Brain Tumor Detection and Classification using Convolution Neural Network

Manjunath S, Sanjay Pande M B, Raveesh B N, Madhusudhan G K

Abstract: Understanding Human activity has lead researchers to work on one of the major organ of human body namely Brain. The smooth function of Human Brain enhances the activities of human body. The systematic working of Human brain is affected by various causes. In the present work, we have taken one such cause that is Brain tumor, which is mainly due to abnormal growth of Cells in Brain.The recognition of Brain is generally done by Magnetic resonance imaging (MRI). The major drawback of this is to find the exact location/position. Hence it becomes important to find the means and methods to detect, identify and classify the disease based upon the image. The proposed work involves Extraction to grading of Tumor to be relevant class. The complexity of the present work is due to conversion of the extracted image as symbolic data and use of Convolution Neural networks. The experimentation were corroborated with BPNN and CNN classifier.

***Keywords—Convolution, Neural Network, Brain Tumor, LBP,Fuzzy-*C**

I INTRODUCTION

The unregulated growth of cancerous cells in human body is a threat to the human life.The tumor growth in the brain is always a dangerous state for a human life.The tumor in the brain is either benign or malignant. The uniform tumor structure is present in the benign type without active cancer cells.The non-uniform structure exists in the malignant structure with active cancer cells. The different grades of tumors are low and high. The scaling is done from grade I to grade IV to differentiate the benign as well as the malignant tumor.Example for the grade I and II glioma is benign tumor. The grade III and IV glioma example is malignant tumor. The low gradetumor is to be treated at the earliest in order to restrict its growth in to the high grade tumor. The use of computed tomography or else the magnetic resonance imaging is needed to identify the grade II gliomas in the patients in a regular interval of 6 months.The low grade I and II glioma brain tumorare regarded as the treatable cases. The grade III and IV malignant tumor are treatable using chemotherapy or by radiotherapy. The mid-gradetumoris called as anaplastic astrocytomas. It shows growth of irregularity with growth index increased in comparison with the other tumors of low grade.

Revised Manuscript Received on May 20, 2019

Manjunath S Research Scholar, Department of Computer Science and Engineering, Jain University, Bengaluru, Karnataka, India'

Dr. Sanjay Pande M B Professor and Head, Department of Computer Science and Engineering, GMIT, Davangere

Madhusudhan G K, Assistant Professor Department of Computer Science and Engineering, VVIET, Mysuru Karnataka,

The glioblastoma is of the highest grade.In order to give suitable treatment it is a necessary step to identify the brain tumor at the earliest stage of its occurrence. After the identification of the brain tumor, it is needed to be evaluated for parameters like location, size with effect on the neighbouring areas. The treatment is decided based on parameters. The early detection of the tumor is necessary to increase the survival chance of the patient with tumor infection [1].

II METHODOLOGY

A. Pre – processing technique

Pre-processing converts the original image to gray scale to reduce the remove the unwanted noise, for image reconstruction along with the image enhancement. At every point the intensity of the gray scale image is either 1 or 0. Every pixel has value of the intensity lacking of color. Intensity values may also be in terms of fractions. It is helpful in the segmentation method by giving the precise information. After converting the image to gray scale, the noise is to be removed from by applying it to the filter. The filter can be either high pass frequency or low pass frequency. The different researchers have used several types of filters for the noise removal of the image. Dr M.Karnan, A.Lakshmi and Dr A.S.Balchandra and many others have used the median filter. R.B.Dubey for MRI image has used the Gaussian filter to remove the noise. Sobel filter was applied by the Deepthimurthy in the research work. Sobel filter works on the principle to calculate the intensity differences of the pixel and it is a derivative mask type [2].

In order to obtain the quality image the primary step in the image analysis is the enhancement of the image. Afterwards it is to minimize the noise for obtaining the good quality image. The main advantage of the image enhancement ensures that the image edges of importance are highlighted with improvement in the sharpness aiding in the detection of presence of the tumor [3].Several types of filters are capable to reduce and remove the noise from the image. The average filters remove the noise from the image but sharpness needs to be compromised. The median filteris applied to remove the noise salt and pepper. Image sharpening is done by applying the high pass filter. The object boundaries can be enhanced by the Gaussian filter. The Gaussian filter is good in the detection of edge and highlighting the tumor [4].

Retrieval Number: A10080581C19/19©BEIESP

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

The image de-noising and Poisson noise can be removed from the. It employs the window sliding across the image. In the window the pixel median intensity value is taken as the pixel output intensity value during the process. It conserves the image edges and also reduces the noise.Each pixel is fixed to the median of the pixel values in the locality of the analogous pixel values [5].

B Detection and identification of lesions

The tumor size real assessment is not determined by the lesion diameter due to the reasons like lesion irregularity, variations in the measurement amongst the inter-observer, intra-observer along with the various levels in the scanning levels which are gathered from various analyses [6].

The tumors of the brain contain different shapes, locations and size. The regular method to detect the lesion is done by tracing manually by the experts. The automatic identification of the lesion helps in minimizing the time delay caused by the practitioners for analysis [7].

C Fuzzy-C Means Clustering Segmentation

The technique of allocating an objects set to groups is defined as clusters. Hard and soft are the two different types of clusters. When the division of data is into different cluster it is called as hard clusters. If the data elements fit into to greater than single cluster it is known as soft cluster. The information extraction should be done in a systematic method from the existing database. In the analysis of the cluster, the N object is partitioned into C - clusters. The objects within the cluster requirebeing same to each other and dissimilar clustershould be different with each other. It is helpful in the dataset compact representation.Cluster is useful to know the patterns in the dataset. Application of the algorithms is done to know the clusters. It is a iterative method to know the data from the dataset. The commonly used clustering algorithm is based on fuzzy.The fuzzy set theory was suggested by Zadeh in 1965. It indicates the uncertainty belonging which was defined by a member function. It basically employs the fuzzy partitioning in which the data point fits to every group with different membership scores ranging between 0 and 1.It is based on the assignment of membership to every data point analogous to every cluster center grounded on the origin of distance amongst the data point in additionto the cluster center. Based on the data number large intensity adjacent to the center of the cluster, its membership is largely nearer to the specific cluster center. Its gains are converging and unsupervised.The various limitations are sensitivity towards the initial guess, more computational time and noise. The various cluster means are Possibilistic C Means (PCM), Fuzzy - Possibilistic CMeans (FPCM) also Possibilistic Fuzzy C-Means (PFCM).The squared norm is applied by the FCM to calculate the data point and prototype resemblance [8].

D GLCM

The features of image shape can be extracted by connected regions and features of image texture can be extracted by Gray Level Co-occurrence Matrix (GLCM). Features give knowledge about the image. The feature set is known as the feature vector. It is the primary input for classification algorithm.

Thedifferent researchers have carried out research work as follows.

Table 1: Research work carried out by different researchers in feature extraction

Sl. No.	Feature extraction	Author
1	Rough set theory with 90 % classification efficiency	Rajesh et al [9]
2	complementary wavelet transform	Hiremath et al [10]
3	dimensionality reduction using sub band grouping and selection	Huang et al [11]
4	statistical texture features from input dataset, Classification efficiency obtained was 80%.	Ramteke et al [12]
5	texture symmetry and intensity based features The average accuracy was 96.02%	Xuan et al [13]
6	Principal component analysis (PCA) for the extraction of features, classification accuracy obtained was 96.33%.	Othman et al [14]

The second order statistic is employed to know the pixel spatial relationship. Several gray level co-occurrences mixtures in the image are represented by GLCM. Calculating frequently the pixel with intensity value 'i ' in an exact spatial relationship to a pixel j GLCM is created.The default spatial relationship is amongst the instantneighbour on its right and its pixel. The linkcan be specified with dissimilar offsets and angles. The GLCM matrix contains of 256*256 matrix; for an 8-bit image the intensity values are given by Ii and Ij.Isotropic GLCM can be calculated by adding the results from diverse directions by this rotation invariant GLCM can be accomplished [15].

E LBP

The Local Binary Patterns (LBP), histograms can be extracted from the small regions of the face and one feature vector is formed. The face is represented effectively by the feature vector which is useful in measuring the resemblances between images.The maximum information about the image is to be extracted in the feature extraction stage. The features after obtaining from the images are compared along with the database images. It is achieved by the classification stage. The classification stage output is the image identity from the database having maximum matching score, therefore minute variations are compared with the input image. Threshold value helps to verify that whether differences are sufficiently lesser or not.The LBP is one the method for feature extraction introduced by the Ojala et al in 1996[16]. The digital image shape and texture can be described by the LBP.

Retrieval Number: A10080581C19/19©BEIESP

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The image is divided into various minor areas for extracting the features.The feature basically contains the binary patterns which refer to the pixel surroundings in the area. The features are concatenated to form a one histogram feature which in turn forms a image representation.The measurement of the similarity amongst the histograms is useful in the comparison of the images. LBP method is very effective in the face recognition with respect to performance in speed and discrimination parameters.

F LBP Principle

Introduced the LBP operatorOjala et al. [16].It basicallyworks byusing the eight neighbours of a pixel. The centerpixel valueis used as the threshold. Whenever the nearby pixel possesses a greater value or same in comparison with the center pixel, the pixel is assigned with one value or else zero value. The center pixel LBP code is generated with concatenating the eight zeros or ones to form the binary code as presented in the figure 1.

1	1	1
1		0
0	0	1

5	4	3
4	3	1
2	0	3

Fig1: The original LBP operator

Binary: 11101001

Decimal: 233

The LBP operator can also be used for different size pixels. From the centerpixel a circle is arranged with R as the radius. On the edges of the circle, sampling points P are used for comparing with the pixel at the center. In order to obtain the sampling points in the neighbourhood values with different radius in addition tomany pixels, the interpolation is required.
In place ofneighbourhoods the representation (P, R) is used. Figure 2.Shows three neighboursets for diverse values for P and R.

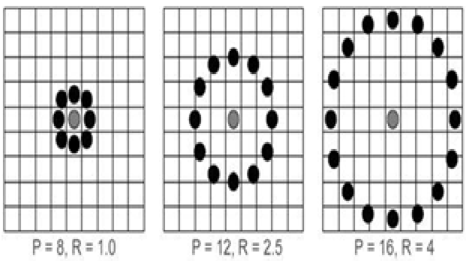

Fig 2: Circularly neighbour-sets for three differentvalues of P and R

Let, the center pixel co-ordinates be (x_c, y_c), and P neighbors coordinates be neighbors (x_p, y_p), with circle having radius R is calculated by using sines and cosines as follows:

$$x_p = x_c + R\cos(2\ p/P) \qquad (1)$$

$$x_p = y_c + R\sin(2\ p/P) \qquad (2)$$

Whenever thecenter pixel value is g_c with the neighborsgray value as g_p, the value of p = 0,...,P - 1, local neighbourhood pixel (x_c, y_c) texture T is calculated as

$$T = t\ (g_c, g_0, \ldots, g_{Pi}) \qquad (3)$$

After getting the point values, the texture can be described by different way. It is done by the subtraction of center located pixel value by the pixel values which is located on the circle. By this method, the local texture is denoted as a combined spreading of the center pixel value as well as the differences:

$$T = t\ (g_c, g_0\text{-}g_c, \ldots, g_{p-1}\text{-}g_c) \qquad (4)$$

The whole luminance of the image is represented by (g_c). It is not related to the texture of the local image. For texture analysis this does not give information.

The textural characteristics information in the basic joint distribution (Eq.3) is conserved in the combined dissimilarity distribution[17].

$$T \approx t\ (g_0\text{-}g_c, \ldots, g_{p-1}\text{-}g_c) \quad (5)$$

G Feature selection – Principal Component Analysis (PCA)

In order to analyze the image, image dimension reduction is performed. The most relevant information is to be preserved while reducing the image. Over all in the pattern recognition as well as general classification problems methods like Independent Component Analysis (ICA), fisher line discriminate analysis and principal component analysis (PCA) are used. The variable selection procedure is extensively used. The multi-layer perception is applied in the variable selection. The regression method lacks the unified optimality criteria. PCA has good optimality property.
Let the random vector for a linear transformation be $X \in R^n$along with zero mean, covariance matrix $\sum_x$ random vector to a smaller dimension $Y \in R^n$, q < n

$$Y = A_q^T X \qquad (6)$$

Along with $A_q^R\ A_q = I_q$,here $I_q$is identity matrix of q x q. $A_q$is the matrix of n x q in which column are q ortho-normal values of Eigen in the covariance matrix $\sum_x$ .Thelinear transformation has ten optimal properties. One of the most significant properties is the expansion of the points spreading in the minor dimensional space of the points in the minor dimensional space, i.e, transformed space points are maintained to the greater extent by keeping the original space changes. The original data and the predicted data mean square error is reduced.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

In order to select a sub-set of the basic variables for the vector in random X, it is visualized as X transformation in linear using the matrix of transformation

$$A_q = \begin{bmatrix} I_q \\ [\mathbf{0}]_{(n-q)\times q} \end{bmatrix} \quad (7)$$

Or else other matrix with $A_q$rows permutation. Without losing the generality, let $A_q$be the matrix of transformation, and covariance matrix is re-written as below,

$$\Sigma = \begin{bmatrix} \{\Sigma_{11}\}_{q\times q} & \{\Sigma_{12}\}_{q\times(n-q)} \\ \{\Sigma_{21}\}_{(n-q)\times q} & \{\Sigma_{22}\}_{(n-q)\times(n-q)} \end{bmatrix} \quad (8)$$

The subset to maximize $|\Sigma_Y| = |\Sigma_{11}|$ is same as the points spread maximization in the lesser dimensional space, which retains the basic data changes. The reduction of the mean square prediction error is same as reducing the face of $\Sigma_{22|1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ .For obtaining the q feature for optimal subset, any one quantity from the above is calculated to obtain every probable group for features of q. The subset discovery is the major drawback. To a larger feature vector it is difficult to get the subset [18].

H Convolutional Neural Network (CNN)

Convolutional Layer**.**Let the convolutional layer be 1. The layer of input 1 contains m 1^ (l−1) maps of feature by the preceding layer, with each size m 2 ^ (l−1) × m3 ^ (l−1).When 1=1, input is nothing but a one image I with either single or many channel. The convolutional neural network receives input of raw images. The layer of the output is m 1^ (l) feature maps with m2 (l) × m3 (l) size. The feature of i[th]feature map

The layer 1 contains m 1^ (l) feature maps with size m2 (l) × m3 (l). The feature map i[th] of layer 1 is represented as Yi^ (l), and calculated as below

$$Y_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{m_1^{(l-1)}} K_{i,j}^{(l)} * Y_j^{(l-1)} \quad (9)$$

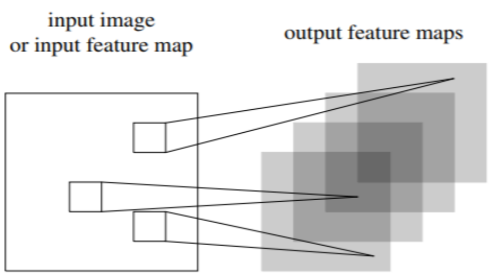

Fig3: Feature maps form the input image

The single convolution layer is shown in the below figure 3. In the layer 1 which is convolutional layer, (if 1=1) in the input image or else previous layer feature map is convolved using several filter gives output feature maps of layer 1.

where $B_i^{(l)}$ is a bias matrix and $K_{i,j}^{(l)}$ is the filter of size $2h_1^{(l)}+1 \times 2h_2^{(l)}+1$ connecting the j^{th} feature map in

In the layer l with i[th] feature map along withthe layer (l-1) [19].The terms, m2^ (l) and m3^ (l) are affected by the effects of the border. In the feature map input valid region the discrete convolution, the pixel in which the sum of equation (9) is denoted, the size of the output of the feature map possess the size

$$m_2^{(l)} = m_2^{(l-1)} - 2h_1^{(l)} \quad \text{and} \quad m_3^{(l)} = m_3^{(l-1)} - 2h_2^{(l)}. \quad (10)$$

In calculating the feature map of fixed type Yi^ (l), the used filters are of the same size, where

K I,j^ (l) =Ki,k^ (l) for j ~= k. (11)

For relating the layer of convolution for its operation as shown in the equation (11) with the multilayer perceptron, the equation is rewritten.

Every feature map Y i^(l) in the layer 1 contains m2^ (l) · m3^ (l) units organized in the 2-D array. The position (r,s) unit calculates the output

$$\begin{aligned}\left(Y_i^{(l)}\right)_{r,s} &= \left(B_i^{(l)}\right)_{r,s} + \sum_{j=1}^{m_1^{(l-1)}} \left(K_{i,j}^{(l)} * Y_j^{(l-1)}\right)_{r,s} \\ &= \left(B_i^{(l)}\right)_{r,s} + \sum_{j=1}^{m_1^{(l-1)}} \sum_{u=-h_1^{(l)}}^{h_1^{(l)}} \sum_{v=-h_2^{(l)}}^{h_2^{(l)}} \left(K_{i,j}^{(l)}\right)_{u,v} \left(Y_j^{(l-1)}\right)_{r+u,s+v}.\end{aligned} \quad (12)$$

The network weights which are trainable established in the K (l) i, j filters also the bias matrices B (1) i. The s (l) 1 and s (l) 2 skipping factors are used for subsampling.The pixel numbers are constant in the directions of vertical and horizontal before applying to the filter second time.

The output feature size of the map with skipping factors as above is shown below,

$$m_2^{(l)} = \frac{m_2^{(l-1)} - 2h_1^{(l)}}{s_1^{(l)} + 1} \quad \text{and} \quad m_3^{(l)} = \frac{m_3^{(l-1)} - 2h_2^{(l)}}{s_2^{(l)} + 1}. \quad (13)$$

Local Contrast Normalization Layer

Let the layer of contrast normalization be 1. The aim of the contrast normalization local is for applying the native effectiveness amongst the neighbouring units along with feature map also with the similar spatial position within several feature maps.

Retrieval Number: A10080581C19/19©BEIESP

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The normalization for subtraction operation computed as follows for a given m (l−1) 1 feature,

$$Y_i^{(l)} = Y_i^{(l-1)} - \sum_{j=1}^{m^{(l-1)}} K_{G(\sigma)} * Y_j^{(l-1)} \quad (14)$$

From equation n (14), the KG (σ) is Gaussian filter. As discussed in the [20], brightness localization is applied along with the rectified linear units. The layer 1 output is shown below,

$$\left(Y_i^{(l)}\right)_{r,s} = \frac{\left(Y_i^{(l-1)}\right)_{r,s}}{\left(\kappa + \mu \sum_{j=1}^{m_1^{(l-1)}} \left(Y_j^{(l-1)}\right)_{r,s}^2\right)^{\mu}} \quad (15)$$

In which, hyper parameters are κ, λ, μ which are set by validation set [20].

In the equation (15), the sum runs abovea subsettowards the layer (l −1). The layers of native contrast normalization represented by NS and NB, respectively.

Feature Pooling and Subsampling Layer

The resolution reduction is done in several methods [21]. In the figure 4 and [21] which shows the pooling along with subsampling layer. Let layer 1 is a pooling along with subsampling layer. For a given m (1-1) 1 = 4 earlier layer feature maps, every features mapare sub-sampled as well as pooled independently. For every unit in the m (1)1 = 4 feature map output signifies the maximum or average in the stable window in the analogous feature map of the layer (l −1).

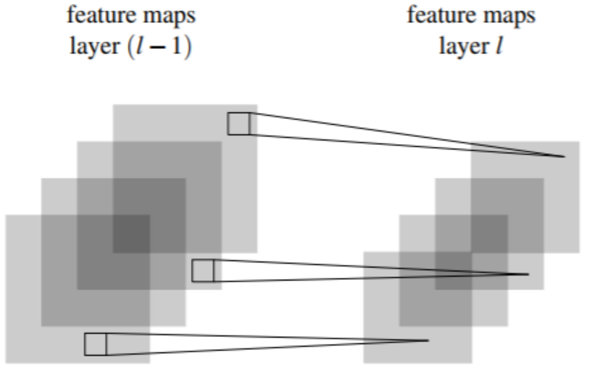

Fig4: Feature map for layer l and (l-1)

In the conventional convolutional neural network, the subsampling is achieved with the application of the skipping factors. In the pooling layer 1, the m (l) 1 = m (l−1) 1 is output features map with size condensed. The pooling gets acted with the windows placement at the non-overlapping locations in every feature-map. By placing single value for every window the feature maps are subsampled. The two pooling types are, average pooling and max pooling.

The usage of the boxcar filter22 operation is known as the average pooling with layer identified by PA. In max pooling, every window the highest value is considered. In this, the layer is identified as PM [22].The max pooling helps in obtaining the faster convergence during convergence. While using the boxcar filter22, the average pooling operation and here layer is denoted by PA. For max pooling, each window maximum value is considered. The layer is represented by PM. As deliberatedin [21]the max pooling helps in obtaining the earlier convergence during training. With the application of the overlapping windows, with size 2p x 2p and the placing of q units apart, both the max and average pooling are applied. Then, if q<p the overlap of the windows occurs. It gives the possibility of reduction in the training set [20].

III EXPERIMENTAL RESULTS

Table 2: Comparison of ANN and CNN for classification false rate

Classifier name	Total no. of Test	No. of correct class	No. of wrong class	Accuracy	Sensitivity	Specificity
ANN	370	273	97	73.6486	0.47297	0.90541
CNN	370	320	50	86.4865	0.72973	0.91892

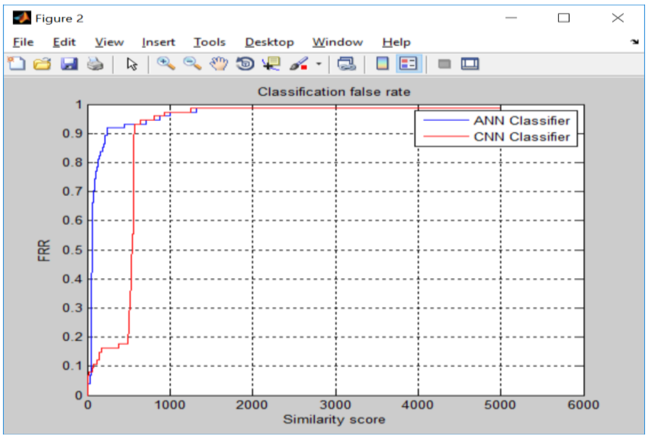

Fig5: Comparison of ANN and CNN for classification false rate

Retrieval Number: A10080581C19/19©BEIESP
Published By:
Blue Eyes Intelligence Engineering & Sciences Publication

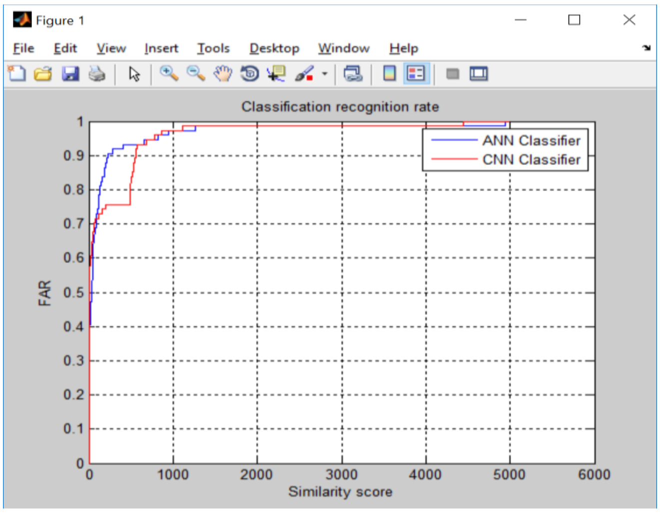

Fig 6: Classification recognition rate

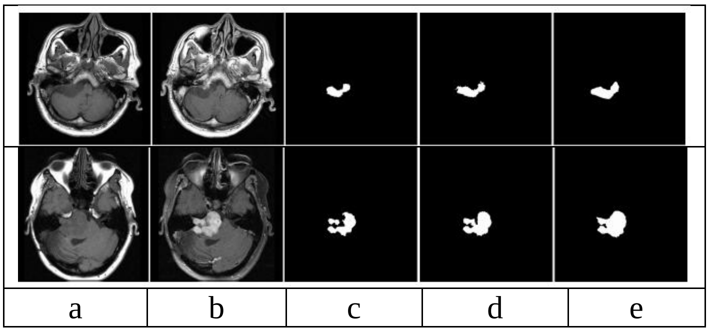

Fig 7: Brain Tumor Segmentation Results

Brain tumor segmentation from two patients and the corresponding ground (a)-(b) as in Figure 7: The original T1W brain MR images before and after contrast enhancement; (c): The segmentation results using Semi supervised fuzzy clustering method; (d): The segmentation results using the proposed fuzzy C means clustering method; (e): The ground truth traced by the radiologist.

IV CONCLUSION

The CNN classification method is found to be more accurate with a percentage of 86.4865, with increased sensitivity of 0.72973 and higher specificity of 0.91892 in comparison with ANN method results. The CNN method found to be better than the ANN method in the brain tumor detection.

REFERENCES

1. NileshBhaskarraoBahadure, Arun Kumar Ray, and Har Pal Thethi, Image Analysis for MRI Based Brain Tumor Detection and Feature Extraction Using Biologically Inspired BWT and SVM, International Journal of Biomedical Imaging Volume 2017, Article ID 9749108, 12 pages https://doi.org/10.1155/2017/9749108.
2. Luxitkappor,SanjeevThankur, A survey on Brain tumor detection using image processing techniques, 7th International conference on cloud computing, Data Science & Engineering – Confluence ,page no. 582 – 585.
3. A. Sindhu, S.Meera, A survey on Detecting Brain Tumor in MRI, IJIRCCE, Vol. 3, Issue 1, January 2015.
4. Malathi R, Dr Nadirabanu Kamal A R ,Braintumor detection and identification using K – Means clustering technique, , IJANA, 27th March 2015.
5. A. Sindhu, S.Meera,A survey on Detecting Brain Tumor in MRI, IJIRCCE, Vol. 3, Issue 1, January 2015.
6. Ed – EdilyMohd. Ansari, Muhd. MuddZakkirMohd. Hatta, ZawZawHtike and Shoon Lee Win, Tumor detection in medical imaging: A survey, IJAIT, Vol 4, No. 1, February 2014.
7. Ana sanjuan et al., Automated identification of brain tumors from single MR images based on segmentation with refined patient-specific priors, frontiers in neuroscience, methods article, 17th December 2013, doi: 3389/fnins.2013.00241
8. R.Suganya, R.Shanthi, Fuzzy C-Means Algorithm – A Review, IJSRP, Volume 2, Issue 11, November 2012, ISSN 2250-3153.
9. T. Rajesh, R. Suja Mani Malar," Rough Set Theory and Feed Forward Neural Network Based Brain Tumor Detection in Magnetic Resonance Images",IEEE International Conference on Advanced Nanomaterials& Emerging Engineering Technologies, 20 13.
10. R. J. Ramteke1, KhachaneMonali Y., "Automatic Medical Image Classification and Abnormality Detection Using K-Nearest Neighbour" , International Journal of Advanced Computer Research,Volume-2 Number-4 Issue-6 December-2012.
11. MohdFauzi Othman, MohdAriffanan, MohdBasri, "Probabilistic Neural Network for Brain Tumor Classification" ,IEEE International Conference on Intelligent Systems, Modelling and Simulation,2011.
12. Xiao Xuan, Qingmin Liao, Statistical Structure Analysis in MRI Brain Tumor Segmentation" ,IEEE International Conference on Image and Graphics, 2007.
13. Hiremath P, Shivashankar S. Wavelet basedfeatures for texture classification. Graphics,Vision and Image Processing Journal;6:55-8.2006.
14. Huang K, Aviyente S. Wavelet featureselection for image classification. IEEE Transaction on Image Processing; 17:1709-20, 2008.
15. Shijinkumar P.S etal.Extraction of texture features using GLCM and shape features using connected regions, IJET, ISSN: 0975 - 4024.
16. T .Ojala, M. Pietikainen and D. Harwood, "A comparative study of texture measures with classification based on feature distributions" Pattern Recognition vol. 29, 1996.
17. Md. Adbul Rahim, Md. NajmulHossain, Tanzillah Wahid & Md. ShafilAzam, Face recognition using Local Binary Patterns (LBP), Global Journal of computer science and technology graphics & vision, volume 12 Issue 4 version 1.0 2013.
18. Ira Cohen, Qi Tian, Xian Sean Zhou, Thomas S.Huang, Feature Selection Using Principle Feature Analysis, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana.
19. Y. LeCun, K. Kavukvuoglu, and C. Farabet.Convolutional networks and applications in vision. In Circuits and Systems, International Symposium on, pages 253–256, 2010.
20. A. Krizhevsky, I. Sutskever, and G. E. Hinton.Image Net classification with deep convolutional neural networks.In Advances in Neural Information Processing Systems, pages 1097–1105, 2012.
21. [K. Jarrett, K. Kavukcuogl, M. Ranzato, and Y. LeCun. What is the best multi-stage architecture for object recognition? In Computer Vision, International Conference on, pages 2146–2153, 2009.
22. [D. Scherer, A. Müller, and S. Behnke.Evaluation of pooling operations in convolutional architectures for object recognition. In Artificial Neural Networks, International Conference on, pages 92 – 101, 2010.

AUTHORS PROFILE

Manjunath S Research Scholar, Department of Computer Science and Engineering, Jain University, Bengaluru, Karnataka, India'

Dr. Sanjay Pande M B Professor and Head, Department of Computer Science and Engineering, GMIT, Davangere, Karnataka, India rakroop99@gmail.com

Dr Raveesh B N Head and Preofessor Department of Psychiatry Mysore Medical College, Mysuru Karnataka, India raveesh6@yahoo.com

Madhusudhan G K, Assistant Professor Department of Computer Science and Engineering, VVIET, Mysuru Karnataka, India rmadhugk06@gmail.com

.