Intensity population based unsupervised hemorrhage segmentation from brain CT images

Soumi Ray [a,*], Vinod Kumar [a], Chirag Ahuja [b], Niranjan Khandelwal [b]

[a] Biomedical Engineering Lab, Department of Electrical Engineering, Indian Institute of Technology Roorkee, Roorkee 247667, Uttarakhand, India
[b] Department of Radiodiagnosis, Post graduate Institute of Medical Education and Research, Chandigarh, India

ARTICLE INFO

ABSTRACT

This article has proposed an intelligent knowledge driven method to segment hemorrhage from brain CT images using the information of pixel intensity population and distribution. A mathematical model is designed to identify the unexpected variation in pixel intensity population in a brain CT image having hemorrhage. Complete batch of multi-slice CT scan images is taken as input. Fusion of knowledge of brain anatomy with intensity distribution information of CT brain image results in a unique solution for hemorrhage segmentation. To test the robustness, segmentation of different types of hemorrhage of different patients is done using the proposed method. The results are accepted and validated by radiology experts. A fully automatic and fast Computer Aided Diagnosis (CAD) is designed, using the proposed method, to segment hemorrhage automatically, in the absence of an expert, for further inspections like checking severity, volume, size, shape and type of hemorrhage. Competence of the CAD is tested against mostly used established clustering methods to demonstrate its potential.

1. Introduction

Brain hemorrhage or intracranial hemorrhage (ICH) is an unfortunate incident. Due to leakage or rupture of one or more blood vessels, blood accumulates in the local vicinity (Quain, 1883). Each year, brain injuries account for thousands of deaths and a significant number of people suffer temporary and permanent disability (Wedro, 2015). Among several reasons, the most common reason for brain injury is head injury due to accidents, though every head injury is not necessarily a brain injury. In accidental cases, bleeding in the brain usually occurs at the time of injury. However, symptoms may develop immediately or progress gradually over time depending on the severity of an impact. Symptoms of general hemorrhage cases can arise anytime and do not last for long in preliminary stages, but with time get worse. Immediate medical care should be sought for the patients not fully awake after an injury. Some hemorrhages can be treated by medication, some may require neurosurgery to remove blood clots and relieve pressure on the brain to save lives.

In detection and diagnosis of hemorrhage, CT imaging modality plays a very important role because of its speed, low cost and capability of capturing reasonably good contrast images for precise investigation (Al-Ayyoub, Alawad, Al-Darabsah, & Aljarrah, 2013). It also has advantages over MRI for the patients having ferromagnetic or electrical implants or claustrophobia. These advantages of CT imaging over MRI make the choice of imaging modality easier (Adams et al., 2007; Balasooriya & Perera, 2012; Chawla, Sharma, Sivaswamy, & Kishore, 2009; Li, Hu, Wu, & Chen, 2009). In CT images the hemorrhage patches are captured as bright homogeneous areas within the brain. Whatever the behavioral symptoms may be, for a radiologist the symptom lies in image texture. Presence of any type of hemorrhage is described in terms of pixel intensity. In brain scan images, the presence of some bright, easily detectable homogeneous spot denotes the hemorrhage. The texture and intensity of different types of hemorrhage are very similar.

The number of hemorrhage cases is significant in a nowadays scenario. The rate of head injury in India is very high and as per Indian Head Injury Foundation, it is the highest in the world. Here, more than 100,000 dies every year and over 1 million suffer from serious head injuries. In India, 1 out of 6 trauma patients dies, whereas in the United States the proportion is only 1 out of 200 (Admin, 2014). It has been experienced that expert radiologist per hospital sometimes is not sufficient to attend all the cases on time. In some remote areas, even availability of radiologist during an emergency is critical. On contrary, accurate measurement and flawless detection of hemorrhage are not easy for non-specialist radiologists and practitioners (Admin, 2014). We also cannot neglect the human error factors due to fatigue, rush or some other reasons.

* Corresponding author. Present address: E101, Saraswti CHS, Thakur village, Kandivali (east), Mumbai 400101, India.
E-mail address: soumidee@iitr.ac.in (S. Ray).

To overcome those difficulties a fast, lightweight and a sufficiently accurate method for detection of hemorrhage is proposed in this article. Using this method a computer aided diagnostic (CAD) system is developed to support clinical diagnosis system.

2. State of the art

Several CAD systems for brain image analysis and diagnosis are designed and presented in different research articles for decades. CT scan is used to see through the brain to know whether there is any ICH or not. Because of significant gray level difference, hemorrhages are detectable by software programs from CT scan images.

Majority of reported research works have more or less similar approach for CT image analysis and brain hemorrhage detection. Region based approach is proposed by several researchers (Bardera et al., 2009; Gong et al., 2007; Liu, Yuan, Liu, Li, & Yin, 2008; Loncaric et al., 1999; Pérez, Valdés, Guevara, Rodríguez, & Molina, 2007), whereas one research has done segmentation of brain into different objects and then analysis of each object is carried out by feature extraction (Balasooriya & Perera, 2012). Liu et al. and few others proposed a hemispheric symmetry detection approach to detect brain lesions (Chan, 2007; Datta, Datta, & Biswas, 2011; Hara et al., 2007; Liu et al., 2003; Liu, Rothfus, & Kanade, 1998). This method checks the symmetry for each slice and any abnormal region at one side referred as hemorrhage. In some literature, midline (the line to segment brain into two hemispheres) shift is considered as an indication of hemorrhage (Chan, 2007; Hara et al., 2007). On contrary, Yuh et al. calculated midline shifting as supportive finding to establish the existence of hemorrhage (Yuh, Gean, Manley, Callen, & Wintermark, 2008).

Hemorrhage finding is mostly done by locating threshold on the histogram. This is done automatically through machines artificial intelligence using different algorithms like fuzzy C-means (Cheng & Cheng, 1998; Li et al., 2009), Hopfield neural network (Cheng et al., 1996; Lin, Cheng, & Mao, 1996), maximum entropy (Li et al., 2009), fuzzy maximal likelihood estimation (Loncaric et al., 1999), probabilistic neural network (Cheng & Cheng, 1998), genetic algorithm (Liu et al., 2008). To detect hemorrhage few more approaches to knowledge based detection (Cosic & Loucaric, 1997), histogram-based K-means clustering (Dhawan, Loncaric, Hitt, Broderick, & Brott, 1993), wavelet-based texture analysis (Liu et al., 2008) are also proposed in different works. Unsupervised clustering and backtracking tree search method are also reported in some papers (Lončarić et al., 1996; Loncaric, Cosic, & Dhawan, 1996). As hemorrhage is detected by searching hyperdense region within the image (Dhawan et al., 1993), no image registration is required; still, one literature has presented images registration (Chan, 2007) between two hemispheres. Pre-processing of images using morphological operation (Chan, 2007), and median filter (Li et al., 2009; Pérez et al., 2007) has also been included in few research works.

In recent time deep learning methods are also in use for segmentation (Brosch et al., 2016; Ghesu et al., 2016; Kalinovskya, Liauchuka, & Tarasaub, 2017; Pereira, Pinto, Alves, & Silva, 2016). Most of them are supervised techniques which require a large number of data. This requirement itself is a disadvantage of these methods (Greenspan, van Ginneken, & Summers, 2016; Ravì et al., 2017). Different deep learning techniques are outperformed by the traditional *k*-means method used in an unsupervised system which involves a limited number of data (Dundar, Kou, Zhang, He, & Rajwa, 2015).

Though there are few CAD systems already reported, room for more research is still open (Maduskar & Acharyya, 2009) due to lack of accuracy in the segmentation of reported systems, complex coding, difficulties in integrity with real-time machines, less user-friendliness. In this article, a simple, fast and robust CAD is proposed for hemorrhage segmentation from intracranial brain re-
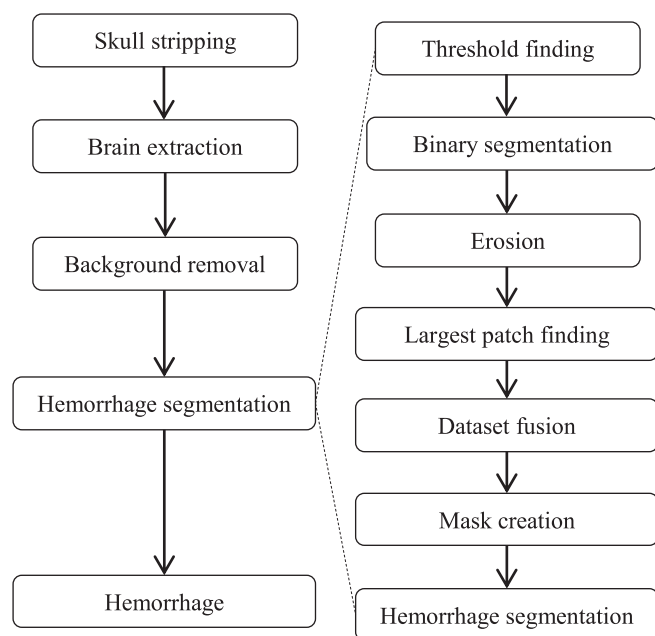

Fig. 1. CAD process flow.

gion. Outline of CAD process flow after reading input dataset is presented in Fig. 1. The intensity and density of the hemorrhage pixels within brain CT images are explored to locate and segment out hemorrhage for further diagnostic requirements.

3. Materials and methods

3.1. Datasets

Datasets of multi-slice brain CT images are taken for hemorrhage detection. Images are scanned by a 64 slice CT scanner with slice thickness 5 mm. The original DICOM (Digital Imaging & Communication in Medicine) images, provided by radiology department of Postgraduate Institute of Medical Education and Research (PGIMER), Chandigarh, India, are stored and saved in Pentium based desktop computer having 64-bit Windows operating system. Stored images are converted to 8-bit TIFF (Tagged Image File Format) images by open source ImageJ software and used for further analysis. Dimension of each image is 512 × 512. Dataset of 22 patients', having total 590 slices, are considered for this work. The number of slices is different for different dataset. Age, sex or gender has not taken into account during segmentation. CAD is designed in 32-bit LabVIEW (Laboratory Virtual Instrumentation Engineering Workbench) platform.

In this article, the word 'dataset' presents all scanned CT slices of a patient in one sitting. The number of slices can vary depending on machine, practitioner's decision and machine capability. And the entire data used in this work is referred as 'database'.

3.2. CT image intensity distribution

CT images are acquired using x-ray exposure. Absorption amount of x-ray is different for different tissues. The complete range i.e. from no absorption to complete absorption is presented using Hounsfield Unit (HU) scale where air is presented by −1000 HU, water as 0HU and bone has the highest value from several hundred to +3000HU. HU is then mapped to gray scale of intensity range [0255] to display image digitally. Depending on practitioners' requirement a range of HU is emphasized and rest is suppressed to nearest boundary value. This method is called windowing. The
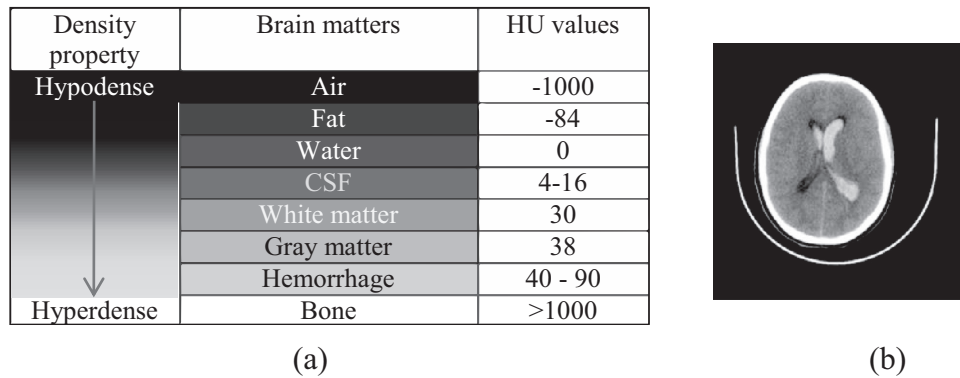
Density property	Brain matters	HU values
Hypodense	Air	-1000
	Fat	-84
	Water	0
	CSF	4-16
	White matter	30
	Gray matter	38
	Hemorrhage	40 - 90
Hyperdense	Bone	>1000

(a)

(b)

Fig. 2. (a) Intensity varies with density in CT image (b) CT image.

window is defined using two parameters window width (ww) and window level (wl). Window width is the complete length of HU which is considered to be mapped into gray level. Window level is the mid-HU value of that width.

All types of hemorrhage introduce bright homogeneous spots in the brain CT image. Depending on the amount of collected blood the brightness varies. More dense layer offers brighter spot in image. On the other hand, the brightness of hemorrhage is always significantly lower than that of bone or skull. To detect hemorrhage, brain window scanning mode is used. In brain window mode default wl value is 40 HU (Ee, Sim, Teh, & Ting, 2016). Value of ww is different (100 or 400) for different datasets in use.

Fig. 2(a) presents the gradual change in image pixels' intensity hence HU values (Ros & Mortele, 2007; Zimmerman, Gibby, & Carmody, 2012) with the change in density of different parts of brain. The example of a brain CT image is shown in Fig. 2(b). The skull and headrest are presented by the brightest part of the image where hemorrhage patches are significantly brighter than surroundings but satisfactorily low from skull intensity. The transition from one part to another part offers gradual change in pixel intensity making the boarder-less significant.

3.3. Pre-processing of data

Before segmentation, pre-processing of CT images is important to increase sensitivity and accuracy of CAD system. For any CT image, the background and skull which offers no clinical information are respectively the darkest and brightest part of the image. The intensity value remains mostly zero for background. Skull, some other external parts like headrest, disease like calcification offers very high intensity region in the image. These areas contain no information about hemorrhage. Sometimes soft tissue edema gets imaged during scan. These parts introduce noise in segmentation and thresholding process due to its matching intensity level with brain matters and/or hemorrhage. Removal of all these non-relevant parts from image is required before processing data for hemorrhage segmentation.

Each digital image is basically a 2D numeric array which can readily be used for mathematical computation. High volume background and high intensity skull removal reduce load in further data computation. Designed CAD dynamically selects threshold for background and skull of each image slice from its intensity population curve. Histogram shown in Fig. 3(a) has a large peak at lower intensity level which represents the highly dense low intensity volume i.e. background and another large peak near maximum available intensity representing the skull. None of these two areas contains any brain or disease relevant information. The 3rd highest peak in histogram represents the population of brain matter.

3.3.1. Skull stripping

Knowledge driven thresholding for skull is done. Observation of histograms of large number of brain CT scan data, yields two important facts, – (i) the population of intensity becomes almost even before high volume skull intensity and (ii) in any 8-bit gray scale – CT image major skull information is confined within the highest 1/3rd intensity range of available gray intensity range of the image. So to compute skull threshold highest 1/3rd intensity range of respective images is considered. In the population curve, absolute slope of each intensity level with respect to the next level is calculated. The most even part nearest to the high volume skull is considered as the boundary between brain and skull. Intensity value that offers minimum value of function $f(x)$, described in Eq. (1), denotes the threshold for skull. $f(x)$ is a function of slope weighted by a distance cost factor, d, where the distance is the difference between the highest available intensity and intensity of corresponding slope.

$$f(x) = \frac{\partial P}{\partial i} * d \tag{1}$$

∂P and ∂i presents deviation in population and intensity respectively at the point of measurement;
at ith instant, $\partial p = |p_{i-1} - p_i|$ and $\partial i = |I_{i-1} - I_i|$

The resulted threshold value is plotted on the graph shown in Fig. 3(a) along with population curve of brain image shown in Fig. 3(b). The skull removed image is also shown in Fig. 3(c). Threshold point is calculated for each slice of a dataset separately.

3.3.2. Anatomy based brain segmentation

After extracting the threshold value, binarization is done to segment brain from the slice. Each slice is converted into binary image as shown in Fig. 3(d) having two levels – bright level 1 is skull and dark level 2 is non-skull parts. In level 2, along with brain, other non-brain parts and image background are also included. As per human brain anatomy, brain must be encapsulated within the skull. So all information at level2 which are not encapsulated by level1 is identified and converted to level1. Remaining area of level 2 can have one or more separate region(s).

A query runs in new binary image to find largest connected encapsulated region. Because of its speed (Cabaret, Lacassagne, & Oudni, 2014), LSL$_{STD}$ method (Lacassagne & Zavidovique, 2009) is used to find target area. LSL$_{STD}$ method uses run length encoding to do line-relative labeling. Each row is checked separately to mark the image segments and backgrounds. Odd numbers are used to label image segments and even numbers for background. After labeling each row adjacent odd labels are searched to tag under one absolute label. All non-adjacent odd labels and all even labels are rejected.
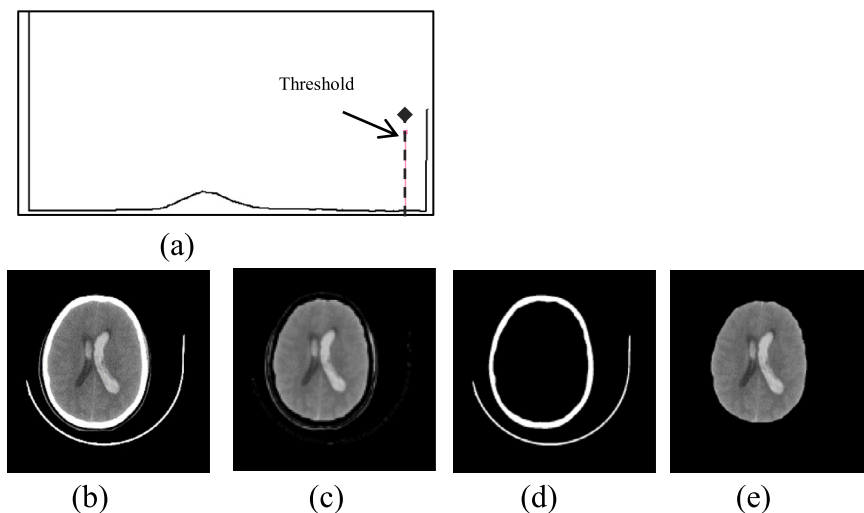
Fig. 3. (a) Intensity histogram with skull threshold, (b) original CT image (c) skull removed image (d) removed skull (binary image) (e) segmented brain.
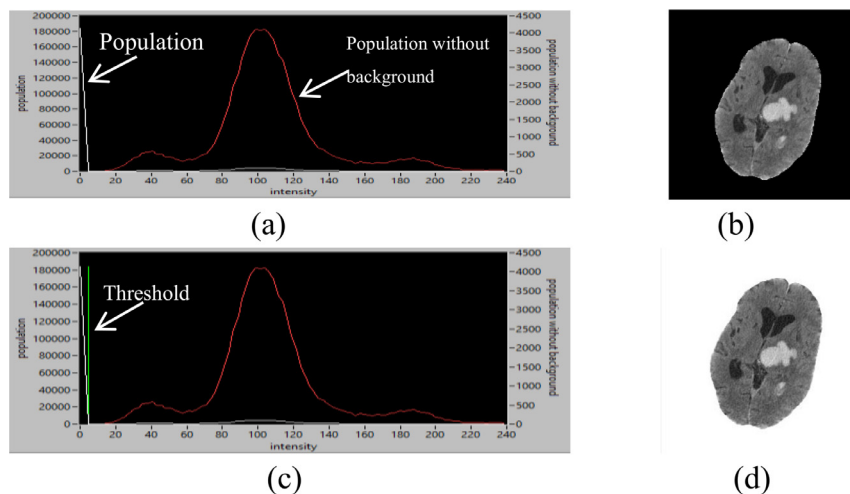

Fig. 4. (a): population of intensity with and without background (b) image under test (c) background threshold (d) background removed image.

Identified connected segment is stored and the search is repeated to find any other such connected segments. Once one segment is identified, before using the image for the next search, the identified segment is converted into the background. This modified image is then used as input of the next search. The method continues until all the segments are identified. Among all identified connected segments the largest connected region represents the brain.

Now a binary mask is created by converting value of pixels of largest connected area – to 1 and rest to 0. When the corresponding skull removed image is multiplied by the mask, brain gets segmented as shown in Fig. 3(e). Each slice thus returns image having brain where all non-brain parts along with skull are converted to background i.e. 0. The new segmented dataset will be in discussion from now onwards. All processing and analysis will be done on this dataset rather than the original dataset.

3.3.3. Background removing

This new dataset has very large pixel density at background intensity because of converting all non-intracranial information to background. The population curve of image shown in Fig. 4(b) is plotted in Fig. 4(a). 1st peak which is the largest one in this population curve represents background. This peak is detected by CAD. In all clinical brain CT images, there is a sudden change in population after this peak. This change offers a very high slope. In between 1st peak and 2nd highest peak, the highest population deviation with respect to 1st peak is calculated and used to determine the background intensity threshold as shown in Fig. 4(c). The resultant image output is Fig. 4(d). In short, the threshold value is the intensity at which the intensity population deviation from highest peak, created by background, is maximum. This value must lie between two highest population peaks of intensity histogram of a skull removed brain image. It has been considered that all intensity values other than background i.e. above background threshold are carrying clinical brain information. In the proposed work, hemorrhage is detected from these background removed pre-processed dataset by image analysis.

3.4. Segmentation of hemorrhage by population histogram

To segment hemorrhage from brain CT images, thresholding technique is used to find the target pixel cluster. We have considered two clusters – brain and hemorrhage. Cut-off between these
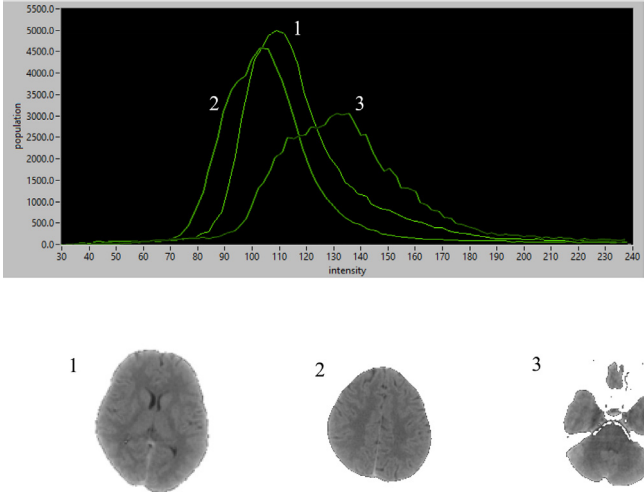
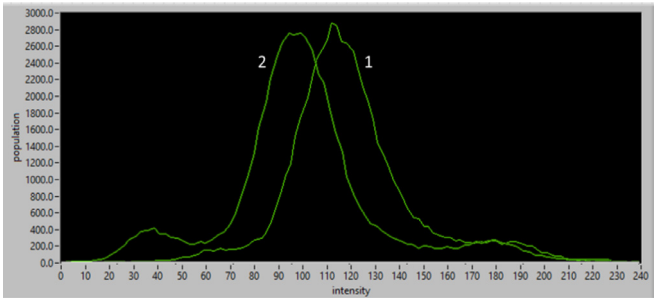**Fig. 5.** Intensity population distribution of image 1, 2, 3.

Fig. 6. Intensity population for slices with hemorrhage.

two clusters i.e. threshold for hemorrhage is computed from image intensity population i.e. histogram analysis.

Pixel distribution of CT brain image is analyzed based on density of pixels at different intensity level. Fig. 5 is presenting population of intensity of three background as well as skull removed normal brain CT images shown in Fig. 5(1–3) acquired from three different subjects. Depending upon the actual slice volume of a CT image, amplitude and position of maximum population can vary; but the pattern of population distribution remains unaltered. The maximum population due to brain matter happens near the middle of the complete intensity range. At higher intensity range the population reduces gradually and become very low near the end. Similar plots of patients' data (shown in Figs. 3(d) and 4(d)) are shown in Fig. 6. It has unchanged pattern for the initial part till the decreasing slope of highest peak, but a small bell shaped new peak is introduced due to hemorrhage which increases population in high intensity range. Threshold for hemorrhage lies near the end of downward slope of highest population peak and before the upward slope of newly introduced population peak. This work has proposed dynamic threshold finding to identify hemorrhage.

Expected population (eP) is the population predicted from past population values. The value is calculated by taking moving average of available past i.e. the lower intensity levels' density information for each instance as described in Eq. (2). Wide change in population pattern occurs twice creating greater slopes around the highest population peak. These changes differ significantly from the expected population change depicted from the history of pop-
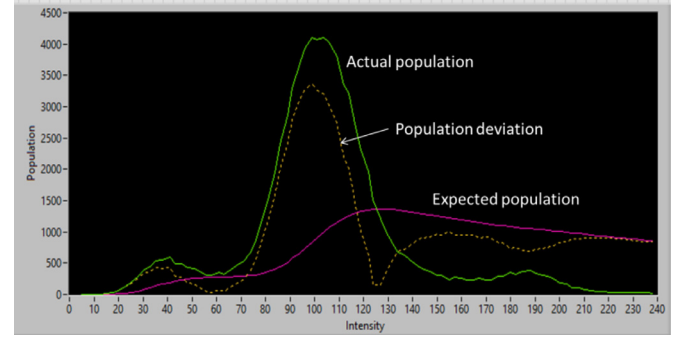

Fig. 7. Actual intensity population, expected intensity population and deviation of actual from expected intensity population.

ulation of previous intensities.

$$expected\ population\ eP_{i+1} = \frac{\sum_0^i P}{n} \tag{2}$$

where value for $(i + 1)$th intensity level is calculated from all n past intensity level [0 to i] values.

The actual population and the expected population are compared in Fig. 7 for CT slice shown in Fig. 4(d). Actual population reaches the highest peak due to large no of pixels of brain matter. The expected curve does not follow the actual curve towards that peak. It increases its value slowly and reaches highest value at an intensity which is higher than the intensity at which actual population reaches highest value. After that intensity value, expected population decreases constantly but with very poor slope. It meets the actual population value and offers zero deviation. The value of expected population remains higher than actual population beyond the zero deviation point, due to very high peak value of actual population.

A deviation curve, calculated using Eq. (3), is plotted to present the difference between actual and expected population. It is shown as a dotted line in Fig. 7. Deviation reduces to minimum value as actual approaches to expected population. Beyond that intensity deviation increases due to continuous fall of actual population and very low change in expected population.

$$deviation\ D_i = |P_i - eP_i| \tag{3}$$

This calculated deviation curve presents the overflow or vacancy in pixel availability in actual population with respect to the predicted or expected population. Higher actual population symbolizes overflow due to higher pixel availability than expected and the reverse represents vacancy due to deficiency in pixel population. The point where actual population meets expected population is saturation point as the pixel density of actual population is satisfied with the expected value. The situation becomes critical where deviation becomes higher than actual population. This is happen due to very low strength of actual population with respect to expected population. That switching point on population graph is mentioned as upset point in this work. This upset point intensity is considered as threshold for hemorrhage in a brain CT scan.

Mathematically that threshold intensity value is calculated by taking *Population to Deviation Ratio (PDR)* stated in Eq. (4). Actual population is divided by deviation of it from its expected value at each point. At the upset point, *PDR* becomes less than 1 as the information available in image is lower than deviation.

$$Population\ to\ deviation\ ratio,\ PDR = \frac{P}{D} \tag{4}$$

Where P presents the population or histogram data and D present the deviation curve. Deviation curve D can be calculated
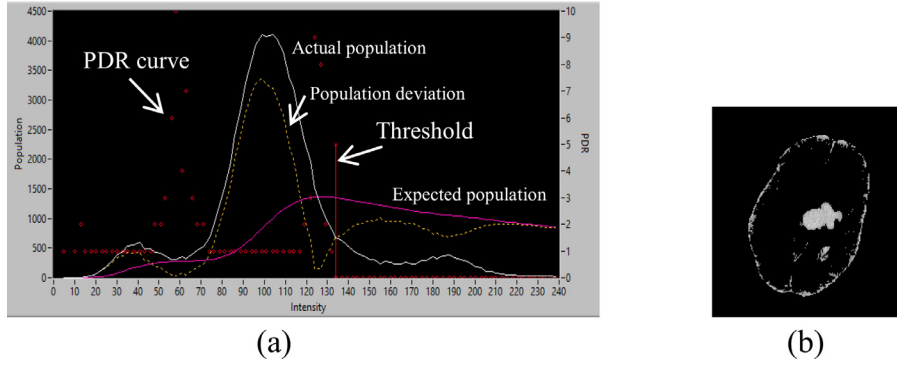
Fig. 8. (a) PDR and threshold (b) segmented image.

as follows –

$$D = |P - eP| \qquad (5)$$

$$\therefore PDR = \frac{P}{|P - eP|} \qquad (6)$$

$$= \frac{1}{\left|1 - \frac{eP}{P}\right|} \qquad (7)$$

The nature of *PDR* is dominated by the factor $\mu = eP/P$ as $\mu > 0$ for $eP < P$, $\mu = 0$ for $eP = P$, $\mu < 0$ for $eP > P$. As absolute of denominator is taken, *PDR* will remain positive non-zero for any other conditions than $eP = P$.

At upset point, *PDR* will switch from $PDR \geq 1$ to $PDR < 1$ as $D > P$ at that point. For $PDR < 1$ Eq. (7) can be written as

$$\frac{1}{\left|1 - \frac{eP}{P}\right|} < 1 \; or, \; \left|1 - \frac{eP}{P}\right| > 1$$

$$\qquad (8)$$

$$so \; it \; can \; be \; rewritten \; as \; \frac{eP}{P} - 1 > 1 \; or, \; eP > 2P$$

So the threshold for hemorrhage in a brain CT image is that intensity at which expected population becomes higher than twice of actual population.

As per anatomy, hemorrhage intensity is always higher than brain matter intensity. So, to reduce computational load and to avoid any unexpected error, the threshold search is limited in the intensity region higher than the highest density peak in histogram. After highest peak in population curve, the 1st upset point intensity is taken as threshold for hemorrhage. PDR along with actual population, expected population and deviation curves are shown in Fig. 8 for Fig. 4(d). The upset point or threshold intensity for hemorrhage is shown here by a straight line. The segmented image with hemorrhage keeps all values above threshold as they were and rest are converted to background as shown in Fig. 8(b).

3.5. Fusion for location finalization

The segmented image shown in Fig. 8(b) contains hemorrhage as well as some non-hemorrhage skull adjacent parts. To improve segmentation result segmented image gone through single iteration erosion. This morphological operation shows high impact on deleting skull left out parts as shown in Fig. 9(a) but also affects the hemorrhage patch. To get rid of this effect, first the largest connected part in morphological output is identified as hemorrhage and then any available non-zero information attached to that area in the pre-erosion image is reconsidered for reconstruction of hemorrhage. The final hemorrhage segmentation result is shown in Fig. 9(b).
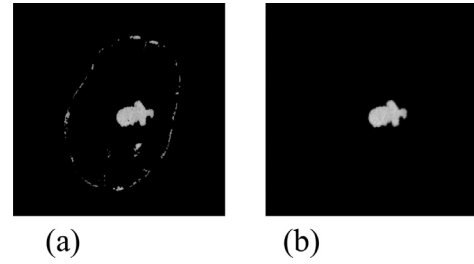

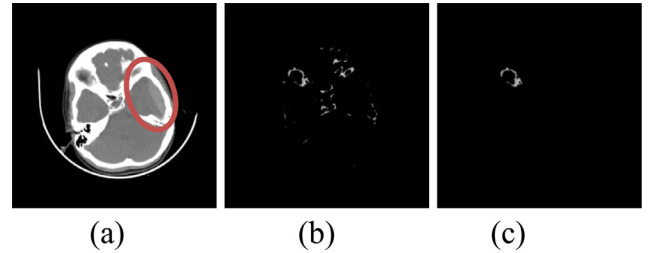Fig. 9. (a) Morphological operation (b) reconstruction.

Fig. 10. (a) Original image (hemorrhage is marked by circle) (b) segmented image (c) output after morphological operation.

Sometimes in some slices, due to other compact larger area of matching intensity, the actual hemorrhage not gets detected by the CAD as shown in Fig. 10. In Fig. 10(a) hemorrhage is marked by a circle. In Fig. 10(d), we can see only a few scattered pixels are segmented form hemorrhage area. The largest connected area shown in Fig. 10(c) is not a part of actual hemorrhage. To solve the problem, binarization of all segmented slices of a dataset is done. Then the images are stacked up one after another and fused linearly. Correctly segmented slices offer overlapping area whereas wrongly segmented slices result into different scattered area.

After the linear fusion, the highest strength area is selected as hemorrhage zone. This hypothesis is elaborated by a 6 × 6 matrix here.

Say, dataset D has total 5 image slices each of which is a 6 × 6 matrix. I_1 to I_5 are the image slices after segmentation and morphological operations, where *0* represents deleted brain, non-brain and background part *f* represents high intensity brain part, *h* is the
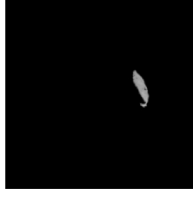
Fig. 11. Segmented hemorrhage.

hemorrhage.

$$
I_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 & f & 0 \\ 0 & 0 & 0 & f & f & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
I_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 & f & 0 \\ 0 & h & 0 & 0 & f & 0 \\ 0 & h & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

$$
I_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & f & 0 \\ 0 & h & h & 0 & 0 & 0 \\ 0 & h & 0 & 0 & f & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
I_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 & 0 & 0 \\ 0 & h & h & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

$$
I_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 & 0 & 0 \\ 0 & h & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

The largest part of each image is highlighted using bold font. Now, all pixels in largest parts are converted to value 1 and rest to 0.These binary images are then added linearly and resultant image I becomes –

$$
I = \sum_1^5 I_{g_i}
$$

$$
= \left\{ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \middle| f = 1 \ for \ I_1, \ h = 1 \ for \ other \ than \ I_1 \right\}
$$

Where Ig presents segmented binary images. The strength of two areas is evaluated as S_h and S_f.

$$
S_h = \sum \{ I | f = 0 \} = \sum \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = 11
$$

and similarly $S_f = 3$.

For final hemorrhage segmentation, pixel locations of highest strength area are saved to create a mask. Each segmented slice is then tested against that mask to see if any information in that area or in adjacent areas is available. All positive response locations are selected as hemorrhage candidate for the respective slice. The success of this backpropagation information collection is presented in Fig. 11 for the image shown in Fig. 10(a). It shows that the largest area selected in Fig. 10(c) is cancelled and the actual area marked in Fig. 10(a) is selected as hemorrhage by mask.

Table 1
Sensitivity, accuracy and dice coefficient analysis of proposed method.

Dataset	%Sensitivity	%Accuracy	%Dice coefficient
1	90	96.43	91.81
2	77.78	93.33	96.72
3	100	92.86	95.37
4	83.33	92.86	98.78
5	100	95.65	89.35
6	100	100.00	90.87
7	100	88.00	91.06
8	100	84.62	94.4
9	90.91	92.31	92.44
10	81.25	88.00	92.17
11	100	76.92	93.78
12	100	100.00	92.5
13	80	80.77	92.83
14	100	100	95.78
15	92.31	97.06	95.58
16	100	96.43	91.81
17	87.5	96.43	96.62
18	89.47	93.75	95.87
19	100	100.00	98.58
20	100	100.00	84.56
21	94.44	96.43	92.72
22	100	72.00	68.13
Average	**93.95**	**92.45**	**92.35**

4. Result

The proposed CAD is unsupervised, fully automatic system. Efficiency of this proposed segmentation technique is tested against popular segmentation methods to establish its potential. In case of data sets segmentation, Liu and Yu (2009) and Jung, Kang, and Heo (2014) have demonstrated the power of *K*-means clustering method over Otsu and Expectation Minimization methods respectively. As discussed in Section 2, Dundar et al. (2015) has shown that unsupervised *K*-means technique outperforms several deep learning techniques when implemented on a small dataset. Fuzzy *C*-means method is an advancement of *K*-means method (Afrin, Al-Amin, & Tabassum, 2015). So, we have compared the performance of proposed method with *K*-means and fuzzy C-means segmentation. In Fig. 12 the segmentation results are shown with ground truth annotation. In *result 2*, hemorrhage segmentation annotation (blue line) is overlapped with ground truth and thus not visible. Non-hemorrhagic segmented areas are not marked but visible in the segmentation result. Two class segmentation offers poor distinguishability for both the techniques. Though with increase in number of classes, the segmentation methods increase its sensitivity, even after segmenting for six classes the proposed method outcome remains ahead in competition.

The segmented results are shown to the radiology experts of PGIMER, Chandigarh, India for comments and validation. As per the expert's opinion, hemorrhage patch detection using proposed CAD satisfies the ground truth for all the cases under test. It also has successfully collected hemorrhage information from dataset having low volume hemorrhage. Though in few slices of a dataset sometimes wrong areas are labeled as hemorrhage, after fusion and final selection those are nullified and the ground truth is satisfied.

For further analysis, quantitative evaluation in term of sensitivity, accuracy and dice coefficient of proposed segmentation method is carried out on 22 datasets of CT brain hemorrhage images. The dice coefficient shows that each dataset is segmented with enough similarity with ground truth marked by experts. Calculated average dice coefficient value is 0.9235. Evaluated results are collated in Table 1. For proposed method sensitivity lies in between 70%−100%, with an average of 93.95%. Average error is 9.43% and average accuracy is 92.45%. Accuracy, sensitivity and dice coefficient of proposed method are compared with that of *K*-means and
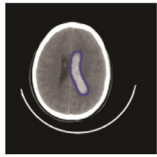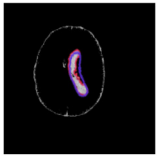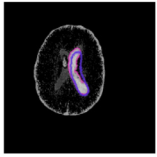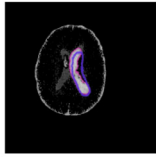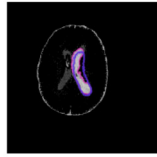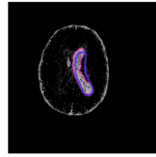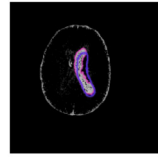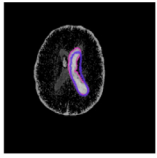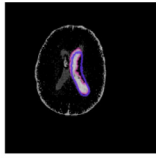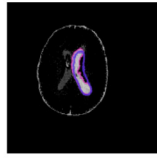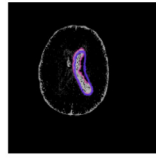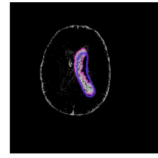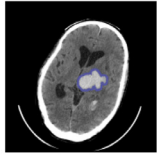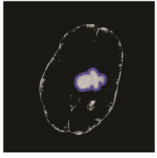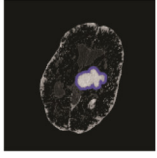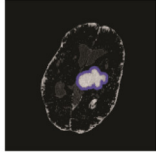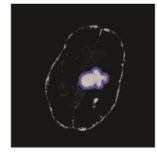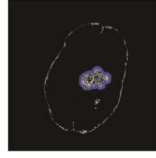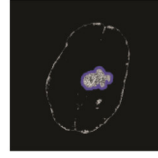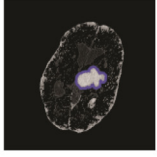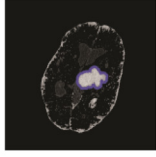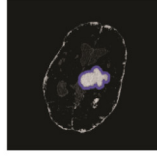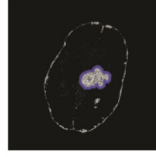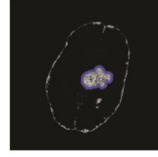
Fig. 12. Thresholded images are compared. The blue line is used to show the ground truth and pink line to show segmented area of hemorrhage. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

fuzzy C-means methods for $k = 6$. To do segmentation using K-means and fuzzy C-means, the program flow described in Fig. 1 is kept unaltered; only 'threshold finding' and 'binary segmentation' steps are replaced by respective method. Average accuracy and sensitivity for K-means is 80.56% and 68.36%, for fuzzy C-means is 88.39% and 78.43% respectively. Comparison graphs shown in Figs. 13–16 are demonstrating the power of proposed method over conventional two.

Regardless of this, no hemorrhage dataset is treated by CAD as normal dataset i.e. no false negative diagnosis by this CAD. This
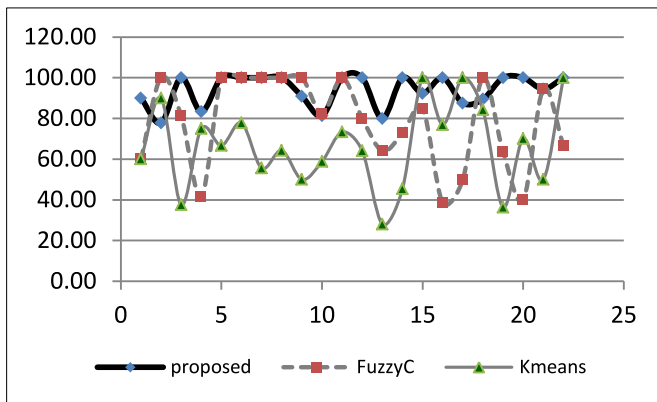
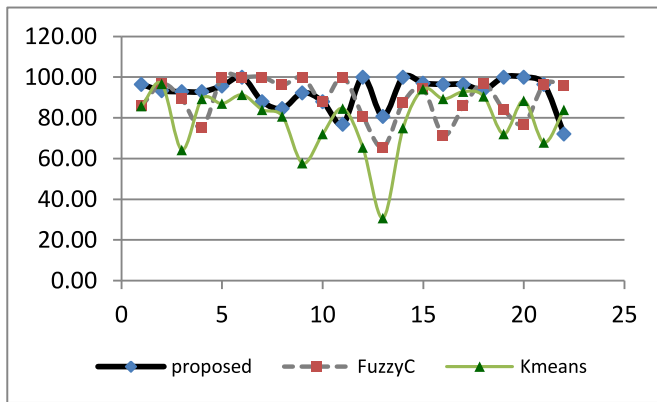**Fig. 13.** Comparison of %sensitivity of three methods.

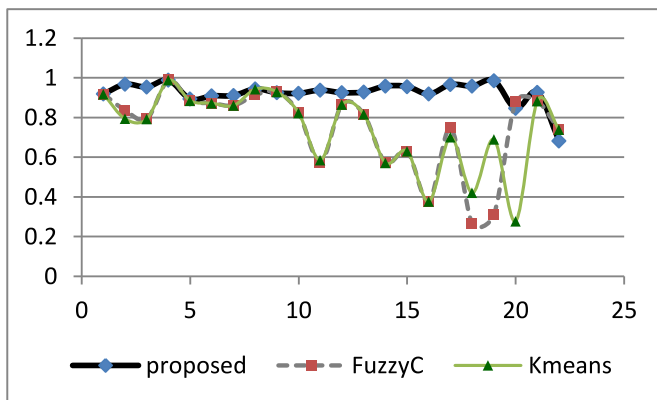**Fig. 14.** Comparison of %accuracy of three methods.

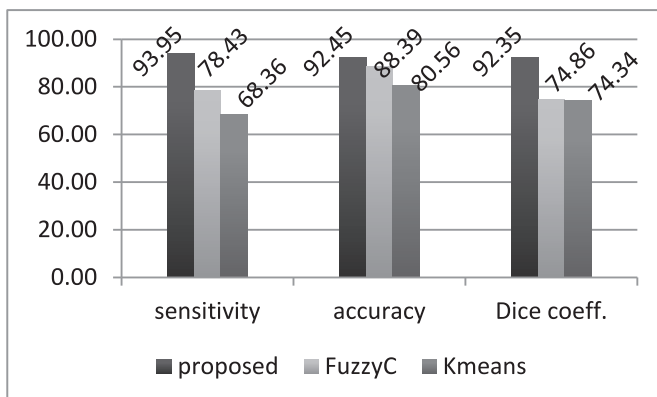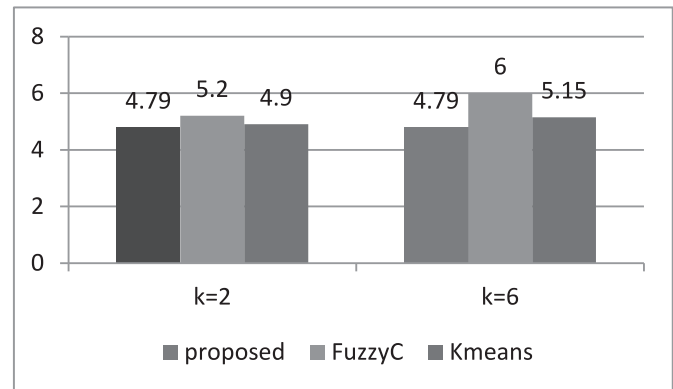**Fig. 15.** comparison of dice coefficient of three methods.

Fig. 16. COMPARISON of average of %sensitivity, %accuracy and %Dice coefficient.

Fig. 17. Comparison of segmentation time (in seconds) of three methods (k value varies for traditional methods, for proposed method it is 2 for both the cases.).

turns the CAD sensitivity (true positive by CAD/true positive as per ground truth) to 100%. As per practitioners, higher sensitivity is most important requirement for any clinical support system to avoid risk of treatment delay or serious loss due to wrong diagnosis.

The proposed thresholding technique takes 4.79 s to segment a 36 slices dataset when runs in a computer having 4 GB RAM, Intel® Core(TM) i3 processor running at 2 GHz speed and 64 bit Windows 10 OS platform. In the same system, segmentation of the same database by K-means and Fuzzy C-means take 4.9 s and 5.2 s, respectively when each slice is segmented in two clusters. Required time increases with increase in number of clusters. Segmentation times are compared in Fig. 17. Time required for K-means and fuzzy c-means clustering for both $k = 2$ and $k = 6$ are compared with time required for proposed thresholding method. As per the result, proposed method has outperformed other two methods.

There are different works already presented by researchers for hemorrhage segmentation as discussed in state of the art section. Two (Bhadauria et al., 2013; Shahangian & Pourghassem, 2013) of such works have significant resemblance with the proposed work. Quantitative analysis of these two works is compared with our proposed work.

Shahangian et al. have presented automatic hemorrhage segmentation from brain CT images with good accuracy. But three major deficiencies are noticed in this research, selection of initial thresholds, brain matter segmentation, and threshold value selection for hemorrhage segmentation. In the article, background and skull are removed by absolute threshold values 100 and 225, respectively. In many cases, parts of brain, as well as parts of hemorrhage, are removed by those fixed threshold values. No restoration method is proposed to recover such parts. An example of such case is shown in Fig. 18. Adaptive threshold selection process of the proposed method has proven advantages over method proposed by Shahangian et al. In the next step, they proposed selection of brain by selecting the largest connected area. At the base of skull, eyes and other facial parts are got scanned and the actual brain area remains very low. Such slices often offer the largest segment which includes eyes but excludes brain and thus reduce the sensitivity of the system. But in proposed method, as the skull encapsulated area is considered, brain area is always inclusive in segmented patch. An example is presented in Fig. 19. Finally, there is no clue given in the article about threshold selection for hemorrhage. Rather it has been stated as "we define an appropriate threshold", though the process is claimed to be automatic. This is the most important part on which automatic hemorrhage segmentation and accuracy
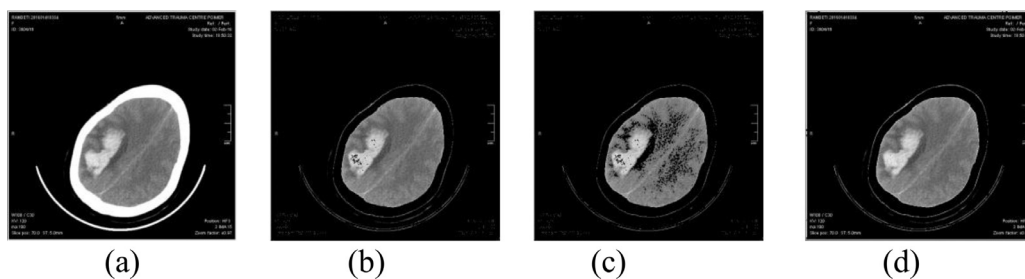
Fig. 18. (a) Original image (b) skull removed by threshold 225 (c) background removed by threshold 100 (d) adaptive thresholding for skull and background.

Table 2
Comparison chart of Shahangian, Bhadauria and proposed methods' results.

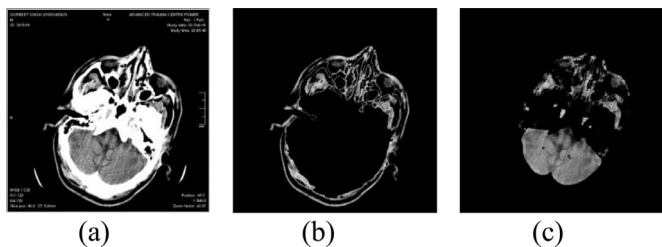|	Sensitivity	Specificity	Dice coefficient	Accuracy	Image handling capacity
Shahangian	85.27	96.05	83.69	**93.80**	Single image
Bhadauria	80.67	99.58	87.58	86.24	Single image
Proposed method	**93.95**	**100**	**92.35**	92.45	**Multislice scanned data**
Advantage observed in	*Proposed method*	*Proposed method*	*Proposed method*	*Shahangian*	*Proposed method*

Fig. 19. (a) Original image (b) brain by largest area selection (c) brain by proposed method.

of the system highly dependent. Hence the quantitative analysis presented in their paper remains questionable.

An integrated process is proposed by Bhadauria et al. It offers better result than Chen and Vese method, region growing technique and fuzzy C means method. They have used active contour method along with FCM to improve the performance of the process. The performance parameters of this method and Shahangian's method are assembled and compared with our proposed method in Table 2.

It has been observed that the potential of the proposed method is much higher in terms of sensitivity, dice similarity and specificity. Only Shahangian's method is calming higher accuracy, justification of which is not found in their article and its dice coefficient is very poor. Another important advantage of the proposed method is its image handling capacity. It can segment hemorrhage from all affected slices of a dataset in single run, whereas, other two methods can segment one slice at a time.

5. Discussion and conclusion

The research target was to design a fast, efficient, light weight, easy to use and reliable medical support system for hemorrhage segmentation from brain CT scan images to initiate further diagnosis and treatment as per requirement. The proposed system's potential lies in zero false negative in final decision. The segmentation results are reviewed and approved by doctors of radiology department, PGIMER, Chandigarh, India. Proposed CAD is a threshold based technique. To find threshold automatically from intensity population, prior knowledge of brain anatomy and CT image intensity distribution for brain and hemorrhage are used. Information outside the skull is removed to avoid error due to overlapping

intensity range. Background and skull are also removed to reduce computational load to make the system faster.

The proposed unsupervised technique based CAD is expected to be helpful for critical situation like remote diagnosis support, absence of expert practitioner and bulk patient management due to an accidental emergency. This method requires no human interaction for segmentation; whereas *K*-means and fuzzy *C*-means need human interpretation to find right cluster among multiple output clusters. For an example, an expert is required to select particular cluster containing hemorrhage for 12 times when the complete dataset contains 12 slices. So these methods are unable to fulfill our requirement of using the CAD in absence of practitioner for hemorrhage segmentation.

More powerful CAD can be designed by further researches. Facilities like a fast but accurate classification of hemorrhage, detecting hemorrhage from a dataset where both normal and hemorrhage data are present, advice about diagnosis depending on the severity of segmented hemorrhage can be included in the CAD in the future.

The limitation associated with the proposed CAD is that it cannot detect all hemorrhages in case of Simultaneous Multiple Intracerebral Hemorrhages (SMICH). It can only detect the presence of largest hemorrhage which is sufficient to trigger medical attention and the requirement of expert's opinion. Advanced research can be done to identify multiple hemorrhages.

Funding

No funding was used for this research work.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

An ethical clearance has been taken from Institute Ethical Committee and all procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

References

Adams, H. P., Del Zoppo, G., Alberts, M. J., Bhatt, D. L., Brass, L., & Furlan, A. (2007). Guidelines for the early management of adults with ischemic stroke a guideline from the American heart association/American stroke association stroke council, clinical cardiology council, cardiovascular radiology and intervention council, and the atherosclerotic peripheral vascular disease and quality of care outcomes in research interdisciplinary working groups: the American academy of neurology affirms the value of this guideline as an educational tool for neurologists. *Circulation, 115*, e478–e534.

Admin. (10 August 2014). *Traumatic brain injury*. Available: http://indianheadinjuryfoundation.org/traumatic-brain-injury/.

Afrin, F., Al-Amin, M., & Tabassum, M. (2015). Comparative performance of using PCA with K-means and fuzzy C means clustering for costomer segmentation. *International Journal Of Scientific & Technology Research, 4*, 70–74.

Al-Ayyoub, M., Alawad, D., Al-Darabsah, K., & Aljarrah, I. (2013). Automatic detection and classification of brain hemorrhages. *WSEAS Transactions on Computers, 12*, 395–405.

Balasooriya, U., & Perera, M. S. (2012). Intelligent brain hemorrhage diagnosis using artificial neural networks. In *Business Engineering and Industrial Applications Colloquium (BEIAC), 2012* (pp. 128–133). IEEE.

Bardera, A., Boada, I., Feixas, M., Remollo, S., Blasco, G., & Silva, Y. (2009). Semi-automated method for brain hematoma and edema quantification using computed tomography. *Computerized Medical Imaging and Graphics, 33*, 304–311.

Bhadauria, H., Singh, A., & Dewal, M. (2013). An integrated method for hemorrhage segmentation from brain CT imaging. *Computers & Electrical Engineering, 39*, 1527–1536.

Brosch, T., Tang, L. Y., Yoo, Y., Li, D. K., Traboulsee, A., & Tam, R. (2016). Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation. *IEEE Transactions on Medical Imaging, 35*, 1229–1239.

Cabaret, L., Lacassagne, L., & Oudni, L. (2014). A review of world's fastest connected component labeling algorithms: Speed and energy estimation. In *Proceedings of the conference on design and architectures for signal and image processing (DASIP)* (pp. 1–6). IEEE.

Chan, T. (2007). Computer aided detection of small acute intracranial hemorrhage on computer tomography of brain. *Computerized Medical Imaging and Graphics, 31*, 285–298.

Chawla, M., Sharma, S., Sivaswamy, J., & Kishore, L. (2009). A method for automatic detection and classification of stroke from brain CT images. In *EMBC 2009. Annual International Conference of the IEEE* (pp. 3581–3584). Engineering in Medicine and Biology Society.

Cheng, D.-C., & Cheng, K.-S. (1998). A PC-based medical image analysis system for brain CT hemorrhage area extraction. In *Proceedings of the eleventh IEEE symposium on computer-based medical systems, 1998* (pp. 240–245).

Cheng, K.-S., Lin, J.-S., & Mao, C.-W. (1996). The application of competitive Hopfield neural network to medical image segmentation. *IEEE Transactions on Medical Imaging, 15*, 560–567.

Cosic, D., & Loucaric, S. (1997). Computer system for quantitative: Analysis of ICH from CT head images. In *Proceedings of the nineteenth annual international conference of the IEEE engineering in medicine and biology society, 1997* (pp. 553–556).

Datta, A., Datta, A., & Biswas, B. (2011). A fuzzy multilayer perceptron network based detection and classification of lobar intra-cerebral hemorrhage from computed tomography images of brain. In *Proceedings of the 2011 international conference on recent trends in information systems (ReTIS)* (pp. 257–262).

Dhawan, A. P., Loncaric, S., Hitt, K., Broderick, J., & Brott, T. (1993). Image analysis and 3-d visualization of intracerebral brain hemorrhage. In *Proceedings of sixth annual IEEE symposium on computer-based medical systems, 1993* (pp. 140–145).

Dundar, M., Kou, Q., Zhang, B., He, Y., & Rajwa, B. (2015). Simplicity of kmeans versus deepness of deep learning: A case of unsupervised feature learning with limited data. In *Proceedings of the 2015 IEEE fourteenth international conference on machine learning and applications (ICMLA)* (pp. 883–888).

Ee, C., Sim, K., Teh, V., & Ting, F. (2016). Estimation of window width setting for CT scan brain images using mean of greyscale level to standard deviation ratio. In *Proceedings of the international conference on robotics, automation and sciences (ICORAS)* (pp. 1–6).

Ghesu, F. C., Krubasik, E., Georgescu, B., Singh, V., Zheng, Y., & Hornegger, J. (2016). Marginal space deep learning: Efficient architecture for volumetric image parsing. *IEEE Transactions on Medical Imaging, 35*, 1217–1228.

Gong, T., Liu, R., Tan, C. L., Farzad, N., Lee, C. K., & Pang, B. C. (2007). Classification of CT brain images of head trauma. In *Proceedings of the IAPR international workshop on pattern recognition in bioinformatics* (pp. 401–408).

Greenspan, H., van Ginneken, B., & Summers, R. M. (2016). Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging, 35*, 1153–1159.

Hara, T., Matoba, N., Zhou, X., Yokoi, S., Aizawa, H., & Fujita, H. (2007). Automated detection of extradural and subdural hematoma for contrast-enhanced CT images in emergency medical care. *Medical Imaging* pp. 651432-651432-4.

Jung, Y. G., Kang, M. S., & Heo, J. (2014). Clustering performance comparison using K-means and expectation maximization algorithms. *Biotechnology & Biotechnological Equipment, 28*, S44–S48.

Kalinovskya, A., Liauchuka, V., & Tarasaub, A. (2017). Lesion detection in CT images using deep learning semantic segmentation technique. In *Proceedings of the IS-PRS-international archives of the photogrammetry, remote sensing and spatial information sciences* (pp. 13–17).

Lacassagne, L., & Zavidovique, B. (2009). Light speed labeling for risc architectures. In *Proceedings of the 2009 sixteenth IEEE international conference on image processing (ICIP)* (pp. 3245–3248).

Li, Y., Hu, Q., Wu, J., & Chen, Z. (2009). A hybrid approach to detection of brain hemorrhage candidates from clinical head ct scans. In *Proceedings of the sixth international conference on fuzzy systems and knowledge discovery, 2009. FSKD'09* (pp. 361–365).

Lin, J.-S., Cheng, K.-S., & Mao, C.-W. (1996). A fuzzy Hopfield neural network for medical image segmentation. *IEEE Transactions on Nuclear Science, 43*, 2389–2398.

Liu, D., & Yu, J. (2009). Otsu method and K-means. In *Proceedings of the ninth international conference on hybrid intelligent systems, 2009. HIS'09* (pp. 344–349).

Liu, B., Yuan, Q., Liu, Z., Li, X., & Yin, X. (2008). Automatic segmentation of intracranial hematoma and volume measurement. In *Proceedings of the 2008 thirtieth annual international conference of the IEEE engineering in medicine and biology society* (pp. 1214–1217).

Liu, R., Tan, C. L., Leong, T.-Y., Lee, C. K., Pang, B. C., & Lim, C. T. (2008). *Hemorrhage slices detection in brain CT images* (pp. 1–4). ICPR.

Liu, Y., Rothfus, W. E., & Kanade, T. (1998). Content-based 3d neuroradiologic image retrieval: Preliminary results. In *Proceedings of the 1998 IEEE international workshop on content-based access of image and video database* (pp. 91–100).

Liu, Y., Lazar, N. A., Rothfus, W., Dellaert, F., Moore, A., Schneider, J., et al. (2003). Semantic-based biomedical image indexing and retrieval. In *Carnegie Mellon University Research Showcase @ CMU*. Robotics Institute, School of Computer Science. http://repository.cmu.edu/cgi/viewcontent.cgi?article=1921&context=robotics.

Loncaric, S., Cosic, D., & Dhawan, A. (1996). Hierarchical segmentation of CT images. In *Proceedings of the eighteenth annual international conference of the IEEE EMBS*. IEEE.

Lončarić, S., Ćosić, D., & Dhawan, A. P. (1996). Segmentation of CT Head Images. In *Proceedings of the international symposium on computer and communication systems for image guided diagnosis and therapy*.

Loncaric, S., Dhawan, A. P., Cosic, D., Kovacevic, D., Broderick, J., & Brott, T. (1999). Quantitative intracerebral brain hemorrhage analysis. In *Proceedings of the medical imaging'99* (pp. 886–894).

Maduskar, P., & Acharyya, M. (2009). Automatic identification of intracranial hemorrhage in non-contrast CT with large slice thickness for trauma cases. In *Proceedings of the SPIE medical imaging* pp. 726011-726011-8.

Pérez, N., Valdés, J. A., Guevara, M. A., Rodríguez, L. A., & Molina, J. (2007). Set of methods for spontaneous ICH segmentation and tracking from CT head images. In *Proceedings of the iberoamerican congress on pattern recognition* (pp. 212–220).

Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging, 35*, 1240–1251.

Quain, R. (1883). *A dictionary of medicine: Including general pathology, general therapeutics*. D. Appleton and Company.

Ravì, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., & Lo, B. (2017). Deep learning for health informatics. *IEEE Journal of Biomedical And Health Informatics, 21*, 4–21.

Ros, P. R., & Mortele, K. J. (2007). *CT and MRI of the abdomen and pelvis: A teaching file*. Lippincott: Williams & Wilkins.

Shahangian, B., & Pourghassem, H. (2013). Automatic brain hemorrhage segmentation and classification in CT scan images. In *Proceedings of the 2013 eighth iranian conference on machine vision and image processing (MVIP)* (pp. 467–471).

Wedro, B. (2015). (11 May 2015, 11 August). *Head Injury (Brain Injury)*. Available: http://www.medicinenet.com/head_injury/article.htm.

Yuh, E. L., Gean, A. D., Manley, G. T., Callen, A. L., & Wintermark, M. (2008). Computer-aided assessment of head computed tomography (CT) studies in patients with suspected traumatic brain injury. *Journal of Neurotrauma, 25*, 1163–1172.

Zimmerman, R. A., Gibby, W. A., & Carmody, R. F. (2012). *Neuroimaging: Clinical and physical principles*. Springer Science & Business Media.