Performance Analysis and Comparison of Machine Learning Algorithms for Classification of Brain Tumor in MRI Images

P.Kavipriya

Sri Ramakrishna College of Arts and Science -Autonomous (Formerly S.N.R Sons College), Coimbatore, Tamil Nadu, India

ABSTRACT

MRI is a progressive imaging system in medical field utilized to make best digital scan images of the internal parts enclosed in the human body. MRIs generate more detailed scan images than CT scans and are the favored way to identify a brain tumor. A brain tumor is defined as the growth of unusual cells in the tissues of the brain, which can be benign/ noncancerous or malignant/cancerous. It signifies a fascinating method for the structural valuation of tumors in brain since also offers high resolution data as well as greater soft tissue contrast. In this research, MRI scan images are taken for process further. So, in this research, the various machine learning techniques utilized for brain tumor detection such as SVM, KNN, NB and ensemble are analyzed. Hence, all these classification techniques are examined for finest results also reach maximum accuracy.

KEY WORDS: BRAIN-TUMOR, ENSEMBLE, GLRM, MAGNETIC RESONANCE IMAGE, ML ALGORITHMS.

INTRODUCTION

Abnormal growth of cells in brain is called as brain tumor. MRI imaging system is the mostly utilized scan device for analysis. The size and location of brain tumor detection is significant role in medical system that is pretty challenging due to the varied forms, presences and sizes of tumors. (K. Machhale et.al, 2015). Therefore, to overcome these challenges, classification process is necessary. Image processing (IP) methods are applied to increase the process of classification of Brain tumor in MRI scan images (Babu et.al, 2020). The below Fig.1 includes 4 basic phases in IP system.

In IP, Image pre-processing, segmentation, feature extraction and image classification are mostly concerned basic operations: Pre-processing: Pre-processing is done reduce noise and enhance the image for further processing. This step improves image quality and increase surety and accuracy in detecting tumor. (J.Rajeesh et.al, 2010). Segmentation: the method of separating an input image into (object and background) various fragments. Feature extraction (FE): It extracts finest features from a segmented images required with the purpose of classification process. Classification: The aim of classification is to classify every pixels into appropriate classes.

The primary objective of the proposed work is to classify the brain tumor using ML techniques in MR scan images. In this research, MATLAB interactive tool is used for doing implementation process. Hence MATLAB has been taken into consideration and all the techniques have been implemented using MATLAB. 2.

159

Figure 1: Essentials steps in IP
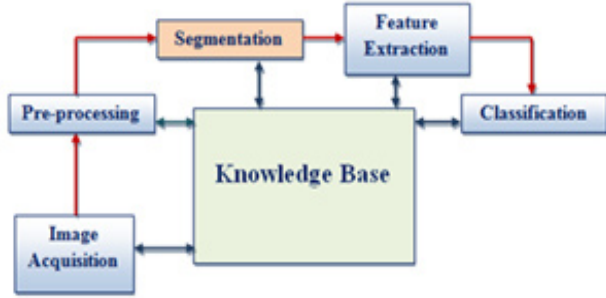
Step1	Choose a 3 by 3 2-Dimension window form input medical image. Let's assume P_{mn} is the current pixel value of the input window.
Step2	Now Sort every values appear in the present window and split those windows as two categories, such as X_{mn}= pixels excluding 0 and 255, Y_{mn}= pixels including 0 and 255.
Step3	In case of the present pixel rate occurs from 0 to 255 i.e. $P_{mn} \in M_{mn}$, then P_{mn} is a noise free pixel that is leave as it is.
Step 4	If all the values are noiseless, then performing pixel is restored with median value of the current window.
Step 4 Step 5	If $P_{mn} \in N_{mn}$ and its adjacent pixels are including the noise values, then increase the dimension of the processing window into 5X5. • In current window, the pixel values are not 0 or 255, then obtain the middle value from current window after that restore P_{mn}. • If pixels lie in the processing window, then restore P_{mn} with preprocessed pixel.
Step6	Shift the window to subsequent element.

Proposed Methodology: IP and ML algorithms are applied to increase the brain tumor classification process in MR scan images. The proposed techniques, Decision Based median filter for noise removal (Priyanka Kamboj et.al, 2013), K-means clustering method for segmentation, Gray-Level Run-Length Matrix (GLRM) to extract features for Brain tumor images and ML techniques for classification are analyzed and compared.

2.1 Hybrid Median Filter (HMF): HMF originally express the firmness of the input image pixel, which is affected with noise depend on the value of pixel. For example, Pepper noise denoted as 0 and salt noise denoted as 255. Finally, the middle pixel values which means the value range from 0 to 255 is measured as a noise free pixel (C. Anjanappa et.al, 2015). So, when the pixel is identifies as noise, it can be interchanged by median value for noise filtering. This procedure has following steps, The Methodology of Proposed Filter system is described in Fig2.

Figure 2: Methodology of Proposed Filter
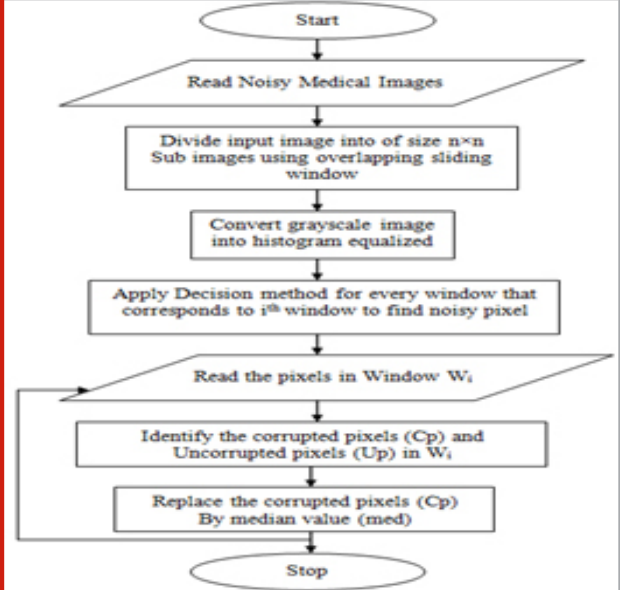
i. K–MC Algorithm: KMC (K-means clustering) method have been developed as a digital image segmentation technique in various fields and applications. Novel KMC algorithm select k initial points as primary clustering centers, various points can acquire various solutions. The K-means algorithm develops a divisive clustering and utilizes a correspondence metric to assign every document to one of k clusters. The clusters are considered as a standard of all pixels contained within the cluster.
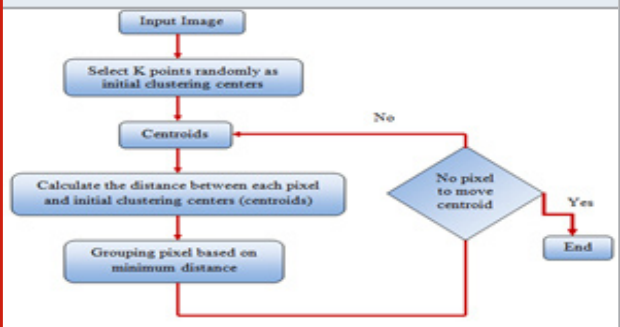

Figure 3: Flow of KMC Algorithm

Fig.3.illustrates the Methodology of K-means clustering Algorithm that has some limitations. For example, the superiority of the results is depend on the initial selection of centroid. That cab be sensibly chosen so that it gets desire segmentation part. (Zexuan, Ji et.al, 2012).

FE (Feature Extraction): It is a dimensionality reduction process in IP. If the input dataset is too large that can be suspected to be extremely redundant then the large dataset will be changed into an abridged exemplification of feature vectors. This process is called as FE. (S.A. Akar et.al, 2016).

GLRM: It is represented in the form of a matrix for geometrical features. It gives a measure of the intensity of the pixels along the given direction mentioned as Run length. It has two dimensions. Here, each element is represented as the number of components 'j' with the intensity 'i', in the specified directions. In that way, every gray level value can be estimated that how many times it occur in that specific direction run. Whether 2 successive pixels have the same intensity value, next time it takes for 3 pixels and compares it and next it goes for 4 and so on (Hongli Guo et.al, 2020). Features extracted by GLRM were.

SRE: "Short Run Emphasis: It processes the scattering of 'short runs ' that is calculated by equation 1. It is extremely based on number of 'short runs' also it is probable to be of little value to the fine texture and of great value to the rough texture.

$$SRE = \frac{\sum_i \sum_j p(i,j)}{j^2} \tag{1}$$

LRE: "Long Run Emphasis: It processes the scattering of 'long runs', which is extremely based on the number of 'long runs' also it is estimated to be of a substantial value of fine textures and small values on coarse textures calculated by equation 2.

$$LRE = \frac{\sum_i \sum_j j^2 p(i,j)}{\sum_i \sum_j p(i,j)} \tag{2}$$

RLU: Run Length non–Uniformity: It extents the equation of run length during the input image and is expected to be of little value if the length of the run is similar to the image that is shown in below equation 3.

$$RLU = \frac{\sum_i (\sum_j p(i,j))^2}{\sum_i \sum_j p(i,j)} \tag{3}$$

GLN: " Gray Level Non–uniformity: It calculates equation of 'gray-level' values in the input image and that estimated to be of little value if the grade of gray is similar to the image described which is described in equation 4.

$$GLN = \frac{\sum_j (\sum_i p(i,j))^2}{\sum_i \sum_j p(i,j)} \tag{4}$$

RP: " Run Percentage: It calculates the compatibility and circulation of runs of an input image in a specific way, which is the greatest value if the 'run length' is 1 aimed at altogether degrees of gray in a particular direction. It is calculated by equation 5.

$$RP = \frac{\sum_i \sum_j p(i,j)}{n} \tag{5}$$

LGLRE: "Low Gray Level Run Emphasis: It calculates the scattering of 'low gray-level' values with a maximum value representing a greater concentration of low gray-level values in the image that is described in equation 6.

$$LGLRE = \frac{\sum_i \sum_j \frac{p(i,j)}{i^2}}{\sum_i \sum_j p(i,j)} \tag{6}$$

HGLRE: "High Gray Level Run Emphasis: HGLRE measures the distribution of the higher gray-level value in the image shown in below equation 7. (Hongli Guo et.al, 2020).

$$HGLRE = \frac{\sum_j \sum_i i^2 p(i,j)}{\sum_i \sum_j p(i,j)} \tag{7}$$

Supervised Classification: This research applies the some Supervised ML algorithms (NB, KNN &SVM).

SVM: "Support Vector Machine: It is a learning machine for 2 group classification problems, which is established to classify the image features as positives or else negatives. It works well for classification task and it has the potential to hold optimal features. One more benefit of SVM is the strongness where there's a thin set of samples because most of the cases are linearly divisible (R. Balakrishna et.al, 2018). In statistical term, SVM builds for separating hyper plane in high-dimensional vector spaces. Assume, feature points are showed as (x, y) 'tuples' where 'x$_j$'is the feature values plus 'y$_j$' is the class. Multi-dimensional feature space consists for hyper plane is described in the equation 8 and 9 below.

$$b.x + b_{-0} = 0 \tag{8}$$

The function for the above equation is calculated as:

$$f(x^*) = b.x^* + b_0 \tag{9}$$

Here, it's to seek out 'b' and 'b_0' in order to determine maximal margin hyper plane. The SVM technique requests to locate the hyperplane that generates the greatest margin for the two classes between training points. It also castigates the total space of points on the erroneous side of their margin when the two groups of data overlap. This helps to tolerating an inadequate number of misclassifications next to the margin. Here, new two parameters '∈' and 'C' are introduced to allow violation. Maximize margin of 'M' is defined as equation 10,11 and 12,

$$\sum_{j=1}^{p} b_j^2 \tag{10}$$

And

$$y_i(b.x + b_0) \geq M(1-\epsilon_i), \forall_i = 1\ldots n \quad (11)$$

Where

$$\epsilon_i \geq 0, \sum_{i=1}^{n} \epsilon_i \leq C$$

Parameter 'C' together controls the percentage of individual 'i' and are frequently customized to violate the margin. In this research, SVM classifier is employed for experimentation.

Algorithm: SVM classifier
Input: Extracted features
Output: Accuracy and Validity
- *Step1:* Start
- *Step2:* Input the dataset
- *Step3:* Classify the dataset
- *Step4:* Apply the SVM Machine learning with kernel function (Radial Based Function (RBF))
- *Step5:* Specify the Hyper-plane
- *Step6:* If obtained Accuracy and validity is NOT acceptable then go to step 4
- *Step7:* End

KNN: "K-Nearest Neighbor: KNN algorithm plays an important role in machine learning system. (Saravanan K1 et.al, 2014). Consider $X_j=\{x_1,x_2,\ldots,x_{iN}\}$ and $X_j=\{x_1,x_2,\ldots,x_{jN}\}$ the sample population, thus to measure the similarity between them and the distance is calculated as given in equation 12.

$$\text{Dist}(X_i, X_i) = \sqrt{\sum_{m=1}^{N}\left(x_{im} - x_{jm}\right)^2} \quad (12)$$

In KNN, K is the quantity of closest neighbors. The quantity of neighbors is the center central factor. K is commonly an odd number if the quantity of classes is 2. When K=1, at that point the calculation is known as the closest neighbor calculation. This is the most straightforward case.

Algorithm: KNN classifier
Input: Extracted features
Output: Accuracy and Validity
Classify (X, Y, x)//X: Training data; Y: Class labels of X, x: Unknown sample
- *Step1:* Start
- *Step2:* Select the K (number of the neighbors) value
- *Step3:* For i=1 to m do
- *Step4:* Compute distance d (Xi, x)
- *Step5:* end for
- *Step6:* Compute set I containing indices for the k smallest distances d (Xi, x)
- *Step7:* Return majority label for {Yi where i ∈ I}
- *Step8:* End

NB: "Naïve Bayes: It is utilized for many operations for example spam filtering also other areas of text classification. Wherein, the joint prospects of features then grades are utilized to approximately compute the probability score of grades of a specified feature subsets.

This classifier applied the simple probabilistic classifier, which assist in classifying a data 'd_r', out of classes $c_i \in$.

$C(C_{i=1}^{m} = c_1, c_2,\ldots.c_m)$. The finest class returns in 'NB' classification is the Maximum Posterior (MAP) class described by below equation 13,

$$\begin{array}{c} C_{map} \\ C1\in C \end{array} = \text{argmax} P(c_i) P \quad (13)$$

Here, the class '$P(c_i)$' can be calculated by dividing the total number of features in class 'c_i' by the entire number of features. $P(d_r|c_i)$ denoted the number of incidence of the feature in data 'dr'belongs to class 'c_i'. The probability value '$P(c_i|d_r)$' will be calculated for every latent class, but '$P(d_r)$' doesn't change for every class. Accordingly, it can drop the denominator. It chooses the maximum probable classes 'c_{map}' of given data 'd' by computing the posterior probability of every class (R. Balakrishna et.al, 2018).

Algorithm: NB classifier
Input: Training dataset T,
F= (f_1, f_2, f_3... f_n) // value of the predictor variable in testing dataset.
Output: A class of testing dataset.
Step1: Read the training dataset T;
Step2: Calculate the mean and standard deviation of the predictor variables in each class;
Step3: Repeat
        Calculate the probability of f_i using the gauss density equation in each class;
        Until the probability of all predictor variables (f_1, f_2, f_3... f_n) has been calculated.
Step4: Calculate the likelihood for each class.
Step5: Get the greatest likelihood.

d.Ensemble KNN–SVM algorithm: In classification algorithms, each one has its own advantage and disadvantage. So, Ensemble SVM-KNN algorithm is compared with above mentioned algorithms to achieve the highest accuracy then others. The working mechanism of proposed algorithm is explained in the below section. In proposed technique, K nearest neighbor technique finds the distance between test sample and training sample.

Algorithm: Ensemble KNN - SVM classifier

Input: MRI brain tumor images with class label (benign or malignant) i.e. $(X_1, C_1),(X_2, C_2)\ldots (X_n, C_n)$;

Feature pool F= $\{f_m, m=1\ldots n\}$; Number of iterations = R

Initialization: Weight of each features calculated with equation 14.

$$\frac{1}{N}; \forall \, i \, (i = 1, ..., N); \sigma = 1000 \qquad (14)$$

For r = 1 to R do:

(a) Generate a training set by sampling with $\{w_i(r)\}$
(b) Train base classifier hr ((Proposed Ensemble SVM - KNN Classifier)) using this training set

1. Apply SVM classifier on MRI data set with K-fold cross-validation and K=10.
2. Update the weights.
3. According to Wolfe dual form (equation 15 and 16), weight minimization is
4. Predict the test KNN-SVM class using the cross validated model with minimum weight.
5. Apply weighted K-Nearest Neighbor Classifier with number of nearest neighbors K=10 on MRI data set.
6. Apply K-fold cross validation with K=10.
7. Weight contribution of each k neighbor
8. Set initial weights of KNN = updated minimum weights of SVM.
9. X_t is test MRI image
10. Predict the test MRI class using the cross validated model with minimum weight.
11. Take weighted average of predictions from both the models.
12. Compute the training error.

$$Minimize : w(\alpha) = -\sum_{i=1}^{N} \alpha_i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j k(X_i, X_j) \qquad (15)$$

$$Subject - to : \sum_{i=1}^{N} y_i \alpha_i = 0, \forall i; 0 \leq \alpha_i \leq C \qquad (16)$$

Figure 4: Denoised images by using various filters
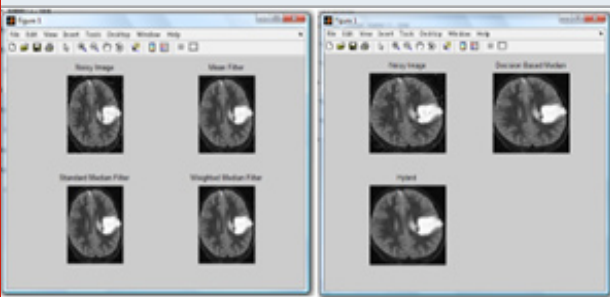
Ensemble KNN-SVM as component classifier for brain tumor classification. Proposed scheme gives classification accuracy of 97.3% for brain tumor MRI classification. Results reveal that proposed Ensemble KNN-SVM outperforms other methods.

RESULTS AND DISCUSSION

163

The efficiency of the ensemble proposed method is demonstrated by the experimental results. The proposed filter can eliminate the noise without deteriorating the original image. Experiment outcomes demonstration that the proposed method can increase the filtering process significantly. In the below Fig.4., the Denoised images by using various filters are presented. Image segmentation refers to clustering the homogeneous pixels into various groups while classification is next hierarchy which labels those clustered pixels as different classes. In this work, a primary task of brain tumor detection is to divide tumor images into segments, which contain different grey values, textures and other characteristics. Although easily segmented by human observer, there exists no robust automated approach that can consistently separate relevant categories in MRI Brain tumor-imagery.

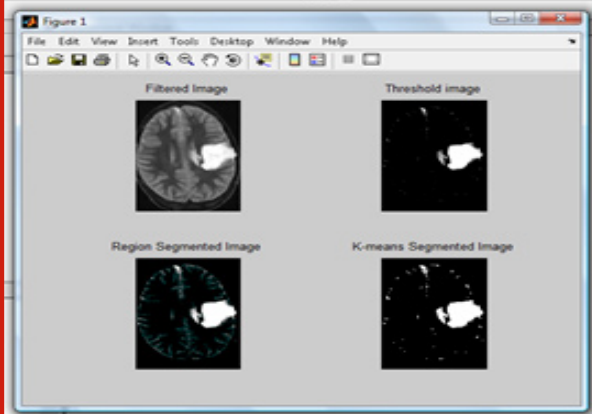Figure 5: Segmentation Outcomes for MRI (Brain Tumor) images

Table 1. Performance comparison of various ML algorithms with various evolution parameters

ML Methods	Precision	Recall	Accuracy
NB	84.21	85.71	85.7
KNN	91.22	92.85	92.4
SVM	96.36	94.64	95.8
Ensemble	98.8	96.44	97.3

In the above Fig.5, the results for Brain Tumor MRI images are illustrated. The Brain tumor data set contains approximately 119 samples extracted from the MRI images. Tumor classification is the important area in machine learning fields. Texture features namely GLN, HGRE, LGRE, LRE, RLN, RP and SRE. The Brain tumor dataset consists of 119 images, out of which 80% were taken as training data and 20% features were taken as testing data. The proposed metrics evaluate the performance of given algorithm by comparing the qualities of input and output images.

3.1 Performance Evaluation: The extracted GLRM features are utilized for training as well as testing system. For valuation, 80% of data were taken for training and then 20% of data were taken for testing. The assessment is accepted out for the different algorithms with the below parameters (S. Radha Priya et.al, 2020).

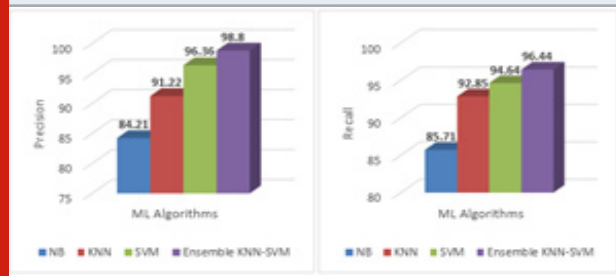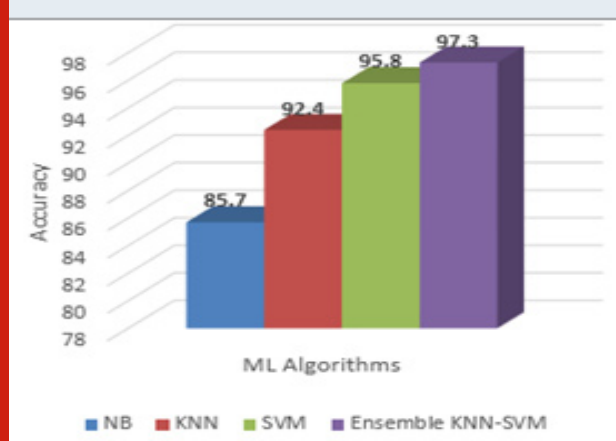Figure 6: Performance evaluation with Precision and Recall

Figure 7: Performance evaluation with Accuracy

$$Accuracy = \frac{DR}{TNI} \times 100 \qquad (17)$$

$$Precision = \frac{TP}{TP+FP} \times 100 \qquad (18)$$

$$Recall = \frac{TN}{TP+FN} \times 100 \qquad (19)$$

In the above equations (17,18 & 19), Where TP: 'True Positive' , TN: 'True Negative' , FP: 'False Positive' and FN: 'False Negative' values and 'DR': 'Detected Results' and 'TNI': 'Total number of Iterations'. The above figures (Fig.10 and Fig.11) clearly shows that Accuracy, Precision and Recall of the ensemble algorithm has a maximum accuracy when compared with the other ML techniques. The proposed ensemble classifier yields better classification accuracy, because it has a regularization parameter, which avoids over-fitting.

CONCLUSION

In normalization process, hybrid median filter shows proficient denoised also a best input image detail-preservative ability. The results confirm good performance, which could be used for the filtering the Brain tumor MRI images. This filter can eliminate the noise without deteriorating the original image. In segmentation process, the Brain tumor images were clearly segmented using region growing segmentation method. The features of the Brain Tumor images were extracted in feature extraction phase by using GLRM technique. The extracted features were used for classification. From the experimental results of classification phase, it is shown that proposed ensemble KNN-SVM gives the better performance in terms of various performance evaluation parameters.

REFERENCES

Babu, R. G., Maheswari, K. U., Zarro, C., Parameshachari, B. D., & Ullo, S. L. (2020). Land-Use and Land-Cover Classification Using a Human Group-Based Particle Swarm Optimization Algorithm with an LSTM Classifier on Hybrid Pre-Processing Remote-Sensing Images. Remote Sensing, 12(24), 4135.

K. Machhale, H.B. Nandpuru, V. Kapur, L. Kosta. (2015) MRI brain cancer classification using hybrid classifier (SVM-KNN). International Conference on Industrial Instrumentation and Control, pp. 60-65.

J.Rajeesh, R.S.Moni, S.Palanikumar and T.Gopalakrishnan. (2010) Noise reduction in magnetic resonance images using wave atom shrinkage. International journal of image processing(IJIP), volume(4): Issue(2).

Priyanka Kamboj and Versha Rani. (2013) A brief study of various noise model and filering techniques. Journal of Global Research in Computer Science, Volume 4, No 4, pp. 166-171.

C. Anjanappa,H.S.Sheshadri. (2015) Development of mathematical morphology filter for medical image impulse noise removal,'' International Conference on Emerging Research in Electronics, Computer Science and Technology, pp-311-318.

Zexuan, Ji., Sun, Q., Xia, Y., Chen, Q., Xia, D., & Feng, D. (2012). Generalized rough fuzzy c-means algorithm for brain MR image segmentation. Computer Methods and Programs in Biomedicine, 108(2), 644–655.

S.A. Akar, (2016) Determination of optimal parameters for bilateral filter in brain MR image De-noising. Applied soft computing 87-96.

Hongli Guo,Bin Li,Youmei Zhang,Yu Zhang,Wei Li,Fengjuan Qiao, Xuewen Rong, Shuwang Zhou.Gait (2020). Recognition Based on the Feature Extraction of Gabor Filter and Linear Discriminant Analysis and Improved Local Coupled Extreme Learning Machine. Hindawi Mathematical Problems in Engineering.

Saravanan K and S. Sasithra. (2014) Review on Classification Based on Artificial Neural Networks. International Journal of Ambient Systems and Applications Vol.2, No.4.

R. Balakrishna, R. Anandan. (2018) Soft Computing Analysis for Detection of Pancreatic Cancer Using MATLAB. International Journal of Pure and Applied Mathematics, Volume 119 No. 18, 379-392. ISSN: 1314-3395.

S. Radha Priya, Dr. M. Devapriya. (2020) Twitter Sentiment Analysis with Diabetic Drugs Using Machine Learning Techniques with Glowworm Swarm Optimization Algorithm. International Journal of Engineering Research & Technology, ISSN: 2278-0181, Vol. 9 Issue 07.