Journal of Engineering Science and Technology
Vol. 14, No. 4 (2019) 2098 - 2115

OBJECT TRACKING BASED 3D MODELLING AND QUANTIFICATION OF ABNORMAL CONTOURS IN BRAIN MRI DICOM STUDY

K. SURESH*, U. SAKTHI

Sathyabama Institute of Science and Technology,
Department of Information Technology, St. Joseph's College of Engineering,
Rajiv Gandhi Road, Chennai-600119, India
*Corresponding Author: sureshk@stjosephs.ac.in

Abstract

DICOM (Digital Imaging and Communications in Medicine) is used in the detection and volumetric analysis of tumour specimens and hence, plays a vital role in medical diagnosis. In this proposed work, non-rigid tumour regions are tracked based on an adaptive weighted scale of the mean shift framework and moment features of CAMShift. Bhattacharyya coefficient metric is used to ensure an optimal fit of target candidate over the target model (tumour), thus, minimizing background interference or pre-processing. The candidate coordinates are determined using the moment features, which enables the segmentation of the tumour region in T2-weighted MR slices. The volume of the tumour region is estimated based on the physical spacing information available in the DICOM grid, which is compared to the volume calculated using frustum model. Similar measurements are done with a re-sampled grid that is developed using associations in spatial directionality. The experiments and their subsequent analysis prove that, in comparison to DICOM grid, the re-sampled grid provides better accuracy and offers a smoother surface (of the segmented tumour region) upon 3D construction of the tumour.

Keywords: 3D construction, CAMShift, DICOM model, Mean shift, Medical image segmentation, Moment features, Object tracking.

1. Introduction

Brain tumour detection and volumetric quantification of abnormalities are important in medical diagnosis and further tracking. The MRI series of a patient is signified as a high-dimensional feature space that includes multimodal intensity features, as well as spatial features [1]. Each MR image is organized as a grid of data points to visualize brain anatomy. Every unit of an MR image corresponds to a voxel, a volume element, whose values represent various tissue classes such as WM, GM, CSF, normal tissues and solid tumour's respectively [2].

The volume of a voxel depends on the MR scan parameters, i.e., slice thickness and pixel spacing. MR images are usually delivered in DICOM format. Besides the MR image modality, DICOM-files contain information about the MR scan and patients' details [3]. DICOM facilitates to create private tags, which provides a better volume of metadata compared to other formats. Normally an MR scan acquires more than one slice, which leads to an image sequence M×N×K with K slices. These slices are then separated and worked upon separately to get a better estimation of the tissue classes present in it.

MR scan acquires a range of intensities related to different tissue classes in the DICOM slices. Tracking of abnormal objects present in the series is vital to understand its deepness and reaction to treatment. Kaushal et al. [4] explained that object tracking is a critical task in many fields of study, and many algorithms have been proposed to overcome the difficulties arising from noise, occlusions, clutters, and changes in the foreground object and/or background environment. Classification approaches like Convolutional Neural Networks (CNN) and R-CNN have better accuracy but result in a computational blow up while training the brain slices.

Real-time object tracking requires two major components, target representation and localization. In object tracking, target model and target candidate models are represented by their probability density functions in the quantized feature space. The histogram-based models are normalized with an isotropic kernel to reduce target localization. A correlation score between the target model and the candidate model is used as a similarity measure [5]. Measures such as Jensen-Shannon (JS) divergence, Point-wise Mutual Information (PMI), Bhattacharyya Coefficient (BC) can be used for similarity analysis. According to Kailath [6], a metric derived from BC is typically used to ensure the similarity between target and the candidates in the next frame. The motion of a target in the successive frames can undergo a scale change, thus, task-specific information combined with tracking leads to reliable performance.

Among various tracking methods, the mean shift-tracking algorithm is a popular one due to its simplicity and efficiency. The mean shift algorithm was originally developed for data analysis and later introduced to the field of computer vision. Mean shift uses static distributions and iterates towards the maximum likelihood of the target intensities. The distributions are not recomputed for every frame [5, 7]. The main disadvantage of a mean shift algorithm was that it had a fixed target window size and hence, was not suitable for objects whose size varied continuously.

The mean shift algorithm was later modified into Continuously Adaptive Mean Shift (CAMShift) algorithm for the purpose of adaptive window tracking. CAMShift uses continuous adaptive probability distributions and uses spatial moments to iterate towards the target distribution [8]. The centroid coordinates, height and width of the candidate are computed using image moments. CAMShift holds the advantage to deal

with partial occlusion and target deformation [9]. The probability estimation of whether a pixel belongs to the specific target is obtained using the weighted histogram of the target model and the isotropic kernel profile. CAMShift with weighted histogram and quantized features show greater performance than the mean shift procedure, in equivalent colour spaces where no assumptions were made about the target. The weight values did not change with respect to the window size and it uses the Eigenvalues directly as its width and height of the object. Ultimately, the assumption proved to be disadvantageous to CAMSHIFT as the weight values change and the Eigenvalues were proportional to the width and height.

In tracking with mean shift and CAMShift techniques, the target position is found using a weight image. In CAMShift tracking, the weight image is constant and depends only on the target model but in a mean shift, the weight image depends on both the target and candidate model. The weight values in mean shift change dynamically when the scale of the target changes [10]. CAMShift does not reflect the scale change of the target with respect to the weight image, it is not accurate to estimate the target location. Hence, a scaled adaptive mean shift tracking will be more reliable than CAMShift algorithm [11].

The abnormal tissue classes in the MR slices possess a steady scale change from the initial slice until the series end. An adaptive target candidate is required to track the scale change in the target model. By integrating the CAMShift adaptive distributions and spatial moment features along with dynamic weight image of the mean shift, the target can be effectively tracked even when there is a change in the scale and orientation of the target.

Based on studies by Suresh and Sakthi [12], segmentation is an obligatory step in image analysis: it is a process of separating an image into different regions or blocks that share common and identical properties, such as colour, texture, contrast, brightness, boundaries and grey level. Brain tumour segmentation involves the process of separating the tumour tissues, such as edema, or solid tumour's from normal brain tissues with the help of MR images or other imaging modalities [13]. Extraction of primary object tumour within the target candidate window employs spatial constraints for the targeted segmentation. It returns a single semantically defined region of interest, which is different from unconstrained segmentation that usually returns homogeneous regions [14]. Many segmentation techniques are used by radiographers to discover anomalies in the MR scan images [15-19].

The inter-slice resolution of MR scan images is coarse in which, the slice thickness attribute is significantly larger than the in-plane pixel size. Subsequently, the resolution in the Z direction is significantly lower than the resolutions in the X and Y directions. For a generic MR slice, the spacing in the Z direction is 1-15 mm and for directions X and Y, the pixel spacing is around 0.5-2 mm [20]. This asymmetry in the resolution causes problems such as step-shaped iso-surfaces and discontinuity in structures in 3D reconstructed models.

Image interpolation methods are widely used in MR images for slice generation. Resampling is necessary to improve image quality, by a geometrical alignment of discrete image intensities in the spatial domain. Generation of 3D visualization and volumetric quantification of tissues in DICOM study involves slices to be arranged in a cubical stack. For reconstructing the cubical stack, various interpolation methods such as truncated and windowed since, nearest neighbour, linear, quadratic, cubic

methods are used in medical image processing. High-resolution cubic spline gives a better smoothing response than other interpolation techniques [21].

The cubical stack provides a volumetric estimation of anatomical brain structures. For the follow-up assessment of tumour growth, volume-based measurements are reliable to evaluate the tumour growth in all directions. Langenberg et al. [22] compared volumetric and two-dimensional measurements on MR images containing vestibular schwannoma. With suitable voxel properties available in the DICOM header, the volume of the tumour could be easily estimated.

2. Proposed Methodology

A patient's DICOM study containing 22 axial T2 weighted MR slices is taken as an input dataset. Initially, the target model is assigned with a set of tumour intensities, which is tracked by a target candidate in the entire DICOM series. The error estimate Bhattacharyya Co-efficient is computed based on the similarity between the target model and the candidate model. The coordinates of the candidate window, which holds tumour intensities are obtained and bi-level clustering is performed using K-means, within the tracked candidate co-ordinates to extract tumour contours. The acquired tumour contours are quantified based on geometrical properties. Finally, 3D visualization and volumetric measurement of the tumour is obtained based on the physical spacing information embedded in the DICOM image slices and the resampled stack. The complete processes are shown in Fig. 1.

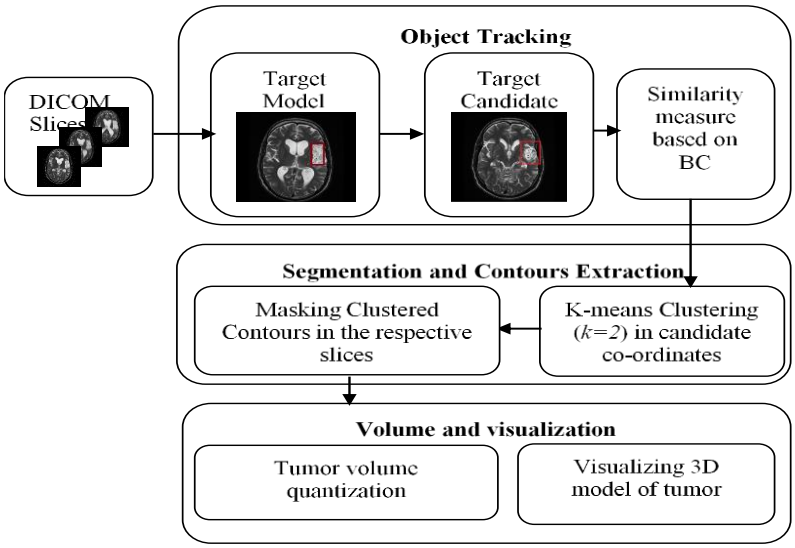

Fig. 1. Proposed architecture.

2.1. Target model and target candidate representation

The target selection is made as a rectangular section in the slice where the tumour intensity is evident. The widely used target representation is the colour histogram-based PDF, which satisfies low computational cost and robustness towards partial occlusions [5]. For sequential processing, discrete pixel intensities of the target

model in the feature space are represented by *m*-bin histograms. Thus, the quantized colour space of the target model is given by:

$$\hat{q} = \{\hat{q}_u\}_{u=1..m} ; \sum_{u=1}^{m} \hat{q}_u = 1 \tag{1}$$

With this target model, target candidates are obtained by the sequential access of the DICOM slices. The target candidate is given by:

$$\hat{p}(y) = \{\hat{p}_u(y)\}_{u=1..m} ; \quad \sum_{u=1}^{m} \hat{p}_u = 1 \tag{2}$$

Let $\{x_i^*\}_{i=1..n}$ be the normalized pixel locations of the target model centred at origin 0 having *n* pixels, *k*(*x*) be the isotropic kernel, *b*(*x**) associates index of the histogram bin to the pixel location x_i^*, in the quantized feature space. The probability of the element $u = \{1 \ldots m\}$ in the target model is then computed as follows:

$$\hat{q}_u = C\sum_{i=1}^{n} k\left(\left\|x_i^*\right\|^2\right)\delta\left[b\left(x_i^*\right) - u\right] \tag{3}$$

where δ the Kronecker delta function and the normalization constant '*C*' is derived by:

$$C = \frac{1}{\sum_{i=1}^{n} k\left(\left\|x_i^*\right\|^2\right)} \qquad ; \sum_{u=1}^{m} \hat{q}_u = 1 \tag{4}$$

Similarly, the probability of the element $u = \{1 \ldots m\}$ in the target candidate at centre position *y* is given by:

$$\hat{p}_u = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)\delta\left[b\left(x_i^*\right) - u\right] \tag{5}$$

where $\{x_i^*\}_{i=1..n_h}$ is the normalized pixel locations of the target candidate, centred at position ***y*** with bandwidth *h* in the current slice and $b(x_i)$ is the colour bin index $u = \{1 \ldots m\}$ of pixel x_i. The normalization process and the kernel profile *k*(*x*) are inherited from the target model. The normalization factor for the candidate model is given by:

$$C_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)} \tag{6}$$

The constant C_h is independent of the centre position ***y*** and is thus, pre-calculated for the given *k*(*x*) and different *h* values [6]. The target model differs in scale and orientation and the corresponding target candidate window is mapped based on the target tumour size. The bandwidth *h* defines the size (number of pixels) of the target candidate.

2.2. Mean shift and similarity metric for target localization

The target model has a similar PDF in the subsequent slice at any location *y*. The estimation of the target location is associated with the target model and candidate distributions. The Bhattacharyya Coefficient (BC), a similarity metric, which maximizes the Bayes error, is represented as:

$$\hat{\rho}(y) = \rho[\hat{p}(y), \hat{q}] \tag{7}$$

where the metric $\hat{\rho}(y)$ calculates the likelihood of the target model $\hat{q}$ and the candidate model $\hat{p}(y)$.

The Bhattacharyya Coefficient's geometric interpretation is based on the cosine of the angle between the candidate unit vector $\left(\sqrt{\hat{p}_1} \ldots \sqrt{\hat{p}_m}\right)$ and the target unit vector $\left(\sqrt{\hat{q}_1} \ldots \sqrt{\hat{q}_m}\right)$ of the dimensional vector space.

$$\hat{\rho}(y) = cos\theta_y = \sum_{u=1}^{m} \hat{p}_u(y)\hat{q}_u \tag{8}$$

In the process of tracking the target, the slice, which has the maximum range of tumour intensities, is selected. The target model and its initial position y_0 are defined in the selected slice. With that, model location y_0, the tumour is tracked from the first slice sequentially.

The Bhattacharyya Coefficient with linear approximation around y_0 is defined by:

$$BC = \rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2}\sum_{u=1}^{m} \sqrt{\hat{p}_u(y_0)\hat{q}_u} + \frac{C_h}{2}\sum_{i=1}^{n_h} w_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \tag{9}$$

In Eq. (9), the first term is independent of candidate location *y* centred at (x_c, y_c) and the second term is maximized to minimize the distance, which estimates the density of candidate location *y* and it is given by:

$$w_i = \sum_{u=1}^{m} \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}}\delta[b(x_i) - u] \tag{10}$$

Ning et al. [11] proposed using a mean shift density estimation procedure, the new position of the target y_1.

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\left\|\frac{y - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} x_i w_i g\left(\left\|\frac{y - x_i}{h}\right\|^2\right)} \tag{11}$$

Since w_i trails multivariate density estimation, we choose Epanechnikov profile for kernel $k(x)$. Therefore, $g(x) = -k(x) = 1$, and Eq. (11) becomes:

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \tag{12}$$

The target candidates are tracked based on mean shift tracking process from the beginning to the end of the DICOM series. The scale of the target candidate depends on w_i to estimate its size and co-ordinates in the image plane.

2.3. Find coordinates of target candidate

The target model has to be tracked using the candidate model over the entire DICOM series. Scale parameter of the target plays a key role in finding the candidate coordinates. Even if the scale of the target model changes in consecutive frames (based on in-stack position attribute), the mean shift merely estimates the position of the target. It fails to adaptively estimate the scale change in the subsequent slices. The target candidate position y_1 is marginally adjusted to encompass the target effectively. The problem is to find the coordinates of the target candidate, which have maximum BC and minimum error.

Ning et al. [11] mentioned the size of the target candidate is as follows:

$$size(\hat{p}(y)) = w_i \hat{q} + w_i \in \tag{13}$$

where $w_i \hat{q}$ gives weighted pixels of the target tumour and $w_i \varepsilon$ gives the weighted pixels, which do not belong to the target (background pixels) within the target candidate $\hat{p}(y)$. In order to reduce error ε, BC has to be maximized.

Equation (12) can be rewritten based on target and background pixel intensities in the candidate window as follows:

$$CI(x,y)=TI(x,y)+BI(x,y) \tag{14}$$

To estimate the target candidate area until BC convergence, calculate the *zero*th-order moment for the target candidate.

$$M_{00} = \sum_x \sum_y CI(x,y) \tag{15}$$

The tracked target is generally enclosed in a bigger target candidate region. The probability of target intensities must be maximized, resulting in a minimized probability of background. Therefore, the weight coefficient w_i in BC, Eq. (9) is employed to increase the weight of the target, which in turn suppresses the background. For this purpose, BC($0 \leq \rho \leq 1$) referred in Eq. (8) is used as a similarity measure by associating target with the target candidate. A maximum BC gives intensities of M_{00} more target features than the background. Therefore, BC is used as a metric for taking M_{00} as the target coordinates.

The target candidate area is estimated as:

$$TCA = c(\rho) M_{00} \tag{16}$$

where $c(\rho)$ is the correction function used to adjust M_{00} close to the real target scale. The candidate function $c(\rho)$ follows the exponential distribution and it is given by:

$$c(\rho) = \exp\left(\frac{\rho - 1}{\sigma}\right) \qquad (0 \le \rho \le 1) \tag{17}$$

When $\rho \approx 1$, $c(\rho)$ approaches 1, such that, the target is completely enclosed within the target candidate window with minimal background features. If $c(\rho)$ decreases, M_{00} is considerably bigger than the real target. When $\rho \approx 0$, $c(\rho)$ returns a minimum value and the tracked tumour is completely lost in the specific slice in the tracking progression.

To estimate the coordinates of final compacted M_{00} in the tumour containing slices, first and second order moments are calculated. The moments offer overall image information for analysing shapes [9].

The first order moment is given by:

$$M_{10} = \sum_x \sum_y xCI(x,y) \qquad M_{01} = \sum_x \sum_y yCI(x,y) \tag{18}$$

Using the zeroth and first-order moments, the centroid position of the candidate window is calculated as:

$$(x_c, y_c) = \left(\frac{M_{01}}{M_{00}}, \frac{M_{10}}{M_{00}}\right) \tag{19}$$

Similarly, the height and width of the candidate are measured as:

$$H = \sqrt{\frac{(a+b) + \sqrt{b^2 + (a-c)^2}}{2}} \quad W = \sqrt{\frac{(a+c) + \sqrt{b^2 + (a-c)^2}}{2}} \tag{20}$$

where

$$a = \frac{M_{20}}{M_{00}} - x_c^2 \;;\; M_{20} = \sum_x \sum_y x^2 CI(x,y)$$

$$b = 2\frac{M_{11}}{M_{00}} - 2x_c y_c \;;\; M_{11} = \sum_x \sum_y xyCI(x,y)$$

$$c = \frac{M_{02}}{M_{00}} - x_c^2 \;;\; M_{02} = \sum_x \sum_y y^2 CI(x,y)$$

Based on the estimation of previous coordinates, the candidate window size is extended to track the target in the succeeding frames. Even if the scale of the target changes, it still persists in the larger candidate region. The scale parameter Vd to enlarge the size of the target candidate is given by:

$$H = H + \nabla d \quad , W = W + \nabla d \tag{21}$$

The algorithmic description of object tracking procedure is depicted in Fig. 2.

Algorithm

Initialize the target model $\hat{q}_u$ in appropriate slice and initialize the position y_0 .

1. Track the target in the successive slice using y_1 with a larger candidate window $\hat{p}(y)$.
2. Calculate zeroth order moment (M_{00}) to estimate candidate area.
3. Compare $\hat{p}(y)$ and $\hat{q}$using similarity coefficient $\rho[\hat{p}(y),\hat{q}]$ in the tracking process of the DICOM series:

   if ($\rho \approx 1$)

   M_{00} has maximum tumor intensities than background. Goto step-5.

   else if ($\rho \approx 0$)

   M_{00} has minimum tumor intensities than background. Reduce candidate window size. Goto step-3.

   else

   Tracked target lost.
4. Compute first and second order moments for finding centroid, height and width of the target candidate.
5. Estimate the position of target candidate in the next slice and trace the candidate coordinates in the consecutive slices.

Fig. 2. Algorithm for tracking tumour intensities in DICOM study.

2.4. Clustering based on tumour coordinates

In order to extract the anatomical structures of various tissue classes in the brain MR slices, segmentation methods are used. Clustering is an unsupervised way of segmentation, which partitions a set of intensities of an image into disjoint clusters. The coordinates of the target candidate window obtained from the above process act as a bounding box to limit clustering within the box boundaries. As the bounding box comprises tumour and background voxels, traditional *k*-means with $k = 2$ is applied to separate the region of interest and background.

2.4.1. *K* means clustering

The *K*-means yields is a fixed number of clusters based on chosen *k*. Let the DICOM slice be S_i and the bounding box *B*, containing a group of pixels p_i such that, $p_i \varepsilon B$. As the number of clusters *k* is chosen as two, the intensities within the box $B = \{p_i | i = 1,2,3,\ldots,n\}$ are segmented by labelling individual voxels in the

bounding box as foreground or background. The cluster means of the two classes within the box is computed as:

$$M = \frac{\sum_{i::c(i)=2} p_i}{2} \tag{22}$$

where *c(i)* represents the cluster centroids initially chosen at random. The Euclidean distance is computed between a voxel and the cluster centroids in order to identify the nearest centroid

$$d(i,j) = \sum_{j=2}^{n} \sum_{i=1} p_i^{(j)} - c_j \tag{23}$$

The cluster centroids are re-computed until the entire tumour and background voxel intensities are converged to their nearest centroid. The clustering yields segmented tumour contours in the respective slices.

2.5. Mask clustered tumour in respective slices

Masking separates foreground, background or probable background/foreground features. The contours generated after the clustering process are masked in order to separate the clustered tumour voxels and thus, assigning the rest of the tissues as background. Contour masking retains the spatial location of clustered tumour voxels and ignores background brain tissues. Mutual information (I) is computed to ensure the similarity between the contour mask and the corresponding slice [23]. The weighted contribution (W_i) of the contour mask (CM_i) to the original slice (S_i) is calculated as:

$$W_i = \frac{1}{CM_i} e^{\frac{-I_i - I_{min}}{I_{max} - I_{min}}} \tag{24}$$

where I_i represents the mutual information between CM_i and S_i I_{max}, I_{min} are the maximum and minimum mutual information for the overall CM_i.

2.6. Resampling and slice interpolation of slices

MR scan usually acquires more than one slice in a study, which forms a slice sequence [$NS \times PS_x \times PS_y$] containing a number of slices (NS) and pixel spacing along X and Y axes (PS_x, PS_y). The DICOM study taken in this work contains 22 slices of axial T2 MR images. Every voxel in the DICOM slice corresponds to a tissue value, which acts as a volume component. The slices have DICOM metadata in which, it is observed that the 'spacing between slices' is 6.5 mm, 'Pixel Spacing (PS_x/PS_y)' is (0.4492/0.4492) and 'slice thickness' is 5 mm. The inter-slice resolution of these slices is coarse, such that interpolation is required to construct the in-between slices from existing slices. Every '5 mm' thick slice comprises five '1 mm' thin slices. This is done by calculating the resize factor of the slices using slice thickness and pixel spacing attribute in the header of DICOM.

The obtained contour mask from the tumour containing slices contributes to tumour voxels in the original slices stack [22×512×512]. This stack, which holds enhanced tumour voxels, is resampled to [110×230×230] slices stack by cubic spline interpolation function. As the original stack migrates to resampled stack, number of slices are contributed for modelling the tumour. Cubic Spline interpolation is a piecewise third order polynomial, used to interpolate every pair of voxels in the adjacent slices. Cubic spline interpolation fits the intermediate voxels based on a cubic polynomial curve that traverses between the real adjacent slices. Let (x_i, y_i, z_i) be the voxel coordinates.

The control points are given by:

$$P_i = (x_i, y_i, z_i),\ k = 0,1,2,...,n \tag{25}$$

The following equations fit the intermediate voxels in all three planes

$$\begin{aligned} x(u) &= a_x X^3 + b_x X^2 + c_x X + d_x \\ y(u) &= a_y X^3 + b_y X^2 + c_y X + d_y \quad ,\ 0 \le X \le 1 \\ z(u) &= a_z X^3 + b_z X^2 + c_z X + d_z \end{aligned} \tag{26}$$

where *a* to *d* represent the coefficients in the polynomial to be determined for each intermediate voxel among the pair of adjacent slices. Boundary conditions are set between curve sections to attain the values for each coefficient.

2.7. Tumour quantization and 3D modelling

The volume of the region of interest is calculated using frustum model. The slices are sorted based on the in-stack position attribute of the DICOM. The whole volume of the tumour is computed by summing up the tumour volumes of the two successive slices in the study. With known top area (A_1), base area (A_2) and slice thickness (T) related with height, it is possible to calculate the volume of the frustum and it is given by:

$$\text{Volume of two consecutive slices } (V_i) = \frac{T}{3}\left(A_1 + A_2 + \sqrt{A_1 A_2}\right) \tag{27}$$

$$\text{Total volume} = \sum V_i \tag{28}$$

The area of each slice A_1 is computed by the relation,

$$A_i = Maskedvoxels * pixelspacing^2 \tag{29}$$

where the masked voxels specify the number of clustered tumour voxels contributed in the mask region. Pixel spacing **(PS_x/PS_y)** can be referred from DICOM header or the resampled values.

The slices considered are elongated after resampling and interpolation. The extended cubical stack holds the brain tissue classes and the tumour voxels. In order to visualize the tumour voxels in 3D, the set of tumour voxels from each slice are made to form a voxel cloud (mesh). This mesh is generated by merging all the three

anatomical planes of the DICOM slices. This results in an interconnected set of vertices and triangular faces.

The spatial relationship between tumour contours of one slice and its adjacent slice is defined by the DICOM header data such as image orientation, pixel spacing, slice thickness and image position. The contoured mesh is structured as:

$$\begin{bmatrix} P_x \\ P_y \\ P_z \\ 1 \end{bmatrix} = \begin{bmatrix} r_x.\lambda x & c_x.\lambda y & 0 & s_x \\ r_y.\lambda x & c_y.\lambda y & 0 & s_y \\ r_z.\lambda x & c_z.\lambda y & 0 & s_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 0 \\ 1 \end{bmatrix} \tag{30}$$

where P_x, Py, P_z represent the point coordinates in mm, the s_x, s_y, s_z values signify the image position attribute. The r_x, r_y, r_z and c_x, c_y, c_z give the direction cosine values of the image orientation attribute and the values Δx, Δy give the row and column spacing of pixel spacing attribute.

3. Results and Discussion

The patient study considered in this work has 22 DICOM slices of size 512×512. The proposed work takes tumour intensities on a slice, which holds maximum tumour portion as a target model. The target is initialized with a desired rectangular region (slice-12) of size 54×112. The spatial information of the target model is used as a reference for tracking in the series. The scale parameter ∇d is assigned as 10 to generate a larger candidate window such that the target model will be completely bounded by the candidate window. The target is tracked from the first slice until the end.

Figure 3 shows the tracking results of the target for slices 9 to 14. The external rectangle coordinates act as a candidate and are used to identify the real target. The target model is compared with the candidate area (M_{00}) using BC thus, minimizes the intra-class difference. Based on studies by Saha et al. [24], minimizing intra-class intensities and intra-class entropy are attempted to qualify bounding box coordinates. Table 1 illustrates the comparative BC estimates and the dice score of the proposed scaled adaptive method against intensity based and entropy-based bounding box techniques. Dice coefficient is computed as $2|G \cap B|/(|G| + |B|)$, where G is the expert's marked boundary and B being the box co-ordinates obtained from scaled adaptive method. With an increased BC and Dice co-efficient of the proposed scheme, a better fit of the candidate window and similarity over the target is obtained.

The key property of the bounding box is assigning the exterior pixels to the background class, which surrounds the box.

If BC is zero then the target model is lost, which occurs in the exact beginning and end of the slice sequence. This procedure certainly tracks the target even if the scale of the target changes. Bounding box restricts the segmentation process and clusters the target intensities and background within the box inland. In order to extract target tumour intensities and to mask tumour intensities in the respective slices, clustering is performed. The intensities within the candidate window are clustered with *k*-means clustering.

Table 1. Comparison with existing bounding box techniques.

Slice	BC estimate			Dice score		
	Intensity based	Entropy based	(Scaled adaptive)	Intensity based	Entropy based	(Scaled adaptive)
9	0.4926	0.4131	0.6952	0.7166	0.5321	0.8177
10	0.5231	0.4395	0.7413	0.7213	0.6113	0.8209
11	0.5497	0.4490	0.8223	0.7427	0.6725	0.8217
12	0.5983	0.4892	0.8339	0.7431	0.6731	0.8248
13	0.5845	0.4821	0.8394	0.7032	0.6608	0.8310
14	0.5902	0.4783	0.7892	0.7156	0.6556	0.7892

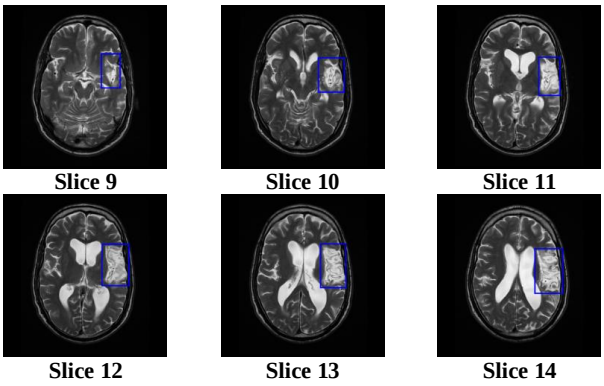

Fig. 3. Tracked tumour boundary by maximizing BC.

The number of clusters (k) is chosen as two since the box comprises merely of tumour intensities and background, thus, achieves tumour masking in the spatial domain. This segments the dense tumour and the background features into two classes. Tumour contours are extracted from the clustering process within the tracked coordinates. The actual intensities and the spatial location of the tumour voxels are obtained by masking the respective slice with the clustered contours.

Figure 4 shows the segmented tumour contours of the corresponding slices. The visualized contours have the real intensities taken from MR slices by masking. Contour area of each slice is measured using masked voxels and pixel spacing attribute. The area computed for consecutive slices is used to estimate tumour volume. The volume rendering and volume calculation will give different results for DICOM and resampled grid because of varying properties such as pixel spacing and slice thickness.

Table 2 comprises DICOM and resampled grid attributes such as dimensions, voxel depth, pixel spacing and voxel size. Frustum based volume calculation is performed for DICOM and resampled grid. In Table 3, tumour volume is calculated from slice 9 to slice 15 using DICOM attributes. The frustum volume of adjacent slices is computed based on contour area and slice thickness. The total volume is obtained by summing all the intermediate volumes. Similarly, the tumour volume of the resampled grid is computed as shown in Table 4. By comparing Table 3 and Table 4, it is evident that the volume of the resampled grid is greater than the volume of DICOM Grid. Since the inter-slice resolution is lower in the DICOM grid having 5

mm slice thickness, cubic spline interpolation utilizes the intermediate intensity values and the inter-slice resolution is improved to '1 mm'. This increase in volume is due to the increase in the number of slices stacked in the resampled grid.

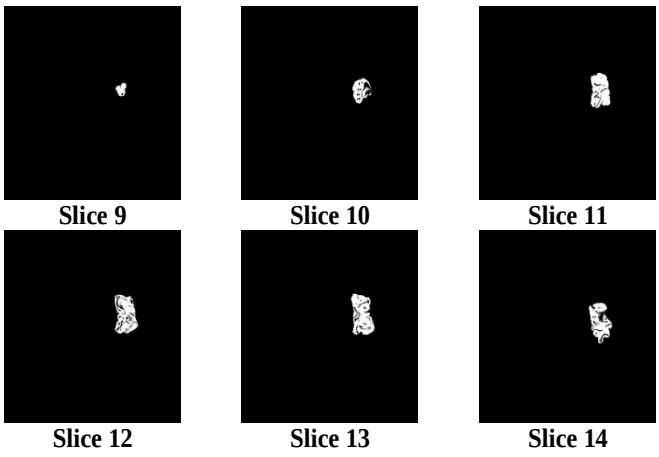

Fig. 4. Segmented tumour contours of slices.

Table 2. DICOM vs. resampled grid.

	DICOM grid	**Resampled grid**
Dimensions	22×512×512	110×230×230
Voxel depth(T)	5 mm	1 mm
Pixel spacing (PS_x /PS_y)	(0.4492 mm/ 0.4492 mm)	(1 mm/1 mm)
Tumor voxels (sum)	21673	26944
Tumor ratio	0.003758	0.004630
Size (per voxel)	1.0089032	0.9998

Table 3. Volume calculation from DICOM header data (frustum).

Slice no.	**Area (mm^2)**	**Volume (mm^3) from area (A_1, A_2)**
9	149.12	1375.89
10	424.75	
10	424.75	3056.90
11	819.43	
11	819.43	4507.41
12	986.10	
12	986.10	5027.55
13	1025.05	
13	1025.05	4627.28
14	829.32	
14	829.32	2181.33
15	139.43	
Estimated volume		20776.35 mm^3

Pixel Spacing (PS_x/PS_y) = (0.4492 mm/ 0.4492 mm); Slice thickness = 5 mm

Table 4. Volume calculation after resampling and interpolation (frustum).

Volume (mm^3) from area (A_1, A_2)	Total volume (mm^3)
$\sum \frac{T}{3}(A_1 + A_2 + \sqrt{A_1 A_2})$	26848.95

Pixel Spacing (PS_x/PS_y) = (1 mm/1 mm); Slice thickness = 1 mm

Table 5 shows a comparison of computed volume between the frustum model and the tumour voxels contributed in DICOM and resampled grid. The resampled grid offers minimum error when compared to that of DICOM grid. This asymmetry in DICOM grid causes a step-shaped iso-surface in 3D reconstruction. Resampled grid offers smooth iso-surface and greater accuracy in 3D visualization.

Figure 5 shows the 3D resampled grid for the patient study. The three anatomical planes are used and their dimensions are set in mm. For the DICOM grid, the dimensions [X×Y×Z] is set as [512×512×22]. For the resampled grid shown in Fig. 5 the dimensions are viewed as [230×230×110]. The model permits to visualize the object interactively in all the three directions specified. The resampled grid offers smoother extracted anatomical structure in 3D.

Table 5. DICOM vs. resampled grid.

	No. of voxels	Voxel size	Tumor voxels *voxel size	Frustum model	Error %
DICOM	21673	1.0089032	21865.9	20776.35	4.9
Resampled	26944	0.9998	26938.6	26848.95	0.3

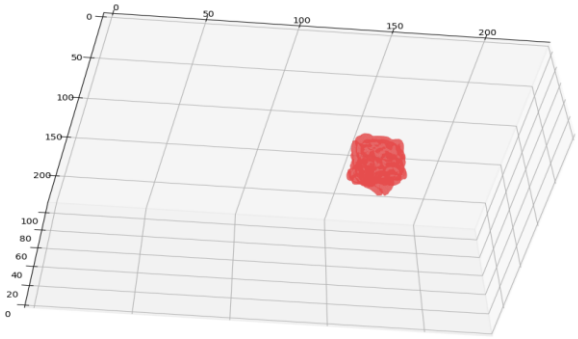

Fig. 5. Resampled grid.

4. Conclusion

The adaptive tracking of vital tumour contours reduced the influence of other brain tissues in the segmentation process. In our work, the mean shift tracking with the Bhattacharya coefficient measure indicated that the target is entirely inside the candidate model with minimal background intensities. A controlled tumour segmentation is done within the obtained candidate box coordinates such that no pre-processing is required for the slices. Volume quantification of cubical stack compared with frustum model provides a much-needed breakthrough in the tumour evaluation process. DICOM grid is converted to a resampled grid based on the resize factor computed using DICOM metadata and the slices are stacked

accordingly. The qualitative and quantitative evaluations of the resampled stack against DICOM stack were analysed and the results proved that the resampled grid exhibits better accuracy. 3D Modelling based on computer-aided procedures helps physicists to locate tumour during initial diagnosis. This proposed framework can be extended by automatic initialization of the target model from a trained classification model of abnormalities.

Nomenclatures

C_h	Normalization constant for candidate model
CM_i	Contour mask
$d(i, j)$	Euclidean distance, mm
I_{min}	Minimum mutual information
I_{max}	Maximum mutual information
M_{00}	Zeroth order moment, mm^2
$\hat{p}(y)$	Candidate model
$\hat{q}$	Target model
T	Slice thickness, mm
w_i	Density estimation of image

Greek Symbols

Δd	Scale parameter
$\in$	Error rate
ρ	Bhattacharya Coefficient metric

Abbreviations

BC	Bhattacharya Coefficient
BI	Background Intensities
CI	Candidate Intensities
DICOM	Digital Imaging and Communications in Medicine
TCA	Total Candidate Area

References

1. El-Dahshan, E.S.A.; Mohsen, H.M.; Revett, K.; and Salem, A.-B.M. (2014). Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm. *Expert Systems with Applications*, 41(11), 5526-5545.
2. Grover, V.P.; Tognarelli, J.M.; Crossey, M.M.; Cox, I.J.; Taylor-Robinson, S.D.; and McPhail, M.J. (2015). Magnetic resonance imaging: Principles and techniques: Lessons for clinicians. *Journal of Clinical and Experimental Hepatology*, 5(3), 246-255.
3. Bidgood Jr, W.D.; Horii, S.C.; Prior, F.W.; and Van Syckle, D.E.V. (1997). Understanding and using DICOM, the data interchange standard for biomedical imaging. *Journal of the American Medical Informatics Association*, 4(3), 199-212.
4. Kaushal, M.; Khehra, B.S.; and Sharma, A. (2018). Soft computing based object detection and tracking approaches: State-of-the-art survey. *Applied Soft Computing*, 70, 423-464.

5. Comaniciu, D.; Ramesh, V.; and Meer, P. (2003). Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5), 564-577.
6. Kailath, T. (1967). The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Communication Technology*, 15(1), 52-60.
7. Yang, C.; Duraiswami, R.; and Davis, L. (2005). Efficient mean-shift tracking via a new similarity measure. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, California, United States of America, 176-183.
8. Wang, Z.; Yang, X.; Xu, Y.; and Yu, S. (2009). CAMShift guided particle filter for visual tracking. *Proceedings of the IEEE Workshop on Signal Processing Systems.* Shanghai, China, 6 pages.
9. Zhong, K.; Zhang, Z.; and Zhao, Z. (2018). Vehicle detection and tracking based on GMM and enhanced CAMShift algorithm. *Journal of Electrical and Electronic Engineering*, 6(2), 40-45.
10. Salhi, A.; and Jammoussi, Y. (2012). Object tracking system using CAMShift, Meanshift and Kalman filter. *World Academy of Science, Engineering and Technology*, 64, 674-679.
11. Ning, J.; Zhang, L.; Zhang, D.; and Wu, C. (2012). Scale and orientation adaptive mean shift tracking. *IET Computer Vision*, 6(1), 52-61.
12. Suresh, K.; and Sakthi, U. (2018). Robust multi-thresholding in noisy grayscale images using otsu's function and harmony search optimization algorithm. *Advances in Electronics, Communication and Computing*, 491-499.
13. Rajinikanth, V.; Raja, N.S.M.; and Kamalanand, K. (2017). Firefly algorithm assisted segmentation of tumor from brain MRI using Tsallis function and Markov random field. *Journal of Control Engineering and Applied Informatics*, 19(3), 97-106.
14. Lempitsky, V.; Kohli, P.; Rother, C.; and Sharp, T. (2009). Image segmentation with a bounding box prior. *Proceedings of the IEEE 12th International Conference on Computer Vision*. Kyoto, Japan, 277-284.
15. Rajinikanth, V.; Madhavaraja, N.; Satapathy, S.C.; and Fernandes, S.L. (2017). Otsu's multi-thresholding and active contour snake model to segment dermoscopy images. *Journal of Medical Imaging and Health Informatics*, 7(8), 1837-1840.
16. Raja, N.S.M.; Rajinikanth, V.; Fernandes, S.L.; and Satapathy, S.C. (2017). Segmentation of breast thermal images using kapur's entropy and hidden markov random field. *Journal of Medical Imaging and Health Informatics*, 7(8), 1825-1829.
17. Shree, T.D.V.; Revanth, K.; Raja, N.S.M.; and Rajinikanth, V. (2018). A hybrid image processing approach to examine abnormality in retinal optic disc. *Procedia Computer Science*, 125, 157-164.
18. Raja, N.S.M.; Suresh, D.; Kirthinigodweena, A. (2017). A novel image processing scheme to evaluate breast thermal images. *Journal of Engineering Science and Technology (JESTEC)*, 9th Eureca 2017 Special Issue, 106-115.
19. Suresh, D.; Raja, N.S.M.; Godweena, A.K. (2018). Kapur/Tsallis entropy guided segmentation of plasmodium species from thin blood smear images.

Journal of Engineering Science and Technology (JESTEC), 9th Eureca 2017 Special Issue, 83-93.

20. Lin, Q.; Du, M.; and Guo, Y. (2014). Slice interpolation for MRI using disassemble-reassemble method. *TELKOMNIKA*, 12(8), 6190-6197.
21. Parker, J.A.; Kenyon, R.V.; and Troxel, D.E. (1983). Comparison of interpolating methods for image resampling. *IEEE Transactions on Medical Imaging*, 2(1), 31-39.
22. Langenberg, R.v.d.; Bondt, B.J.d.; Nelemans, P.J.; Baumert, B.G.; and Stokroos, R.J. (2009). Follow-up assessment of vestibular schwannomas: Volume quantification versus two-dimensional measurements. *Neuroradiology*, 51(8), 517-524.
23. Re, E.C.d.; Gao, Y.; Eckbo, R.; Petryshen, T.L.; Blokland, G.A.M.; Seidman, L.J.; Konishi, J.; Goldstein, J.M.; McCarley, R.W.; Shenton, M.E.; and Bouix, S. (2016). A new MRI masking technique based on multi-atlas brain segmentation in controls and schizophrenia: A rapid and viable alternative to manual masking. *Journal of Neuroimaging*, 26(1), 28-36.
24. Saha, B.N.; Ray, N.; Greiner, R.; Murtha, A.; and Zhang, H. (2012). Quick detection of brain tumors and edemas: A bounding box method using symmetry. *Computerized Medical Imaging and Graphics*, 36(2), 95-107.