SURF serves better results for Brain Tumor Classification

Nitish Zulpe[1] and Vaijinath V. Bhosle[2]
1 Computer Science,
[1]College of Computer Science and Information Technology, Latur-413512
Maharashtra, India
nitish.Zulpe@gmail.com
2 Computer Science,
College of Computer Science and Information Technology (COCSIT),
Latur, Maharashtra, India,
vvbhosle80@gmail.com

Abstract
Automatic recognition system for medical images is challenging task in the field of medical image processing. Medical images acquired from different modalities such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), etc which are used for the diagnosis purpose. In the medical field, brain tumor classification is very important phase for the further treatment. Human interpretation of large number of MRI slices (Normal or Abnormal) may leads to misclassification hence there is need of such a automated recognition system, which can classify the type of the brain tumor. In this research work, we used four different classes of brain tumors and extract feature using SURF. this algorithm can detect and define local features for any interest object and extract features or descriptor points from it and compare these features/ descriptor by the features that extracted from origin image, matching process has been done among features and decision made based on similar features found, this algorithm called Speeded Up Robust Features (SURF) algorithm. Speeded up Robust Features from each class, and applied to two-layered Feed forward Neural Network, which gives 97.5% classification rate.

Keywords: *MRI, CT, GLCM, Neural Network*

I. INTRODUCTION

Abnormal growth of cell in the brain causes the brain tumor and may affect any person almost of any age. Brain tumors can have a variety of shapes and sizes; it can appear at any location and in different image intensities [1]. Brain tumor classification is very significant phase in the medical field. The images acquired from different modalities such as CT, MR that should be verified by the physician for the further treatment, but the manual classification of the MR images is the challenging and time consuming task [2]. Human observations may lead to misclassification and hence there is need of automatic or semiautomatic classification techniques to make the difference between different tumor types.

We found many classification techniques have been given for the determining the tumor type from the given MR images such as, Matthew C. Clarke et al. [3] developed a method for abnormal MRI volume identification with slice segmentation using Fuzzy C-means (FCM) algorithm. Chang et al. [4,5] reported the SVM is an best tool in sonography for the diagnosis of

Breast cancer. W. Chu *et. al.* [6] proposed that LS-SVM generally are able to deliver higher classification accuracy than the other existing data classification algorithms.

In medical image analysis, the determination of tissue type (normal or abnormal) and classification of tissue pathology are performed by using texture. MR image texture proved to be useful to determine the tumor type [7]. H. Bay. [8] Suggested a Speeded up Robust Features "SURF" algorithm is a local feature and descriptor algorithm that can be used in many application such as object recognition

The process of Speeded UP Robust Features "SURF" algorithm can be divided into three main steps. First step is "Detection step", in this step interest points are selected at distinctive locations in the origin image, such as corners, blobs and T-junctions and this process must be robustly.

Second step is "Description step", in this step interest points should have unique identifiers does not depend on features scale and rotations which are called descriptor, the information of interest points represented by descriptor which are vectors that contain information about the points itself and the surroundings. Third step is "Matching step", in this step descriptor vectors are compared between the object image and the new input or origin image, the matching score is calculated based on the distance between vectors e.g

SURF algorithm was first presented by [1] in 2006, Use an integer approximation of determinant of Hessian blob detector, which can be computed with three integer operations using an integral image. SURF features descriptor are calculated by the sum of Haar wavelet response around interest points. And these can be computed by the concept of integral image.

II. METHODS AND MATERIALS

A. Magnetic Resonance Imaging

A magnetic resonance imaging instrument (MRI Scanner) uses powerful magnets to polarize and excite hydrogen nuclei i.e. proton in water molecules in human tissue, producing a detectable signal which is spatially encoded, resulting in images of the body. MRI uses three electromagnetic fields

1) A very strong static magnetic field to polarize the hydrogen nuclei, called the static field.
2) A weaker time varying field(s) for spatial encoding called the gradient field.
3) A weak radio frequency field for manipulation of hydrogen nuclei to produce measurable signals Collected through RF antenna.

Class I (Astrocytoma)

The patient was a 35-year-old man; MR demonstrates an area of mixed signal intensity on proton density (PD) and T2-weighted (T2) images in a left occipital region. Contrast enhancement shows the lesion to contain cystic elements. Thallium images show an anterior border of high uptake, consistent with a small region of tumor recurrence.

Class II (Meningioma)

The patient was a 75-year-old man who had an 8 - 10 month history of progressive difficulty walking. He had noted some left lower extremity weakness and some difficulty with memory and concentration. He was alert and oriented, but had slow and hesitating speech. He could recall only 1 of 3 objects at five minutes.

Class III (Metastatic bronchogenic carcinoma)

This 42 year old woman with a long history of tobacco use began having headaches one month before these images were obtained. Brain images show a large mass with surrounding edema, and compression of adjacent midbrain structures. The MR demonstrates the tumor as an area of high signal intensity on proton density (PD) and T2-weighted (T2) images in a large left temporal region.

Class IV (Sarcoma)

The patient was a 22 year old man who was admitted for resection of Ewing's sarcoma (peripheral/primitive neuroepithelial tumor- PNET). Vaguely described visual difficulty was noted retrospectively to have begun approximately one month prior to admission.

B. Surf Algorithm

Speeded UP robust Features "SURF" algorithm is considered a robust local feature detector and extractor algorithm and can be used in many computer vision application like object recognition, 3D reconstruction and its one the best approaches suitable for real-time application [1]. The interest point detection which is represented by a vector call descriptor in SURF algorithm is based on scale space theory , SURF algorithm use an Integer approximations as the determinant of Hessian blob detector which can be computed fast with an integral image [1]. The integral image is an image where each point in this image X which equal to (x, y) T Stores the sum of all pixels in the input image (I) within a rectangular region which is formed by the origin and X, see Eq. 1:

$$I(X)\sum_{i=0}^{i\leq x}\sum_{j=0}^{j\leq y} I(i,j) \qquad (1)$$

The integral images are used in Hessian matrix approximation which reduce the time of computation effectively. Since Hessian matrix has good performance and also has good accuracy, in image I, X=(x, y) is a given point in an origin image, the Hessian matrix H(X, σ) in X at scale σ is defined in Eq. 2:

$$H(X,\sigma)=\begin{bmatrix} L_{xx}(x,\sigma)Lxy(x,\sigma) \\ L_{yx}(x,\sigma)Lyy(x,\sigma) \end{bmatrix} \qquad (2)$$

Where Lxx(x, σ) is the convolution result of the second order derivation of Gaussian Filter [(∂^2)/ (∂x^2)] g (σ) with the image in point X, and similarity for Lxy(x, σ) and Lyy(x, σ). To reduce the computation time, set of 9×9 box filter is used for approximations for Gaussian second order derivatives with σ =1.2 and this value represent the lowest scale (i.e. highest spatial resolution) for computing the blob response maps. We will denote them by Dxx, Dyy and Dxy. The weights applied to the rectangular regions are kept simple for computational efficiency, see Eq. 3:

$$\det(H_{approx}) = D_{xx}D_{yy} - (\omega D_{xy})^2 \qquad (3)$$

Where ω is the relative weight of filter response and given by this formula for box filter 9×9 and σ=1.2 [3], for ω see Eq. 4:

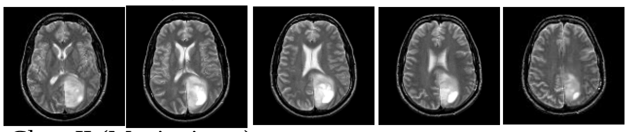
$$\omega = \frac{|L_{xy}(1.2)|D_{yy}(9)|F}{|L_{yy}(1.2)|F|D_{xy}(9)|F} \qquad (4)$$

The relative weight of filter ω is used in this formula to keep the balance for the Hessian determination process and the weight can be changes depending on the scale by default this weight did not have a significant impact on the result [9] because we are using a metric for counting the percentage of inliers points founded as well as outliers. SURF applies different sizes from box filters to search and compare interest points,

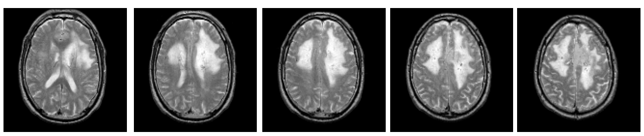
C Dataset

Four different classes of Brain tumor MR images are used for the experimental work in which each class contains 20 samples; total 80 samples are collected from the Whole Brain Atlas (WBA). Every image is having the exact size of 256x256 in axial view.

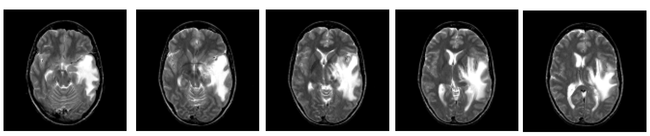
Class I (Astrocytoma)



Class II (Meningioma)



Class III (Metastatic bronchogenic carcinoma)



Class IV (Sarcoma)

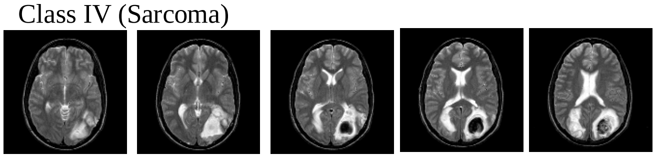


Fig. 1 Five sample MR images of four classes

D. Preprocessing

Medical image analysis requires the preprocessing because the noise may be added to the MR images due to imaging devices. We used the Gaussian filter to improve the quality of the Image by Noise suppression, contract enhancement, intensity equalization, outlier elimination. The Gaussian distribution in 1-D has the form:

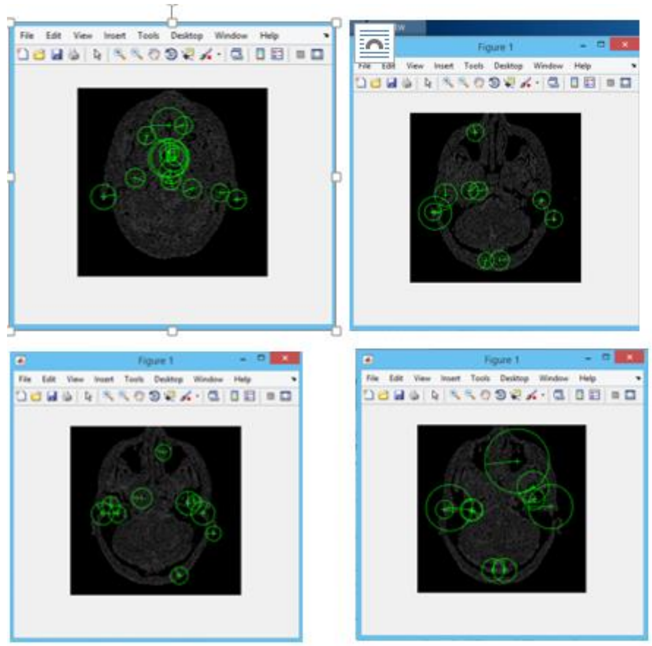
$$G(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} \qquad (5)$$

Where σ is the standard deviation of the distribution. We have also assumed that the distribution has a mean of zero. In 2-D, an isotropic (*i.e.* circularly symmetric) Gaussian has the form:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (6)$$

E. Features extraction

Following Image show features extracted by surf technique.



III. CLASSIFICATION

For the classification, purpose we used the two layers feed forward neural network in which learning assumes the availability of a labeled (i.e. ground-truthed) set of training data made up of N input and output

$$T = \{(X_i, d_i)\}_{i=1}^{N} \qquad (7)$$

Where

X_i is input vector for the i_{th} example

d_i is the desired output for the i_{th} example

N is the sample size.

A two layer feed forward network with sigmoid activation function is designed with 75 input neurons, 10 hidden neurons and 4 output neurons for the classification.

Network architecture

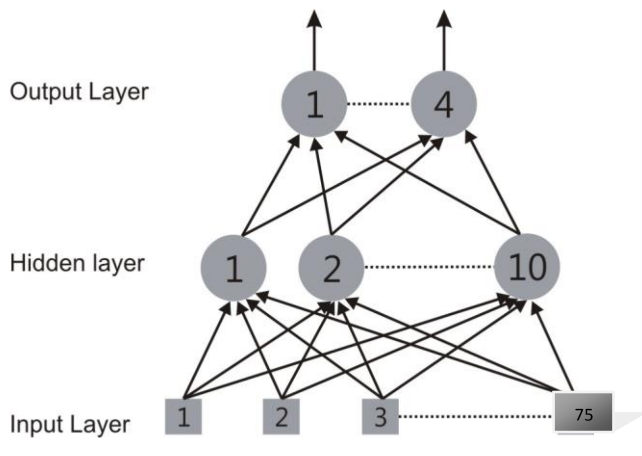


Fig. 2 Network Architecture of Two-Layer FF Network

The Levenberg-Marquardt algorithm [10] is used for the training the neural network which is a very simple, but robust, method for approximating a function. Basically, it consists in solving the equation:

$$(J^t J + \lambda I)\delta = J^t E \tag{8}$$

Where J is the Jacobian matrix for the system, λ is the Levenberg's damping factor, **δ** is the weight update vector that we want to find and E is the error vector containing the output errors for each input vector used on training the network. The δ tell us by how much we should change our network weights to achieve a (possibly) better solution[11].

The J^tJ matrix can also be known as the approximated Hessian. The **λ** damping factor is adjusted at each iteration, and guides the optimization process. If reduction of E is rapid, a smaller value can be used, bringing the algorithm closer to the Gauss–Newton algorithm, whereas if iteration gives insufficient reduction in the residual, **λ** can be increased, giving a step closer to the gradient descent direction. For the training purpose we used the 56 samples, 16 samples for validation and 8 samples for the testing. The training stops when a classifier gives a higher accuracy value with minimum training and testing errors[12].

IV. EXPERIMENTAL RESULT

Near about 75 SURF features of each slice is calculated. 80 samples of four classes (Class I, Class II, Class III, and Class IV) which form the input vector of size 80x75. Target output is designed with the size of 80x4. The LM training algorithm outperformed in this experiment by classifying the input data in 35 epochs with the average training time of 12 seconds. The performance measured and outcome of the network are as follows:

Table 2: Performance Measure of LM training algorithm

Classes	No of classified Images	No of Miss-classified Images	Accuracy (%)
Class-I	20	Nil	100%
Class-II	20	Nil	100%
Class-III	19	01	95%
Class-IV	19	01	95%

Table 3: Performance Measure of LM training algorithm

Number of epochs	35
Performance	0.0124
Training Performance	0.0306
Validation Performance	0.0157
Testing Performance	0.0424
classification Rate	97.5%

The error measures like Mean Squared Error (MSE) and Percent Error (PE) are recorded. MSE is the mean of the squared error between the desired output and the actual output of the neural network.

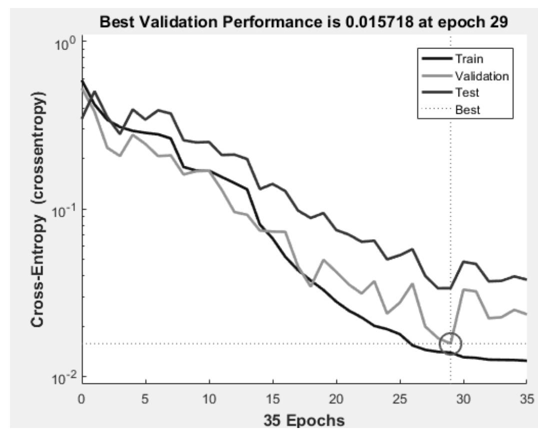


Fig. 3 Performance Graph of the classifier

MSE is calculated using following equation

$$MSE = \frac{\sum_{j=0}^{P}\sum_{i=0}^{N}(d_{ij} - y_{ij})^2}{NP} \tag{9}$$

Where, P = number of output processing elements

N = number of exemplars in the training data set

y_{ij} = estimated network emissions output for pattern i at processing element j

d_{ij} = actual output for emissions exemplar *i* at processing element *j*.

Percent Error indicates the fraction of samples, which are misclassified. Value 0 means no misclassifications.

$$\%error = \frac{100}{NP}\sum_{j=0}^{P}\sum_{i=0}^{N}\frac{|\,dy_{ij} - dd_{ij}\,|}{dd_{ij}} \tag{10}$$

Where
P = number of output processing elements
N = number of patterns in the training data set
dy_{ij} = demoralized network emissions output for pattern *i* at processing element *j*
dd_{ij} = demoralized desired network emissions output for exemplar *i* at processing element *j*.

A. Confusion Matrix

The confusion matrix gives the accuracy of the classification problem. 80 MR images of 4 different classes are classified with 97.5% accuracy with Levenberg-Marquardt algorithm. The diagonal elements of the confusion matrix shows classified groups. The following confusion matrix illustrates that, Class I and Class II are exactly recognized in which all images fall in the same class. But for the Class III and Class IV only one image fall outside the class therefore average error rate is 2.5%.

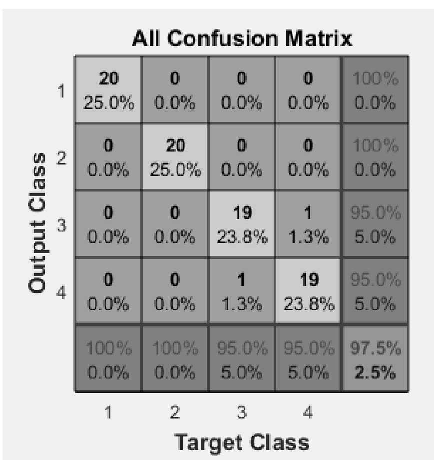
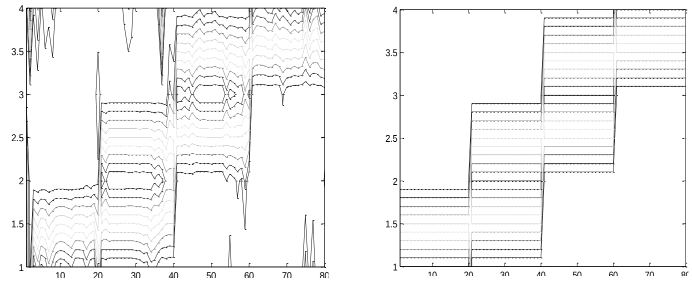


Fig. 4 Confusion Matrix showing the recognition rate

B. Contour Plot

A contour plot is a graphical technique for representing a 3-dimensional surface by plotting constant *z* slices, called contours, on a 2-dimensional format. That is, given a value for *z*, lines are drawn for connecting the (x, y) coordinates where that *z* value occurs. Target contour plot represents desired output vector and output contour plot represents actual output. Following contour plot shows the target and output contours.



(a) Output contour (b) Target contour

Fig. 5 Contour plot of Target and actual Output

V. CONCLUSION

In this research paper, we try to classify the four different classes of tumor types such as Astrocytoma, Meningioma, Metastatic bronchogenic carcinoma, and Sarcoma. All the MRI slices collected from the WBA and after preprocessing the SURF features used to train the feed forward neural network with Levenberg Marquart (LM) nonlinear optimization algorithm which gives the better recognition rate of 97.5%. May this work will assist the physician to make the final decision for the further treatment.

REFERENCES

[1] Ricci P.E., Dungan D. H., "Imaging of low- and intermediate-grade gliomas", Seminars in Radiation Oncology, 2001, 11(2), p. 103-112.
[2] Kaus *et.al.*, " Automated segmentation of MR images of brain tumors", Journal of radiology, vol.218, no.2, pp: 586-591, 2001.
[3] M. C. Clark, L. O. Hall, D. B. Goldgof, L.P. Clarke, R. P. Velthuizen, and M. S.Silbiger, "MRI Segmentation using Fuzzy Clustering Techniques", IEEE Engineering in Medicine and Biology, pp. 730-742, 1994.
[4] Chang R.F., Wu W.J., Moon W.K., Chou Y.H., Chen D.R., "Support vector machines for diagnosis of breast tumors on US images", Academic Radiology, 2003, 10(2), p. 189-197.
[5] Chang R.F., Wu W.J., Moon W.K., Chou Y.H., Chen D.R., "Improvement in breast tumor discrimination by support vector machines and speckle-emphasis texture analysis", Ultrasound in Medicine and Biology, 2003, 29(5), p. 679-686.

[6] W. Chu, C. J. Ong, and S. S. Keerthi "An improved conjugate gradient scheme to the solution of least squares SVM", IEEE Transactions on Neural Networks 16(2): pp. 498-501, 2005.

[7] Schad L.R., Bluml S., "MR tissue characterization of intracranial tumors by means of texture analysis", Magnetic Resonance Imaging, 1993, 11(6), p. 889-896.

[8] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346-359, June 2008.

[9] Jin Zhao , Sichao Zhu , Xinming Huang , " Real-Time traffic sign detection using SURF features on FPGA " , HPEC , 10.1109 , IEEE pp.2-3, Sep 2013.

[10] Martin T. Hagan and Mohammad B. Menhaj, "Training Feedforward Networks with the Marquardt Algorithm", IEEE Trans. on Neural Networks, Vol. *5*, No. *6*, November 1994.

[11] N S Zulpe and V P Pawar, "GLCM textural features for Brain Tumor Classification", International Journal of Computer Science Issues (IJCSI) 9 (3), 354

[12] Vaijinath V Bhosle, Vrushsen P. Pawar, "Automatic Logo Extraction and Detection for Document Verification using SIFT and SURF", International Journal of Engineering Research & Technology, 555-560

Dr. N. S. Zulpe, pursed Master of Science (Computer Science) from Swami Ramanand Teerth Marathwada University, Nanded in 2004. He has also completed Ph.D. from Savitribai Phule Pune University and currently working as In-charge Principal at College of Computer Science and Information Technology (COCSIT). He is a member of IEEE & Member of Computer Society of India (CSI). He has published more than 18 research papers in reputed National and International Journals. His main research work focuses on Medical Image Processing, Pattern Recognition and Neural Network. He has 14 years of teaching experience and 10 years of Research Experience.

Dr. Vaijinath V Bhosle received M.Sc.(Computer Management) degree from SRTM University, Nanded in the year 2008. He has also completed Ph.D. from SRTMU, Nanded and currently working as Assistant Professor in the College of Computer Science and Information Technology, Latur, Maharashtra.