Journal of Network and Innovative Computing
ISSN 2160-2174 Volume 2 (2014) pp. 022-029

Machine learning approach to track the progress of lesions using Change Detection

Ankita Mitra[1], Arunava De[2], Anup Kumar Bhattacharjee[3]

[1]Department of Electronics and Communication,
National Institute of Technology, Durgapur, India,
ankimitra.2009@gmail.com

[2] Department of Information Technology,
Dr. B.C. Roy Engineering College, Durgapur, India,
arunavade@yahoo.com

[3]Department of Electronics and Communication,
National Institute of Technology, Durgapur, India,
akbece12@yahoo.com

Abstract. **Brain MR image requires analysis of complex data. In this article use of machine learning approaches together with the Change detection algorithms is aimed at identifying changes in sets of images or image sequences at different times. In other words it is the process of identifying the changes in a state of an object over time. The phenomena of change detection are used to detect the progress of lesions in MR Images. Two class classification models are used to identify the object, which are the lesions and the background of the MR image. De-noising of MR image is done using Peaks over Threshold for Shrinking the Wavelet coefficients. Change detection algorithms are used on a sequence of MR images to find the extent of progression of lesions in an patient.**

Keywords: Region of Interest, Magnetic Resonance Imaging, Multi Resolution Wavelet Analysis, Change Detection, Peaks over Threshold,Wavelet Thresholding.

I. Introduction

This paper proposes to study the progression of lesions in MR images of brain. The lesion can be a simple blot clot to cancerous formation. MR images are multimodal images. They have more than one object. Different thresholds are used to segregate different objects of the multimodal images. The histogram of multimodal images has multiple peaks as opposed to two peaks of bi-modal image.

Segmentation is subjective in nature because it depends on where the emphasis is given. Therefore it is not proper to always minimize or maximize a certain objective function as discussed by [1,2]. Using two class classification model the MR image of brain is segmented into lesions and background.

The existing de-noising methods involves hard and soft thresholding. In soft thresholding absolute values of coefficients whose values are greater than a threshold is kept whereas other coefficients are shrunk that is they are scaled towards zero. The only difference between the hard and soft thresholding procedures is in the choice of the nonlinear transform that is applied on the detailed wavelet coefficients. The following nonlinear transform is used:

$$S(x) = sign(x)(|x| - \tau)I(|x| > \tau) \qquad (1)$$

Where τ is a threshold. The wavelet shrinkage method used here is based on the Peaks Over Threshold [3,4] modelling of extreme values.

Change detection is used to identify the changes occurring in a sequence of MR image of brain. The image of brain is compared with another image of brain from another time. The differences in the two images show the progress of the lesions. This method of identifying the changes occurring in MR images of brain may find wide use in quantifying the progress of treatment in patients.

II. Related Work

An energy criterion formulated by intensity based class uncertainty and region homogeneity was introduced by [5]. The between class variance is maximized in histogram based thresholding by [6]. The method shows satisfactory results in various applications but it has the disadvantage that it tends to split the larger part when the sizes of object and background are unequal [7]. Knowledge about the range of background proportion to the ROI is used to confine the range of threshold selection and achieved reliable results in segmenting MR images [8]. By optimizing the weighted sum of the within-class variance and the intensity contrast a threshold can be obtained [9].

Identifying the differences in a sequence of images over time is the objective of change detection algorithms[10]. Image Ratio-ing and image differencing techniques were developed in the 1970's [11]. Image ratio-ing is similar to image differencing and is done between two images.

III. Proposed Algorithm

A. Pre-processing the MR Image

The input images must be co-registered otherwise this leads to incorrect results. Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. It geometrically aligns two images – the reference and sensed images. The input image is pre-processed to a data range between 0 and 1.

B. ROI Generation using Machine learning approach

The evaluation of segmentation is based on similarity and proximity. It is based on the idea that if two regions are close to each other they are likely to be part of one grouping. The principle of good continuation states that natural objects have a smooth boundary line rather than radical points. The principle of similarity is fundamental to determine the groups. There are two principles-

a) Intra- region similarity: The region has similar brightness and texture.

b) Inter-region dissimilarity: The elements in separate regions are dissimilar in brightness and texture.

A classifier is constructed with the features stated above and it is trained on MR image of brain. A supervised learning approach is employed for the segmentation task. The classifier is used to separate the brain MR image into lesions and background. The MR image is a multi-modal image. There are multiple objects and hence multiple thresholds in an MR image. A lesion has a specific threshold. Hence the region of lesions will have a different intensity as that of the rest of the MR image of brain.

A supervised training data set is developed taking the intensity of the pixels as reference and classifying the MR image of brain as binary (either lesion or background). K-Nearest Neighbour (KNN) classifier is used to segment the MR image of brain having lesions.

Precision and Recall is used to evaluate segmentation accuracy as discussed in [1]**.** The irrelevant MR image data without lesions are discarded because they are of no use in the change detection algorithm. Only images with lesions are considered as the input to the algorithm. This step gives the ROI of the MR image which contains the lesions.

C. De-Noising the ROI using Peaks over Threshold method

The generic methodology of Wavelet shrinkage or Wavelet thresholding involves three main steps: (1) The Wavelet transform is computed from the data, (2) Large coefficients are kept or reduced, others are removed (Shrinkage); (3) The inverse DWT is applied to the shrunken set of coefficients. Wavelet has a great appeal in separating signal from noise by thresholding the wavelet coefficients. A key job in wavelet shrinkage method is the choice of threshold. The threshold is data driven and chosen using Peaks over Threshold method of modelling extreme values.

1) Applying the Wavelet Shrinkage Method

a) Apply the Discrete Wavelet Transform

Wavelet signals are modelled by combining translations and dilations of an oscillatory function with a finite duration called “Wavelet”. The resulting transform is easy to interpret and valuable for time frequency analysis.
The continuous Wavelet transform [12] of a continuous, square-integrable function *f(x)*, relative to a real valued wavelet $\psi(x)$, is

$$W\varphi(s,\tau) = \int_{-\infty}^{\infty} f(x)\varphi_{s,\tau}(x)dx \tag{2}$$

Where

$$\varphi_{s,\tau}(x) = \frac{1}{\sqrt{s}}\varphi\left(\frac{x-\tau}{s}\right)$$

and s and τ are *translation* and *scale* parameters respectively. Given $W\varphi(s,\tau)$, *f(x)* can be obtained using the *inverse continuous wavelet transform*.

$$f(x) = \frac{1}{C_\varphi}\int_0^\alpha \int_{-\alpha}^{\alpha} W_\varphi(s,\tau)\frac{\varphi_{s,\tau}(x)}{s^2}d\tau\, ds \tag{3}$$

Where

$$C_\varphi = \int_{-\infty}^{\infty}\frac{|\psi(u)|^2}{|u|}$$

where (u) is the Fourier Transform of ψ(x).
DWT is the implementation of the Wavelet Transform using discrete set of the wavelet translations and scales which obey predefined rules. DWT decomposes the signal into mutually orthogonal set of wavelets which is the basic difference between DWT and CWT. A 6 level Haar Wavelet transform is used for the decomposition purpose.
Apply the discrete Wavelet Transform DWT:

$$(\omega) = W(y) \tag{4}$$

Where W is an ortho-normal transformation and $(w)=(w_{j.k})$ is a vector of wavelet coefficients. The output vector $(w)=w_{j.k}, j\geq 0, k=0,\ldots,2^j$ has length n and is indexed by both location parameter k (corresponding to frequency $\omega = 2^{-j}$).

b) Use of Peaks Over Threshold approach for Shrinking

An unknown signal *f(x)* is to be estimated from some noisy data $n = (n_1, n_2, \ldots n_n)$.
The following model is assumed [3]

$$n_j = f_j + r_j,\ j=1,\ldots,m. \tag{5}$$

where $f_j = f(j/m)$, j=1,….,m is a sample of the original signal and r_j are independent and identically distributed random variables with zero mean and finite variance.
In an unknown distribution F consisting of independent and identically distributed observations $D_1, D_2, D_3,\ldots D_n$. The excess losses above a threshold value τ are of particular interest. Let d_0 be the finite or infinite right endpoint of the distribution F.

This means $d_0 = \sup\{d \in R: F(d) < 1\} \leq \infty$. The distribution function which defines the excesses over threshold τ is given by

$$F_\tau(d) = P\{D - \tau \leq d \mid D > \tau\}$$

$$= \frac{F(d+\tau) - F(\tau)}{1 - F(\tau)} \qquad (6)$$

for $0 \leq d < d_0 - \tau$. $F_\tau(d)$ is the probability that a loss exceeds the threshold τ by an amount not more than d, given that the threshold is exceeded [4].

The distribution which models the excesses is the generalized Pareto Distribution (GPD) which is usually expressed as a two parameter distribution with the distribution function given as

$$G_{\xi,\sigma}(d) = \begin{cases} 1 - \left(1 + {}^{\xi d}/_{\sigma}\right)^{-1/\xi} & if\ \xi \neq 0, \\ 1 - \exp\left({}^{-d}/_{\sigma}\right) & if\ \xi = 0 \end{cases} \qquad (7)$$

where $\sigma > 0$, and the support is $d \geq 0$ when $\xi \geq 0$ and $0 \leq x \leq -\sigma/\xi$ when $\xi < 0$. The GPD consists of three other distributions

when

$\xi > 0$ Re-parameterized version of usual Pareto Distribution

$\xi < 0$ Type − 2 Pareto Distribution

$\xi = 0$ gives the Exponential Distribution

If a location parameter μ is added then GPD $G_{\xi,\mu,\sigma}(x)$ is defined as $G_{\xi,\sigma}(x - \mu)$.

The method of shrinking the wavelet coefficients involves the discarding the in-significant coefficients whereas keeping the coefficients that are statistically significant (H_0: f is smooth *Vs* H_1 : f has some discontinuities). After leaving intact the wavelet coefficients at low resolution levels corresponding to the 'macro features' of the signal, the thresholding is done at all the levels (level >3) of resolution of 6 level Haar Wavelet coefficients.

The advantage of this method lies in the use of GPD (7) on the DWT. The value of shape and scale parameter is estimated using the DWT at the highest resolution level. The test of H_0 *Vs* H_1 based on GPD is then used on wavelet coefficients at an arbitrary resolution level (level >3). The threshold is set at a value $\tau = \tau_i$ in the following equation using soft thresholding which is not rejected by the test of H_0 *Vs* H_1.

$$(w') = (w'_{j,k}) = sgn\,(w_{j,k}) \times (|w_{j,k}| - \tau) \times I(|w_{j,k}| \geq \tau) \qquad (8)$$

Estimating shape and scale parameters-

i)At the highest resolution level take the absolute value of the DWT.

$$\left(D_1, \ldots \ldots, D_{\frac{n}{2}}\right) = (|w|) = (|w_{j,k}|, k = 0, \ldots\ldots, 2^{J-1}) \qquad (9)$$

ii) Define excesses over the threshold τ>0:

$$\tilde{D}_i = (D_i - \tau)_+ \; i = 1, 2, \ldots \qquad (10)$$

Also $(D - \tau)_+ = \max(d - \tau, 0)$

iii) Compute the Maximum Likelihood estimates of shape and scale parameters from the $\tilde{D}_i$'s.

c) Wavelet Reconstruction

Wavelet reconstruction is performed based on the original approximation coefficients and the modified detail coefficients for each level of decomposition.

Apply the inverse DWT to get an estimate of f :

$$\hat{f} = W^t(w') \qquad (11)$$

The ROI and de-noising of Fig. 1(a) and (b) is depicted in Fig. 2 (a) and (b).

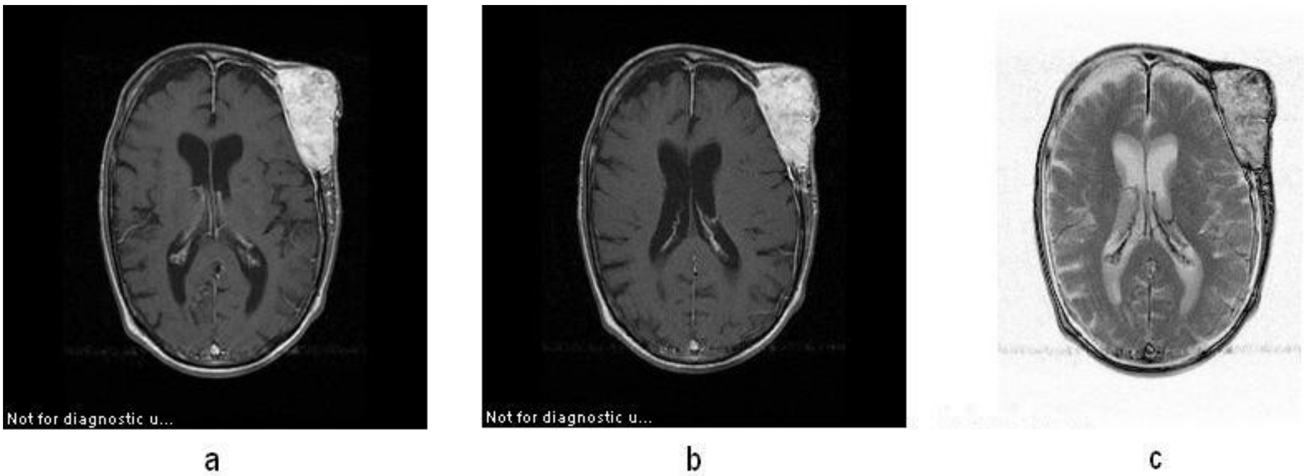
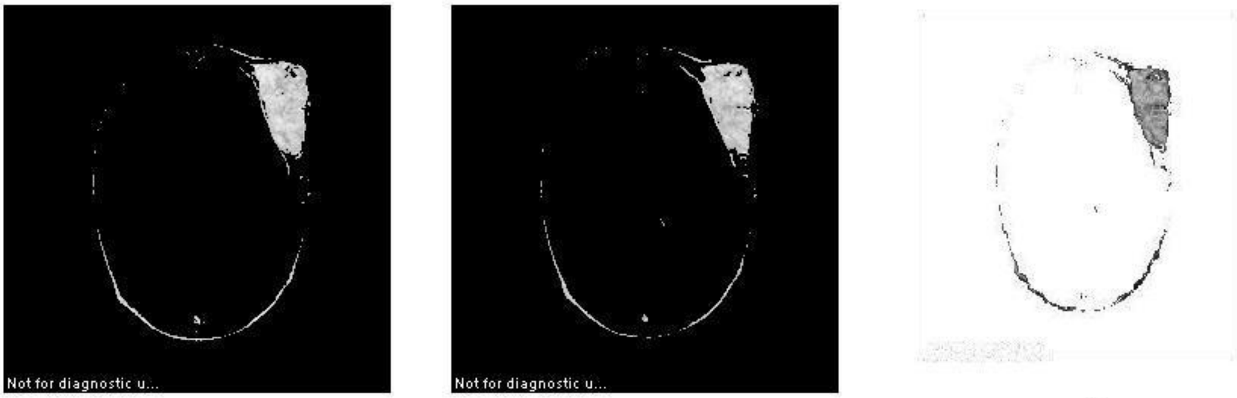


Fig.1 a) Original un- segmented brain MRI-1 b) Original un- segmented brain MRI-2 c) Change detection output using differencing.



a b c

Fig.2. a) ROI-1 after De-noising b) ROI-2 after De-noising c) Change detection output using differencing

D. Change detection using image differencing

A supervised training data set is developed taking the intensity of the pixels as reference and classifying the MR image of brain as binary (either lesion or background). K-Nearest Neighbour (KNN) classifier is used to segment the MR image of brain having lesions.

The ROI is generated after the segmentation. The ROI is filtered to get the lesion devoid of the background information. The lesions of the MR images taken at different times are compared using change detection algorithms.

Image differencing is based on the signed difference image.

$$E(x) = I_2(x) - I_1(x). \quad (12)$$

Histogram based thresholding is done on the difference image. A threshold value is chosen which is then applied to the difference image. The Change Detection difference map is used to produce a difference image which characterizes the difference between the initial and final state of the MR image of the brain.

The change mask is generated according to the following equation [13,14]:

$$D(x)=\begin{cases} 1, if\ |D(x)| > \tau \\ 0, otherwise \end{cases} \quad (13)$$

The threshold τ is determined empirically. Histogram based thresholding is applied. Figure 3 depicts the flow diagram of the proposed technique.

Journal of Network and Innovative Computing
ISSN 2160-2174 Volume 2 (2014) pp. 022-029

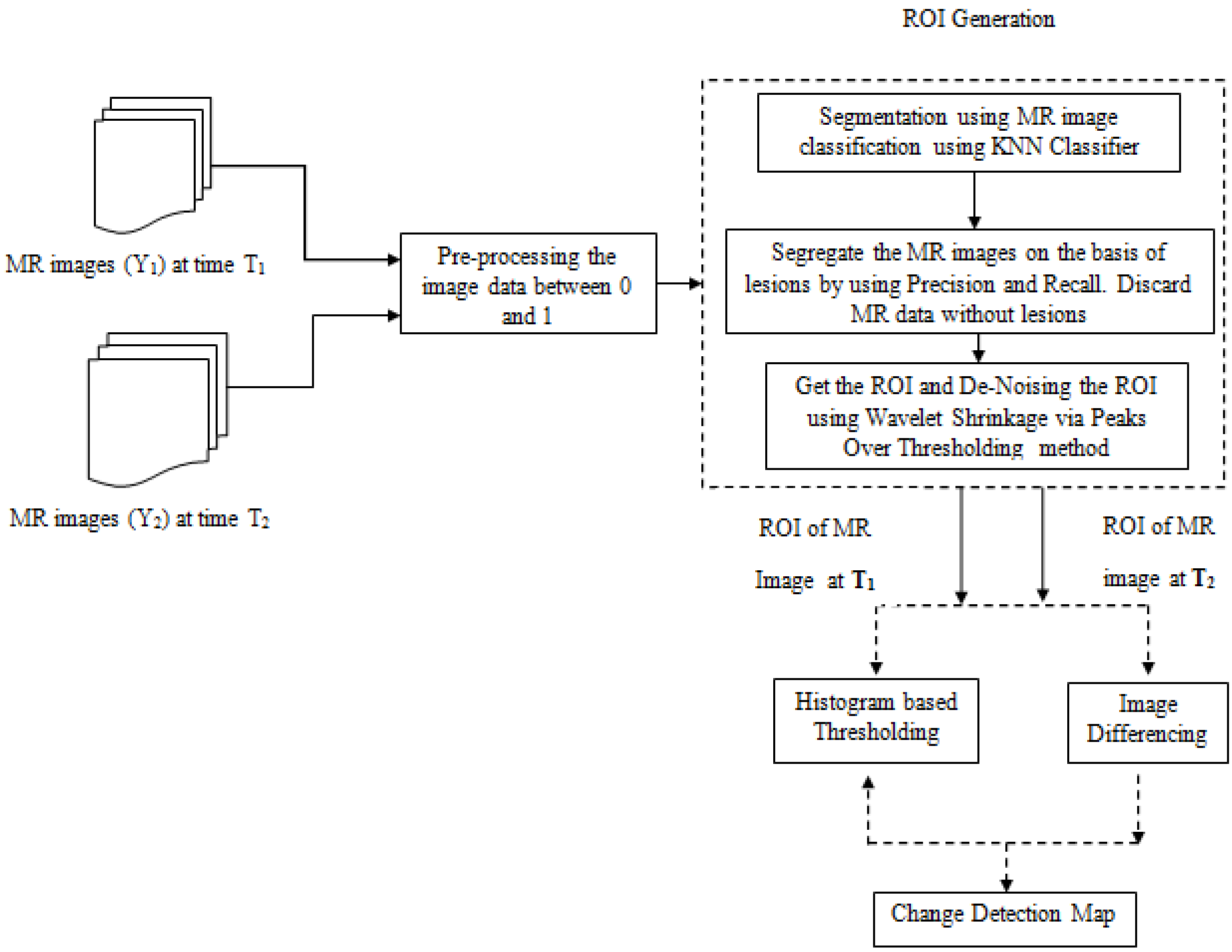


Figure.3 Flow diagram of the proposed Algorithm.

IV Results and Analysis

Precision and recall [2] is used to evaluate segmentation accuracy which segments the MR image into those images which contain the lesions and those which do not. This metric is used to segregate the data into the MR image that contain lesions and those which do not contain lesions. The precision of a class is the number of true positives (the number of items correctly labelled as belonging to the positive class) divided by the total number of elements labelled as belonging to the positive class (the sum of true positives and false positives defined as items incorrectly labelled as belonging to the positive class). Recall is the number of true positives divided by the total number of elements that actually belong to the positive class (that is the sum of true positives and false negatives, which are items which are not labelled as belonging to the positive class but should have been).

The slices that are depicted fig.4 and fig.5 are labelled as per the presence of lesions. The "Relevant Case" is the slices that contain the lesions whereas "Not Relevant Case" is the slices which do not contain the lesions. With the above expert knowledge we proceed to calculate the accuracy of the segregation of the input images which contain lesions and those which do not contain lesions. Table 1 lists the values needed for evaluating segmentation accuracy.

TN/True Negative: Case was negative and predicted negative.

TP/True Positive: Case was positive and predicted positive.

FN/False Negative: Case was positive but predicted negative.

FP/False Positive: Case was negative but predicted positive.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{14}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{15}$$

The Precision and Recall values for Patient Dataset-1(Fig.4) are 87.5% and 77.78 % respectively whereas for Patient Dataset-2(Fig.5) the values are 88.89% and 72.72% respectively. The images which do not contain lesions are discarded. The relevant values required for the calculation is enumerated in Table 1.

TABLE – 1

Patient	Expert Knowledge	Total Cases		Predicted Negative	Predicted Positive
Patient Dataset 1	Presence of Lesions in some or all of the MRI's	19	Negative Cases	9	1
			Positive Cases	2	7
Patient Dataset 2	Presence of Lesions in some or all of the MRI's	19	Negative Cases	7	1
			Positive Cases	3	8

Total Cases: Total 19 MRI slices are tested for each patient
Positive Cases: MRI's with Lesions.
Negative Cases: MRI's without Lesions

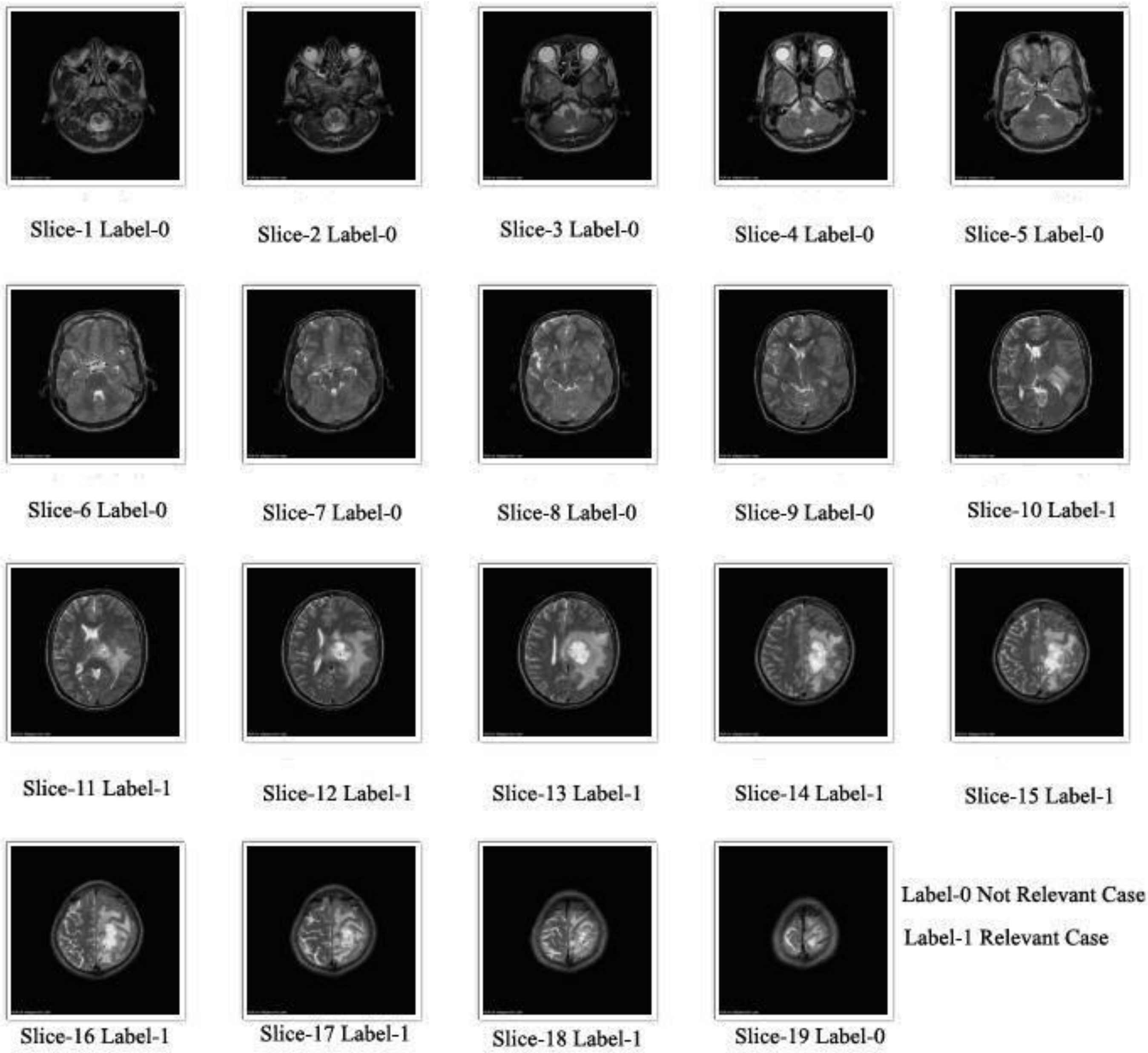


Fig.4 Patient Dataset-1

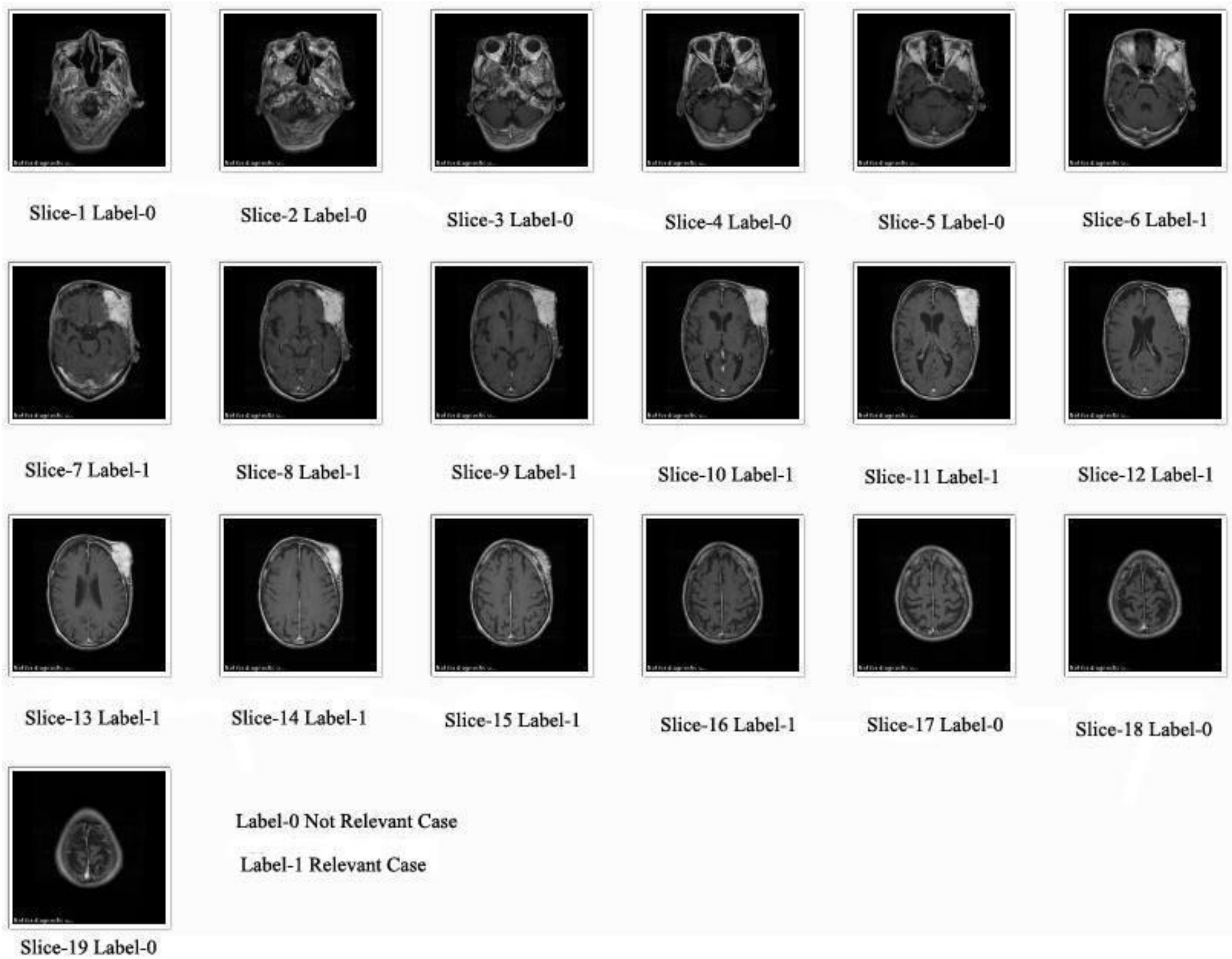


Fig.5 Patient Dataset-2

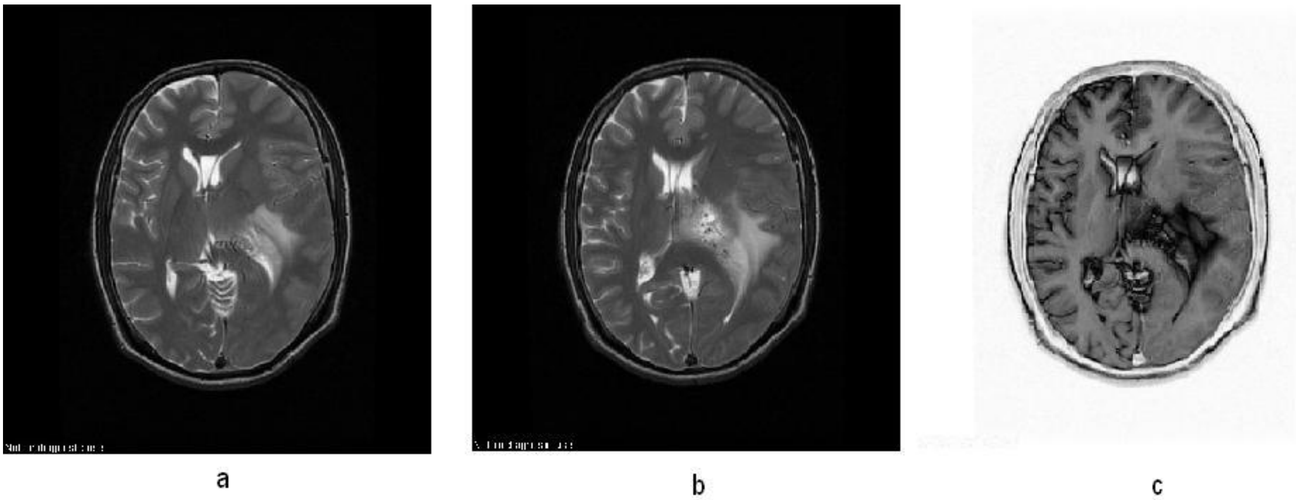


Fig.6**.** a) Un- segmented brain MRI-1, b) Un- segmented brain MRI-2 , c) Change detection output using Differencing

Due to space constraints only a few segmentation results are displayed. A classification accuracy of 92% was obtained when the MR image was segmented into two classes of lesions and background, which forms the ROI (fig 2) using KNN classifier. The ROI and de-noising of Fig. 1(a) and (b) is depicted in Fig. 2 (a) and (b). Fig.6 belongs to Dataset-1(Fig.4) and Fig.1 belongs to Dataset-2 (Fig.5).

The difference is computed by subtracting the initial state image from the final state image.The initial and final images are the ROI, containing the lesions of the MR image. Here the initial and final state image is the MR image of brain at two different time intervals. If the pixels in the final state MR image become brighter than the initial state MR image than it leads to positive change else it leads to negative change (final state image is dimmer). If there is no change then the output of the image is zero otherwise the changed regions are visible in the image. Figure 6 and 1 involves the change detection of un-segmented MR images. Whereas in Fig. 2 the ROI consists of the lesions and everything else is the background.

V CONCLUSION

Medical MR imaging is an emerging area and is very relevant in clinical applications. The machine and patient incorporated noise degrades the ability of human and computer interpretation of the MR images. Hence correct diagnosis requires efficient de-noising techniques. In this paper Peaks over Threshold approach is used for Wavelet based thresholding.

Historically people learned all types of information by doing change detection. They have found wide use in remote sensing applications. It has found wide use in the field in Medical Image Processing. The aim of this work is to use change detection approaches to identify and study the changes occurring in the human brain MRI's of patients suffering from both malignant and benign tumours. Basic change detection approaches to identify changes in brain MRI's of patients are used. Favourable results are obtained and the result of this algorithm may be incorporated in development of change detection tools for brain MRI.

References

[1]Arunava De, Anup Kumar Bhattacharjee, Chandan Kumar Chanda, Bansibadan Maji, Hybrid Particle Swarm Optimization with Wavelet Mutation based Segmentation and Progressive Transmission Technique for MRI Images, in: *International Journal of Innovative Computing, Information and Control*, 8(7(B)) (2012), pp. 5179-5197.

[2] Arunava De, Anup Kumar Bhattacharjee, Chandan Kumar Chanda, Bansibadan Maji ,Entropy Maximization based Segmentation, Transmission and Wavelet Fusion of MRI Images, in: *International Journal of Hybrid Intelligent Systems*, IOS press, Netherlands, 10 (2013), pp.57-69.

[3]Marc Raimondo, "*Tutorial on Wavelet Shrinkage Via Peaks Over Threshold*", The University of Sydney, Australia,2002.

[4] A.J.McNeil, T.Saladin," *Tutorials on The Peaks over Thresholds Method for Estimating High Quantiles of Loss Distributions*", Department Mathematik, ETH Zentrum,Zurich,1997.

[5] P.K. Saha and J.K. Udupa, Optimum image thresholding via class uncertainty and region homogeneity,in: *IEEE Trans. Pattern Anal. Mach. Intell.*, 23, no .7 (2001),pp . 689–706.

[6] N. Otsu, A thresholding selection method from gray- level histograms, in: *IEEE Trans. Syst. Man Cybern*, 9,no . 1 (1979), pp . 62– 66.

[7] J. Kittler and J. Illingworth, On threshold selection using clustering criteria,in: *IEEE Trans. Syst. Man Cybern.*, 15, no . 5(1985), pp . 652–655.

[8]Q. Hu,Z. Hou and W.L. Nowinski, Supervised range-constrained thresholding, in: *IEEE Trans. Image Process,* 15,no . 1 (2006) ,pp . 228– 240.

[9]Y. Qiao, Q. Hu, G.Qian, S. Luo, and W. L. Nowinski,Thresholding based on variance and intensity contrast, in:*Pattern Recognition*,40(2007), pp. 596 – 608.

[10] A.Singh: Digital change detection techniques using remotely-sensed data, in: *Int.J.Remote Sens.*10 (1989), pp-989-1003.

[11] R.S.Lunetta, C.D. Elvidge: Remote Sensing change Detection: Environmental Monitoring Methods and Applications ,*Ann Arbor Press, Chelsea*, 1998.

[12]R.C.Gonzalez and R.E.Woods, Digital Image processing, Second edition, Prentice Hall India, 2005, pp.376 – 394.

[13] R.J.Radke, S.Andra, O.Al-Kofahi, B.Roysam, Image Change Detection Algorithms: A Systematic Survey,in: *IEEE Trans. Image.Process*, 14(2005),pp-294-307.

[14] Ankita Mitra, Arunava De, Anup Kumar Bhattacharjee, Detection of Progression of Lesions in MRI Using Change Detection, in: *Proc. of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA), Springer Advances in Intelligent Systems and Computing* Vol.247, pp.467-473, Nov. 2013