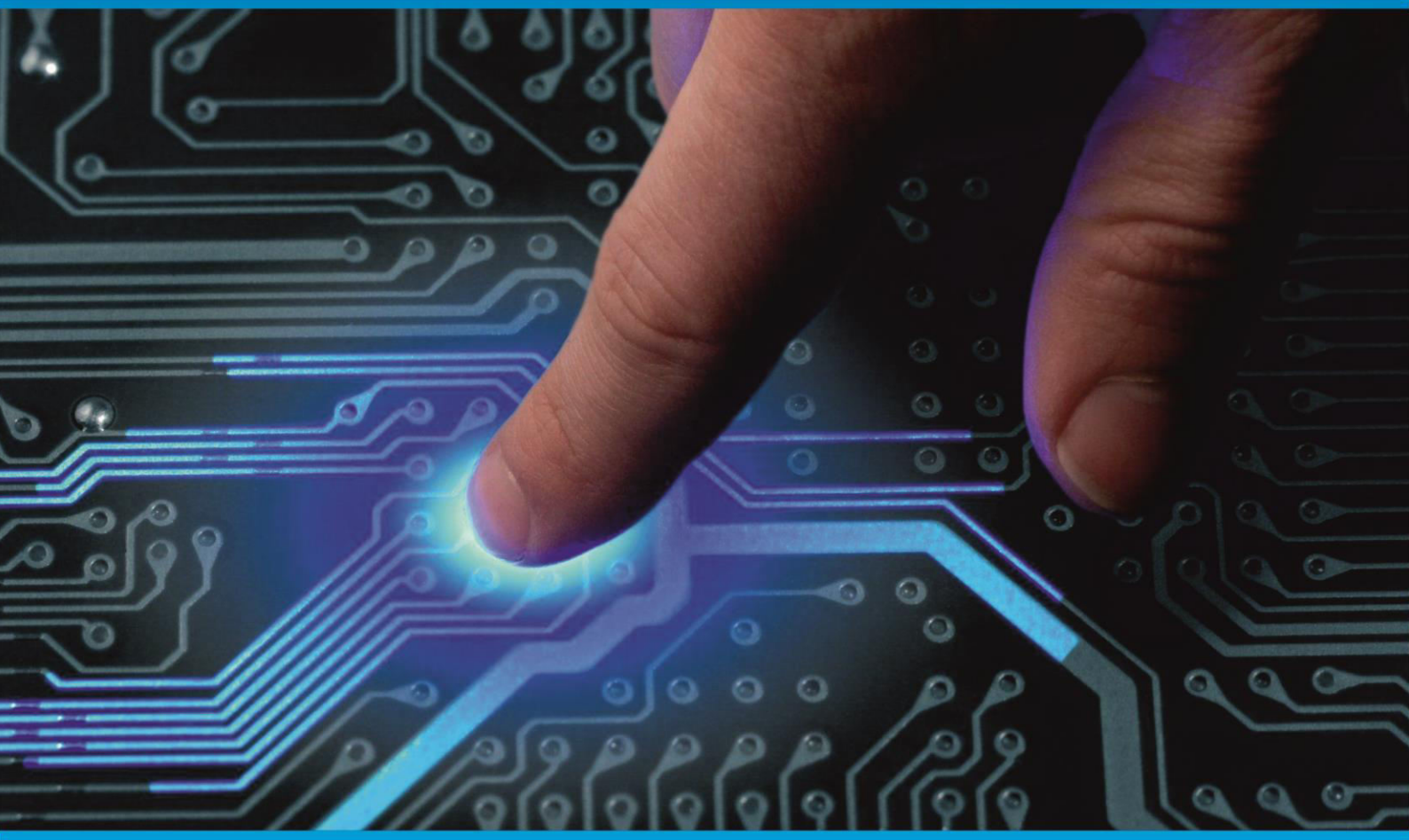
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 2, February 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902032 |

Classification of Brain MR Images into Malignant and Benign using Texture Features and Machine Learning Algorithm

Priya Kishor Chiwande

M.Tech, Dept. of CSE, Tulshiramji Gaikwad Patil College of Engineering and Technology, Mohgaon, Wardha Road, Nagpur, India

ABSTRACT: This research aim to classify brain MRI into malignant and benign class. First, the brain MRI is preprocessing to remove out the skull area called skull stripping. Texture and statistical features will be extracted from the skull stripped image and machine learning algorithm is then used to train the features for malignant and benign class. A Trained model can use further to classify the brain MRI into Malignant and Benign.

KEYWORDS: Benign, GLCM, high grade glioma, KNN, Low grade glioma, malignant, MRI, RBF kernel, SVM

I. INTRODUCTION

Day by day with the rapidly increasing population, cancer has one of the major global public health issues. Imaging plays a very vital role in analysis of diagnosis of patients having brain tumors. Tumors are unwanted groups of cells (tissue) which growth by uncontrolled cell division. Brain tumors are depends on the cell type from which they grow. Brain tumors classified into two types such as a primary and secondary. Primary tumors are collected of cells just like those that belong to the organ or tissue where they start. A primary brain tumor starts growing from cells in the brain. Day by day malignant tumors growing fast and can extend over large surrounding tissues. Secondary tumors are generated of cells from another part of the body that has spread to one or more areas. Secondary brain tumors are actually undistributed of cancer cells from somewhere else in the body that have metastasized, or spread to the brain. Radiologists inspect MRI images based on the visual clarification to identify the presence of tumors.

The goal of this paper is to classify the brain MRI into malignant and benign class. Algorithm used SVM, KNN algorithm. With the help of a web scrapping technology, website data could be collect in a format. It can used Machine learning method for collecting data. Then it is used data cleaning and data pre-processing method. So it will be showing more accuracy for Brain Tumor detection.

Many diagnostic imaging methods, including Computed Tomography (CT), Positron Emission Tomography (PET), and MRI may be used for early diagnosis of brain tumors. Magnetic Resonance Imaging is one of the green in the utility of brain tumor identification and diagnosis relative to other imaging approaches due to high spatial judgment, high soft tissue assessment. MRI, it does not emit any harmful radiation and is a non-invasive solution. There are 4 main steps in the method for tumor classification as shown in Figure 1. (a) Image pre-processing – removes noises in the image; (b) segmentation - find the tumor in an MR photograph; (c) feature extraction- removing redundant Features; (d) type identification - getting to know a type version of the usage of capabilities. Although MRI tends to be green in offering statistics on the location and size of tumors, MRI is not capable of classifying tumor forms, therefore, the implementation strategies include spinal and biopsy faucet procedures, which can be time-consuming and painful.

Our purpose of this project is to gain a high accuracy in discriminating the two kinds of tumors through an aggregate of several strategies for photograph segmentation, function extraction, and class. Hence, the proposed method has the capability of assisting scientific diagnosis.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902032 |

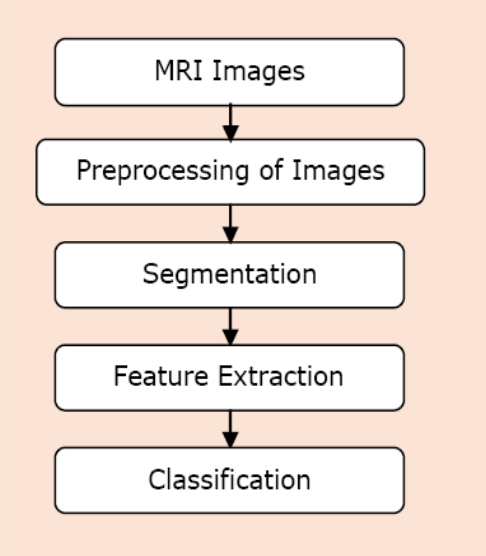


Fig. 1. Steps in Image Processing

II. RELATED WORK

R. J. Ramteke et al. [9] present an automatic medical image Classification technique. KNN classifier is used to classify the medical image into a normal and abnormal image. KNN is a very simple method which required low computational cost. Khushboo Singh et al. [8] Proposed the MRI image classification technique based on SVM classifier. Advanced classification techniques based on Support Vector. Support vector machine is a supervised learning algorithm. Quadratic programming performed the classification in SVM. Shweta Jain et al.[6] extract a feature using the GLCM technique and extracted features were classified using the artificial neural network. The biological nervous system inspired a mathematical problem in ANN. Priyanka et al.[7] proposed survey on the brain tumor detection algorithm and its location in the brain. Classifiers such as SVM, KNN, ANN, etc. have a large number of applications such as handwritten character recognition, face detection, iris detection, text classification, etc. Hari Babu Nandpura et al.[5] proposed a classification technique to identify normal and abnormal MRI brain image using techniques based on Support Vector Machines(SVM). In this paper, MRI Images will be carried out for gray scale, symmetrical and texture features for feature extraction. The proposed classifier gives 84% accuracy. A new hybrid technique based on the support vector machine (SVM) and fuzzy c-means the brain tumor classification are proposed in [12]. In this paper, the image is enhanced using enhancement techniques such as contrast, improvement, and mid-range stretch. Morphological operations and double thre-sholding are used for skull striping. In the brain MRI image, clustering (Fuzzy c-means (FCM)) is used for the segmentation of the image to detect the suspicious region. The MRI technique contains many imaging modalities that scan and captures the internal structure of the human brain. In paper [14], noise removal technique, extraction of gray-level co-occurrence matrix (GLCM) features, DWT based brain tumor region techniques are mainly concentrated to reduce the complexity and improve the performance. In [18], an automated method is proposed to differentiate easily between cancerous and non-cancerous MRI of the brain. Different techniques have been applied for the segmentation of candidate lesions. Then, the Support Vector Machine (SVM) classifier is applied with different cross-validations on the features set to compare the performance of the proposed framework. From the above literature survey, it is observed that, the most of the methods used individual feature extraction techniques. The individual features are insufficient to describe the brain MRI classification. There is need of a different type of feature extraction technique (we can say hybrid) to a robustly classification of the brain MRI into malignant and benign.

III. OBJECTIVES

This research aim to classify the brain MRI into malignant benign class. First the brain MRI is preprocessing to remove out the skull area called skull stripping. Texture and statistical features will be extracted from the skull stripped image and machine learning algorithm is then used to train the features for malignant and benign class. Trained model can use further to classify the brain MRI into Malignant and Benign. As a test of its effectiveness, we applied it to Parametric Test data; feedback from engineers has been encouraging. Future work should explore methods to improve

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902032 |

the accuracy of the classifiers by including the feature selection algorithms in the process, to find feature subsets that reliably result high accuracies. The future industrial work is the inline deployment of the system.

IV. PROPOSED RESEARCH METHODOLOGY

We have used advanced data extraction and analysis. The critical aspect of the code is presented in the Appendix with full details on data processing, analysis, results and interpretations.

In our case, the raw data of each attribute has been integrated into a complete data set. It was written into a CSV file to storage. Pandas Library for Python provides perfect data management and abundant analysis methods. We then prepare raw data before computational analysis.

This system has four steps; Pre-processing, morphological filtering, feature extraction and classification.

3.1 Datasets Description

Define In this method, the datasets is data collection, data processed and data clean are available from Web Scrapping Technology

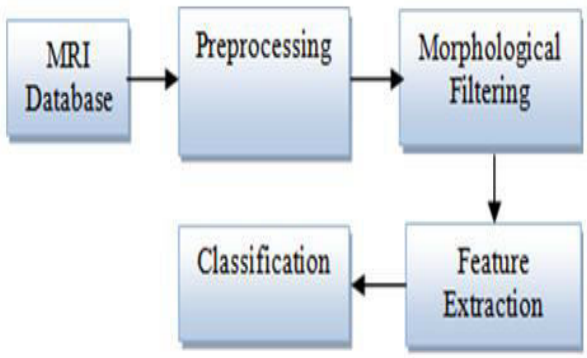


Fig. 2. Proposed work models

A. ***Database image:***

The MRI images for this approach are taken as a clinical database from Tata memorial hospital, Mumbai, the database contains malignant and benign tumor images and also from a BRATS 2012 database, we are taken a standard database images, contain glioma high and low grade images.

B. ***Preprocessing:***

All the input images are in RGB format. First RGB images are converted into grayscale. The captured medical images are noisy. Rician noise and pepper and salt noise suffered most of the medical images. Pre-processing operations include the median filter and power law transformation. Median filtering is a nonlinear filtering technique. It is important to remove the noise from an earlier stage to getting the accuracy at the last stages. The Median filter is used to suppress the pepper and salt noise from the image while preserving the edges. Considering a middle pixel of the mask as updating the pixel used the 3x3 mask is to remove the noise. Medical images are poor in contrast [10].

Neighbouring pixels get merged into one another in the lower contrast image. To improve the contrast of the image, this system uses power law transformation [11].

$$S = CT^{y}$$

S Is the power law transformation (S) of the given image,

Where, C is input intensity

γ is the output intensity

C. *Feature Extraction:*

Feature extraction is a technique to represent images in the feature set of an object of interest. There are different types of features such as texture, color, shape. The MRI image can be better distinguished by texture features. GLCM is a know texture feature extraction technique [12]. Different distances and different direction produces texture

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902032 |

classification, which produces classified output of input images where identified each texture region. For this approach of feature extraction single distance and four directions method is used to extract the GLCM feature.

D. ***Classification:***

The classification is the concluded step of the proposed work. In this proposed work, Machine learning algorithms: SVM and KNN classified brain MRI.

a.SVM

Support vector machine is a flawless method to find out the hyper plane between two different particular classes in high dimensional feature space which can be used for classification.

Supervised machine learning algorithm is also a Support vector machine is a [12]. Supervised learning techniques processed through two steps: Training and Testing. In the training phase, databases considered the two types first is the 251 (85 malignant and 166 benign) MRI clinical database images and second is the 80 (50 low grade glioma and 30 high grade glioma) standard MRI image is considered for training and 100 (50 malignant and 50 benign) images of clinical database and 40 (25 low grade glioma and 15 high grade glioma) images for testing respectively.

Sr. No.	Features	Formulas
1	Contrast	$\text{Contrast} = \sum_{i,j} \lvert i-j \rvert^2 p(i,j)$
2	Homogeneity	$Homogeneity = \sum_{i,j} \frac{1}{1+(i-j)^2} p(i,j)$
3	Correlation	$\text{Correlation} = \sum_{i,j} \frac{(i,j)p(i,j) - \mu_i \mu_j}{\sigma_i \sigma_j}$
4	Sum of Average	$\text{Sum of Average} = \sum_{i=0}^{2(N-1)} i \times p_{x+y}(i)$
5	Sum of Variance	$SOV = \sum_{i,j} (i-\mu)^2 P(i,j)$
6	Standard Deviation	$SD = \sqrt{\sum_{i,j=0}^{N-1} p(i,j)(i-\mu_i)^2}$
7	Autocorrelation	$\text{Autocorr} = \sum_{i,j} p(i,j) \log(p(i,j))$
8	Dissimilarity	$\text{DissIMILARity} = \sum_{i,j} \lvert i-j \rvert e(i,j)$
9	Energy	$\text{Energy} = \sum_{i,j} p(i,j)^2$
10	Entropy	$\text{Entropy} = -\sum_{i,j} p(i,j) \log(p(i,j))$
11	Difference of variance	$\text{Di}^{\dagger\dagger}\text{var} = \left(\sqrt{\frac{\sum \lvert p(i,j) - p(i,j)^2 \rvert}{n}}\right)^2$
12	Difference of Entropy	$\text{Diff Entroey} = \sum_{i,j} \lvert i-j \rvert^2 e(i,j)$
13	INV	$INV = -\sum_{i,j} e(i,j) \log(1 + \lvert e(i,j) \rvert)$
14	INN	$INN = \sum_{i,j} e(i,j)/(1 + \lvert e(i,j) \rvert)^2$

Table 1. The GLCM FEATURES

. SVM classification algorithm is depending on different kernel methods i.e. linear, radial basic function (RBF) and quadratic kernel function. The radial basic function By using linear function the SVM classifier classifies the image as

$$f(x) = W^T X + b \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Where, X is the training samples,

W is the weight assigned,

b is bias or offset

SVM classified into two types such as linear and Non-linear classification. The linear SVM classifier is importance to nonlinear classifier for mapping the input pattern into higher dimensional feature space. The data which can be linearly separable can be examine using hyper plane and the data which is linearly non separable those data are examine methodically with kernel function like higher order polynomial kernel, is apply on two samples **x** and **x'**, which indicate as feature vectors in some input space, and it can be defined as,

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902032 |

$$K(x, x') = \exp\left(\frac{||x-x'||^2}{2a^2}\right) \quad \text{.....(3)}$$

The value of kernel function is decreases according distance and ranges between zero (on the limit) and one (when x = x').

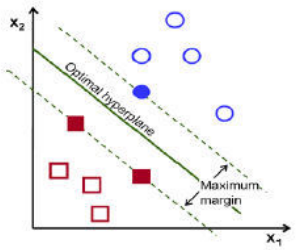


Fig. 3. Optimal hyper plane margin

b.KNN

KNN is a simple and robust classification technique. In this classifier, finding the k nearest training neighbour vector classify the testing feature vector. The distance between the training and testing vector is calculated by different distance, cityblock, chebychev, Minkowski, Mahalanobis, cosine, correlation, Spearman, hamming, Jaccord etc. In this method, Euclidean, cosine, cityblock, correlation distances are measured between testing and training data vector. The Euclidean distance between testing and training vector is given by

$$d(a,b) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2} \quad \text{..... (4)}$$

The label of the smallest distance feature vector is conveying to the testing vector.

As feature extracted for training and testing set of images, the different dimension is got by we in some space and these value of extracted feature take as an observation, its coordinate in that dimension from the characteristic, so set of Points in a space. So we can now consider the similarity of two different points to the distance between them in a space under some suitable metric.

Such a way in which the points are decided the applied algorithm from the training set are enough similar to the point considered when choosing the class to estimate the k closest data points picked to the new observation, and to take the most common class among these. In this way the k Nearest Neighbour algorithm performed.

The K- nearest neighbour algorithm is as follow [12]:

1. A positive integer value k is defined, along with the new sample.
2. Select the k values in our database which are close to the new testing sample.
3. We find out the most similar classification of these Entries.
4. This is the classification we give to the new sample using the value of k.
5. If the satisfactory results not obtained, changed value of k till correct results not obtained.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

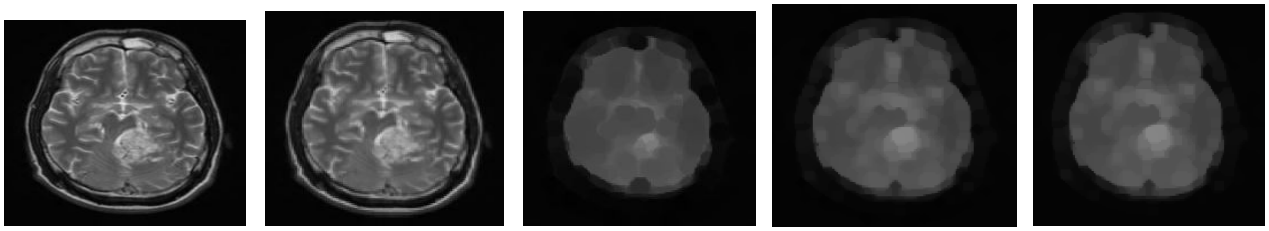
|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902032 |

V. DISCUSSION ABOUT RESULT

In the proposed work, the brain MRI is classified into four main steps: Image preprocessing, Morphological filtering, feature extraction and supervised classification. The each step results are shown below.

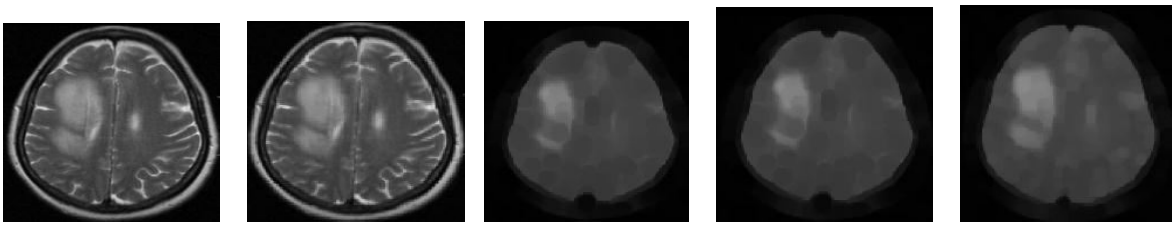
Case I: ***Results of malignant brain MRI***



(a) (b) (c) (d) (e)

Fig. 4. Pre-processing and morphological operation result of malignant brain MRI (a) Database Image (b) Median filter output (c) Erosion output (d) Dilation Output (e) Power law transformation output

Case II: **Results of benign brain MRI**



(a) (b) (c) (d) (e)

Fig. 5. Pre-processing and morphological operation result of Benign brain MRI (a) Database Image (b) Median filter output (c) Erosion output (d) Dilation Output (e) Power law transformation output

Kernels / Parameters	RBF	Linear	Quadratic
TP	21	20	26
TN	28	16	21
FP	1	10	2
FN	1	7	0
Sensitivity	95%	74%	100%
Specificity	96%	61%	91%
Accuracy	96%	67%	95%

Table II. Performance Analysis of SVM Classifier on Clinical Database

Distance / Parameters	Eculidean	Cityblock	Cosine	Correlation
TP	13	14	24	21
TN	24	22	21	20
FP	1	1	0	0
FN	12	11	5	4
Sensitivity	52%	56%	88%	84%
Specificity	96%	95%	100%	100%
Accuracy	74%	75%	93%	91%

Table III. Performance Analysis of KNN Classifier on Clinical Database

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902032 |

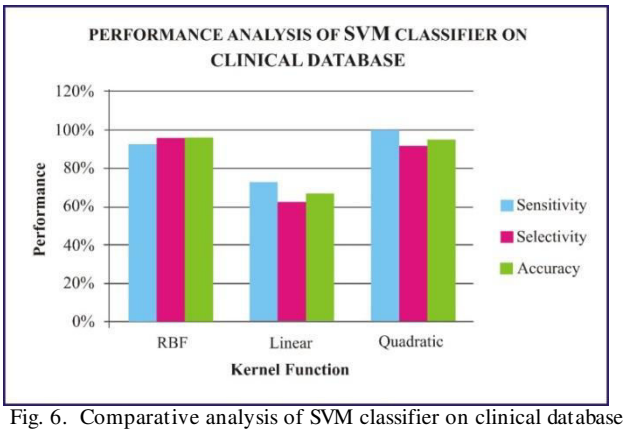


Fig. 6. Comparative analysis of SVM classifier on clinical database

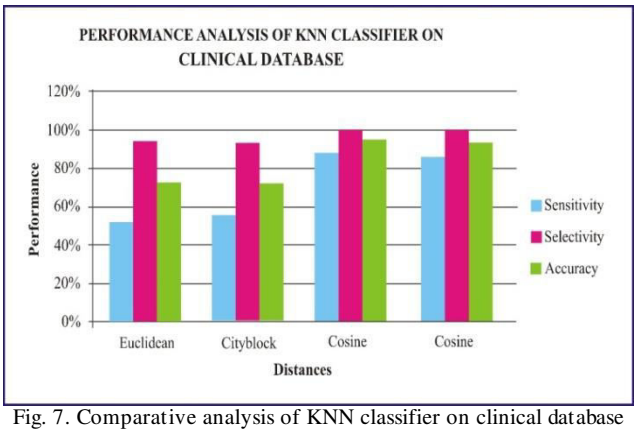


Fig. 7. Comparative analysis of KNN classifier on clinical database

The performance of the Classifier is calculates based on the three performance parameter; Sensitivity, Specificity and accuracy of the system. The formulae for the metrics are

$$Sensitivity = \frac{TP}{TP+FN} * 100\%$$

$$Specificity = \frac{TN}{TN+FP} * 100\%$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} * 100\%$$

Where,

TP = Malignant image is detect as malignant
TN = Benign image is detected as benign
FP = Benign image is detected as Malignant
FN = Malignant image is detected as benign

The performance of the proposed system has been compared with existing method described by Hari Babu Nandpuru. The existing system used DWT for feature extraction and features were classified by SVM and KNN algorithms. Comparison of the proposed system with the system proposed by [5] is shown in Table VI and graphically represented in fig. 12.

Methods	Precision(%)	Recall(%)	F Measure(%)
Proposed (SVM)	100	76	86.36
Proposed (KNN)	88	73.33	79.99
SVM	89.87	74.67	85.84
KNN	92.64	92.84	93.42

Table IV. Comparison of Proposed System with Existing System on BRATS Database

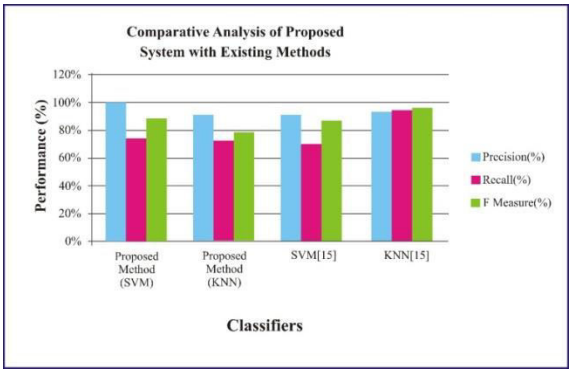


Fig. 8. Comparative analysis of proposed system with existing methods

VI. CONCLUSION

The brain MRI images classify into malignant and benign type using supervised SVM and KNN classifiers. In the training phase, databases considered the two types first is the 251 (85 malignant and 166 benign) MRI clinical

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902032 |

database images and second is the 80 (50 low grade glioma and 30 high grade glioma) standard MRI image is considered for training and 100 (50 malignant and 50 benign) images of clinical database and 40 (25 low grade glioma and 15 high grade glioma) images for testing respectively. The accuracy of the proposed system is 96% and 86% for SVM and KNN respectively for Brats database. From the results of proposed system, it is concluded that the KNN the accuracy of the SVM classifier is greater than the KNN classifier. It is also found that as we increases the number of training images the performance of SVM classifier increases.

In future, the accuracy of the proposed system can be increased by using the hybrid SVMKNN classifier.

REFERENCES

1. Classification of Brain Tumors: https://www.aans.org/en/Media/Classifications-of -Brain-Tumors. Access on 29 June 2020.
2. Gumaei, A., Hassan, M. M., Hassan, M. R., Alelaiwi, A., & Fortino, G. (2019). A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification. IEEE Access, 7, 36266-36273.
3. Usman, Khalid, and Kashif Rajpoot. "Brain tumor classification from multi-modality MRI using and machine learning", Pattern Analysis and Applications 20.3 (2017): 871-881.
4. Janki Naik, Prof. Sagar Patel, "Tumor Detection and Classification from using Decision Tree in Brain MRI", Tumor Detection and Classification using Decision Tree in Brain MRI, ISSN: 2321-9939.
5. Hari Babu Nandpuru, Dr. S. S. Salankar, Prof. V. R. Bora, "MRI Brain Cancer Classification Using Support Vector Machine", 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science, 978-1-4799-2526-1.
6. Shweta Jain, "Brain Cancer Classification Using GLCM Based Feature Extraction in Artificial Neural Network", International Journal of Computer Science and Mobile Computing (IJCSMC), ISSN: 2229-3345, Vol. 4 No. 07 Jul 2013.
7. Priyanka Balwinder Singh, "A REVIEW ON BRAIN TUMOR DETECTION USING SEGMENTATION", International Journal of Computer Science and Mobile Computing (IJCSMC), ISSN: 2229-3345, Vol. 4 No. 07 Jul 2013.
8. Khushboo Singh, Satya Verma, "Detecting Brain Mri Anomalies By Using Svm Classification", International Journal of Engineering Research and Applications, ISSN: 2248-9622, Vol. 2, Issue-4, June-July 2012.
9. Dr. R J Ramteke and Khachane Monali Y, "Automatic Medical Image Classification and Abnormality Detection Using KNearest Neighbor", International Journal of Advanced Computer Research, Volume-2 Number-4 Issue-6 December-2012.
10. Andac, Hamamci et al. "Tumor-Cut: Segmentation of Brain Tumors on Contrast Enhanced MR Images for Radiosurgery Applications", IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 31, No.3, MARCH 2012,pp.790-804.
11. Dr. R J Ramteke and Khachane Monali Y, "Automatic Medical Image Classification and Abnormality Detection Using KNearest Neighbor", International Journal of Advanced Computer Research, Volume-2 Number-4 Issue-6 December-2012.
12. 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.
13. Andac, Hamamci et.at., "Tumor-Cut: Segmentation of Brain Tumor on Contrast Enhanced MR Images for Radiosurgery Application", IEEE TRANSACTIONS ON MEDICAL IMAGING,VOL.31, NO.3, MARCH 2012, pp.790-804.
14. Nan Zhang, Su Ruan, Stephane Lebonvallet, Qingmin Liao and Yuemin Zhu. Kernel Feature Selection to Fuse Multi-spectral MRI Images for Brain Tumor Segmentation. Computer Vision and Image Understanding, 2011, 115(2):256-269.
15. Paterek, A. Padma et. Al., "Automatic Classification and Segmentation of Brain Tumor in CT Images using Optimal Dominant Gray level Run length Texture Features", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 10, 2011, pp.-53-59.
16. EI-Sayed Ahmed EI-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.
17. Ahmed KHARRAT, Karim GASMI, Mohamed BENMESSAOUD, Nacera BENAMRANE and Mohamed ABID, "A Hybrid Approach for Automatic Classification of Brain MRI Using Genetic Algorithm and Support Vector Machine", Journal of Science, Issue17, July-December 2010, pp.71-82, ISSN 1583-0233.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 2, February 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0902032 |

18. H.Men, Y. Gao,Y. Wu and X.Li, " study on Classification Method Based on Support Vector Machine", 2009 First International Workshop on Education Technology and Computer Science, Wuhan, Hubei, 2009, pp. 369-373.
19. Ali Reza Fallahi, Mohammad Pooyan, Hojat Mohammad nejad, "Application of Morphological Operations in Human Brain CT Image with SVM", IEEE journal, 978-1-4244-2902-8/09.
20. V. Vapanik, "The Nature of Statistical Learning Theory", Springer, N. Y., 1995. ISBN 0-387-94559-8.

BIOGRAPHY

Ms. Priya Chiwande, currently pursuing M.Tech (CSE) from Nagpur University, received degree in B.E (CSE) from Nagpur University and Professional Software Engineer.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details