Semi-automated classification of CT Scans in Traumatic Brain Injury Patients

Adnan N. Qureshi
Institute for Research in Applicable Computing,
University of Bedfordshire, UK

ABSTRACT

A 'silent epidemic' affecting millions worldwide every year is the Traumatic Brain Injury. Management of these patients essentially involves neuroimaging and noncontrast Computed Tomography (CT) scans are the first choice amongst doctors. However, inter-observer variability, considered 'Achilles heel' amongst radiologists, can lead to missed diagnoses and grave consequences. This paper presents a hybrid approach for semi-automated classification of CT features according to Marshall CT Scheme. The proposed method uses template matching, artificial neural networks and active contours for segmentation of significant anatomical landmarks and estimation of haematoma volume on brain CT scans. The proposed method is efficient and robust in segmenting cross-sectional, noncontrast CT scans and has been evaluated on images from subjects with different ages and both genders. The hybrid method has an average ICC $\geq$ 0.97 and Jaccard Index $\geq$ 0.86 compared with the manual demarcations by radiology experts and performs better than the state of the art. Hence, the approach can be used to provide second opinions very close to the experts' intuition.

General Terms:

Computed Tomography Scans, Segmentation

Keywords:

Traumatic brain injury, pattern recognition, neural networks, segmentation, active contours

1. INTRODUCTION

Traumatic Brain Injury (TBI) is one of the serious public health problems affecting millions worldwide every year. Also called a 'Silent Epidemic' it affects all age groups and both genders [6, 17]. While the severity of the injury can range from mild to severe, the high mortality and morbidity rates due to TBI can be attributed to the fact that the extent of injury is often difficult to be accurately reported by doctors making prognosis nearly impossible to predict [15]. Management of TBI cases essentially involves suggesting neuroimaging and noncontrast Computed Tomography (CT) scans are the first choice amongst doctors due to widespread availability and the less time taken to perform the scan.
The emergency clinical approach includes neurophysiological examination and assessment of the severity of the trauma in using CT classification schemes such as one proposed by Marshall et al.

(Fig. 1) to help establish diagnosis, prognosis, mortality, morbidity and hospitalization time after trauma [10]. The initial noncontrast CT carries significant prognostic value and combined with other clinical parameters, can be helpful in prediction of long-term outcomes [17].
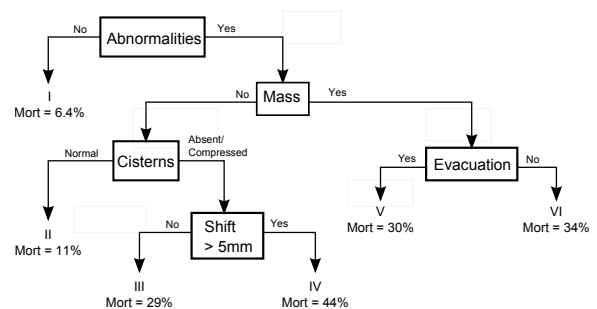

Fig. 1: Marshall et al. CT Rating representing categories of abnormalities seen on CT scans. [10]

However, depending upon the perceptual skills, deduction and knowledge of the observers, the interpretation results are often fraught with variability and can lead to errors. Postmortem studies show that up to 20% of fatal illness can be misdiagnosed [9]. A study conducted to investigate performance of junior doctors in accident and emergency (A&E) to diagnose significant x-ray reported that only 32% of junior doctors correctly diagnosed the illness while the average performance of the senior doctors was 80% [11]. Errors in interpretation of medical images can be avoided with double reporting and to this end, the potential of using computer-aided diagnosis (CAD) in mainstream radiology is quite encouraging [14, 13]. The results from CAD are frequently used as "second opinion" by the observing clinicians, however, their usage requires manual intervention (and seeding) and most of the proposed techniques are not fast enough to be used on a large volume of data.
Computer aided quantitative radiology and volumetry studies of the anatomical structures have been undertaken by many researchers utilizing techniques from image processing and machine learning domains [13]. An important first step in such systems, segmentation, provides a pertinent representation of image information and a surfeit of methods have been proposed [18]. Manual demarcation of regions of interest are still considered 'gold standard', however,
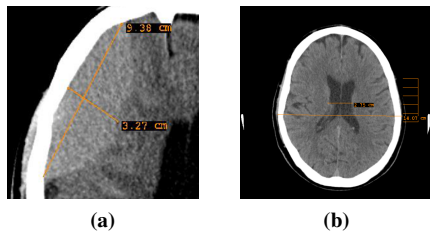
(a)	**(b)**

Fig. 2: Measurements on Axial CT Scans manually annotated by radiology experts.

these are time consuming and a robust and efficient method is required to assist the users.

Segmentation algorithms that are not spatially guided, usually cluster the pixels intensities to generate regions of interest. The usefulness of spatially blind methods is not at par with the performance of spatially guided algorithms because these can generate disjoint, quasi-homogeneous regions and extensive user interaction and parameter tweaking required by these is not favourable in clinical settings [14]. Techniques such as simple linear iterative clustering (SLIC), adapt K-means approach and cluster pixels into superpixels [2]. The input image is divided into small user-defined sections, which are the initial superpixels, and then the centres of these are used for computing the mean of clusters. Iterations update the regions and the location of centres. Based on the distance of the pixels from the cluster centre, a label is assigned and the process is repeated until the error between successive cluster centres falls below the defined threshold [1]. The algorithm, however, intrinsically depends upon the selection of size of the initial superpixels which needs to be changed based on the size of the respective region of interest as reported by [16].

On the contrary, the neighbourhood information can result in better segmentation taking into account the spatial characteristics of medical images [13]. Techniques such as region growing and merging, active contours and level sets preserve the spatial relationship between the pixels and neighbourhood information [18]. The application of active contours for segmentation of the neuroimages has gained significant popularity in the recent years [20, 19, 5, 16]. The underlying principle of active contours is to deform a boundary around the region of interest to achieve the required segmentation. The implementation by Chan-Vese is considered state of the art when the edges are not used to differentiate between the region of interest (ROI) and non-ROI [4]. Medical images, such as brain CT, mostly present regions of interest without hard, distinct edges and active contours not relying on gradient based edges can produce better results of segmentation. However, placement of the initial seed curve still poses a constraint on the evolution of the curve and has to be performed manually by the user. Similarly, if distinct edges are not considered, the evolving curve can creep into the surrounding regions based on pixel intensity.

Several hybrid methods have been proposed to segment regions in brain CT images. Bhadauria et al. proposed combination of Fuzzy C-Means clustering and region-based active contour to segment the haemorrhage region [3]. They used clustering to find an initial mask and then propagate it to the haemorrhage boundaries using region-based active contour. An automatic procedure based on modified distance regularized level set (DRLSE) has been proposed by Prakash et al. to segment haemorrhage in brain CT images [12]. Their method first creates a general intensity mask defined by a set of DRLSE parameters to localize the region rapidly. Then the

region boundaries are identified in another level set step by choosing the parameters so that more accurate boundaries are detected. This method is sensitive to initial threshold values and can creep beyond boundaries when the normal brain tissue and trauma have very close intensity values.

An improvement of the DRLSE based segmentation was proposed by [16]. In that method, the pixels are first classified using SLIC and then the results are fed into the DRLSE algorithm to delineate the final segmentation. The proposed method has been implemented on brain trauma CT and results show improvement from the DRLSE algorithm. However, the performance of SLIC is highly dependant on the selection of the parameters of the algorithm.

Machine learning and computational intelligence has recently seen an extensive and tenable integration in medical problems based on the premise that these systems can adaptively learn and optimize the relationship between inputs and outputs [7]. An artificial neural network (ANN), such as the feedforward net, can fit a given finite input-output mapping problem and their use in medical image processing highlights their efficiency in automatic delineation of the ROI as proposed by [14, 13]. Medical images e.g., CT scans, usually have inhomogeneity of background and noise; and recognition of ROI in 2-D images with such characteristics using ANN has given plausible results. Furthermore, ANN's can be trained using a few representative images which have been manually annotated and labelled in order to learn to recognize the ROI, and the process later on becomes automated requiring no interaction from the user to provide initial location of the ROI [13].

The significance of classifying TBI CT scans has been highlighted by related literature, however, there is still a gap requiring a computer-aided system for performing segmentation, measurements and classification of CT scans in a semi-automated approach. In this paper, a hybrid approach using template matching, artificial neural networks and active contours is presented. The proposed method efficiently segments and robustly performs linear measurements on axial CT scans of brain based on the approach proposed by [14, 16, 13], and classifies CT according to Marshal et al. Scheme. The hybrid method can identify cerebrospinal fluid (CSF) spaces i.e., cisterns and ventricles, detect midline shift and estimate volume of intracranial haematomas. The rest of the paper is organized as follows. Section 2 describes the methodology of pre-processing, processing and analyses steps. Section 3 presents and discusses the results and the paper is concluded in section 4.

2. METHODOLOGY

The proposed method uses ANN to identify ROI for segmentation from the input images. The binarized pixels from the ANN's output act as initial seeding point for active contour based segmentation. Measurements of segmented ROI are then used for calculating the length and breadth of the haematoma or the ventricles (Fig. 2) on the respective axial slice. The input images are DICOM (Digital Imaging and Communications in Medicine) files and the training set (with PNG files) has been manually segmented and measured by radiology experts. The dataset used for training, validation and testing consists of cross-sectional, noncontrast CT studies from subjects of both genders and different ages.

The proposed algorithm only requires the clinician to manually select the key slice, hence, it is semi-automated. Thereafter, all the subsequent steps are fully automatic and do not require any user interaction.

2.1 Estimation of brain midline

Midline in the brain is an important reference point for assessing the severity of injuries and pathological conditions. It is an anatomical imaginary line formed by the falx cerebri and passes through the septum pellucidum.
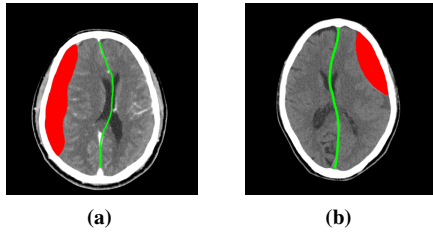

(a)   **(b)**

Fig. 3: Manual annotation of the haematoma (red) and midline shift (green) on brain CT scans by radiology expert.

The first step is to align the head into proper orientation on the CT image to ensure that the reference points for measurements are properly identified. The skull boundaries are extracted using Canny edge detection with $\sigma = 2$ and $threshold = [0.5\ 0.2]$ in the experiments as these suppress detection of spurious edges. The centroid of the skull boundary is detected which should lie on an ideal midline. The CT scan containing nose of the patient is selected as reference and using morphological dilation with 'disc' structuring element and grey level thresholding, the skull is extracted and subjected to 'Canny' edge detection (Fig. 4). The dilation of X by Y is defined as:

$$X \oplus Y = \bigcup X_y \qquad (1)$$

The region properties give the convex hull enclosing the skull and the orientation of the major axis of the edge θ_1 with respect to the ideal midline passing through the centroid. In addition, demarcating the bony protuberance of the anterior and posterior falx cerebri attachments provide the orientation angle θ_2 as shown in Fig. 5. These bony landmarks are usually not affected in traumatic and pathological cases in which the brain parenchyma may show abnormalities. The rotation angle Θ is then estimated by ensuring that the directions of θ_1 and θ_2 agree. The rotation angle Θ has to be corrected with reference to the skull centroid as shown in Fig. 5.

The brain midline passes through the septum pellucidum. A set of 32 template images is created using the experts' annotations. Using cross correlation and sum of squared difference for template
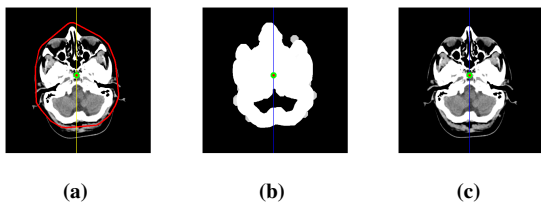

(a)   **(b)**   **(c)**

Fig. 4: Estimation of tilt of head in CT Scans. Image (a) shows rotation of the head during CT scan procedure with reference line passing through centroid. Image (b) is the result of morphological dilation and gives estimation of tilt (θ_1) towards right of patient. Image (c) is the final rotated CT.

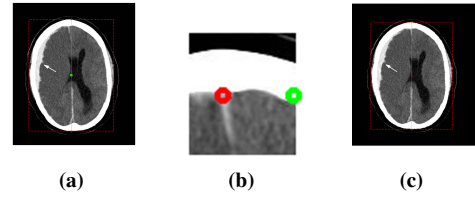

(a)   **(b)**   **(c)**

Fig. 5: Correction of orientation of head in CT Scans using anterior bony protuberance. Image (a) shows tilted head during CT scan procedure. Image (b) shows the estimation of anterior bony protuberance giving θ_2 and image (c) is correction of orientation ($\Theta = -3.35°$).
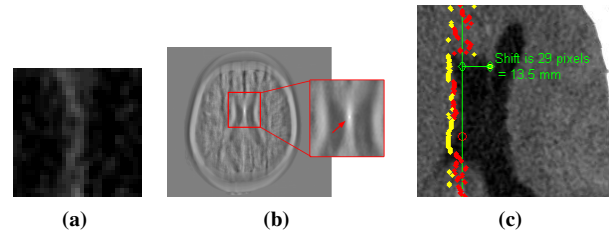

(a)   **(b)**   **(c)**

Fig. 6: Template matching for Septum Pellucidum. Image (a) is a template of the septum pellucidum region. Images (b) shows the identified location of septum pellucidum using NCC. Image (c) also shows the estimation of midline shift ($13.5mm$ towards left) in the respective case.

matching, the septum pellucidum is located and used as a landmark for locating the midline (Fig. 6). The normalised cross-correlation (NCC) of a template, $t(x, y)$ with a subimage $f(x, y)$ can be calculated as given in equation (2).

$$NCC = \frac{1}{n} \sum_{x,y} \frac{(f(x, y) - \bar{f})(t(x, y) - \bar{t})}{\sigma_f \sigma_t} \qquad (2)$$

Where n is the number of pixels, $\bar{f}$ and σ_f are the average and standard deviation of f respectively and $\bar{t}$ and σ_t are the average and standard deviation of t respectively.

Identifying the septum pellucidum and its location with reference to skull can be used for measuring the midline shift which is an important clinical feature in assessment of traumatic or pathological conditions and a shift $> 5mm$ is a factor for differentiating between Rating III and Rating IV (and above) in Marshal et al. scheme. Similarly, the appearance of haematoma changes with time and at some point between 3 and 21 days, it can become isodense to the adjacent cortex, making identification potentially tricky. In such cases, visualising a number of indirect signs such as mass effect, sulcal distortion and midline shift are significantly important, life-saving indicators.

2.2 Pixel classification using ANN

An input image is represented as an $(n \times m)$ matrix $\mathbf{X} = \{x_{ij}\}, i = 1 \ldots n, j = 1 \ldots m$, whose pixels can either belong to a region of interest class Ω_0 or to a non region Ω_1. A rectangular window $\mathbf{P}$, which is a $(k \times k)$ matrix containing the central pixel x_{ij} and $r - 1$ nearest pixels, where $r = k^2$, is slid through the image $\mathbf{X}$. The sliding window transforms the image $\mathbf{X}$ into an $r \times q$ matrix where $q = (n - k + 1)(m - k + 1)$ and the central

element of a column vector, hence, represents the pixel x_{ij} of the input image. In the experiments, a sliding window with $k = 5$ gave the optimum results. In addition, statistical features (Mean, Median, Standard deviation, Shannon's Entropy and Discrete Fourier Transform (DFT) are extracted and constitute the feature vectors in **Z**. For the identification of cisterns, the spatial location of pixels of region of interest also constitutes the feature vectors. The results in Fig. 7 represent performance in terms of receiver operator characteristic (ROC) curves. It is observed that ANN trained on these features give better results when a combination of the features constitutes the feature vector as compared to using them individually.
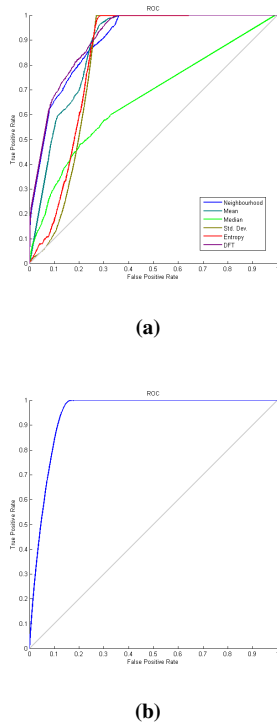

(a)

(b)

Fig. 7: Receiver Operator Characteristic (ROC) curves of ANN's performance based on difference image features. Image (a) represents performance with individual statistical features extracted from the images: neighbourhood, mean, median, std. dev., entropy and DFT. The image (b) shows the performance when the feature vector comprises of their combination.

The matrix **Z** is fed to the Feedforward ANN with one hidden layer of 10 neurons and the output neuron makes a decision $y_c = \{0, 1\}, c = 1 \dots q$, thus, classifying the image pixels to classes Ω_0 or Ω_1 which represent non-region of interest and region of interest, respectively. The Levenberg-Marquardt backpropagation function, which is highly recommended as a first-choice supervised algorithm, is used in addition to scaled conjugate gradient, Bayesian regulation, BFGS quasi-Newton and Gradient descent with momentum and adaptive learning rate backpropagation, however, Jacobian training is not supported on GPU in Matlab.
The ANN is trained with the aim to minimize an error function ϵ for a given number of training epochs. The error is calculated as mean square error (MSE) as given in Eq. 3 with n samples, Y_O outputs and Y_T targets.

$$\epsilon = \frac{1}{n} \sum_{i=1}^{n} (Y_O - Y_T)^2 \qquad (3)$$

(a)   **(b)**   **(c)**   **(d)**
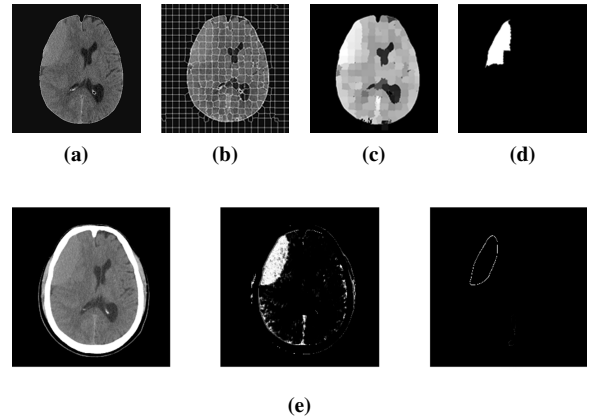

(e)

Fig. 8: Comparison of the classification of pixels into probable ROI using SLIC (a)-(d) and ANN (e). It is observed that the ANN based classification results in better approximation of the haematoma as compared to the jagged results from SLIC.
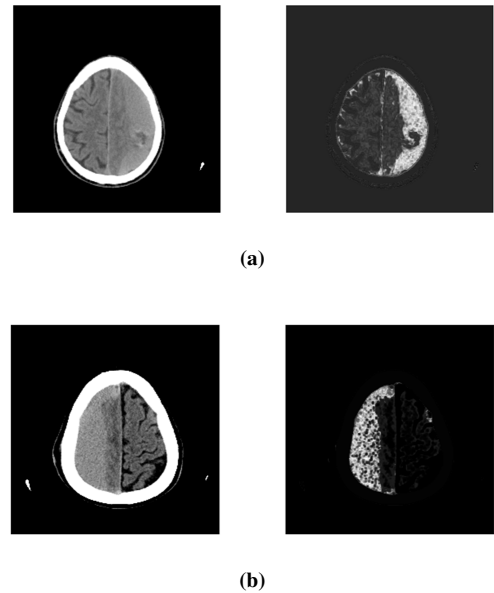

(a)

(b)

Fig. 9: Pixel classification with ANN. Images (a) and (b) show the input image (left column) and the pixels identified within the ROI (right column). However, there are some misclassified pixels, which are spatially disjoint with the ROI, because of similar pixel intensities.

The CT images, whose pixels are manually annotated by radiology expert (e.g., Fig 3), are used and the ANN is trained to approximate these. The pixel classification by ANN gives better results as compared to SLIC superpixels which require a lot of parameter setting (Fig.8). The different ROI's require different ANN's because of their intensity characteristics. The Fig. 9 depicts pixels of haematomas classified by respective ANN. The approach is implemented to identify ROI which represent subdural or epidural haematomas. The output from the ANN can then be differentiated into classes Ω_0 or Ω_1 by the squashing (transfer) function. In the experiments, pure linear, saturated linear and the hyperbolic tangent sigmoid functions produced better differentiation than the log sigmoid functions. For further experiments, pure linear function was used as squashing function.

2.3 Segmentation of CSF spaces and Haematomas

The training images have their pixels manually labelled and assigned either to class Ω_0 or Ω_1 by a radiology expert (Fig. 3) and the ANN is trained to approximate these. The approach is implemented to identify regions which represent subdural or epidural haematomas, ventricles and cisterns. The ANN classifies pixels belonging to either the ROI or to the non-ROI. Due to the characteristics of the imaging modality, some pixels with similar intensity values could get classified as false positives although they are spatially disjoint from the pixels of the ROI. These regions are further refined using active contours as subsequently described.

The active contour method (using 'Edge' or 'Chan-Vese' implementations) doest not give plausible results and segmentation deteriorates due to the size and placement of the initial seed curve. The initial curve has to be manually placed by the user inside the ROI in the form of a polygon and it then evolves to fit the ROI, which is a tedious and time consuming approach. It can also make the segmentation results vary when there is slight difference in the shape of the initial polygon curve entered by user. The 'Chan-Vese' and 'Edge' implementations were run for 100, 200 and 400 iterations for each input image and there was no curve evolution observed beyond 400 iterations. The maximum Jaccard Index (JI) of 'Edge' based method is 0.3312 and with 'Chan-Vese', it is 0.6626 and the segmentation is shown in Fig. 10. The maximum Dice's coefficient observed is 0.4976 and 0.7970 with 'Edge' and 'Chan-Vese' implementations respectively.

In the proposed approach, the initial curve for the active contour is extracted from the binarized ANN output using edge detection. Hence, the initial curve automatically gets seeded and provides the necessary seeding mask. The result of haematoma segmentation using hybrid approach is shown in Fig. 10d which gives better delineation of the ROI. The Fig. 11 presents segmentation of three cases using the proposed hybrid approach.

It should be noted, however, that the CT is a mirror image of the anatomy of the patient as the CT is visualized from below upwards. Hence, the presentation of haematoma on the left on CT is actually right side of the patient's brain and vice versa. Also, For most of the TBI cases, the location of haematoma is usually 60% temporoparietal, 20% frontal and 20% parieto-occipital.

2.4 Measurements and Estimation of Volume

The CT scan studies in DICOM format are viewed using 'brain window' (width = 80, length = 40) with image size of 512×512 pixels. A radiology expert then provided the measurements which are stored for comparison with automatic measurements by the algorithm.

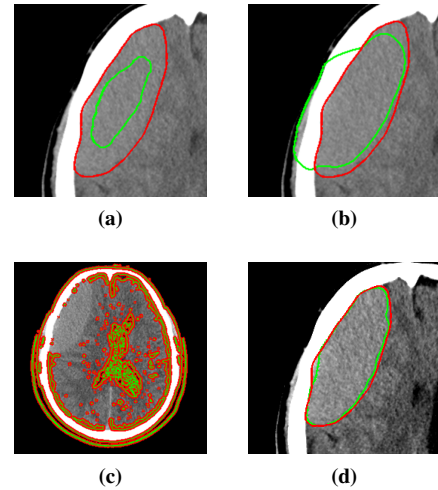**(a)**   **(b)**   **(c)**   **(d)**

Fig. 10: Segmentation results from Active contour based approaches. Green curve shows the active contour while red is the manual demarcation. Image (a) shows output with edge based approach which undersegments, image (b) is the result of 'Chan-Vese' method in which the curve creeps into the adjacent skull bone outside the ROI and image (c) is result of DRLSE which fails to recognise haematoma as ROI. Image (d) is the result of hybrid method closely approximates the ROI.

(a)   **(b)**   **(c)**
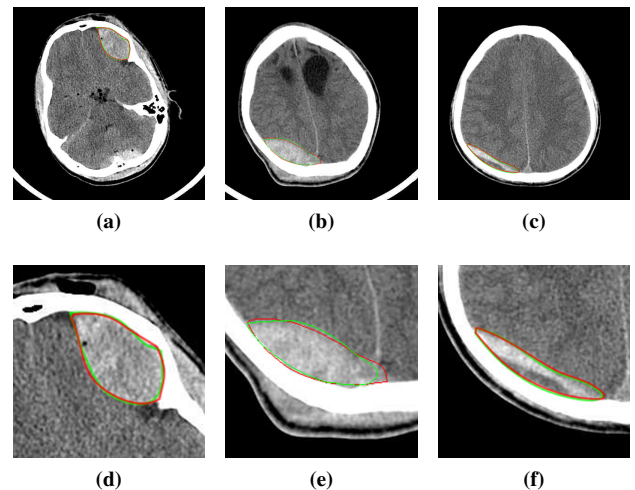
(d)   **(e)**   **(f)**

Fig. 11: Segmentation results from hybrid approach. Green curve shows the active contour while red is the manual demarcation. Bottom row shows magnified images corresponding to top row.

The contiguous image regions from step 2.3 are used by the algorithm to determine measurements such as extrema, area, orientation and axes lengths. This identifies and counts the number of pixels in respective linear measurements and this is divided by spatial resolution of the respective study to give the measurements. The volume of intracerebral hematoma was calculated according to the formula $V = \frac{a \times b \times c}{2}$, where a is the largest diameter of haematoma, b the largest diameter normal to a , and c is the approximate number of

CT slices with haemorrhage multiplied by the slice thickness. This equation of volume of ellipsoid assumes $\pi \approx 3$ for simplification [8]. The brain CT scans used in experiments were obtained using *brain protocol* on Toshiba Activion 16, with $1mm$ slice thickness and pixel spacing of $0.468/0.468$ in the x, y axes in the 16-bit, 512×512 pixel DICOM images and from Philips Brilliance 16 with $5mm$ slice thickness and pixel spacing of $0.354/0.354$.
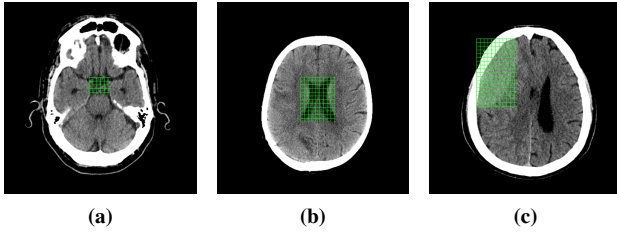

Fig. 12: Output of the proposed method. Image (a) shows a grid overlaid on the supra cellar cistern, (b) demarcates the ventricles and (c) identifies the haematoma.

3. RESULTS AND DISCUSSION

The technique has been applied on 109 anonymized, cross-sectional, noncontrast CT studies from different subjects with ages between 5 to 85 years and both genders. Out of these, 51 cases with haematomas are used for volume estimation and a total of 866 CT sclices with intracranial bleeds are analysed. Two radiology experts had independently performed manual segmentation and pixel labelling of 20 cases to be used for training. The experts also provided measurements and judgements for all cases.

Midline in the brain is an important reference point for assessing the severity of injuries and pathological conditions and in Marshall et al. CT Classification, it is the significant characteristic of rating IV and above. There can be mild to moderate shift but the resulting manifestations can be quite serious and a shift $> 5mm$ is considered a surgical emergency. The proposed method shows robust performance in detecting tilted head of the patient during the scan procedure and corrects it for proper subsequent measurements (table 1). For evaluation, true positives (TP), true negatives (TN), false positives (FP) and the false negatives (FN) are calculated for estimating sensitivity and specificity.Sensitivity reflects the algorithm's ability to identify the given condition correctly and is calculated as:

$$Sensitivity = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (4)$$

Similarly, specificity relates to the ability to exclude a condition correctly and is calculated as:

$$Specificity = \frac{\text{TN}}{\text{TN} + \text{FP}} \qquad (5)$$

The sensitivity and specificity of the approach used for correction of orientation of the head are 83% and 67%, respectively, while considering the orientation as either tilted or normal (boolean). The method uses the anterior and posterior bony protuberance of the skull as landmarks and rotates the image accordingly so that the ideal midline would be approximately vertical. The midline shift is classified as either $> 5mm$ or none (boolean), and the average

Table 1. : Performance of Proposed Method

CT Feature	Sensitivity	Specificity
Correction of Orientation	83%	67%
Midline Shift, Class≥ IV	89%	80%

performance of the proposed method is 89% and 80% in terms of sensitivity and specificity, respectively.

The Jaccard Index (JI) or coefficient is used to measure similarity between ground truth and segmentation results. JI measures similarity between finite sets and is defined as:

$$J(M, A) = \frac{\mid M \cap A \mid}{\mid M \cup A \mid} \qquad (6)$$

Where M is the set of pixels from manual segmentation, A is the output of the proposed algorithm and $0 \leq J(M, A) \leq 1$. JI, hence, is 0 when there is no overlap between the experts' demarcation (M) and algorithm's approximation (A) and is equal to 1 when there is perfect overlap. Consequently, values close to 1 indicate better performance. Another measure which not very different from JI is Dice's coefficient and is also used to quantify the performance of the proposed method and is calculated as:

$$D = \frac{2J}{(1 + J)} \qquad (7)$$

where J is the Jaccard index calculated earlier. Compared to the performance of state of the art active contours which had JI of 0.5667 ± 0.2052 with 'Chan-Vese' method, the hybrid approach shows JI of 0.8689 ± 0.042 when a feedforward net is used for initial classification of pixels for subsequent segmentation. The results are given in Table 2 which shows JI and Dice's coefficient based evaluations.

Table 2. : Performance of proposed approach using Feedforward Neural Net with active contours compared with Chan-Vese approach, DRLSE and SLIC with DRLSE.

Method	Jaccard Index	Dice's coefficient
Active Contour (Chan-Vese)	0.5667 ± 0.2052	0.7271 ± 0.156
DRLSE	0.518 ± 0.11	0.675 ± 0.099
SLIC + DRLSE	0.807 ± 0.067	0.892 ± 0.043
Proposed method	0.8689 ± 0.042	0.9169 ± 0.02

The intraclass correlation coefficient (ICC) assesses rating reliability by comparing the variability of different ratings of the same subject (e.g., CT scan) to the total variation across all ratings and all subjects. It is a descriptive statistic used for the assessment of consistency or reproducibility of quantitative measurements made by different observers measuring the same quantity and generally focuses on how well the observers' scores matched the ground truth. For pooled variance $\sigma^2(w)$ within subjects and variance of traits $\sigma^2(b)$, the ICC is given as:

$$ICC = \frac{\sigma^2(b)}{\sigma^2(b) + \sigma^2(w)} \qquad (8)$$

The intraclass correlation coefficient (ICC) was calculated using MedCalc®to compare the results with experts to ascertain reliability of measurements. Linear measurements from experts and our algorithm show ICC range from 0.91 to 0.99 and the results are given in table 3. The identification of the cisterns showed an ICC of

≥ 0.91 which can be improved by adding shape constraint to identify supra cellar and the quadrigeminal cisterns more accurately.

Table 3. : Identification of ROI on the axial CT scans.

Region of Interest	ICC
Supra cellar cistern	0.91
Lateral ventricles	0.97
Intracranial haematoma	0.99

The optimum testing performance of the feedforward ANN was 88.9% to segment the image. Increasing or decreasing the number of hidden layers or neurons in ANN did not significantly improve the performance and the average performance was $87.4 - 88.9\%$ in training and testing. The technique was also evaluated using pattern net with $87.2 - 88.0\%$ performance, while the cascade-forward net had average performance of $84.1 - 87.2\%$.

A *t-test* is a statistical hypothesis test following Student's t distribution if the null hypothesis is supported. The *paired t-test* considers the difference between paired values in observations from the algorithm and expert, estimates the variation of values within each and produces a single number known as a *t-value*. It represents if two sets of observations are significantly different from each other resulting in the acceptance or rejection of the null hypothesis. It is calculated as:

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad (9)$$

where $\bar{x}_1, s_1$ and n_1 are mean, standard deviation and number of observations in algorithm's output and $\bar{x}_2, s_2$ and n_2 are mean, standard deviation and number of observations in expert's annotations, respectively. The Δ is the hypothesized difference between the means. In the experiments, $n_1 = n_2$ and the $\Delta = 0$ are used.

Table 4. : Estimation of Haematoma Length with different methods in brain CT Scans. The $\Delta = 0$ at $\alpha = 0.05$.

Groups	Mean	Std. Dev.	Std. Err.	t
Expert	7.849	2.053		
Proposed Method	7.876	1.970		
Difference	-0.027	0.336	0.047	-0.583
Chan-Vese	9.014	2.344		
Difference	-1.165	1.211	0.170	-6.866
Edge	7.179	2.607		
Difference	0.670	1.736	0.243	2.754

Table 5. : Two Tail t-test on estimation of Haematoma Length with different active contour methods.

Method	p-value	t-crit	lower	upper	Sig.
Chan-Vese method	9.7E-09	2.0E+00	-1.5E+00	-8.2E-01	yes
Edge method	0.008	2.009	0.181	1.158	yes
Proposed method	0.563	2.009	-0.122	0.067	no

The paired t-test evaluation of the proposed method against experts' annotation and state of the art methods is given in Tables 4 - 7. The null hypothesis assumes that there is no statistically significant

Table 6. : Estimation of Haematoma Breadth with different methods in brain CT Scans. The $\Delta = 0$ at $\alpha = 0.05$.

Groups	Mean	Std. Dev.	Std. Err.	t
Expert	3.745	1.206		
Proposed Method	3.755	1.241		
Difference	-0.010	0.259	0.036	-0.270
Chan-Vese	4.380	1.112		
Difference	-0.635	0.751	0.105	-6.039
Edge	3.341	1.213		
Difference	0.404	0.676	0.095	4.266

Table 7. : Two Tail t-test on estimation of Haematoma Breadth with different methods.

Method	p-value	t-crit	lower	upper	Sig.
Chan-Vese method	1.9E-07	2.0E+00	-8.5E-01	-4.2E-01	yes
Edge method	8.8E-05	2.0E+00	2.1E-01	5.9E-01	yes
Proposed Algorithm	0.788	2.009	-0.083	0.063	no

difference ($\Delta = 0$) in the measurements between algorithms and experts. The results of estimating the measurements of the major and minor axes of the haematoma ellipse show that the state of the art Chan-Vese and Edge models have statistically significant difference in the measurements. This rejects the null hypothesis for the state of the art. The performance of the proposed method has no significant difference from the hypothesized mean and, hence, the null hypothesis is not rejected.

The experiments were performed using Matlab® R2013b on Mac OS X and Windows 7 platforms. The Jacobian training of ANN is not supported on GPU in Matlab R2013b and in those cases, only CPU was used.

3.1 Correlation to the Marshall CT Classification

The information gathered from the clinical and imaging profile of the patient is used by the doctors to ascertain the diagnosis, plan treatment and most importantly, establish the prognosis. From the tables 1 and 3, it can be seen that the method can efficiently detect ROI for use in Marshall et al. CT Classification. The significant features: cisterns and ventricles, midline shift and haematomas are identified and measured automatically by the method. The compressed or absent cisterns are seen in Marshall et al. ratings II and III while midline shift classifies IV and above. The presence of mixed or high density lesion $> 25cm^3$ is classified as rating VI. When these characteristics are not identified, the CT would fall in rating I or normal. The performance evaluation is given in Table 8.

Table 8. : Classification of the brain trauma CT according to Marshall et al. CT Rating.

Marshall CT Class	Feature	ICC
I or II	Cisterns and Ventricles	0.91 and 0.97
III or IV	Midline Shift	0.98
V or VI	Lesion $> 25cm^3$	0.99

Since the experiments were conducted on the cross-sectional, non-contrast CT studies performed at the time of admission of within

early hours of injury, none of the patients had the mass lesions removed at that time. Hence, no CT with Rating V were used in the experiments.

4. CONCLUSION

Manual demarcation of ROI by expert radiologists is still considered gold standard, however, the proposed approach has shown plausible results with $87.4 - 88.9\%$ accuracy with ANN in segmenting the different ROI's rquired for Marshal et a. CT classification. JI of 0.8689 ± 0.042 was achieved with feedforward ANN and active contours in segmenting the haematoma. Correction of orientation of head and detection of midline shift $> 5mm$ showed sensitivity of 85% and 80% and specificity of 75% and 86% respectively. The midline shift is the characteristic of Marshall et al. CT classes IV and above and, hence, quantifying it can significantly help the doctors in the emergency department.

The ICC results were ≥ 0.91 compared with experts in identifying cisterns as regions of interest and ≥ 0.97 for ventricles and haematomas. As the manual annotations were performed be expert radiologists, the ICC performance of the method can be quite useful for the junior doctors in the reducing the inter-observer variability. It can also be seen that the hybrid approach adapts well to diverse image characteristics and the performance is very close to the doctor's intuition compared to the state of the art active contours method as shown by the paired t-test results. The combination of ANN with active contours gives significantly improved segmentation than using SLIC with level sets.

Interpretation of the medical images by the radiologists is fraught with errors and variations which represent the weakest aspect of clinical imaging. Research efforts to correlate image features to clinical outcomes are imperative for differential diagnosis, prognostic assessment, pre-surgical mapping and treatment planning in patients, these endeavours are fettered by the time and effort required to manually report clinical findings. Hence, it is proposed that integration of automated systems in clinical and research settings can significantly reduce the radiologists' workload and inter-observer and intra-observer variability.

5. REFERENCES

[1] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels. Technical report, 2010.

[2] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Susstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(11):2274–2282, 2012.

[3] HS Bhadauria, Annapurna Singh, and ML Dewal. An integrated method for hemorrhage segmentation from brain ct imaging. *Computers & Electrical Engineering*, 39(5):1527–1536, 2013.

[4] Tony F Chan and Luminita A Vese. Active contours without edges. *Image Processing, IEEE Transactions on*, 10(2):266–277, 2001.

[5] Anastasia Dubrovina and Ron Kimmel. Active contours for multi-region image segmentation with a single level set function. Technical report, Technion technical report CIS-2012-06, 2012.

[6] Murray Goldstein. Traumatic brain injury: a silent epidemic. *Annals of neurology*, 27(3):327–327, 1990.

[7] Igor Kononenko. Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in medicine*, 23(1):89–109, 2001.

[8] Rashmi U Kothari, Thomas Brott, Joseph P Broderick, William G Barsan, Laura R Sauerbeck, Mario Zuccarello, and Jane Khoury. The abcs of measuring intracerebral hemorrhage volumes. *Stroke*, 27(8):1304–1305, 1996.

[9] David Leonhardt. Why doctors so often get it wrong. *The New York Times*, 2006.

[10] Lawrence F Marshall, Sharon Bowers Marshall, Melville R Klauber, Marjan van Berkum Clark, Howard M Eisenberg, John A Jane, Thomas G Luerssen, Anthony Marmarou, and Mary A Foulkes. A new classification of head injury based on computerized tomography. *Special Supplements*, 75(1S):S14–S20, 1991.

[11] CA McLauchlan, K Jones, and HR Guly. Interpretation of trauma radiographs by junior doctors in accident and emergency departments: a cause for concern? *Journal of accident & emergency medicine*, 14(5):295–298, 1997.

[12] KN Bhanu Prakash, Shi Zhou, Tim C Morgan, Daniel F Hanley, and Wieslaw L Nowinski. Segmentation and quantification of intra-ventricular/cerebral hemorrhage in ct scans by modified distance regularized level set evolution technique. *International journal of computer assisted radiology and surgery*, 7(5):785–798, 2012.

[13] Adnan N Qureshi. A framework for segmentation and estimation of intracranial measurements in ct scans. In *Biomedical Engineering Conference (CIBEC), 2014 Cairo International*, pages 129–132. IEEE, 2014.

[14] Adnan N. Qureshi and Vitaly Schetinin. Computer-aided segmentation and estimation of indices in brain CT scans. In *Medical Image Understanding and Analysis - MIUA 2014. 18th Annual Conference, Egham, UK, July 9-11, 2014. Proceedings.*, pages 161–166, 2014.

[15] Saatman, Duhaime, Bullock, Maas, Valadka, Manley, Workshop Scientific Team, and Advisory Panel Members. Classification of Traumatic Brain Injury for Targeted Therapies. *Journal of Neurotrauma*, 25:719–738, 2008.

[16] Mohammadreza Soltaninejad, Tryphon Lambrou, Adnan N. Qureshi, Nigel M. Allinson, and Xujiong Ye. A hybrid method for haemorrhage segmentation in trauma brain CT. In *Medical Image Understanding and Analysis - MIUA 2014. 18th Annual Conference, Egham, UK, July 9-11, 2014. Proceedings.*, pages 99–104, 2014.

[17] Hilaire J Thompson, Wayne C McCormick, and Sarah H Kagan. Traumatic brain injury in older adults: epidemiology, outcomes, and future implications. *Journal of the American Geriatrics Society*, 54(10):1590–1595, 2006.

[18] Sreenath Rao Vantaram and Eli Saber. Survey of contemporary trends in color image segmentation. *Journal of Electronic Imaging*, 21(4):040901–1, 2012.

[19] Li Wang, Chunming Li, Quansen Sun, Deshen Xia, Chiu-Yen Kao, et al. Active contours driven by local and global intensity fitting energy with application to brain mr image segmentation. *Computerized Medical Imaging and Graphics*, 33(7):520, 2009.

[20] Anthony Yezzi Jr, Satyanad Kichenassamy, Arun Kumar, Peter Olver, and Allen Tannenbaum. A geometric snake model for segmentation of medical imagery. *Medical Imaging, IEEE Transactions on*, 16(2):199–209, 1997.