Brain Tumor Classification Using Learning Machine Approach And Brain Structure Analysis

M.A.Darshana
Department of VLSI Design
Sethu Institute of Technology
Pullor,Karyapatti.

M.Lakshmishree
Department of VLSI Design
Sethu Institute of Technology
Pullor,karyapatti.

C.Sujatha
Associate Professor,
Department of VLSI Design
Sethu Institute of Technology

ABSTRACT- **The project proposes an automatic support system for stage classification using artificial neural network (learning machine) and to detect Brain Tumor through fuzzy clustering methods for medical application. The detection of the Brain Tumor is a challenging problem, due to the structure of the Tumor cells. This project presents a segmentation method, fuzzy clustering algorithm, for segmenting Magnetic Resonance images to detect the Brain Tumor in its early stages and to analyze anatomical structures. The artificial neural network will be used to classify the stage of Brain Tumor that is benign, malignant or normal. Here fast discrete curvelet transformation is used to analysis texture of an image. The segmentation results will be used as a base for a Computer Aided Diagnosis (CAD) system for early detection of Brain Tumor which will improves the chances of survival for the patient. In brain structure analysis, the tissues which are WM and GM are extracted. Probabilistic Neural Network with radial basis function will be employed to implement an automated Brain Tumor classification. Decision making was performed in two stages: feature extraction using GLCM and the classification using PNN-RBF network. The performance of this classifier was evaluated in terms of training performance and classification accuracies. The simulated results will be shown that classifier and segmentation algorithm provides better accuracy than previous method.**

Keywords—fuzzy clustering algorithm, curvelet transformation, GLCM, PNN RBF.

I. INTRODUCTION

Automated classification and detection of tumors indifferent medical images is motivated by the necessity of high accuracy when dealing with a human life. Also, the computer assistance is demanded in medical institutions due to the fact that it could improve the results of humans in such a domain where the false negative cases must be at a very low rate. It has been proven that double reading of medical images could lead to better tumor detection. But the cost implied in double reading is very high, that's why good software to assist humans in medical institutions is of great interest nowadays. Conventional methods of monitoring and diagnosing the diseases rely on detecting the presence of particular features by a human observer. Due to large number of patients in intensive care units and the need for continuous observation of such conditions, several techniques for automated diagnostic systems have been developed in recent years to attempt to solve this problem. Such techniques work by transforming the mostly qualitative diagnostic criteria into a more objective quantitative feature classification problem. In this project the automated classification of brain magnetic resonance images by using some prior knowledge like pixel intensity and some anatomical features is proposed. Currently there are no methods widely accepted therefore automatic and reliable methods for tumor detection are of great need and interest. The application of PNN in the classification of data for MR images problems are not fully utilized yet. These included the clustering and classification techniques especially for MR images problems with huge scale of data and consuming times and energy if done manually. Thus, fully understanding the recognition, classification or clustering techniques is essential to the developments of Neural Network systems particularly in medicine problems.

Segmentation of brain tissues in gray matter, white matter and tumor on medical images is not only of high interest in serial treatment monitoring of "disease burden" in oncologic imaging, but also gaining popularity with the advance of image guided surgical approaches. Outlining the brain tumor contour is a major step in planning spatially localized radiotherapy (e.g., Cyberknife, iMRT) which is usually done manually on contrast enhanced T1 weighted magnetic resonance images (MRI) in current clinical practice. On acquiring T1 MR images after administration of a contrast agent (gadolinium), blood vessels and parts of the tumor, where the

contrast can pass the blood–brain barrier are observed as hyper intense areas, there are various attempts for brain tumor segmentation in the literature which use a single modality, combine multi modalities and use priors obtained from population atlases.
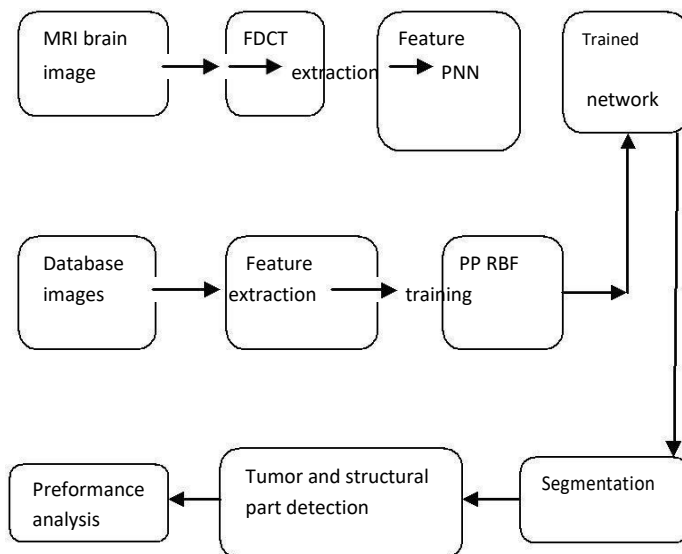
II. BLOCK DIAGRAM

Figure 1. Block diagram

III. PROBABLISTIC NEURAL NETWORKS

Neural networks are predictive models loosely based on the action of biological neurons. The selection of the name "neural network" was one of the great PR successes of the Twentieth Century. It certainly sounds more exciting than a technical description such as "A network of weighted, additive values with nonlinear transfer functions". However, despite the name, neural networks are far from "thinking machines" or "artificial brains". A typical artifical neural network might have a hundred neurons. In comparison, the human nervous system is believed to have about $3x10^{10}$ neurons. We are still light years from "Data".

The original "Perceptron" model was developed by Frank Rosenblatt in 1958. Rosenblatt's model consisted of three layers, (1) a "retina" that distributed inputs to the second layer, (2) "association units" that combine the inputs with weights and trigger a threshold step function which feeds to the output layer, (3) the output layer which combines the values. Unfortunately, the use of a step function in the neurons made the perceptions difficult or impossible to train. A critical analysis of perceptrons

published in 1969 by Marvin Minsky and SeymorePapert pointed out a number of critical weaknesses of perceptrons, and, for a period of time, interest in perceptrons waned. Interest in neural networks was revived in 1986 when David Rumelhart, Geoffrey Hinton and Ronald Williams published "Learning Internal Representations by Error Propagation". They proposed a multilayer neural network with nonlinear but differentiable transfer functions that avoided the pitfalls of the original perceptron's step functions. They also provided a reasonably effective training algorithm for neural networks.

A. Working of PNN network

Although the implementation is very different, probabilistic neural networks are conceptually similar to *K-Nearest Neighbor* (k-NN) models. The basic idea is that a predicted target value of an item is likely to be about the same as other items that have close values of the predictor variables.
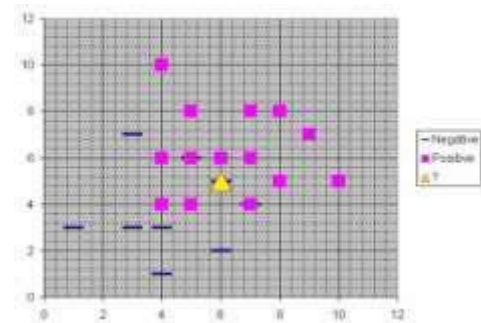

Figure 2. Example to understand the working of PNN network.

Assume that each case in the training set has two predictor variables, *x* and *y*. The cases are plotted using their *x,y* coordinates as shown in the figure. Also assume that the target variable has two categories, *positive* which is denoted by a square and *negative* which is denoted by a dash. Now, suppose we are trying to predict the value of a new case represented by the triangle with predictor values *x*=6, *y*=5.1. Should we predict the target as positive or negative? Note that the triangle is position almost exactly on top of a dash representing a negative value. But that dash is in a fairly unusual position compared to the other dashes which are clustered below the squares and left of center. So it could be that the underlying negative value is an odd case.

The nearest neighbor classification performed for this example depends on how many neighboring points are considered. If 1-NN is used and only the closest point is considered, then clearly the new point should be classified as negative since it is on top of a known negative point. On the other hand, if 9-NN classification is used and the closest 9 points are considered, then the effect of the surrounding 8 positive points may overbalance the close negative point.

A probabilistic neural network builds on this foundation and generalizes it to consider all of the other points. The distance is computed from the point being evaluated to each of the other points, and a *radial basis function* (RBF) (also called a *kernel function*) is applied to the distance to compute the weight (influence) for each point. The radial basis function is so named because the radius distance is the argument to the function.

Weight = RBF (*distance*)

The further some other point is from the new point, the less influence it has
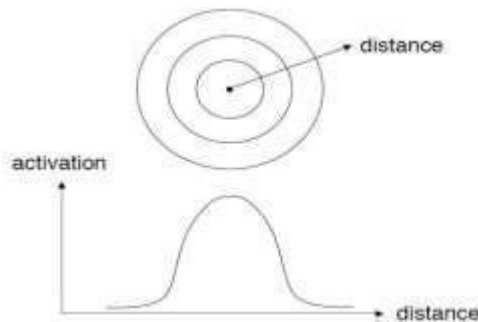

Figure 3: Radial Basis Function

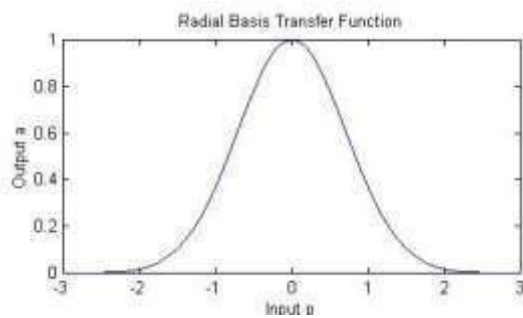Different types of radial basis functions could be used, but the most common is the Gaussian function:

Figure 4. radial basis transfer function

B. Fast Discrete Curvelet Transformation

Curvelet implementations are based on the original construction which uses a pre-processing step involving a special partitioning of phase-space followed by the ridgelet transform which is applied to blocks of data that are well localized in space and frequency.

In the last two or three years, however, curvelets have actually been redesigned in a effort to make them easier to use and understand. As a result, the new construction is considerably simpler and totally transparent. What is interesting here is that the new mathematical architecture suggests innovative algorithmic strategies, and provides the opportunity to improve upon earlier implementations. The two new fast discrete curvelet transforms (FDCTs) which are simpler, faster, and less redundant than existing proposals:

1. Curvelets via USFFT, and
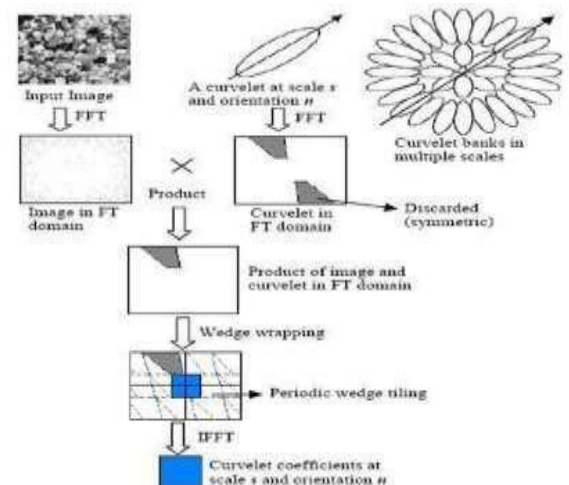2. Curvelets via Wrapping.

Fig. 1.   FDCT via wrapping.

Figure 5: FDCT via wrapping

The block size can be changed at each scale level. The wrapping construction is shown in Fig. 1. If $f[t_1, t_2], 0 \leq t_1, t_2 < n$ is taken to be a Cartesian array and ˆ$f[n1, n2]$ denotes its 2-D discrete Fourier transform, then the architecture of the FDCT via wrapping is as follows.

1.Apply the 2-D FFT and obtain Fourier samples,

$$\hat{f}[n_1, n_2], \qquad -\frac{n}{2} \le n_1, \qquad n_2 < \frac{n}{2}.$$

2.For each scale *j* and angle *l*, form the product

$$\tilde{U}_{j,l}[n_1, n_2]\hat{f}[n_1, n_2]$$

where $U_{j,l}[n1, n2]$ is the discrete localizing window

3.Wrap this product around the origin and obtain

$$\tilde{f}_{j,l}[n_1, n_2] = W(\tilde{U}_{j,l}\hat{f})[n_1, n_2]$$

Where the range for n1 is now $0 <= L1,j$ and $0 <= n2 < l2,j$; $l1,j \sim 2^{j}$ and $l2,j \sim 2^{j/2}$ are constants.Apply the inverse 2-D FFT to each ˜ *fj,l*, hence collecting the discrete coefficients *CD* (*j, l, k*).

IV. SPATIAL FUZZY CLUSTERING MODEL

Fuzzy clustering plays an important role in solving problems in the areas of pattern recognition and fuzzy model identification. A variety of fuzzy clustering methods have been proposed and most of them are based upon distance criteria. One widely used algorithm is the fuzzy c-means (FCM) algorithm. It uses reciprocal distance to compute fuzzy weights. A more efficient algorithm is the new FCFM. It computes the cluster center using Gaussian weights, uses large initial prototypes, and adds processes of eliminating, clustering and merging.

Spatial Fuzzy C Means method incorporates spatial information, and the membership weighting of each cluster is altered after the cluster distribution in the neighborhood is considered. The first pass is the same as that in standard FCM to calculate the membership function in the spectral domain. In the second pass, the membership information of each pixel is mapped to the spatial domain and the spatial function is computed from that. The FCM iteration proceeds with the new membership that is incorporated with the spatial function. The iteration is stopped when the maximum difference between cluster centers or membership functions at two successive iterations is less than a least threshold value.The fuzzy c-means (FCM) algorithm was introduced by J. C. Bezdek . The idea of FCM is using the weights that minimize the total weighted mean-square error:

$$J(w_{qk}, z^{(k)}) = \Sigma_{(k=1,K)} \Sigma_{(k=1,K)} (w_{qk})\| X_{(q)} - Z_{(k)}\|_2$$

$$\Sigma_{(k=1,K)} (w_{qk}) = 1$$

$$w_{qk} = (1/(D_{qk})^2)^{1/(p-1)} / \Sigma_{(k=1,K)} (1/(D_{qk})^2)^{1/(p-1)} , p > 1$$

The FCM allows each feature vector to belong to every cluster with a fuzzy truth value (between 0 and 1), which is computed using Equation (4). The algorithm assigns a feature vector to a cluster according to the maximum weight of the feature vector over all clusters.
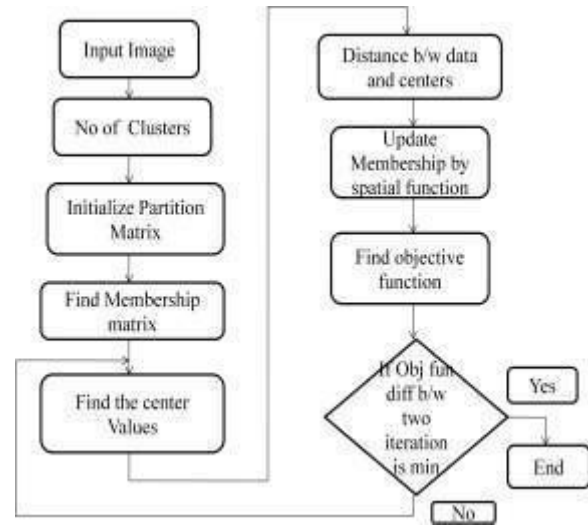
New Fuzzy c-means Implementation

Figure 6: Algorithm flow

1.Initialize the Fuzzy Weights. In order to comparing the FCM with FCFM, our implementation allows the user to choose initializing the weights using feature vectors or randomly. The process of initializing the

weights using feature vectors assigns the first K_{init} (user-given) feature vectors to prototypes then computes the weights by Equation (4).

2.Standardize the Weights over Q**.** During the FCM iteration, the computed cluster centers get closer and closer. To avoid the rapid convergence and always grouping into one cluster, we use

$$w[q,k] = (w[q,k] - w_{min})/(w_{max} - w_{min})$$
                before standardizing the

weights over Q. Where w_{max}, w_{min} are maximum or minimum weights over the weights of all feature vectors for the particular class prototype.

3.Eliminating Empty Clusters. After the fuzzy clustering loop we add a step (Step 8) to eliminate the empty clusters. This step is put outside the fuzzy clustering loop and before calculation of modified XB validity. Without the elimination, the minimum distance of prototype pair used in Equation (8) may be the distance of empty cluster pair. We call the method of eliminating small clusters by passing 0 to the process so it will only eliminate the empty clusters.After the fuzzy c-means iteration, for the purpose of comparison and to pick the optimal result, we add Step 9 to calculate the cluster centers and the modified Xie-Beni clustering validity κ :

The Xie-Beni validity is a product of compactness and separation measures. The compactness-to-separation ratio ν is defined by Equation (6).

$$\nu = \{(1/K)\Sigma_{(k=1,K)}\sigma_k^2\}/D_{min}^2$$

$$\sigma_k^2 = \Sigma_{(q=1,Q)}\ w_{qk}\ \|\ \mathbf{x}^{(q)} - \mathbf{c}^{(k)}\ \|^2$$

D_{min} is the minimum distance between the cluster centers.

The Modified Xie-Beni validity κ is defined as

$$\kappa = D_{min}^2 / \{\Sigma_{(k=1,K)}\sigma_k^2\}$$

The variance of each cluster is calculated by summing over only the members of each cluster rather than over all Q for each cluster, which contrasts with the original Xie-Beni validity measure.

$$\sigma_k^2 = \Sigma_{\{q:\ q\ is\ in\ cluster\ k\}}\ w_{qk}\ \|\ \mathbf{x}^{(q)} - \mathbf{c}^{(k)}\ \|^2$$

The spatial function is included into membership function as given in Equation

$$u_{ij}^{'} = \frac{u_{ij}^{p}\ h_{ij}^{q}}{\Sigma_{k=1}^{c}\ u_{kj}^{p}\ h_{kj}^{q}}$$

Morphological Process

Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. Morphological operations rely only on the relative ordering of pixel values, not on their numerical values, and therefore

are especially suited to the processing of binary images.Morphological techniques probe an image with a small shape or template called a **structuring element**. The structuring element is positioned at all possible locations in the image and it is compared with the corresponding neighbourhood of pixels. Some operations test whether the element "fits" within the neighbourhood, while others test whether it "hits" or intersects the neighbourhood:
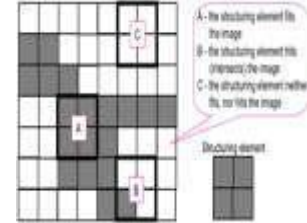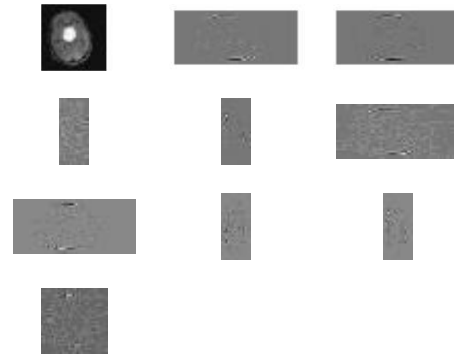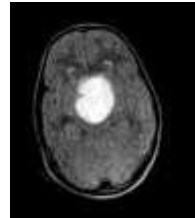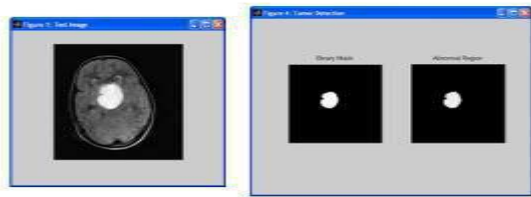

Fig 7: Probing of an image with a structuring element

V. SIMULATED RESULTS

Segmentation: Snapshot- Benign

[8] Mammadagha Mammadov, Engintas ,"An improved version of back propagation algorithm with effective dynamic learning rate and momentum " ,Inter.Conference on Applied Mathematics ,pp:356-361, 2006.
[9] Messen W, Wehrens R, Buydens L, "Supervised Kohonen networks for classification problems",Chemometrics and Intelligent Laboratory Systems, vol.83,pp:99-113,2006.
[10] Orr M.J.L., Hallam J., Murray A., and Leonard .T, "Assessing rbf networks using delve," International Journal of Neural Systems, vol. 10, issue 5, pp. 397-415, 2000

VI. CONCLUSION

The project presented that automated brain image classification for early stage abnormality detection with use of neural network classifier and spotting of tumor was done with image segmentation. Pattern recognition was performed using probabilistic neural network with radial basis function and pattern will be characterized with the help of fast discrete curvelet transform and haralick features analysis.Here. Spatial fuzzy clustering algorithm was utilized effectively for accurate tumor detection to measure the area of abnormal region. From an experiment, system proved that it provides better classification accuracy with various stages of test samples and it consumed less time for process.

REFERENCES

[1].N. Kwak, and C. H. Choi, "Input Feature Selection for Classification Problems", IEEE Transactions on Neural Networks, 13(1), 143–159, 2002.
[2].E. D. Ubeyli and I. Guler, "Feature Extraction from Doppler Ultrasound Signals for Automated Diagnostic Systems", Computers in Biology and Medicine, 35(9), 735–764, 2005.
[3].D.F. Specht, "Probabilistic Neural Networks for Classification, mapping, or associative memory", Proceedings of IEEE International Conference on Neural Networks, Vol.1, IEEE Press, New York, pp. 525-532, June 1988.
[4].D.F. Specht, "Probabilistic Neural Networks" Neural Networks, vol. 3, No.1, pp. 109-118, 1990.
[5].Georgiadis. Et all , " Improving brain tumor characterization on MRI by probabilistic neural networks and non-linear transformation of textural features",Computer Methods and program in biomedicine, vol 89, pp24-32, 2008
[6] Kaus M., "Automated segmentation of MRI brain tumors", Journal of Radiology,vol. 218, pp. 585-591,2001
[7] Kornel, P., Bela, M., Rainer, S., Zalan, D., Zsolt, T. and Janos, F.,"Application of neural network in medicine", Diag. Med. Tech.,vol. 4,issue 3,pp: 538-54 ,1998.

AUTHOR'S PROFILE

M.A.DARSHANA received her **B.E.** degree in Electronics and Communication Engineering from Prince.Dr.K.Vasudevan college of Engineering,Anna University Chennai, India, in 2014. Pursuing **M.E**. degree in VLSI Design from Sethu Institute of Technology, Anna University, Chennai,
India. Her research interest includes digital electronics and VLSI design.

M.LAKSHMISHREE received her **B.E.** degree in Electronics and Communication Engineering from Pandian Saraswathi Yadav engeneering college, AnnaUniversity Chennai, India in 2013. Pursuing **M.E**. degree in VLSI Design from Sethu Institute of
Technology, Anna University, Chennai, India. Her research interest includes digital electronics and VLSI design.

Mrs.C.SUJATHAreceived **B.E** degree in Electronics and Communication Engineering from P.S.N.A College of Engineering and Technology, Dindigul, Tamil Nadu,India in 1997 and **M.E** degree in Applied Electronics from P.S.N.A College of Engineering and Technology, ,Dindigul, Tamil Nadu, India in
2004. She has more than 13 years of teaching experience. Presently she is working as Associate Professor, Department of Electronics and Communication Engineering, Sethu Institute of Technology, Kariapatti, Tamil Nadu, India. She is doing research in the area of Image Processing based on FPGA under the guidance of Dr.D.Selvathi, Prof. /ECE Dept., Mepco Schlenk Engineering College, Sivakasi,. Tamilnadu, India.