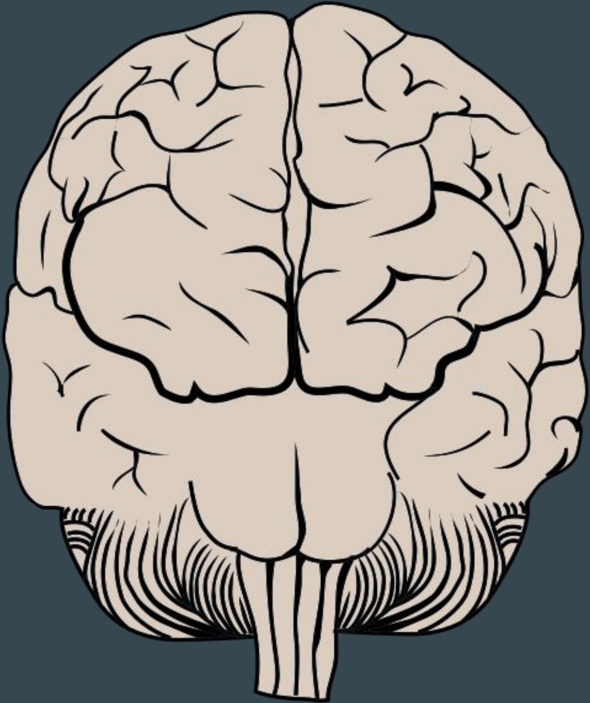
Brain Tumor Detection from MR Image of Brain

BY

Tamal Baroi (Roll: 3201105 No.: 03903)
(Reg No.: 1241511100107 Of 2015)
Shreyashi Pal (Roll: 3202105 No.: 03902)
(Reg No.: 1241721300111 Of 2017)
Suny Dutta (Roll: 3201105 No.: 03907)
(Reg No.: 1241711400112 Of 2017)

Under the Supervision of
Prof. Probir Mondal



Introduction

- The brain is the most important part of the Human body.
- A brain tumor is a result of abnormal growth of brain cells or the migration of cancerous cells of other body parts.
- Brain tumor patients may survive if the tumor is detected in the early stage of the tumor.
- But the diagnosis of a brain tumor is a time-consuming and complex job.
- This time can be saved if we automate the process of brain tumor detection through computers.
- Using a computer, we can extract different features of tumors in less time with more accuracy.
- In our project, we use MR images of the brain to detect tumor.
- There are also different types of problems for identifying tumor regions in images.
- We use some basic and simple algorithms to identify tumor regions in MR images.

Steps to Identify tumor region in MR images of brain

- We review different papers and related works on this problem.
- Most of them have some common steps such as pre-processing, noise removal, segmentation etc.
- We use following steps to detect brain tumor –

    1) Artefact Removal
    2) De-noising
    3) Enhancement
    4) Segmentation
    5) Image with red-boundary tumor region

1. Artefact Removal

- Artefacts in are those pixels or objects in image that not related to or not require in the operation e.g. patient information.
- To remove artefacts, we perform following steps -
  a. Calculate a threshold value for the image using the standard deviation method.
  b. Binarize the image using that threshold value.
  c. Perform connected component analysis and find the maximum connected area.
  d. Assign pixel intensity 1 to that max areas' pixels and set other areas' pixel intensity to 0.
  e. Multiply the main image intensity value of any position(x, y) with the manipulated image's corresponding position(x, y).

Connected Component Analysis is performed as follows in images -

Consider a 2D array with 256 rows and 2 columns. 1st col represent old labels and 2nd col represent new labels called equivalence list.

1. For each non-zero pixel of the image -
   a. Consider 3 upper neighbor pixels and one left side neighbor pixel.
   b. Find minimum value(label) pixel which must not be zero.
   c. If all 4-neighbour are zero then increment the current label by one. Add new labels to the equivalence list. Else re-assign that minimum label to the labels of non-zero neighbor in equivalence list which aren't minimum.
   d. Assign that label to that current pixel.
2. Relabel the image using the equivalence list. For a pixel, search for the label in 1st column and assign the corresponding label from 2nd column.
3. Pixels with the same label will represent connected components.

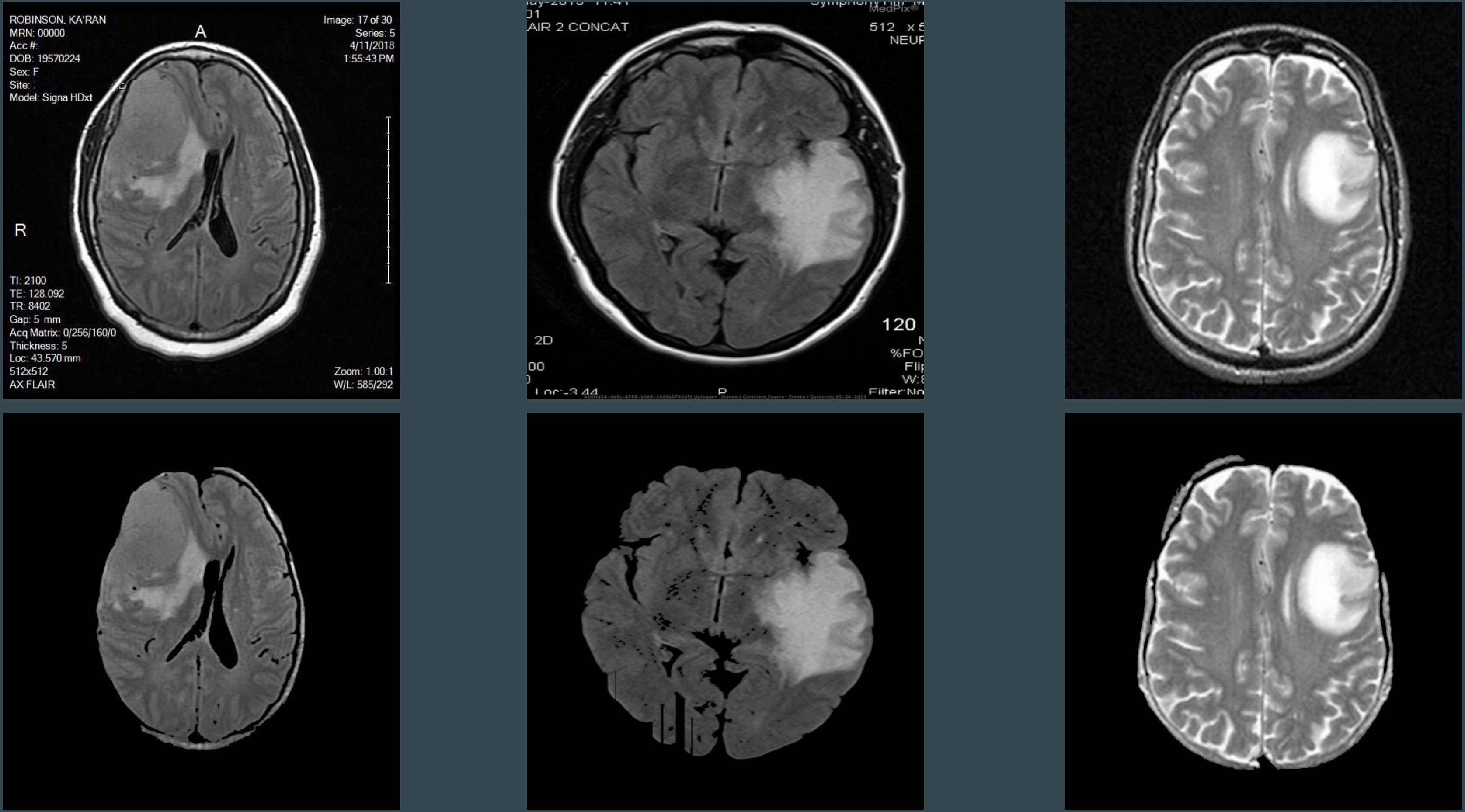


Fig 1 : [Upper Row] Image with Artefacts [Lower Row] Image without Artefacts

2. Denoising

- Noise is abnormal pixel values that are not related to the image subject.
- We use Gaussian smoothing as a noise suppressing method.
- As working with images, we use the 2D Gaussian function -

$$G(x,y)=\frac{1}{2\pi\sigma^2}e^{\frac{-x^2+y^2}{2\sigma^2}} \quad \dots (2.1)$$

where σ is the standard deviation of the distribution.

- A graphical representation of Gaussian distribution with mean 0 and σ = 1 is shown in Fig-2.1.
- For convolving an image with 2D Gaussian function, we take discrete values from continuous Gaussian distribution and store them in a kernel (N x N). e.g. Fig-2.2.

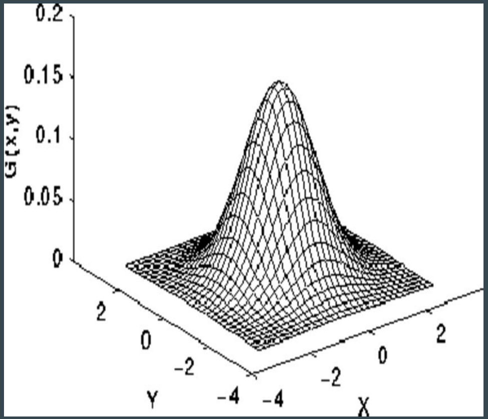


Fig-2.1

1	4	7	4	1
4	16	26	16	4
7	26	41	26	7
4	16	26	16	4
1	4	7	4	1

Fig-2.2

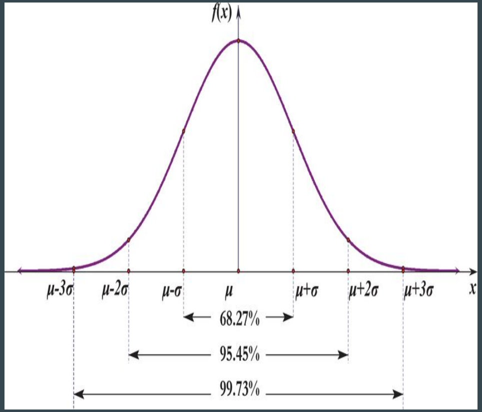


Fig-2.3

❖ Some important properties of Gaussian smoothing filter

- The central pixel is given the highest weight than neighbor pixels and weight decreases (tends to zero) with increasing distance from the center.
- The kernel is separable which allows faster computation.
- A larger value of σ produces a wider peak of distribution (greater blurring).

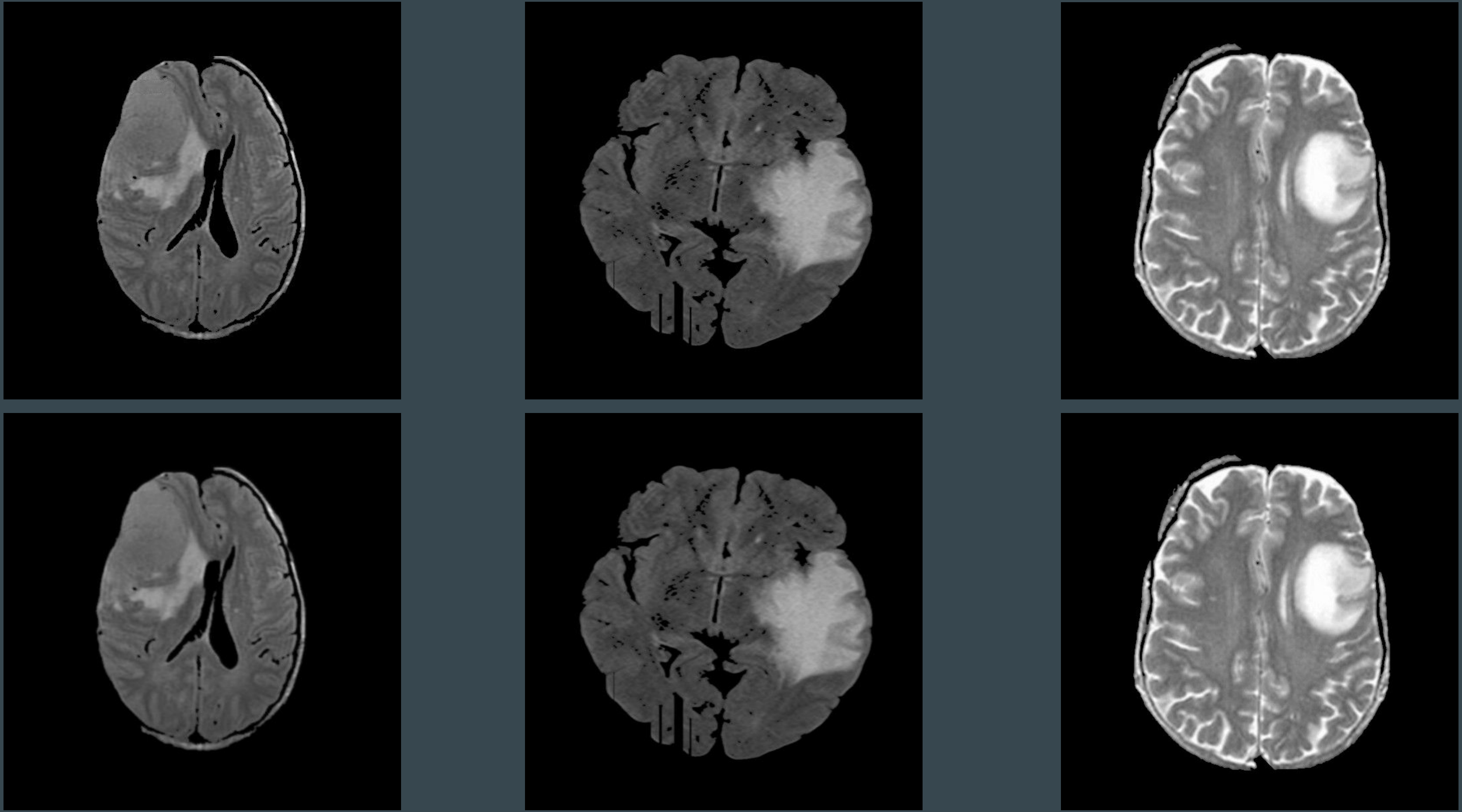


Fig-2.4 : [Upper Row] Image with Noise [Lower Row] Image without Noise

3. Enhancement

- Enhancement means enhancing the pixel values for better results.
- Enhancing is done using the Histogram Equalization method.

Histogram Equalization : The algorithm is as follows -

1. Calculate the total number of pixels for an intensity k = 0,1,2,.....,255.
2. Compute the probability appearance of a pixel intensity.
3. Calculate the cumulative distribution process.
4. Multiply the results of the cumulative distribution process by 255 and round off to the nearest integer.

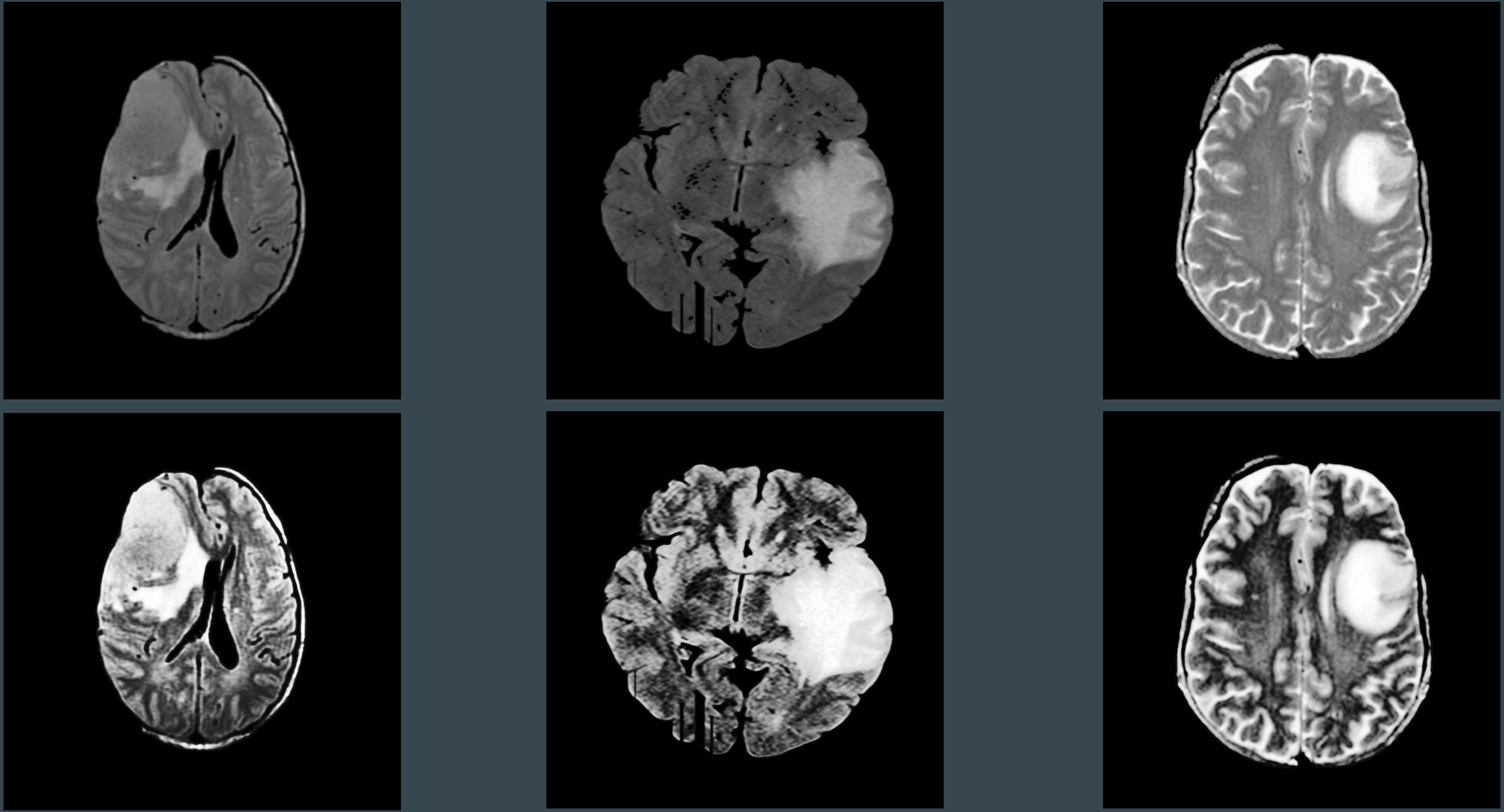


Fig-3.1 : [Upper Row] Image Before enhancement [Lower Row] Image after Enhancement

4. Segmentation

Image segmentation is a process to partition an image into discrete sets that ease the job of further processing.

Fuzzy C-Means algorithm:

- It is a segmentation algorithm developed by Dann and improved by Bezdek.
- This algorithm calculates the likelihood of a pixel for a cluster/set.
- One pixel may belong to multiple clusters.
- The number of clusters must be pre-defined.
- The degree of membership must be between zero and one for each data point for each cluster and depends on the distance between the cluster center and data point.
- Membership value close to one implies a high degree of similarity and close to zero implies a degree of dissimilarity between the point and that cluster.

Before understanding the algorithm, we need to understand some terms used in the algorithm -

✓ **Iteration** - In each iteration of this algorithm, the following objective function is minimized.

$$j = \sum_{i=1}^{N} \sum_{j=1}^{C} d_{ij} \left\| x_i - c_j \right\|^2$$

... (4.1)

- Here, N is the number of data points, C is the number of clusters required.
- c_j is the center vector for cluster j.
- d_{ij} is the degree of membership for the i-th data point x_i in cluster j.
- The norm, || x_i − c_j || measures the similarity (or closeness) of the data point x_i to the centre vector c_j of cluster j.

✓ **Degree Of Membership** - For a given data point x_i , the degree of its membership to cluster j is calculated as follows :

$$d_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - c_k \right\|} \right)^{\frac{2}{m-1}}} \qquad \ldots\ (4.2)$$

- where, m is the fuzziness coefficient and the center vector c_j is calculated as follows:

$$c_j = \frac{\sum_{i=1}^{N} d_{ij}^{m} . x_i}{\sum_{i=1}^{N} d_{ij}^{m}} \qquad \ldots\ (4.3)$$

- In above eq^{n} , d_{ij} is the value of the degree of membership calculated in the previous iteration.
- Note that at the start of the algorithm, the degree of membership for data point i to cluster j is initialized with a random value t_{ij} , $0 \le t_{ij} \le 1$ such that sum of degree of membership of any data point for a all cluster is 1.

✓ **Fuzziness coefficient** -

- The value of the coefficient must be between 1 < m < ∞.
- The fuzziness coefficient 'm' measures the tolerance of the required clustering.
- This value determines how much the clusters can overlap with one another. The higher the value of m, the larger the overlap between clusters.
- In most of the case, its value is 2.

✓ **Termination Condition** - Iteration in the FCM algorithm can be stopped in the following ways.

- First, we can specify no. of iterations. So that algorithm will stop executing after completing that no. of iterations.
- For a small positive number €, one of the following eq^{n} can be checked -

$$\max_{i,j} \left| d_{ij}^{(t)} - d_{ij}^{(t-1)} \right| < \epsilon \quad \text{... (4.4)}$$

$$\max_{1 \le i \le c} \left\| c_{i}^{(t)} - c_{i}^{(t-1)} \right\| < \epsilon \quad \text{... (4.5)}$$

❖ The Fuzzy C-Means algorithm -

Assume that, $X = x_1, x_2, x_3, ..., x_n$ be the set of data points and $V = v_1, v_2, v_3, ..., v_n$ be the set of centers.

1. Assign the initial value of the number of clusters C(>1), fuzziness index m(=2), maximum iteration I_max and threshold €.
2. Initialize the fuzzy partition $d^{(0)}$ randomly according to the constraints of the degree of membership.
3. At the t-step, calculate c cluster centroids $v^{(t)}$ according to eq^n (4.3).
4. Calculate the objective function $J^{(t)}$ according to eq^n (4.1). If $| J^{(t)} - J^{(t-1)} |$ or $t > I_{max}$, then stop; otherwise continue to step-5.
5. Calculate $d^{(t+1)}$ according to eq^n (4.2) and return to step-3.

Cluster Validity Index -

- As mentioned earlier, the number of clusters must be predefined.
- Choosing no. of the cluster for better result isn't a manual task and need to consider different parameter about the dataset.
- First Xie and Beni, in their paper, propose a cluster validity index concept. But it has some drawbacks.
- Bensaid define another validity index which is as follows -

$$V_B(d,c)=\sum_{k=1}^{C}\frac{\sum_{i=1}^{N} d_{ik}^{m}\left\|x_i-c_k\right\|^2}{n_k\sum_{j=1}^{C}\left\|c_j-c_k\right\|^2} \qquad \text{... (4.6)}$$

Here, n_k is the fuzzy cardinality of k-th cluster and calculated as follows -

$$n_k=\sum_{i=1}^{n} d_{ik} \qquad \text{... (4.7)}$$

- The nominator term denotes the variation of the k-th cluster and the denominator term represents compactness.

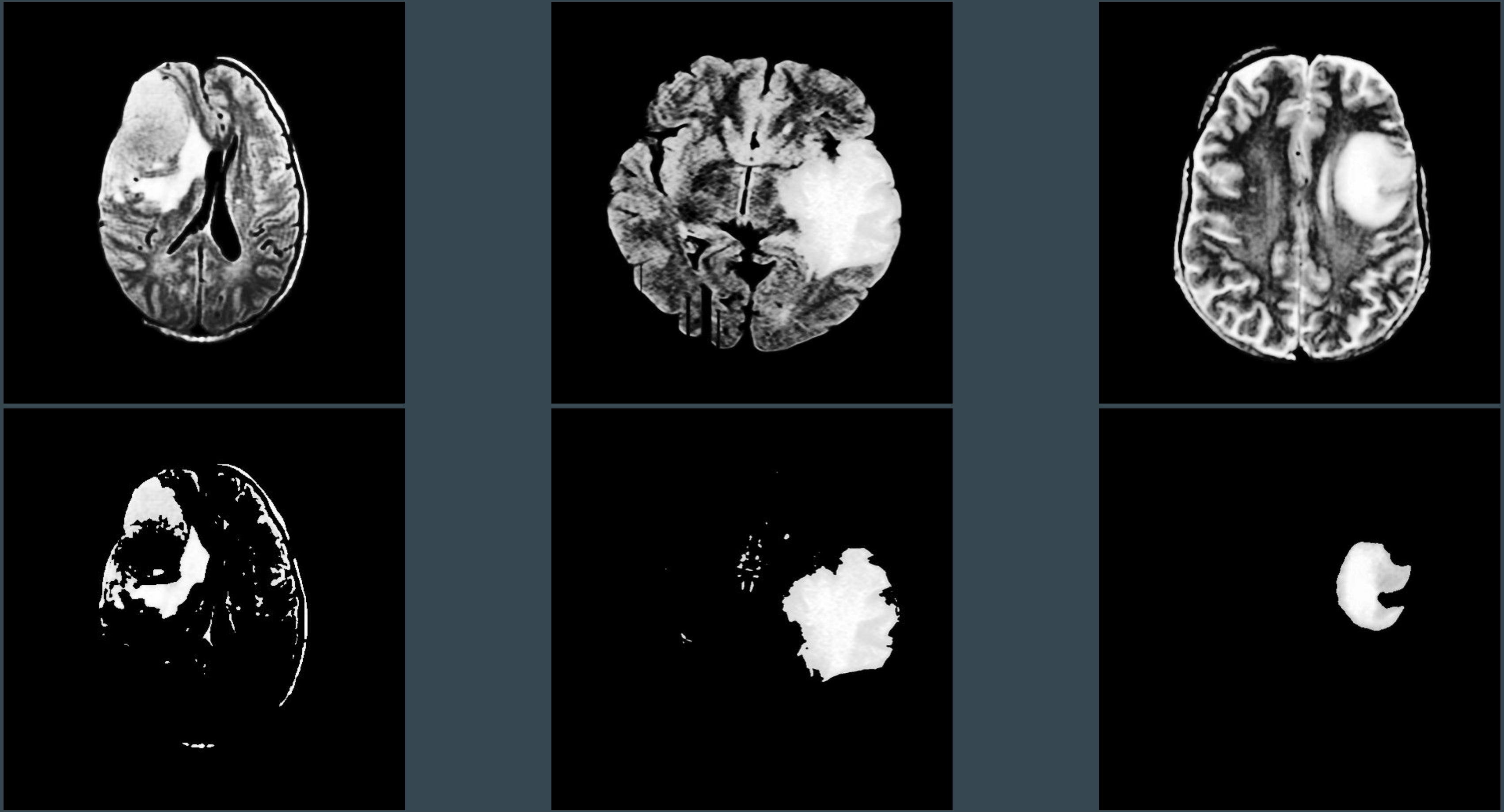


Fig-4.1 : [Upper Row] Image before Segmentation [Lower Row] Image after Segmentation

5. Image with Red Boundary Tumor

We find the tumor area by identifying the max segmented connected area and then we use the Sobel operator to get red-boundary of the tumor and then we clap that red-boundary with the main image. The result is as follows -

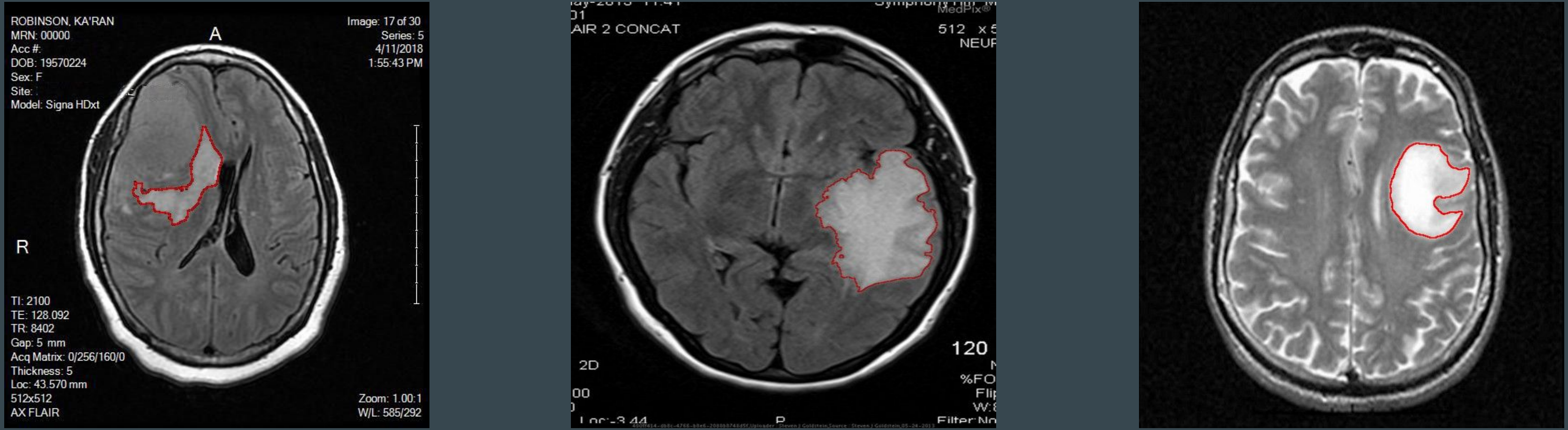


Fig - 5.1 : Image with Red Boundary Tumor

Conclusion

- We successfully detect most of the tumor area and surround the area with red pixels.
- In this project, we use simple approaches to conventional algorithms used widely.
- Anyone can understand the approaches to detect a tumor.
- Using a simple version of algorithms, different problems such as accuracy, inefficiency, less good result, limited uses, etc. becomes understandable.
- As the future goal of this project, we use an enhanced version of the algorithms.
- For an effective understanding of tumor, a 3D model can be built which will specify size, position, surrounding areas, etc.

-: Thank You :-