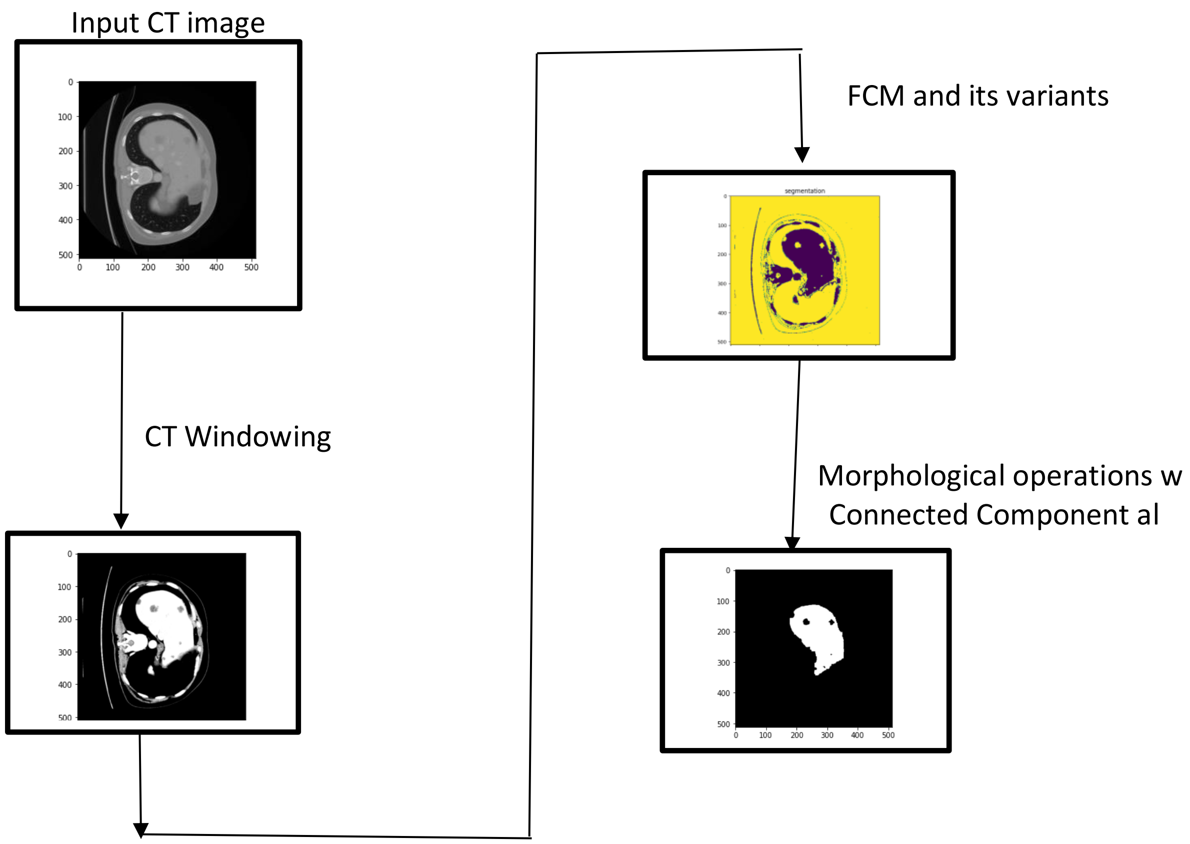


Figure 3-1: workflow diagram of liver segmentation

Detailed description of the approach:

A. Input Image

3Dircadb-01 dataset is used to get the CT image. In this dataset 20 folders are present. Each folder corresponds to each patient. In each folder slice range varies from 74 to 260. Each CT image slice is of 512 × 512 resolution. The CT slice is in Digital Imaging and Communications in Medicine (DICOM) format. Mask images from liver mask folder are used as ground truth for validating the result.

B. Selecting CT scans slices

For our experiment we have used central slice of each patient.

C. Windowing

Windowing [16] is grayscale mapping, contrast stretching, histogram correction, or contrast enhancement. It manipulates grey level component of CT image using Hounsfield unit. After applying windowing on CT images, the view of the image got changed and specific structured of the image is focused. of the image and enhance the texture. Window level is used to regulate the illumination of the image. Window Width is used to regulate the contrast of the image.

Hounsfield unit or CT number: Hounsfield units are named after Sir Godfrey Hounsfield which was inventor of CT. The CT number is also known as Hounsfield unit. The Hounsfield Unit (HU) is a relative quantitative measure of radiodensity used by radiologists when interpreting computed tomography (CT) images. Absorption/attenuation coefficients of radiation in tissue are used to generate grayscale images during CT reconstruction. Hounsfield units represents the x-ray attenuation in the corresponding voxel. Hounsfield units are dimensionless units widely used in CT scans to express CT values in a standardized and convenient format. The Hounsfield unit is proportional to degree of attenuation by tissue. Hounsfield units is calculated using a linear transformation of the measured attenuation coefficient.

Therefore, for a voxel with average linear attenuation coefficient $\mu_{material}$,the equivalent HU value is given by:

$$\frac{(\mu_{material}-\mu_{water})\times 1000}{(\mu_{water})}$$

Where μwater is the attenuation coefficient of water and μair is the attention coefficient of air and μmaterial is the attenuation coefficient of material. And HUwater=0, HUair=-1000 on HU unit. The upper limit can be up to 1000 for bones, 2000 for dense bones, and more than 3000 for metals like steel or sliver.

Bone	1000 HU
Gall Stone	+30 to +120 HU
Clotted blood	+50 HU to +75 HU
White matter	46 HU
liver	40 to 60 HU
Grey matter	43 HU
Unclotted blood	+13 HU to +50 HU
muscle	10 to 40 HU
kidney	30 HU
Cerebrospinal fluid	15 HU
water	0 HU
fat	-50 to -100 HU
air	-1000 HU

Table 3-1: Hounsfield unit value of substances

Window level (WL): Central CT number is called the window level that determines brightness. It is the middle value of the CT values for the range of CT values in an image.

Window Width (WW): It is defined as the range of CT value greater and lower than window level. It determines the contrast. This is the total range of CT values contained in the image. There are two types of window wide window and narrow window. Wide window is defined in the range 400-2000 HU. It is frequently used when dealing with an area where different tissue density is there. Example includes lungs or cortical tissue. Narrow window is defined in the range 50-350 HU. It is the best when examining areas of similar density. Example of narrow window includes soft tissue. For our liver segmentation we have used narrow window.

Upper and lower grey level calculation: if WW and WL are given then,

the upper grey level (u) = WL + (WW ÷ 2)

the lower grey level (l) =WL - (WW ÷ 2)

Typical window width and level values for liver is 150 and Window Level 30 in HU unit.

D Segmentation:

On windowed image we apply the gaussian noise with mean=128 and standard deviation=20. Then we Apply the FCM, FCM_S and SIFCM each one at a time to do segmentation. Here we are using two cluster one containing the liver region and another for non-liver region. The liver containing region is decided by Average intensity of region. As the liver containing region has higher average pixel intensity.

E Post-processing

From the segmented image after getting the liver containing region, we apply morphological operations(erosion) on it to eliminate which strips away the extrusion and strips apart the joined object. After that we apply the connected Component algorithm to extract the largest which is liver. Then we apply dilution to fill the holes and intrusion.

Morphological operations: Morphological operations [17] is like to spatial filtering. Morphological operations are based on structuring element, fit, hit and miss. Structuring element is a matrix that is used to traverse the image pixels. When all the pixels of structuring element cover the object then it is called hit, if at least a single pixel is covered then it is called fit and if not, a single pixel is covered then it is called flop. The morphological operations are of four types but mostly two are used.

(1) Erosion: It cuts the image pixel, removes object joining pixel and removes extrusion.

(2) Dilution: It adds image pixel, fills the holes, and removes intrusion.

Connected component algorithm [18]: It is used to find the total number of connected regions in the image. From that least of regions we can easily find the largest region.

Chapter 4

Experimental Results and Discussion

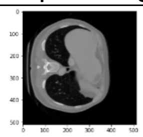
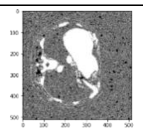
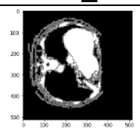
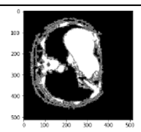
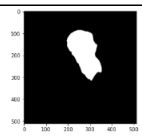
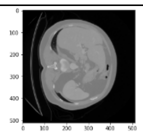
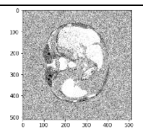
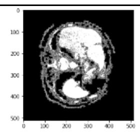
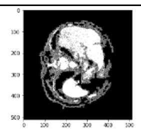
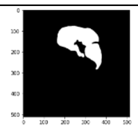
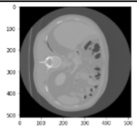
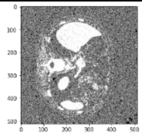
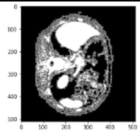
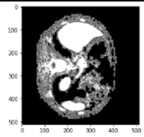
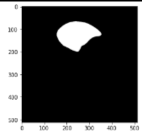
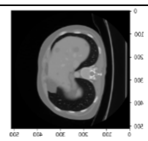
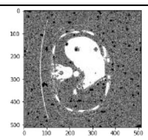
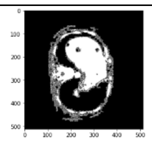
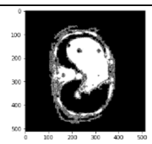
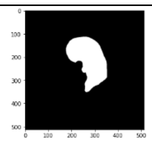
Serial no.	Input image	FCM	FCM_S	SIFCM	Ground Truth
1					
2					
3					
4					

Table 4-1: segmentation result with FCM and its variants on gaussian noisy image

To compare the performance of the FCM based methods discussed in chapter 3 Jaccard similarity coefficient (JSC) and Dice similarity coefficient (DSC) are calculated for each 20 patient's central slices with added gaussian noise. JSC and DSC calculate the likeness between FCM based segmented liver with corresponding ground truth which is manually segmented by radiologists.

FCM, FCM_S and SIFCM method for liver segmentation are applied on the 3D-IRCADb-01 database. For each patient we select the central slice. After adding gaussian noise we do segmentation of each slice using FCM, FCM_S and SIFCM. Then with corresponding ground truth liver we calculate JSC and DSC for each slice. At last, we calculate average JSC and average DC for each method.

Let I = {complete image pixels set}, L = {the segmented liver pixel set}, and M = {ground truth pixels set of liver}; where $L \in I$ and $M \in I$.

True positive (TP):TP is the result where our models accurately determine the positive class. For our liver segmentation problem, it is the set of all pixels that were accurately indicated as liver by models. It is defined by:

$$TP = L \cap M$$

True negative (TN): TN is the result where our models accurately determine the negative class. For our liver segmentation problem, it is the pixels set that were accurately identified as non-liver by models. It is formulated by:

$$TN = L' \cap M'$$

False positive (FP): FP is the result where our models inaccurately determine the positive class. For our liver segmentation problem, it is the pixels set that are inaccurately identified as liver. It is formulated by:

$$FP = L \cap M'$$

False negative (FN): FN is the result where our models inaccurately determine the negative class. For our liver segmentation problem, it is the pixels set that are inaccurately identified as non-liver. It is formulated as:

$$FN = M \cap L'$$

Jaccard similarity score (JSC): JSC is a measure of similarity between two datasets. The value ranges from 0 to 1. If value is high, then the two datasets are more similar. It is defined as:

JSC = TP / (TP + (FP+FN))

Dice similarity coefficient (DSC): DSC is also a measure of similarity between two datasets. It is formulated as:

DSC = 2×TP / (2×TP + (FP+FN))

The range of JSC and DSC lies in the range of 0 and 1. If it is 0, it means no similarity between the datasets, while a value of 1 means both the datasets are similar. For our liver segmentation problem value 0 means inaccurate segmentation and value of 1 means accurate segmentation of liver.

Methods / Patients	FCM	FCM_S	SIFCM
IRCADb-1-01	0.3041	0.4493	**0.4551**
IRCADb-1-02	0.1235	0.3205	**0.3409**
IRCADb-1-03	0.2104	**0.3773**	0.3753
IRCADb-1-04	0.1514	**0.3036**	0.3031
IRCADb-1-05	0.2367	0.2400	**0.2521**
IRCADb-1-06	0.1202	0.1203	**0.1205**
IRCADb-1-07	0.2660	0.5517	**0.5537**
IRCADb-1-08	0.3755	0.5140	**0.5240**
IRCADb-1-09	0.2017	0.5138	**0.5214**
IRCADb-1-10	0.2881	0.4935	**0.4976**
IRCADb-1-11	0.1943	0.3676	**0.3840**
IRCADb-1-12	0.1534	0.3125	**0.3133**
IRCADb-1-13	0.2505	0.3612	**0.3734**
IRCADb-1-14	0.2382	0.4348	**0.4363**
IRCADb-1-15	0.1666	0.4030	**0.4066**
IRCADb-1-16	0.3487	0.4121	**0.4180**

IRCADb-1-17	0.2815	0.4584	**0.4602**
IRCADb-1-18	0.3043	0.3907	**0.4006**
IRCADb-1-19	0.2952	0.6438	**0.6457**
IRCADb-1-20	0.1143	0.3599	**0.3656**
Average DSC score	0.2312 ± 0.08	0.4014 ± 0.12	**0.4074 ± 0.12**

Table 4-2: Dice scores of Liver segmentation on central slice with added gaussian noise

Methods / Patients	FCM	FCM_S	SIFCM
IRCADb-1-01	0.1793	0.2897	**0.2946**
IRCADb-1-02	0.06583	0.1909	**0.2055**
IRCADb-1-03	0.1175	**0.2325**	0.2310
IRCADb-1-04	0.0819	**0.1790**	0.1786
IRCADb-1-05	0.1342	0.1364	**0.1369**
IRCADb-1-06	0.0639	0.0640	**0.0641**
IRCADb-1-07	0.1539	0.3809	**0.3828**
IRCADb-1-08	0.2311	0.3459	**0.3550**
IRCADb-1-09	0.1123	0.3457	**0.3526**
IRCADb-1-10	0.1683	0.3276	**0.3312**
IRCADb-1-11	0.1076	0.2252	**0.2377**
IRCADb-1-12	0.0831	0.1852	**0.1857**
IRCADb-1-13	0.1432	0.2204	**0.2296**
IRCADb-1-14	0.1352	0.2778	**0.2790**
IRCADb-1-15	0.0909	0.2524	**0.2552**
IRCADb-1-16	0.2112	0.2595	**0.2642**
IRCADb-1-17	0.1638	0.2973	**0.2989**
IRCADb-1-18	0.2101	0.2428	**0.2505**
IRCADb-1-19	0.1731	0.4747	**0.4768**
IRCADb-1-20	0.0606	0.2194	**0.2237**
Average JSC score	0.1344 ± 0.05	0.2574 ± 0.09	**0.2617 ± 0.09**

Table 4-3: JSC scores of Liver segmentation on central slice with added gaussian noise

Clearly from the result SIFCM performs better than FCM and FCM_S for liver segmentation. And FCM_S performs better than FCM.AS, average JSC score and DC score for Liver segmentation using SIFCM on 3D-IRCADb-01 dataset is greater than FCM and FCM_S. This is because in SIFCM spatial features of pixel is taken into consideration for handling the noise and Intuitionistic fuzzy sets (IFS) takes into consideration the uncertainty i.e., hesitation degree that may appear due to the inadequate knowledge in describing the membership degree. So SIFCM is better approach than FCM and FCM_S for liver segmentation with noisy images.

Chapter 5

Conclusion and Future Direction

Liver Segmentation from CT scans is an important step for detecting liver anomaly. For finding the liver tumors in liver first step is the segmentation of liver from CT scans. Then in segmented liver we find tumors. In Manually way for finding liver tumors in CT scans radiologist has to do hard and tedious work. Firstly, they segment the liver from abdominal CT scans as it contains more organs apart from liver manually. Then in the segmented liver they find tumors. This task is very complicated as there are lots of organs attached with almost similar intensity level. So, our automated work of liver segmentation in CT scans will play a vital role in automated way of finding the liver tumors. We have investigated three methods such as (i) Fuzzy c-means (FCM) method (ii) Fuzzy c-means with local spatial information (FCM_S) (iii) Spatial Intuitionistic Fuzzy c-means (SIFCM) method to do Liver segmentation. The performance of these three methods is compared with available ground truth of liver which is manually segmented by radiologist on public ally available 3D-IRCADb-01 dataset.

The performance of SIFCM is found better than FCM and FCM_S for liver segmentation with noisy images.

As part of our Future work, we would like to find the liver tumors in the segmented liver. This will speed up the task of finding liver disease by radiologists. If we identify the liver disease early, then we can also save the life of people. So, it will be a great work to humanity.

References

[1] Li, B. N., Chui, C. K., Ong, S. H., & Chang, S. (2009). Integrating FCM and level sets for liver tumor segmentation. In 13th international conference on biomedical engineering (pp. 202-205). Springer, Berlin, Heidelberg.

[2] Kumar, S. S., Moni, R. S., & Rajeesh, J. (2013). Automatic segmentation of liver tumour using a possibilistic alternative fuzzy C-means clustering. International Journal of Computers and Applications, 35(1), 6-12.

[3] Obayya, M., & Rabaie, S. E. (2015). Automated segmentation of suspicious regions in liver ct using fcm. International Journal of Computer Applications, 975, 8887-8890.

[4] Das, A., & Sabut, S. K. (2016). Kernelized fuzzy C-means clustering with adaptive thresholding for segmenting liver tumors. Procedia Computer Science, 92, 389-395

[5] Yugander, P., & Reddy, G. R. (2017, May). Liver tumor segmentation in noisy CT images using distance regularized level set evolution based on fuzzy C-means clustering. In 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT) (pp. 1530-1534). IEEE.

6] Rela, M., Rao, S. N., & Reddy, P. R. (2020). Liver Tumor Segmentation using Superpixel based Fast Fuzzy C Means Clustering. International Journal of Advanced Computer Science and Applications (IJACSA), 11(11).

[7] Khan, Z., & Loganathan, R. (2020, October). AutoLiv: Automated Liver Tumor Segmentation in CT Images. In 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE) (pp. 151-156). IEEE.

[8] Al-Saeed, Y., Soliman, H., & Elmogy, M. (2020, October). Liver Segmentation using Fast-Generalized Fuzzy C-Means (FG-FCM) from CT Scans. In *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)* (pp. 1-6). IEEE.

[9] Pohle, R., & Toennies, K. D. (2001, September). A new approach for model-based adaptive region growing in medical image analysis. In *International conference on computer analysis of images and patterns* (pp. 238-246). Springer, Berlin, Heidelberg.

[10] Zhao, Y., Zan, Y., Wang, X., & Li, G. (2010, May). Fuzzy C-means clustering-based multilayer perceptron neural network for liver CT images automatic segmentation. In 2010 Chinese control and decision conference (pp. 3423-3427). IEEE.

[11] Song, H., Zhang, Q., & Wang, S. (2014, November). Liver segmentation based on SKFCM and improved GrowCut for CT images. In *2014 IEEE international conference on bioinformatics and biomedicine (BIBM)* (pp. 331-334). IEEE.

[12] Yang, M. S. (1993). A survey of fuzzy clustering. *Mathematical and Computer modelling*, *18*(11), 1-16.

[13] Ahmed, M. N., Yamany, S. M., Mohamed, N., Farag, A. A., & Moriarty, T. (2002). A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE transactions on medical imaging*, *21*(3), 193-199.

[14] Tripathy, B. K., Basu, A., & Govel, S. (2014, December). Image segmentation using spatial intuitionistic fuzzy C means clustering. In *2014 IEEE International Conference on Computational Intelligence and Computing Research* (pp. 1-5). IEEE.

[15] Atanassov, K. (2016). Intuitionistic fuzzy sets. *International journal bioautomation*, *20*, 1.

[16] Kaluva, K. C., Khened, M., Kori, A., & Krishnamurthi, G. (2018). 2D-densely connected convolution neural networks for automatic liver and tumor segmentation. *arXiv preprint arXiv:1802.02182*.

[17]Chudasama, D., Patel, T., Joshi, S., & Prajapati, G. I. (2015). Image segmentation using morphological operations. *International Journal of Computer Applications*, *117*(18).

[18] Di Stefano, L., & Bulgarelli, A. (1999, September). A simple and efficient connected components labeling algorithm. In *Proceedings 10th international conference on image analysis and processing* (pp. 322-327). IEEE.