INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 07 ISSUE: 11 | NOVEMBER - 2023 SJIF RATING: 8.176 ISSN: 2582-3930

SVM for Classification of Brain MRI with K-Means for Detection

D Kalpana
Assistant Professor
Dept.of CSE(DS)
Malla Reddy Engineering College
Hyderabad, India
kasamkalpana@gmail.com

K Sudhakar Reddy
Associate Professor
Dept.of Information Technology
Malla Reddy College of Engineering and Technology
Hyderabad, India
kasamsudha86@gmail.com

Abstract— The segmentation of MR brain images using this method is based on K-means clustering, feature extraction using a discrete wavelet transform (DWT), and feature selection using a grey level co-occurrence matrix (GLCM). For this procedure, we have used a Perfect Radial Basis Function (RBF) - Support Vector Machine (SVM) Classifier. Based on fractions of selectivity and sensitivity, the classifier's performance was measured in terms of accuracy. The proposed classifier was found to be 93% accurate. Additionally, the Histogram methodology was applied in this proposed method in place of randomly choosing the cluster centres.

Keywords**: K-means clustering, Histon creation, Discrete Wavelet Transform (DWT), GLCM feature selection, and RBF-SVM Classifier.**

INTRODUCTION:

The support vector machine for categorising brain magnetic resonance imaging(MR) pictures. The suggested method divides MR brain pictures into two categories: normal and pathological. Using a dataset of 52 MR brain imaging scans, we evaluated the suggested method. The distance between a data point and its nearest centre can be decreased using clustering approaches with K-Means, according to Tapas Kanungo et al. (2002) analysis and implementation [1]. Unsupervised segmentation of magnetic resonance imaging (MRI) brain pictures was the subject of a 2003 study by Jing-Hao Xue et al. The image was segmented using that method with the least amount of error possible [2] after being de-noised using the wavelet filter. A concise explanation of the combined c-means clustering for the adjustment of intensity inhomogeneity was provided by Laszlo Szilagyi et al. (2011)[3]. A flawless picture segmentation using pulse coupled neural networks was presented by Kuntimad et al. in 1999 [4]. A noval technique was used by Jussi Tohka et al. (2010) to classify the tissues in magnetic resonance imaging (MRI) brain pictures [5]. With K-Means clustering, Li-Hong Juang and Ming-Ni Wu (2010) presented a method for tracking tumour items in MR brain images [6].

It was noted by Dongxiang Chi (2011) that both the SOM- K and SOM- KS segmentation algorithms can produce superior segmentation results with less computing time [7]. K-means clustering was chosen for the suggested method due to its simplicity in terms of computing.

Cancer has become more common as the population is growing quickly.significant public health concern on a worldwide scale. When determining a patient's diagnosis, imaging plays a crucial role. a brain tumour. When cells divide uncontrollably, they form tumours, which are unwanted collections of cells (tissue). The cell type that a brain tumour originates from determines its name. Secondary and primary brain tumours are distinguished. Exactly like the cells that make up the organ or tissue where they originate, primary tumours are made up of these same cells. Brain cells are where a first brain tumour develops.

Malignant tumours have a quick rate of growth and can spread to include a lot of the surrounding tissues. Cells from one area of the body that have spread to another or several other parts of the body generate secondary tumours.Using the visual clarity of MRI scans, radiologists examine them to determine the presence of tumours. Because the sensitivity of the human eye declines with an increase in instances, most commonly when only a few slices are impacted, there may be a chance that radiologists will make a mistaken diagnosis while analysing a huge volume of MRI data. Because of this, effective automated systems are required for using medical image analysis and categorization. In order to prevent brain damage or perhaps unprotective death, early detection and appropriate treatment of brain tumours are crucial. The two most crucial pieces of information for an accurate and successful treatment are the tumor's location and size. The brain can be imaged using magnetic resonance spectroscopy (MRI).

Tumours in the brain are observed and divided into sections. For medical diagnosis, particularly in brain imaging, magnetic resonance imaging (MRI) is widely employed in hospitals and clinics today. Given that MRI is noninvasive and offers the benefit of highlighting soft tissues. Ionising radiation is not employed during MRI. Because it is a painless, non-aggressive, and non-radioactive approach, MRI is employed in brain imaging.

LITERATURE SURVEY:

The technique for classifying brain tumours is presented by H. B. Nandpuru et al. Using SVM, a supervised machine learning technique, the normal and malignant brains on MRI are separated. The texture, symmetry, and grey features were first extracted. The suggested classifier provides 84% accuracy.

Using an image mining technique, Janki Naik et. proposed classifying brain tumours seen in brain MRI scans. The texture feature extraction technique has been used to extract features from the median-filtered, preprocessed MRI images. Text categorization using decision trees and interclass relationships has more efficiency than using classic mining techniques. Support vector machine (SVM) classifier was utilised by the system, and it provides 83% accuracy.

Ahmed Kharrat et al. have suggested classifying brain tissues in MRI scans using a hybrid technique combining GA and SVM. The spatial grey level dependence technique, or SGLDM, is used to extract the features. The proposed approach provides good accuracy of about 85.22%.The image was subjected to morphological operations suggested by Ali Reza Fallahi et al. [4] before the characteristics were retrieved.

MLPNN and SVM analysis results demonstrate that these methods can enhance classification results in symmetry and grey scale characteristics but decrease results in texture features.SVM allows the system to perform better than MLPNN and RBFNN. Because of the symmetry of the brain, symmetrical features are more accurate than textural features.

Mehadi Jafari and colleagues [5] presented a hybrid strategy combining c a genetic algorithm and SVM. The classifier was given three extracted feature sets as input. During the first

such as entropy, energy, mean, standard deviation, kurtosis, skewness, momentum, and correlation. The second set took into account wavelet-based features, and the third set retrieved information based on frequency transformation. SVM categorises the brain MRI according to normal or abnormal in this procedure, while genetic algorithms serve as feature reduction techniques. The accuracy of the classifier is increased by this method by up to 83.22%.

El-Dahshan et al. [6] present another another hybrid strategy. The brain MRI serves as an input to the system; features were retrieved using the discrete wavelet transform, reduced using the principal component analysis method, and classified using the feed-forward back propagation artificial neural network (FF-BPNN) and KNN. On both training and test datasets, these classifiers deliver 99% accuracy.

For the classification of brain MRIs, Hong Men et al. [7] introduce the neural network and SVM machine learning algorithms. For various parametric values, they used two different types of support vector machines based on polynomial kernel and radial basis function. The outcome of this experiment shows that the support vector machine method is superior to the neural network algorithm.

PROPOSE METHODOLOGY:

In the suggested method, three clusters are created to extract the Region Of Interest (ROI) using a histogram-based centre initialization approach after the input MR brain image has been pre-processed with a median filter. After using the three-level DWT technique to extract features in terms of their eigen values and eigen vectors, the dimensionality of the features is reduced using the Grey Level Co-Occurrence Matrix (GLCM) method, and the salient features are then chosen. To categorise the input as normal or abnormal, these chosen features are provided as inputs to the Supervised Radial Basis Function - SVM classifier. The steps that will be taken for this job are shown in the flowchart below (Fig. 1).

CLUSTERING TECHNIQUES:

A.Initialization of Centroids:

The initialization of centroids is a crucial step in the histogram, a potent tool for segmenting and enhancing images. It is a visual depiction of the pixel intensity values in a picture. An 8-bit grayscale image may have 256 different intensity levels, and the histogram of that image provides information on how those 256 pixels are distributed in terms of intensity. By reducing the intensity difference between a pixel and its neighbours, the median filtering principle is applied in the suggested way to reduce noise. Thamaraichelvi B et al.'s (2015) [8] description of the Histon's construction is given below Fig. 1.

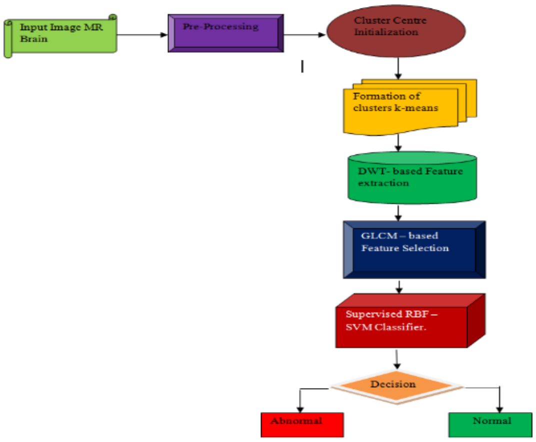


Fig. 1. Schematic diagram for the proposed method

B. K-means cluster

The least complex method in terms of computation is K-Means clustering. When the number of clusters is sufficiently known, an unsupervised approach is utilised to segment biomedical images. [9] Radha et al. (2011). It is also known by the name partitioning algorithm.

C.Clustering algorithm using K-Means

1. Initialising cluster centres is necessary.

2. A 'k' number of clusters are formed by grouping the data points, and the cluster assignments are also random.

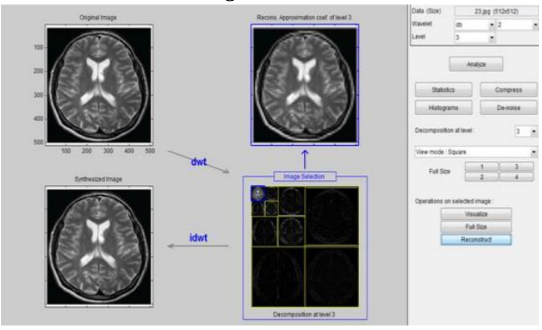


Fig.2. ACommon Method to find 3 Level DWT Decomposition

3. The Euclidean distance between each cluster's mean and each data point was determined.

4. The nearest cluster receives the data points.

5. The procedure was carried out until each data point was assigned to the closest cluster.

EXTRACTION AND SELECTION OF FEATURES

According to S. Chaplot et al. (2006), a Fourier transform represents an image by its frequency contents while a Wavelet transform separates the image into many spatially localised frequency levels. For this project, the Doubechesis-2 Wavelet at level 3 is chosen, and the images are 512x512 in size and each have a 1mm thickness. Figure 2 depicts the original image's three level breakdown.

The sub-band LL is used for the following level of the Wavelet decomposition method. With the use of the db-2 Wavelet tool, 4096 coefficients were ultimately retrieved from the LL sub-band.

The features are then chosen using the GLCM method, a statistical method used to choose the second order textural characteristics. The purpose of this method, as mentioned, is to determine the spatial relationship between a pixel and its immediate neighbours, who are separated by a distance d in the direction of [11].

The features that can be extracted using GLCM, according to Haralick et al. (1973), include entropy, correlation, energy, contrast, Mean, Standard -deviation, Variance, etc. The classifier is then given the chosen features [12].

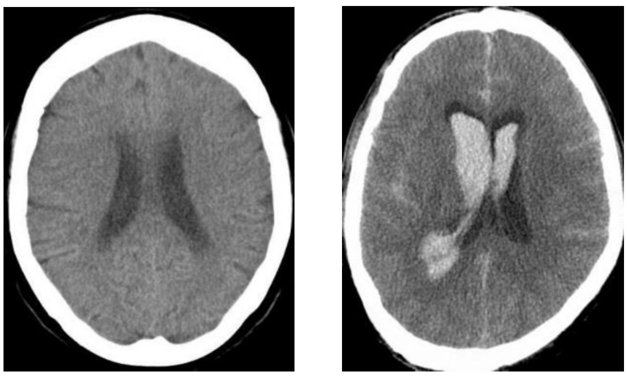


Figure . Normal and abnormal brain CT image

CLASSIFICATION:

A.SVM

Two machine learning algorithms—SVM and KNN—are proposed to be used in this work to classify brain MRI data.

The best way to identify the hyperplane between two distinct classes in a high-dimensional feature space for classification purposes is with a support vector machine. A supervised machine learning algorithm is the support vector machine. Training and testing are the first two steps in the supervised learning approach.

In the training phase, two types of databases were taken into consideration: 251 (85 malignant and 166 benign) MRI clinical database images, 80 (50 low grade glioma and 30 high grade glioma) standard MRI images, and 100 (50 malignant and 50 benign) clinical database images, as well as 40 (25 low grade glioma and 15 high grade glioma) MRI images for testing, respectively.

B.KNN

The classification method KNN is straightforward and reliable. The k nearest training neighbour vector is used in this classifier to classify the testing feature vector. Different distance measuring methods, such as the Euclidean distance, cityblock, chebychev, Minkowski, Mahalanobis, cosine, correlation, Spearman, hamming, and Jaccord, are used to determine the distance between the training and testing vectors. The distances between the training and test data vectors are calculated using the Euclidean, cityblock, cosine, and correlation measures.

As a result of the feature extraction from the training and test sets of images, we obtain various dimensions in a space and the values of the extracted features are taken as observations for the characteristic to be its coordinate in that dimension, resulting in a collection of points in the space. Then, using a proper metric, we may relate the separation between two distinct places in space.

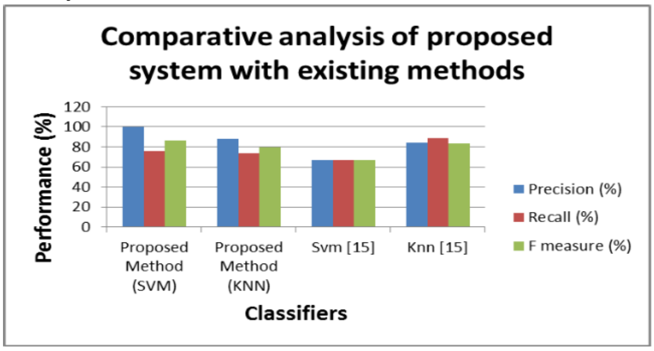
The method by which the applied algorithm determines which of the points from the training set are sufficiently similar to the point examined when choosing the class to estimate for a new observation is to choose the k closest data points to the new observation, and to take the most common class among them. In this manner, the k Nearest Neighbour algorithm

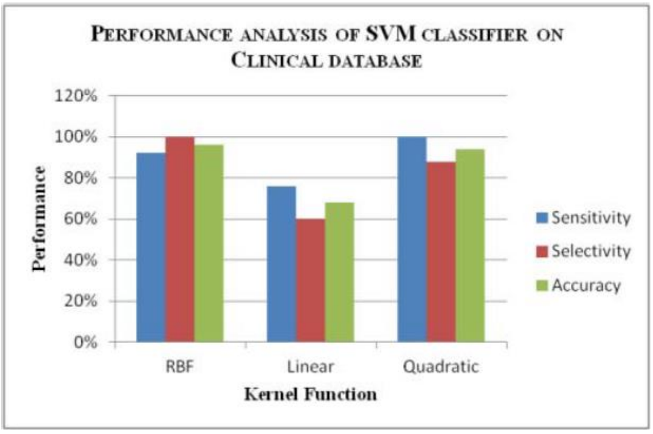
Following is a description of the K-nearest neighbour algorithm:

1. The new sample and an integer with the positive sign k are defined.
2. Decide which k values in our database are most similar to the fresh testing sample.
3. We determine which of these entries is most closely related.
4. Using the value of k, we categorise the new sample as shown in figure 4.
5. Change the value of k till the desired results are not attained if the satisfying results were not obtained.

RESULT AND DISCUSSION:

The chosen picture has dimensions of "512x512" and a thickness of "1" mm per side. Utilising a median filter, the photos are preprocessed. To gather information about the surrounding pixels, a kernel of preference of size "3x3" has been applied to the entire image. For the different types of brain tissues—GM, WM, and CSF—three intensity levels were developed. The terms TP, TN, FP, and FN were used to evaluate the segmentation and classification's accuracy. The results of using k-means clustering are shown in the next image. Fig. 3(i) depicts the actual input photographs, Fig. 3(ii) displays the results of the median filtering procedure, and Fig. 3(iii) displays the final results. depicts the k-means clustering result for typical brain photos in a clear manner. In Table 1, the proposed clustering technique is quantitatively evaluated. Table 2 displays the outcomes of a feature selection method using glcm for typical brain pictures. Fig. 4(i) displays the actual input images, Fig. 4(ii) displays the results of the median filtering procedure, and Fig. 4(iii) amply illustrates the results of the preferred clustering for aberrant images. The chosen clustering method for aberrant MR images is evaluated in Table 3. The outcome of the feature selection approach for aberrant brain imaging using a glcm is shown in Table 4. We examined a total of 50 aberrant cases and 40 normal cases. The Table 5 analysis of the recommended RBF-SVM classifier's results revealed that it has an accuracy of about 93% and can accurately identify 47 abnormal and 36 normal cases.





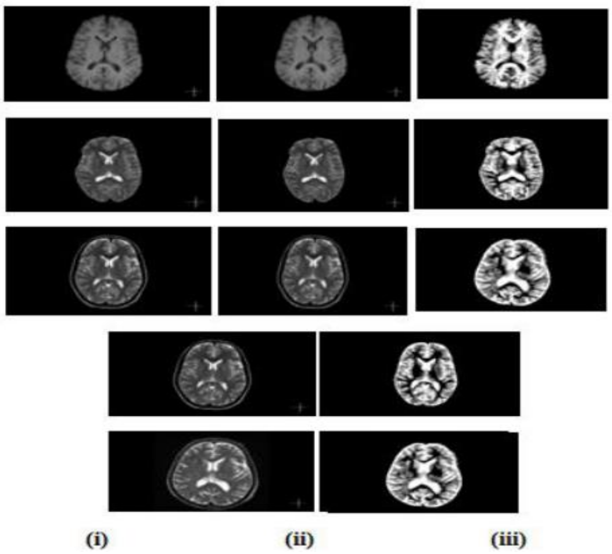


Fig. 3. Normal MR Brain Image Analysis (i) Five Real Input Images (ii) Median Filtered Images (iii) Segmented Images.

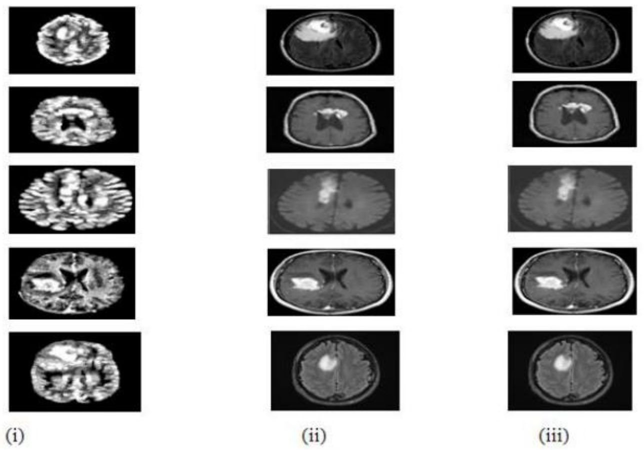


Fig. 4. Abnormal MR Brain Image Analysis (i) Five Real Input Images (ii) Median Filtered Images (iii) Segmented Images

CONCLUSION:

The current study is concentrated on the segmentation and classification of brain tissues for tumour deduction. There are five stages to this project. Clustering, feature extraction, feature selection, and classification using the median filter. The categorization accuracy produced by the suggested work is approximately 93%. This research can also be expanded to use a blend of swarm evolutionary approaches to detect disorders in MR brain pictures.

REFERENCES

1. Tapas Kanungo, David M.M ount, Nathan S.Netanyahu, Christine D.Piatko, Ruth Silverman, Angela Y.Wu,"An Efficient K-Means Clustering Algorithm: Analysis and Implementation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24, No. 7, pp. 881- 892, 2002.

2. Jing- Hao Xue, Aleksandra Pizurica, Wilfried Philips, Etienne Kerre, Rik Van De Walle, Ignace Lemahieu , "An Integrated Method of Adaptive Enhancement for Unsupervised Segmentation of MRI Brain Images", Pattern recognition letters, Vol.24, pp. 2549-2560, 2003.

3. László Szilágyi, Sándor M. Szilágyi, Balázs Benyó, and Zoltán Benyo, " Intensity Inhomogeneity Compensation and Segmentation of MR Brain Images using Hybrid *C*-Means Clustering Models", Biomedical Signal Processing and Control, Vol. 6, Issue 1, pp. 3-12, 2011.

4. Kuntimad, G and Ranganath, H.S, " Perfect Image Segmentation using Pulse Coupled Neural Networks", IEEE Transactions on neural networks, Vol. 10, No.3, pp. 591-598, 1999.

5. Jussi Tohka, Ivo D, Dinov, David W, Shattuck, Arthur W, and Toga, "Brain MRI Tissue Classification Based on Local Markov Random
Fields", Magnetic Resonance Imaging, Vol. 28, Issue 4, pp. 557-573, 2010.

6. Li-Hong Juang, and Ming-Ni Wu, "MRI Brain lesion Image Detection based on Color-Converted K-Means Clustering Segmentation", Measurement, Vol. 43, Issue 7, pp. 941-949, 2010.

7. Dongxiang Chi, "Self- Organizing Map- Based Color Image Segmentation with K-Means Clustering and Saliency Map", International Scholarly Research Network Signal Processing, (Research article id- 393891), pp. 1-18, 2011.

8. Thamaraichelvi B and Yamuna G, "A Complexity Reduced FCM Based Segmentation Technique for Brain MRI Image Classification", Journal of Medical Imaging and Health Informatics, Vol. 5, pp.202-209, 2015.

9. Radha Chitta, Rong Jin, Timothy C.Havens, Anil K.Jain, "Approximate Kernel K-Means: Solution to Large Scale Kernel Clustering", KDD'11, August 21-24,SanDiego,California, USA, 2011.

10. Chaplot S, Patnaik L.M and Jagannathan N.R, "Classification of Magnetic Resonance Brain Images using Wavelets as Input to Support Vector Machine and Neural Network", Biomedical signal processing and control, Vol. 1, pp. 86-92, 2016.

11. Mohanaiah P, Sathyanarayana P and Gurukumar L, "Image Texture Feature Extraction using GLCM Approach", International journal of scientific and research publications, Vol.3, Issue5, pp. 1-5, 2013.

12. Haralick R.M, Shanmugam K and Dinstein L, "Textural Features for Image Classification", IEEE Transactions on Systems, Man and Cybernetics, Vol. 3, No. 6, pp. 610-621, 1973.

13. Hari Babu Nandpuru, Dr. S. S. Salankar, Prof. V. R. Bora, "MRIBrain Cancer Classification Using Support Vector Machine", 2014IEEE Students' Conference on Electrical, Electronics and ComputerScience, 978-1-4799-2526-1.

14. Janki Naik, Prof. Sagar Patel, "Tumor Detection and Classificationusing Decision Tree in Brain MRI", Tumor Detection andClassification using Decision Tree in Brain MRI, ISSN: 2321-9939.

15. Ahmed KHARRAT, Karim GASMI, Mohamed BEN MESSAOUD,Nacera BENAMRANE and Mohamed ABID, "A Hybrid Approachfor Automatic Classification of Brain MRI Using Genetic Algorithmand Support Vector Machine", Leonardo Journal of Sciences, Issue17, July-December 2010, pp. 71-82, ISSN 1583-0233.

16. Ali Reza Fallahi, Mohammad Pooyan, Hojat Mohammad nejad,"Application of Morphological Operations in Human Brain CT Imagewith SVM", IEEE journal, 978-1-4244-2902-8/09.

17. Mehadi jafari, raza, "A hybrid approach for automatic tumor detection of brain MRI using support vector machine and genetic algorithm", global journal of science, engineering and technology ISSN 2322-2441, issue 3, 2012, pp 1-8.

18. El-Sayed Ahmed El-Dahshan, Tamer Hosny, Abdel-Badeeh M.Salem "Hybrid intelligent techniques for MRI brain image classification" ELSEVIER Digital Signal Processing, vol. 20, pp-433–441, 2010.

19. H. Men, Y. Gao, Y. Wu and X. Li, "Study on Classification Method Based on Support Vector Machine," 2009 First InternationalWorkshop on Education Technology and Computer Science, Wuhan, Hubei, 2009, pp. 369-373..