



## BOUNDARY OPTIMIZATION OF GROUND GLASS OPACITY IN CT IMAGES OF LUNG CANCER

Saravanan S.<sup>1</sup>, Selvakumar G.<sup>2</sup>, Amarnath C.<sup>3</sup> and Manikandan S.<sup>4</sup>

<sup>1</sup>Department of Electronics and Communication Engineering, Sathyabama University, Chennai, India

<sup>2</sup>Electrical Sciences, Muthayammal Engineering College, Rasipuram, India

<sup>3</sup>Department of Radiology, Stanley Medical College, Chennai, India

<sup>4</sup>Department of Engineering, LocToc Pte Ltd., Singapore

E-Mail: [saravananamie@gmail.com](mailto:saravananamie@gmail.com)

### ABSTRACT

In computer aided diagnosis, pre-processing, segmentation, feature extraction and classification are the steps involved. For segmentation the boundary must be defined to get regional information inside the boundary. Ground Glass Opacity (GGO) has ill-defined boundary. Hence, there is a necessity to optimize the boundary of GGO. Once the boundary is optimized, feature extraction and classification of malignant and benign of the particular GGO becomes easy. Using Distance Regularized Level Set Evolution (DRLSE) and Active contour without edges independently, the contour is grown and compared with the model image. The model image was created using the expertise of a Radiologist, applying Mattes Mutual information method. The contour which gives maximum mutual information is concluded as the optimized boundary. Wavelet transform has already been proven its application in identifying pits and cracks of corrosion metals. The same analogy is applied in GGO so that the partly solid and liquid GGOs can be precisely classified as malignant or benign. After optimizing the ill-defined boundary of GGO features of wavelet transformation were extracted along with textural features of mean and variance. Skewness and kurtosis were neglected since they were negligibly small. It is shown that on comparing the growing contour with model image using mattes mutual information, the DRLSE method shows greater results, without leaking compared to Active contour without edges. After extraction of features from wavelet transformation and textural features classification as malignant and benign was done using learning vector quantization (LVQ). On finding the optimized boundary, it is easier to classify the ground glass opacity as diffuse finding or local finding. Hence taking two images where one is taken as malignant and other as benign classification was done as benign and benign by the classifier. The malignant have been identified as benign due to minimum number of images used in training.

**Keywords:** computer tomography, active contour, mutual information, DRLSE, boundary, ground glass opacity, feature extraction, LVQ, learning vector quantization, orian, wavelet transform.

### INTRODUCTION

A solitary pulmonary nodule (SPN) is determined as a discrete approximately rounded opacity which is less than or equal to 3 centimeters in diameter and ringed by lung parenchyma and the remaining lung must be without pathology [1-2]. SPN can be benign or malignant. There are certain features which can help to differentiate between them in computer tomography (CT). Those features are as follows:

1. Size
2. Shape
3. Margin
4. Growth
5. Calcification
6. Solid and GGO
7. Air Bronchograms sign
8. Contrast Enhancement

Even though CT can demonstrate all the above features, to classify an SPN into benign or malignant it is still inconclusive to define a nodule with significant specificity that it is a benign or malignant. For the purpose of minimizing lung cancer mortality and improving the prognosis of patients suffering from lung cancer, low-dose computed tomography (LDCT) is accepted as an effective

screening method in high risk individuals. For the success of the screening program, the outcome of the classification of malignancy and benignity is important. Among the pulmonary nodules Ground Glass Opacity (GGO) can be restricted to partly solid and Non-Solid. They appear very frequently in CT lung images and possess greater liability of being malignant compared with the solid ones [2], Normally non-solid GGO nodules show fuzzy boundaries, having highly identical intensity with its background, and part solid GGO nodules showing highly uneven variations in intensity (intensity inhomogeneity) and boundary shapes.

Owed to the fuzzy boundary, the low contrast and intensity inhomogeneity of GGO pulmonary nodules, widely used methods for segmentation such as purely intensity thresholding or model based segmentation methods, may not segment GGO nodules, resulting in boundary leakage [3-5]. By virtue of the above quoted issue, authentic segmentation for GGO nodules continues to be an ongoing research topic. Amidst all the segmentation methods portrayed hitherto, Active Contour Models (ACM) remains to be one of the ultimate methods for image segmentation [6-7] even in the segmentation of pulmonary nodules.



There are numerous fascinating factors of ACM over traditional image segmentation methods in region growing and edge detection such as:

a) Sub-pixel accuracy of object boundaries can be achieved through ACM.

b) The ACM can be comfortably formulated under a fair energy minimization framework, and permit incorporation of distinct prior observations, such as shape and intensity distribution, for powerful image segmentation.

c) The ACM can provide easy and closed contours as the results of segmentation, which are crucial and could be readily used for other application, like multi-features analysis and recognition.

The boundary leakage problem usually occurs in boundary segmentation using ACM. Edge based ACM is introduced through the integrated ACMs. This couple the influence of the edge based ACM and the region based ACM. Commonly, the representation with local information provides improved performance outcome compared with that of global statistic information in determining the segmentation problems of intensity in homogeneity. When the local domain is either too large or too small, local information dependent ACM is incapable of segmenting curve of fuzzy and zigzag edge. If the ACM is based on non-parametric, autonomous and equally distributed image data, the image may be segmented as per the appropriate global to regional strategy.

Even though, the segmentation techniques based on wavelet transform have been suggested to augment the disparity found between the objects and its background [8], it is challenging to bring in an ideal segmentation effect just by relying on general image-data-driven segmentation methods. Lung nodules are incorporated in a complex and structured background. Their identification and segmentation is normally affected by the surrounding anatomical objects. Hence, in a wider sense, the feature space helpful in the recognition of nodules should be embedded with extra prior information, which includes the intensity information, intensity variation, and texture along with many more. Because of the malignant potential and heterogeneous characteristics [9] of the GGO on CT are particularly challenging, the ill-defined boundary of the GGO is the major problem for segmentation of region of interest. Since feature extraction and classification depends on segmentation of region of interest, optimization of ill-defined GGO boundary becomes mandatory for accurate development of CAD. This paper focus on boundary optimization of segmented GGO using the existing active contour method. Further it is extended to classify malignant and benign nodules using LVQ after extracting energy features from wavelet transform of the image and textural features like mean and variance.

### Radiologic features of GGO nodules

GGO nodules are defined as focal areas of increased CT attenuation through which underlying normal lung parenchyma, and pulmonary vessels are seen. Increased CT attenuation of the opacity is due to the reduction in air space volume of the lung as well as

partial/total replacement of air by cells or fluids or blood. As far as GGO is concerned air space volume in lung is only partially replaced by cell or fluid.

Radiologically GGO can be divided into

a) Pure GGO.

b) Partially solid GGO.

Partly solid GGO are otherwise called as mixed GGO. GGO is a nonspecific radiological feature which can be seen in numerous clinical conditions involving various pathological processes. Hence, GGO can not only represents alveolar changes, but also represent surrounding interstitial changes. GGO can be rather a diffuse finding or can be a focal finding.

Diffused GGO can be due to

a) Infection

b) Volume overload status (Pulmonary oedema, Cardiac failure, Renal failure)

c) Interstitial diseases

Focal GGO is more towards space occupying lesion probably a malignant pathology. Likewise, pure GGO without any solid component is more in favor of benign pathology, whereas mixed GGO is more in favor of malignant pathology. While training the neural network for classification, following the management of SPN, growth and proportional solidity of the GGO can be mapped as an input vector for the neural network. Since non-solid ground glass lesions have a malignancy rate of 18%, a precise classification on GGO for malignancy and benignity becomes essential. In this paper methods of segmentation of GGO namely Active Contour without edges [10] and DRLSE [11] method are compared. This paper compares these two methods and its relevance to image segmentation using mutual information on the intermediate contours as moving image. The GGO model image is prepared as a fixed image as per guidelines of Radiologist.

### Active contour without edges

The idea behind ACMs or snakes is to detect objects in the given image  $u_0$  by evolving a curve, subject to conditions from the given image, such that the curve shift towards its interior normal and has to halt at the object boundary. Let  $\Omega$  be a bounded open subset of  $\mathbb{R}^2$ , with  $\partial\Omega$  as its boundary. Let  $u_0: \bar{\Omega} \rightarrow \mathbb{R}$  be a given image, and  $C(S): [0, 1] \rightarrow \mathbb{R}^2$  be a parameterized curve.

In the ACM, confiding on the gradient of the image  $u_0$ , an edge detector helps in stopping the curve evolving on the boundary of the observed object [12-15]. For the models discussed in [13-16], the stopping function 'g' is never zero at the edges, and the curve may pass over the boundary since the discrete gradients are bounded.



Active Contour without edges uses a method, in which the stopping process is not dependent on the gradient of the image  $u_0$ . The stopping term is based on Mumford Shah segmentation techniques. Active Contour without edges can detect contours either with or without the gradient, for example, objects with very smooth boundaries or even with discontinuous boundaries. Additionally, active contour without edges has a level set formulation. Further interior contours are detected automatically, and the initial curve can be found anywhere in the image.

### Distance regularized level set method

Recently, active contour based methods are extensively studied and many researchers use this to address a wide range of image segmentation problems especially in medical imaging. The ACMs can be categorized into two such as parametric active contour models [17] and geometric active contour models [18]. If the active contour model is based on Lagrangian frame work, it is a parametric active contour model, geometric active contour models use curve evolution theory and represented as a level set function (LSF).

A Zero level set, a higher dimensional function is used to refer the active contour [19]. The change in topology and shape of contour can be represented as the curve and level set function, which is an evolution in the geometric active contour models.

The evolution of curve [20] can be expressed as Equation (1).

$$\frac{\partial C(t)}{\partial t} = \beta \vec{N} \quad (1)$$

Where the speed function  $\beta$  controls the evolution speed of the curve and  $\vec{N}$  is the inward normal vector of contour  $C(t)$ . The inward normal vector is expressed as Equation (2).

$$\vec{N} = -\frac{\nabla \phi}{|\nabla \phi|} \quad (2)$$

When the level set is used to express the contour.

Converting the curve evolution in Equation (1) to a partial differential equation (PDF), Equation (3) which is termed as level set evolution equation is obtained.

$$\frac{\partial \phi}{\partial t} = \beta |\nabla \phi| \quad (3)$$

The velocity field of the image plane is unpredictable; because of the irregularities of the LSF during the evolution course controlled by the level set evolution equation. As a result the LSF will lose the signed distance function (SDF) features. This issue can be rectified through re-initializing the LSF. However, some problems are associated along with re-initialization [21].

Whenever re-initialization occurs, there is a reduction in the efficiency of level set evolution. Because of these aforesaid problems, it is essential to seek a level set regularization minus re-initialization. The progression in Li's formulation [11] is labeled as distance regularized level set evolution (DRLSE). The DRLSE can implement a straightforward and powerful narrow band implementation to enhance the computational efficiency. With DRLSE as reference, the segmentation of GGO has been described [11].

### Mutual information

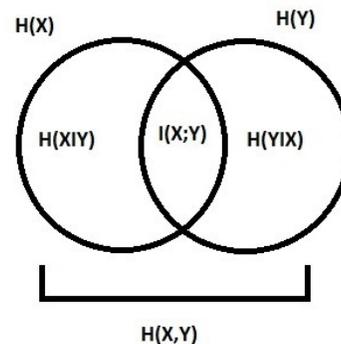


Figure-1. Venn diagram of mutual information.

The mutual dependence between the two random variables is termed as mutual information in probability theory and information theory. Figure-1 represents the mutual information Venn diagram. The volume of information gained about one random variable, with the help of the other random variable is quantified in mutual information. The concept of entropy and mutual information is intricately linked with each other both defining the amount of information in a random variable. Mutual Information is more generic and fix upon how identical the joint distribution  $P(X, Y)$  is to the products of factored marginal distribution  $P(X) P(Y)$ .

### METHODOLOGY

#### Optimizing the ill-defined boundary

Five images of GGO are represented through Figures (2)-(6). With the help of a Radiologist the approximate ill-defined boundary of GGO is marked and the images are cropped manually. The average of the cropped image has been found by adding the pixel intensities in the respective frames of same spatial coordinates and then dividing it by the number of images. Thus the model image is formed.

Active Contour is allowed to grow by giving initial seed points and every growing contour is compared with the model image evaluating mutual information. As per the stopping condition of Active Contour without Edges and also with DRLSE methods the contour will terminate and the slice which consists of maximum mutual information is concluded as the optimized boundary for GGO. Thus the GGO is segmented.



Mutual Information metrics are information theoretic approach to measure how related information is. These methods practice the joint probability distribution of a pixel sampling from a couple of images to measure the definiteness that the values of one set of pixel map to corresponding values on the other image. This information is a quantitative measure of how related the images are. High mutual information signifies an extensive cut in the ambiguity between the two distributions, indicating that the images are reasonably better aligned.

The Mattes Mutual information algorithm practices a single set of pixel locations for the optimization period, rather than drawing a new set of iteration. Both the number of samples employed to figure out the probability density estimates as well as the number of bins useful in computing the entropy are selectable by the users in MATLAB. The marginal and joint probability density function is assessed at the evenly spaced bins using the samples. Entropy values are computed by summing over the bins. Zero-order and third order B Spline kernels are used to calculate the probability density functions of the fixed and moving images appropriately [22-23]. After optimizing the ill-defined boundary of GGO, features of wavelet transformation were extracted along with textural features of mean and variance. Skewness and kurtosis were neglected since they were negligibly small.

### Feature vector

From the Wavelet transformed Image, Energy and Orian are derived as feature vector. A wavelet transform of a 2D image  $I(x,y)$  can be implemented by employing the filters high pass and low pass consequently along the rows and columns of the image. The sub-image emerging from one such operator can be reported through Equations (4-7).

$$L^1(m,n) = [L_x * [L_y * I]](x,y) \quad (4)$$

$$D_1^1(m,n) = [L_x * [H_y * I]](x,y) \quad (5)$$

$$D_2^1(m,n) = [H_x * [L_y * I]](x,y) \quad (6)$$

$$D_3^1(m,n) = [H_x * [H_y * I]](x,y) \quad (7)$$

'\*' represents the convolution operator. The first convolution is executed along the columns of the image, the second along the rows;  $L^1$  is a smoothed rendition of the initial image  $I$ . The thorough image referred as  $D_1^1$ ,  $D_2^1$  and  $D_3^1$ . has all details about the vertical, horizontal and diagonal directions, thus holding distinct orientation data. While  $L_x$  and  $L_y$  is low pass filter and  $H_x$  and  $H_y$  are high pass filters. By emphasizing this operation on consecutive low pass images  $L^{i-1}$ , sub images ( $L^i$ ,  $D_1^i$ ,  $D_2^i$  and  $D_3^i$ ) on various levels are developed.

The outcome of this development is a tree structure with detail illustration for various scales and orientations that is termed as standard (pyramidal) wavelet decomposition (StW). An exhaustive quad tree of images

is obtained while every sub images are decomposed along with  $L^1$ . This is termed as wavelet packet decomposition (WP) or tree structured wavelet transform.

### Energy feature

The information of different scales is effortlessly disintegrated by the decomposition. By this is easier to extract a smaller feature set, by figuring out a single number for every sub-image. The conventional energy is chosen as Equation (8).

$$E_j^i = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N (D_j^i(m,n))^2 \quad (8)$$

Here  $M$  and  $N$  signify the size of the sub-image. The feature vector components comprise of the energies of the sub-images which resulted from wavelet decomposition. For decomposition up to a level  $d$ , the StW yields  $3d+1$  features and the wavelet packet (WP) of  $4^d$ .

### Rotational invariance

The definitive orientation of the textures can be of prime concern in countless texture analysis utilization, along with those where wavelet features are popularized. However for the current application, it is common to claim rotational invariance. The feature vectors and decomposition of energy preserve oriental information.

By commonly combining the three features into a single energy feature per scale, the confinement of oriental information can be avoided. However, this leads to a much raw depiction of energy, overlooking the directionality which accommodates crucial information too. The energy related to a detail sub image  $D_j^i$  can be viewed as the energy for a single direction on a scale  $I$ . This can also be construed locally, where the squared pixel values  $E_j^i(m,n) = (D_j^i(m,n))^2$  (denoted as  $\epsilon_j^i$ ) represents the local energy for a single direction for a sub image. The total energy per pixel for the three local energies  $\epsilon_j^i = 1, 2, 3$  can be extracted from Equation (9). They also provide the measure for the anisotropy of the energy as well [25]. Equation (10) gives the Orian.

$$E_{total}^i(m,n) = \frac{\epsilon_1^i + \epsilon_2^i + \epsilon_3^i}{MN} \quad (9)$$

$$Orian^i(x,y) = \frac{\sqrt{(\epsilon_1^i - \epsilon_2^i)^2 + (\epsilon_1^i - \epsilon_3^i)^2 + (\epsilon_2^i - \epsilon_3^i)^2}}{E_{total}^i(m,n)} \quad (10)$$

The complete feature vector on wavelet decomposition of depth 'd' will be of dimension '2d+1'. For each and every scale 'i', it will have an  $E_{total}^i$  and an Orian. Also it has another energy component of the low pass image  $L^d$ . On taking wavelet transform decomposition the regional information and boundary information of the ill-defined GGO is embedded as total



energy and orian as feature vectors for precise classification of GGO as malignant or benign.

### Moments of a distribution

Central tendency is the tendency to cluster around some particular value. While a set of values has an adequately strong central tendency, it may be appropriate to portray the set by a few numbers that are relevant to its moments and the aggregate of integer powers of the values.

Mean

$$Mean = \frac{\sum_{m=1}^M \sum_{n=1}^N x(m,n)}{N1} \quad (11)$$

Equation 11, which indicates the mean, estimates the value over which the central clustering develops. Having identified a distribution's central value, one can conventionally describe its 'width' or 'variability' around the identified value as shown in Equation (12).

$$Variance = \frac{\sum_{m=1}^M \sum_{n=1}^N (x(m,n) - Mean)^2}{N1} \quad (12)$$

### Standard deviation = $\sqrt{Variance}$

Standard Deviation is the root of variance.

### Skewness

The Skewness otherwise termed as third moment describes the degree of asymmetry of a distribution around its mean. While the mean, standard deviation, and average deviation are dimensional quantities, the skewness is commonly defined in a way to make it non dimensional. It is a plain number that outlines only the shape of the

distribution. Definition of skewness is given in Equation (13).

$$Skewness = \frac{1}{N1} \frac{\sum_{m=1}^M \sum_{n=1}^N (x(m,n) - Mean)^3}{\sqrt{\sum_{m=1}^M \sum_{n=1}^N (x(m,n)^2 - Mean^2)}} \quad (13)$$

A positive value of skewness implies a distribution, with an asymmetric trial ranging out towards more positive x; a negative value implies a distribution whose tail approaches out close to more negative x.

### Kurtosis

Kurtosis is a non-dimensional quantity. It denotes the relative peak or flatness of a distribution, relative to Normal Distribution as found in Equation (14).

$$Kurtosis = \frac{\sum_{m=1}^M \sum_{n=1}^N (x(m,n) - Mean)^4}{N1 \sqrt{\sum_{m=1}^M \sum_{n=1}^N (x(m,n)^2 - Mean^2)}} \quad (14)$$

N1 is the overall number of non-zero pixels.

Lepokurtic is the distribution with positive Kurtosis, for which the outline of the Matterhorn is an example. Platykurtic is the distribution with negative Kurtosis, for which the outline of a loaf of bread is an example. The in-between distribution is describes as mesokurtic.

### Experimental Results

**Table-1.** Mattes mutual information.

CT Sample figures of subjects	MATTES Mutual information										Selection of slice
	1	2	3	4	5	6	7	8	9	10	
2a	0.3429	0.3535	0.3569	0.3535	0.3703	0.3735	0.3766	0.3689	0.3655	0.3729	7
2b	0.2516	0.2516	0.2516	0.2516	0.2516	0.2516	0.2516	0.2516	0.2516	0.2516	1
3a	0.2112	0.2902	0.3175	0.2113	0.2113	0.2113	0.3215	0.3212	0.3222	0.3197	9
3b	0.2922	0.3008	0.3198	0.3120	0.3203	0.3674	0.3188	0.3206	0.3169	0.3105	6
4	0.4451	0.4380	0.4397	0.4287	0.4044	0.3765	0.3753	0.3766	0.4271	0.3750	1
5a	0.3748	0.3911	0.3900	0.3973	0.3943	0.3970	0.3981	0.3931	0.3893	0.4058	10
5b	0.3792	0.4138	0.3666	0.3781	0.3870	0.3790	0.3729	0.3713	0.3520	0.3441	2
6a	0.4133	0.4092	0.4035	0.4015	0.3988	0.3937	0.3931	0.3909	0.3623	0.3665	1
6b	0.2486	0.2496	0.2494	0.2462	0.2462	0.2497	0.2492	0.2485	0.2499	0.2499	10

Table-1 gives the list of the mattes mutual information got on the sample images. The maximum value of the mattes mutual information corresponds to the

slice which has an optimized boundary of the GGO. Table-1 depicts that selection of slice column gives the slice number which has maximum mutual information.

**Table-2.** Feature extraction.

Image# of subject	Total energy		Energy of LL	Orion level 1	Orion level 2	Mean	Variance
	Level 1	Level 2	Level 2				
1a	0.0031	3.36e <sup>3</sup>	3.3887e <sup>5</sup>	8.7121e <sup>4</sup>	2.1564e <sup>4</sup>	247.4869	1698.5
1b	0.0032	3.4679e <sup>3</sup>	4.3981e <sup>5</sup>	8.1012e <sup>4</sup>	1.1842e <sup>4</sup>	250.0131	1091.3
2a	0.0054	3.7478e <sup>3</sup>	4.2214e <sup>5</sup>	5.0174e <sup>4</sup>	1.2484e <sup>4</sup>	247.5128	1725.0
2b	0.0029	3.3791e <sup>3</sup>	4.0901e <sup>5</sup>	10.258e <sup>4</sup>	1.4099e <sup>4</sup>	250.1311	1108.9
3	0.0051	3.7252e <sup>3</sup>	4.0826e <sup>5</sup>	5.3331e <sup>4</sup>	1.328e <sup>4</sup>	249.4564	1277.6
4a	0.0004086	2.9731e <sup>3</sup>	4.1772e <sup>5</sup>	2.8444e <sup>4</sup>	1.2859e <sup>4</sup>	249.7392	1166.3
4b	0.00271	4.8038e <sup>3</sup>	4.1732e <sup>5</sup>	9.8591e <sup>4</sup>	1.3428e <sup>4</sup>	247.5370	1721.4
6a	0.0050	3.6992e <sup>3</sup>	4.0949e <sup>5</sup>	5.4134e <sup>4</sup>	1.3358e <sup>4</sup>	248.5588	1436.8
6b	0.0031	3.4182e <sup>3</sup>	4.2325e <sup>5</sup>	9.4390e <sup>4</sup>	1.2913e <sup>4</sup>	250.2070	1099.4

Table-2 depicts the energy features like total energy and energy of LL adding orion as a special feature along with mean and variance as textural features.

**Table-3.** Input vectors for LVQ after normalization.

Image# of subject	Total energy		Low pass energy	Orion		Mean	Variance	Classification
	Level 1	Level 2	Level 2	Level 1	Level 2			
1a	0.0031	0.699	0.77	0.85	1.0000	0.9891	0.9846	Malignant
1b	0.0032	0.7219	1.00	0.89	0.5492	0.9992	0.6326	Malignant
2a	0.0054	0.78	0.9596	0.49	0.579	0.9892	1.000	Benign
2b	0.0029	0.702	0.9299	1.00	0.654	0.9997	0.6428	Benign
3	0.0051	0.7755	0.9283	0.52	0.6158	0.997	0.7406	Benign
4a	0.000141	0.6189	0.9498	0.277	0.596	0.997	0.6761	Malignant
4b	0.0027	1.000	0.9489	0.9611	0.6223	0.989	0.998	Malignant
6a	0.005	0.77	0.939	0.53	0.6195	0.9934	0.833	Malignant
6b	0.0031	0.7115	0.9623	0.92	0.5986	1.0000	0.6373	Malignant

Considering Table-2 and normalizing the feature vectors Table-3 is derived. Total Energy of Level1, Level 2, Low Pass Energy of Level 2, Orion of Level 1 and Level 2, Mean and Variance, were considered as feature vectors.

### Classification

#### Training algorithm

The objective of the algorithm for the LVQ net is to ascertain the output unit having a matching pattern with the input vector. As the process ends, if both 'x' and 'w' belong to a similar class, weights are displaced toward the new input vector and if they are found to belong to a divergent class, the weights are moved off from the input vector. In this study the winner unit is established as done to Kohonen self-organizing feature map. In comparison of winner unit with the target, the updating of weight is performed. This is explained in the algorithm outlined below. The iterations are repeated by scaling down the learning rate.

The algorithm is as follows:

- Step 1:** Initialization of weight (reference) vectors.  
Learning rate is initialized.
- Step 2:** If STOP is FALSE perform Steps 3-7
- Step 3:** Repeat Steps4-5 for every training input vector x.
- Step 4:** Using squared Euclidean distance Compute J

$$D(j) = \sum (w_{ij} - x_i)^2$$

**Step 5:**  $W_j$  is updated as

If  $t=c_j$ , then

$$w_{J(new)} = w_{J(old)} + \alpha[x - w_J(old)]$$

If  $t \neq c_j$  then

$$w_{J(new)} = w_{J(old)} - \alpha[x - w_J(old)]$$

**Step 6:** Lower the learning rate.

**Step 7:** Check for STOP.

The STOP condition may be a fixed number of iterations or the learning rate reaching a sufficiently small value [26].

On training the network the weight matrix is obtained as follows:

```
weight Matrix after 100 epochs
1.0e+03 *
```

```
0.0000    0.0000
0.0009    0.0012
0.0009    0.0008
0.0009   -0.0021
4.9535    0.0006
0.0010    0.0009
0.0010    0.0039
```

```
>> LVQ_TEST_GGO_DEMO
initial weight matrix
1.0e+03 *
```

```
0.0000    0.0000
0.0009    0.0012
0.0009    0.0008
0.0009   -0.0021
4.9535    0.0006
0.0010    0.0009
0.0010    0.0039
```

```
The class is
2
```

```
The class is
2
```

Seven input features and two clusters are present to classify. After training the given two input vectors which were malignant and benign have been classified as benign and benign. The malignant is classified as benign due to availability of smaller number of images on training the LVQ.

#### Input images

Input images of various subjects are shown in the following figures from Figure-2a to Figure-6b.

Figures 2a and 2b are from the same subjects with different slices one above the other.



**Figure-2(a).** Wedge shaped partly solid GGO noted in the Right middle lobe (mixed) of subject#1.



**Figure-2(b).** Wedge shaped partly solid GGO noted in the Right middle lobe (mixed) of the subject #1.

Figures 3(a) and 3(b) are from same subjects with different slices one above the other.

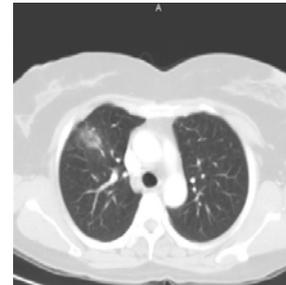


**Figure-3(a).** Partly filled liquid GGO in the left apico posterior segment of subject#2.



**Figure-3(b).** Partly filled liquid GGO in the left apico posterior segment of subject#2.

Figure shows the right upper lobe anterior segment with partly fluid filled GGO of subject#3

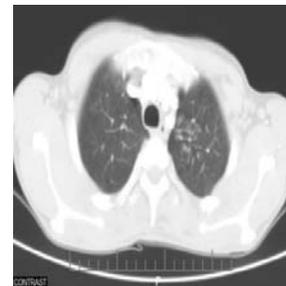


**Figure-4.** Right upper lobe anterior segment shows partly fluid filled GGO of subject#3.

Figures 5(a) and 5(b) are from the same subjects with different slices one above the other.



**Figure-5(a).** Left apical segment solid GGO (mixed) of subject#4.

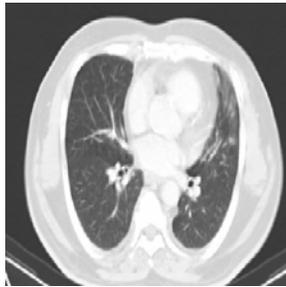


**Figure 5(b).** Left apical segment solid GGO (mixed) of subject#4.

Solid GGO is represented in Figure-6.



**Figure-6(a).** Left Lingula solid GGO (mixed) of subject#5.



**Figure-6(b).** Left Lingula solid GGO (mixed) of subject#5.

**Model image**

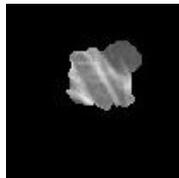
Figure-7 illustrates the input image.



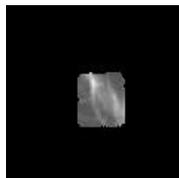
**Figure-7.** Model image.

**Boundary optimized active contour of input images**

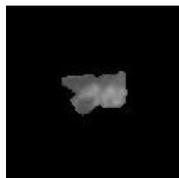
Boundary optimized active contour of Input Images of various subjects are shown in the following figures from Figure 8(a) to Figure-12(b).



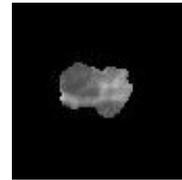
**Figure-8(a).** Boundary optimized active contour of subject#1.



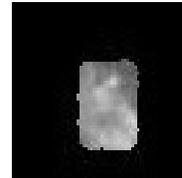
**Figure-8(b).** Boundary optimized active contour of subject#1.



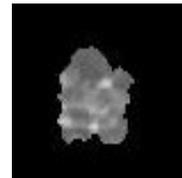
**Figure-9(a).** Boundary optimized active contour of subject#2.



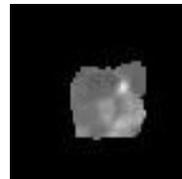
**Figure-9(b).** Boundary optimized active contour of subject#2.



**Figure-10(a).** Boundary optimized active contour of subject#3.



**Figure-10(b).** Boundary optimized active contour of subject#4.



**Figure-11.** Boundary optimized active contour of subject#4.



**Figure-12(a).** Boundary optimized active contour of subject#5.

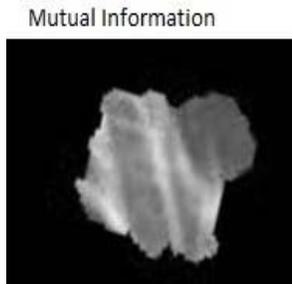


**Figure-12(b).** Boundary optimized active contour of subject#5.

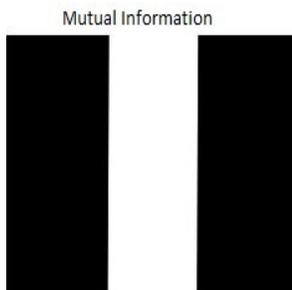


### Image after optimization contained in the moving register

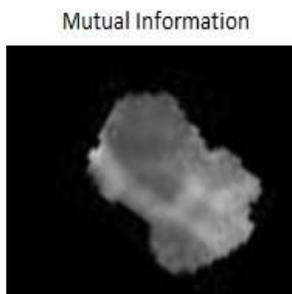
Images after optimization contained in the moving register of subject #1 to subject #5 are shown in figures from Figure-13(a) to Figure-17(b).



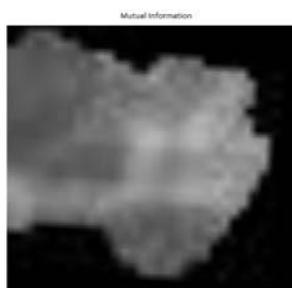
**Figure-13(a).** Image after optimization contained in the moving register of Subject#1.



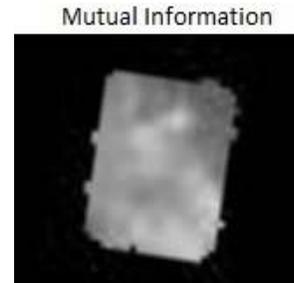
**Figure-13(b).** Image after optimization contained in the moving register of subject#1.



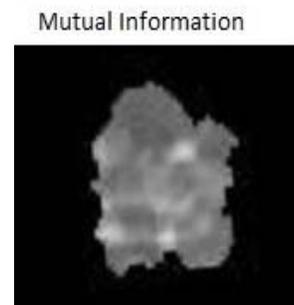
**Figure-14(a).** Image after optimization contained in the moving register of subject#2.



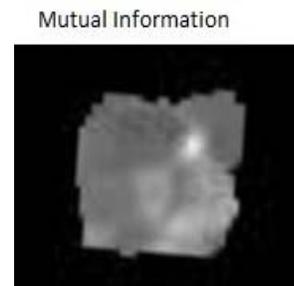
**Figure-14(b).** Image after optimization contained in the moving register of subject#2.



**Figure15.** Image after optimization contained in the moving register of Subject#3.



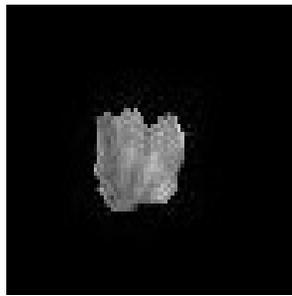
**Figure-16(a).** Image after optimization contained in the moving register of subject#4.



**Figure-16(b).** Image after optimization contained in the moving register of subject#4.



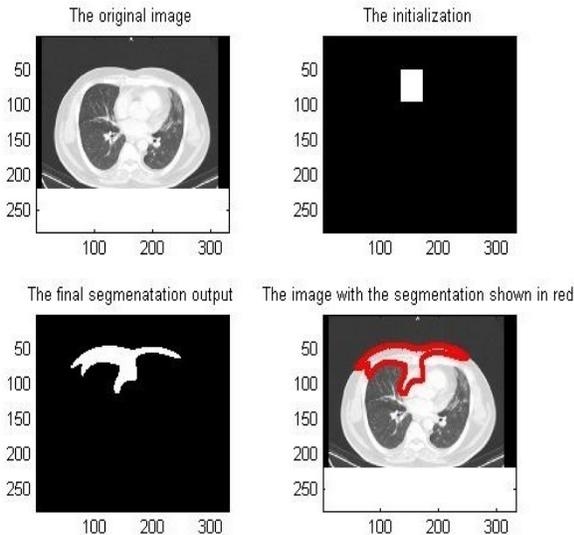
**Figure-17(a).** Image after optimization contained in the moving register of subject#5.



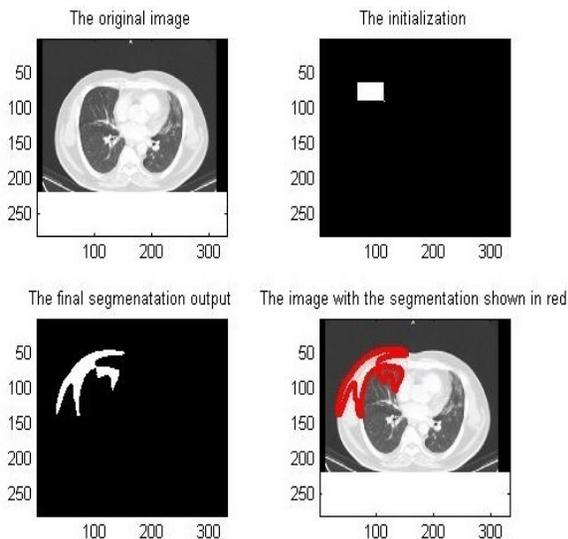
**Figure-17(b).** Image after optimization contained in the moving register of subject#5.

**Results for Chanveese model**

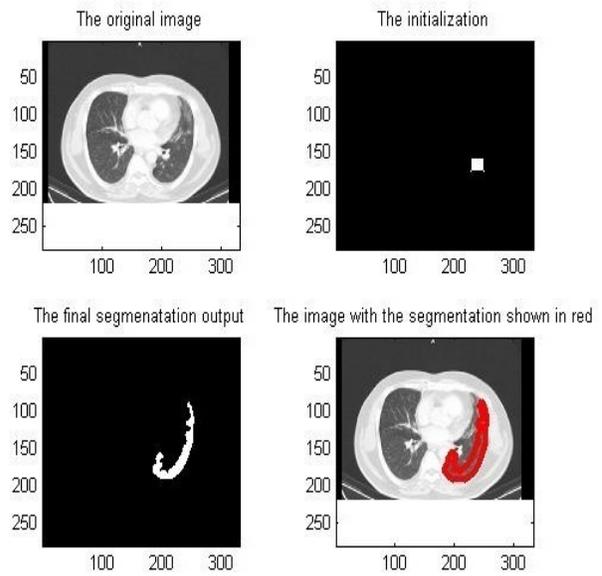
Results for Chanveese model for subjects #1 to #5 are shown from Figures 18(a) to Figure-22(b).



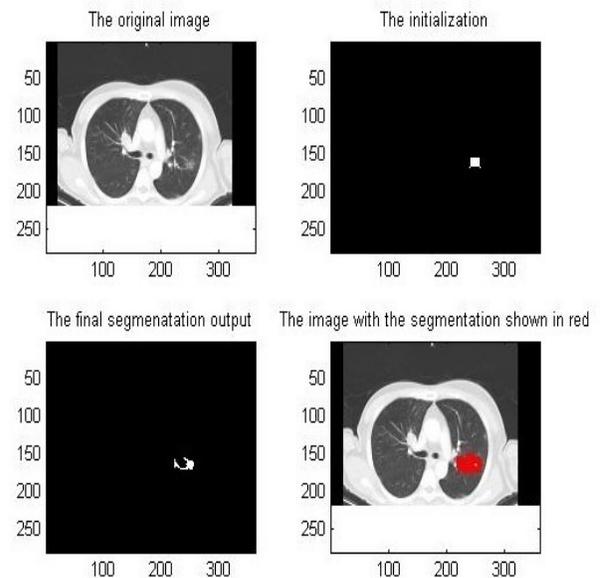
**Figure-18(a).** Results for Chanveese model for subject#1.



**Figure-18(b).** Results for Chanveese model for subject#1.



**Figure-19(a).** Results for Chanveese model for subject#2.



**Figure-19(b).** Results for Chanveese model for subject#2.

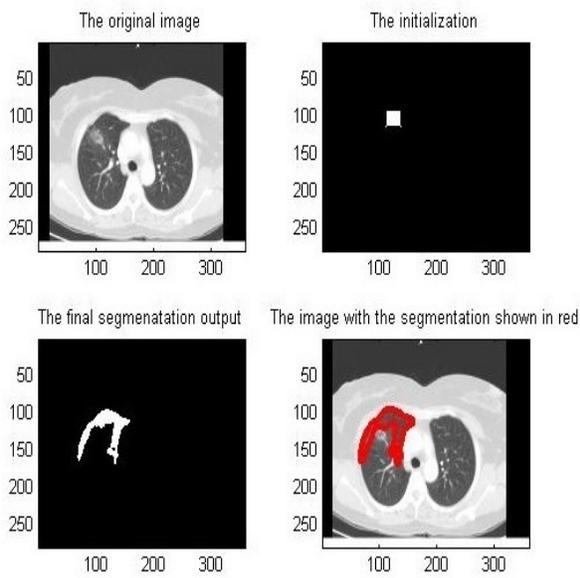


Figure-20. Results for Chanveese model for subject#3.

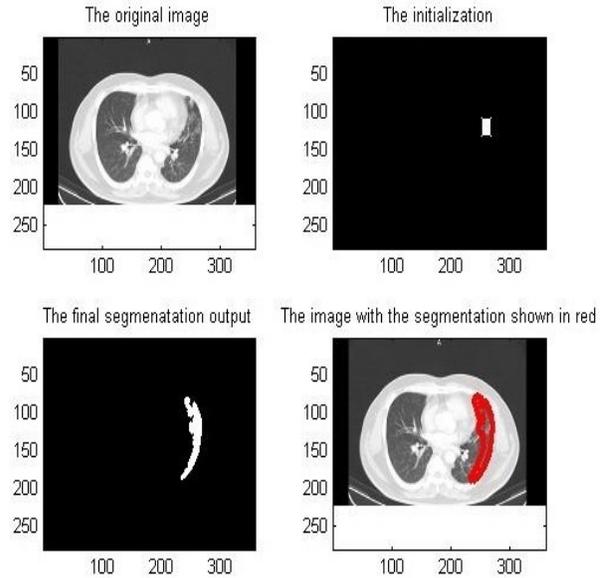


Figure-21(b). Results for Chanveese model for subject#4.

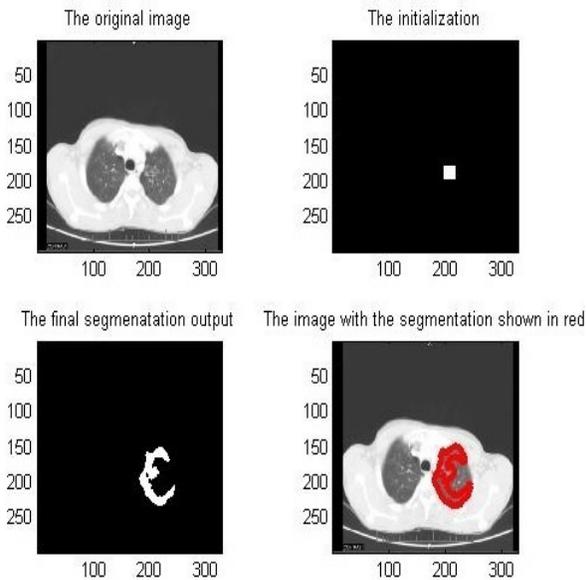


Figure-21(a). Results for Chanveese model for subject#4.

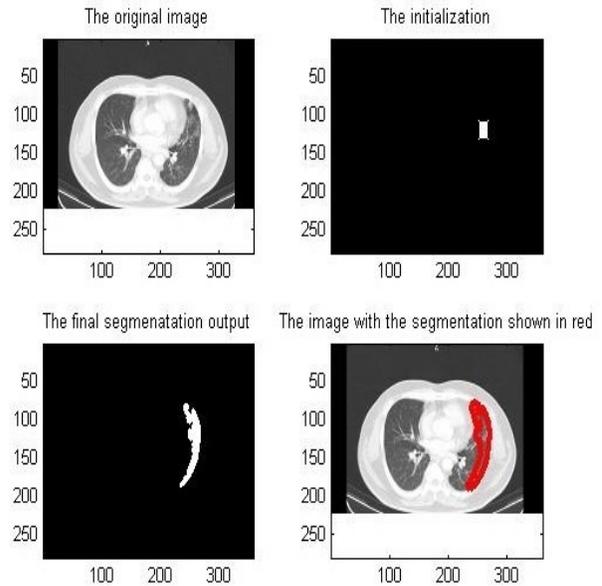


Figure-22(a). Results for Chanveese model for subject#5.

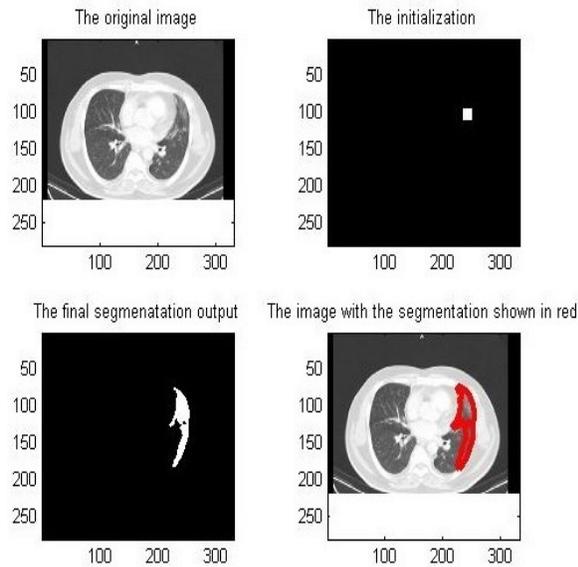


Figure-22(b). Results for Chanveese model for subject#5.

## DISCUSSIONS

The table below shows that the DRLSE method gives correct matching with model image and has maximum mutual information of the growing active contour.

Table-4. Comparison of active contour without edges with DRLSE Method.

Subject #	Active contour without edges	DRLSE
1	Non matching	Matching
1	Non matching	Non matching
2	Non matching	Matching
2	Matching	Matching
3	Non matching	Matching
4	Non matching	Matching
4	Matching	Matching
5	Non matching	Matching
5	Non matching	Matching

## Future work

As per the experience of the Radiologist the nodules containing a ground glass component are more likely to be malignant. Partly solid lesions with ground glass components had a malignancy rate of 63%. Non-solid only ground glass lesions had a malignancy rate of 18%. Only solid lesions had a malignancy rate of 7% [24]. Future work is proposed to be carried out on extracting features and classifying GGO as malignant and benign. The classification made by the Neural Network can be back tracked to explain how it has decided a nodule to be Malignant and Benign. This work will be a notable work for research.

## CONCLUSIONS

The contour was grown by Distance Regularized Level Set Evolution method and compared with model image. The contour giving maximum Mutual information will be the optimized boundary of the GGO. The same methodology is followed in Active Contour without edges also. On comparing the two methods, DRLSE method was found to be giving better results. The wavelet transform have been taken on the input images to extract Low pass image energy at level2, Total Energy as the sum of individual energies (for each direction) of detailed images (Vertical, Horizontal and Diagonal) at level1 and level2, Oriant at level1 and level2. These features along with Mean and Variance have been fed to the LVQ for training and testing. On testing the network one Malignant and one benign nodule was given as input and the network classified as Benign and Benign. The Malignant have been classified as benign is due to inadequate samples to train the LVQ.

## List of abbreviations

GGO	Ground Glass Opacity
LVQ	Learning Vector Quantization
DRLSE	Distance Regularized Level Set Evolution
SPN	Solitary Pulmonary Nodule
CT	Computer Tomography
LDCT	Low-Dose Computed Tomography
ACM	Active Contour Models
LSF	Level Set Function
PDF	Partial Differential Equation
SDF	Signed Distance Function

## ACKNOWLEDGEMENTS

a) We thank Prof. Dr. N. Kumaravel, Emeritus Professor, College of Engineering, Guindy, Anna University, Chennai, India, for his guidance in making use of Mattes Mutual Information as a criterion for comparing the grown Active Contour with Model Image developed by Radiologist.

b) We thank Dr. P.S. Panneerselvam, Medical Director, Shanmuga Hospitals and Salem Cancer Institute, for his support and guidance to optimize the ill-defined boundary of Ground Glass Opacity of Solitary Pulmonary Nodule.

## REFERENCES

- [1] Malik PS, Vinod Raina. 2015. Lung Cancer: Prevalent Trends and Emerging Concepts. Indian Journal of Medical Research. 141(1): 5-7.
- [2] Armstrong P. 2007. Computed Tomography of Solitary Pulmonary Nodules. Diseases of the Heart Chest & Breast. pp. 11-16.
- [3] Diciotti S, Picozzi G, Falchini M. 2008. 3-D segmentation algorithm of small lung nodules in



- spiral CT images. *IEEE Transactions on Information Technology on Biomedicine*. 12(1): 7-19.
- [4] Kubota T, Jerebko AK, Dewan M, Salganicoff M, Krishnan A. 2011. Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. *Medical Image Analysis*. 15: 133-154.
- [5] Sluimer I, Schilham A, Prokop M, Ginneken BV. 2006. Computer analysis of computed tomography scans of the lung: a survey. *IEEE Transactions on Medical Imaging*. 25(4): 385-405.
- [6] Chan T, Vese L. 2001. Active contours without edges. *IEEE Transactions on Image Processing*. 10(2): 266-277.
- [7] Li C, Kao CY, Gore JC, Ding Z. 2008. Minimization of Region-Scalable Fitting Energy for Image Segmentation; *IEEE Transactions on Image Processing*. 17(10): 1940-1949.
- [8] Bin Li, Chen K, Guangming P, Yuanxing G, Tian L, Shanxing O, Wang L. 2016. Segmentation of ground glass opacity pulmonary nodules using an integrated active contour model with wavelet energy-based adaptive local energy and posterior probability-based speed function. *Materials Express (American Scientific Publishers)*. 6(4): 317-327.
- [9] Pedersen JH, Saghir Z, Wille MMW, Thomsen LH, Skov BG, Ashraf H. 2016. Ground Glass Opacity Lung Nodules in the Era of Lung Cancer CT screening: Radiology, Pathology, and clinical management, *Oncology Journal*. 30: 266-274.
- [10] Tony FC, Vese LA. 2001. Active Contours without Edges. *IEEE Transactions on Image Processing*. 10(2): 266-277.
- [11] Li C, Xu C, Gui C, Martin D. Fox. 2010. Distance Regularized Level set evolution and its application to Image Segmentation. *IEEE Transactions on Image Processing*. 19(12): 3243-3254.
- [12] Kass M, Witkin A, Terzopoulos D. 1988. Snakes Active Contours Models. *Int. Journal of Comp Vision*. 1: 321-331.
- [13] Caselles V, Catté F, Coll T. 1993. A geometric model for active contours in image processing. *Numer. Math*. 66: 1. doi: 10.1007/BF01385685.
- [14] Malladi R, Sethian JA, Vemuri BC. 1993. A topology independent shape modeling scheme. *Geometric Methods in Computer Vision II*. 2031: 256-258.
- [15] Caselles V, Kimmel R, Sapiro G. 1997. On Geodesic Active Contours. *International Journal of Computer Vision*. 22(1): 61-79.
- [16] Malladi R, Sethian JA, Vemuri BC. 1994. Evolutionary fronts for topology-independent shape modeling and recovery. *Proceedings of the third European conference on Computer vision*. 1: 3-13.
- [17] Osher S, Sethian JA. 1988. Fronts Propagating with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations. *Journal of Computational Physics*. 79: 12-49.
- [18] Dervieux A, Thomasset F. 1980. A finite element method for the simulation of Rauleigh-Taylor instability. In: Reimund Rautmann, editors. *Approximation Methods for Navier-Stokes Problems*. Germany: pp. 145-158.
- [19] Dervieux A, Thomasset F. 1980. Multifluid incompressible flows by a finite element method. *Proceedings of the Conference, Stanford University, Stanford, California and NASA/Ames (U.S.A.)*. 141: 158-163.
- [20] Caselles V, Catta F, Coll T, Dibos F. 1993. A geometric model for active contours in image processing. *Numer Math*. 66: 1-31.
- [21] Xu C, Yezzi A, Prince J. 2000. On the Relationship between Parametric and Geometric Active Contours. *Proc. of 34th Asilomar Conference on Signals, Systems, and Computers*. 483-489.
- [22] Mattes D, Haynor DR, Vesselle H, Lewellen T, Eubank W. 2001. Non-rigid multimodality image registration, *Medical Imaging 2001, Image Processing*, SPIE publications. 1609: doi:10.1117/12.431046.
- [23] www.mathworks.com.
- [24] Agnieszka C, Katarzyna JM. 2012. Evaluation of solitary pulmonary nodule detected during computed tomography examination, *Polish Journal of Radiology*. 77(2): 22-34.
- [25] Stefan L, Paul S, Gert VW, Dirk VD, Hilde S, Johan WM and Walter B. 1996. A texture analysis approach



to corrosion image classification. Microscopy, Micro analysis, Micro structures. 7(2): 1-10.

- [26] Sivanandam SN, Sumathi S, Deepa SN. 2006. Introduction to Neural Networks using Matlab 6.0, The Mc Graw Hill Companies.