International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 09 | Sept -2020 ISSN: 2582-3930

Detection and Classification of Brain Hemorrhage Using Ensemble Learning

Ashwini kadam[1], Dr.K.B.khanchandani[2]
Ashwini Kadam entc dept, SSGMCE
Dr.K.B.Khanchandani entc dept, SSGMCE

Abstract:

Brain Hemorrhage are one of the diseases caused in the brain. In this project we have designed a software tool in MATLAB for locating brain hemorrhage using segmentation based on supervised methods. We have successfully implemented the different supervised methods like Classification, Regression and Ensemble method for locating brain hematoma. The algorithms tested on database of 50 images and found good results. Also we have given the performance analysis and comparisons of these methods based on no. pixels. The Random forest algorithm gave better result than the other Regression method. Random forest achieved almost 95% result. Accuracy and stability, sensitivity is almost 95%.

INTRODUCTION

In this project we use the ensemble method. This project analyses various techniques to locate hemorrhage objects in Magnetic Resonance (MR) brain images. The input is the MRI image of the axial view of the human brain. In supervise algorithm there are various technique in this project, the Ensemble algorithms used. bagging is a simple and very powerful ensemble method. In the training phase of the input images taken from a given location to extract input features and the known output will be found by naming the images from the type of hemorrhage. Then the net file can be generated using a train tool for the first time after going through few testing iterations by providing the saved input and output files. Once the input

Diagnosing the Subarachnoid hemorrhage can be done efficiently by various machine learning techniques. Purpose of using Machine learning technique is to focus on factors that influence the prediction performance.

C. Amutha devi, Dr. S. P. Rajagopalan has presented, image classification is a critical step for high-level processing of automatic brain stroke classification [7].

features are calculated and the vector is created, to add the image to train, the output will be defined according to the value that has been received as the output

result. Once the input and output files are saved, system can be trained with them. This logic can be used to train the tested images as well. The performance analysis is conducted by taking a MRI Brain hemorrhage image as the input and applying all the algorithms to the image.

LITERATURE REVIEW

Rupali Mahajan, Dr. P. M. Mahajan has presented a diagnosis of brain hemorrhage in more refined manner by feeding CT images & identifies the type of brain hemorrhage using watershed algorithm [2] along with artificial neural network (ANN)

In this paper, presented a method to detect and classify the type of brain hemorrhage from given CT image. This method consists of several stages: image preprocessing, image segmentation, feature extraction and classification. The watershed algorithm is used to segment the image. The features are extracted by using Grey Level Co-occurrence Matrix (GLCM). The feed forward Back Propagation Neural Network is used to classify the type of hemorrhage.

C.Dheeba, S.Vidhya had a survey on Prediction of Brain Hemorrhage Using Various Techniques [1].The main objective of this work is to predict Subarachnoid hemorrhage (SAH) using machine learning techniques and analyzing the classification performance of various existing machine learning algorithms.
In this paper, a method is proposed for classifying the MRI images into stroke and non-stroke images. Features are extracted using Watershed segmentation and Gabor filter. The extracted features are classified using Multilayer Perceptron (MLP). Experiments have been conducted to evaluate the efficiency of the proposed method with varying number of features. . To evaluate the efficiency of the proposed method, a

dataset of 52 DWI scan images are complied. The images were procured from the MRI Department of Vijaya Health Centre, India. Of the 52 images, 25 images are of positive stroke images. This study proposes a method for classification of MRI brain image as stroke and non-stroke.

SYSTEM ARCHITECTURE AND PROPOSED WORK

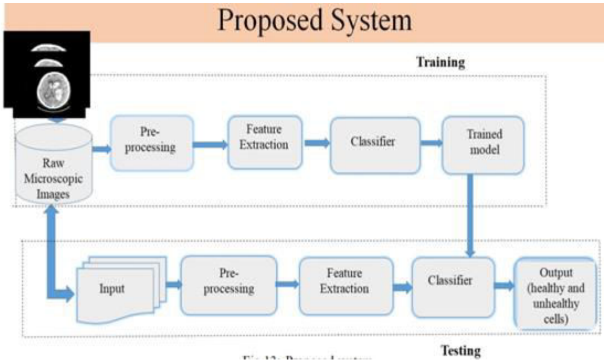


Fig 4.1: block diagram

A brain hemorrhage is bleeding in or around the brain. It is a form of stroke. Causes of brain hemorrhage include high blood pressure (hypertension), abnormally weak or dilated (aneurysm) blood vessels that leak, drug abuse, and trauma. Many people who experience a brain hemorrhage have symptoms as though they are having a stroke, and can develop weakness on one side of their body, difficulty speaking, or a sense of numbness. Difficulties performing usual activities, including problems with walking or even falling, are not uncommon symptoms.

1) Dataset :

The dataset consists of 50 MRI/ CT images of human brain. These images include 25 images of normal brain type, 25 images of abnormal brain type (**Intracerebral Hemorrhage (ICH), Subdural hemorrhage (SDH), Extradural Hemorrhage (EDH), Subarachnoid hemorrhage (SAH)**).

CT/MRI images helps in identifying image bone, soft tissues and blood vessels all at the same time. These images are converted in jpeg form to be uploaded to the system for preprocessing.

2) Pre-processing:

Pre-processing improves the quality of an image. In this system, preprocessing techniques are developed to remove the skull portion surrounding the tissues.

1. Conversion of Image

The CT image is converted into gray scale image to make it contrast. The contrast image helps in giving exact information about the tissues.

Pre-processing is an operation on images at the lowest level of abstraction which improves the image data that suppress unwanted distortions and image features are enhanced that are important for further analysis and processing. It does not increase any kind of image information. First the images are transformed to grey scale image, then they are converted to double precision, images are cropped for satisfactory feature extraction and then the images

2.Morphological operation:

Dilation and erosion operators are further used in complex sequences of opening and closing

3) Feature Extraction:

In our project we are have extracted the feature following steps are the first step is wavelet, second step is discrete wavelet transform, third step is PCA and the last one statistical feature.

3.1) Wavelet texture :

Texture is a property that represents the surface and structure of an Image. Generally speaking, Texture can be defined as a regular repetition of an element or pattern on a surface. Image textures are complex visual patterns composed of entities or regions with sub-patterns with the characteristics of brightness, color, shape, size, etc. An image region has a constant texture if a set of its characteristics are constant, slowly changing or approximately periodic . Texture can be regarded as a similarity grouping in an image. The dimensions of texture cannot be same; hence to represent texture with a single method is not possible. Texture analysis is a major step in texture classification, image segmentation and image shape identification tasks.

3.2) Principal component analysis (PCA)

The main linear technique for dimensionality reduction, principal component analysis, performs a linear mapping of the data to a lower-dimensional space in such a way that the variance of the data in the low dimensional representation is maximized. In practice, the covariance (and sometimes the correlation) matrix of the data is constructed and the eigenvectors on this matrix are computed. The eigenvectors that correspond to the largest Eigen values (the principal components) can now be used to reconstruct a large fraction of the variance of the original data. Moreover, the first few eigenvectors can often be interpreted in terms of the large-scale physical behavior of the system, because they often contribute the vast majority of the system's energy, especially in low-dimensional systems. Still, this must be proven on a case-by-case basis as not all systems exhibit this behavior.. The original space (with dimension of the number of points) has been reduced (with data loss, but hopefully retaining the most important variance) to the space spanned by a few eigenvectors.

Texture features are determined from the observed combination of statistical distribution. As stated the number of intensity of pixels are categorized into first, second and higher order statistics. One way to extract second order statistical features is through (GLCM) Gray level co-occurrence matrix

To create a GLCM, use the graycomatrix function. The function creates a gray-level co-occurrence matrix (GLCM) by calculating how often a pixel with the intensity (gray-level) value i occurs in a specific spatial relationship to a pixel with the value j. By default, the spatial relationship is defined as the pixel of interest and the pixel to its immediate right (horizontally adjacent), but you can specify other spatial relationships between the two pixels. Each element (i,j) in the resultant glcm is simply the sum of the number of times that the pixel with value i occurred in the specified spatial relationship to a pixel with value j in the input image. The number of gray levels in the image determines the size of the GLCM.

1) **Contrast**

Contrast measures the quantity of local changes in an image. It reflects the sensitivity of the textures in relation to changes in the intensity

$$Contrast = \sum_{n=0}^{N_g-1} n^2 \sum_{|i-j|=n} P_d(i,j)$$

2) **Correlation**

This feature measures how correlated a pixel is to its neighborhood. It is the measure of gray tone dependencies in the image.

$$Correlation = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-\mu_i)P_d(i,j)}{\sigma_i\sigma_j}$$

3) **Energy**

Energy also means uniformity, or angular second moment (ASM). The more homogeneous the image is, the larger the value. When energy equals to 1, the image is believed to be a constant image

$$Energy = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} P_d^2(i,j)$$

.4)

Entropy

Entropy is a measure of randomness of intensity image.

$$Entropy = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} P_d(i,j)\log(P_d(i,j))$$

Extraction of statistical moments

We can find many dissimilar aspects of image processing, unchangeable pattern recognition and image encoding pose estimation for which moments can be applied. Statistical moment helps in describing the image content (or distribution) with respect to axis

Statistical moments	Formula
1.Mean	μ
2.Variance	$\sigma = E(x-\mu)^\wedge 2$
3.Standard Deviation	$SD = \sqrt{E(x-\mu)^\wedge 2}$
4.Smoothing index	$smi = 1 - \frac{1}{(1+\sigma^2)}$
5.Skewness	$Skewness = \frac{(X-\mu)^3}{\Sigma}$
6.Kurtosis	$k = n\frac{\sum_{i=1}^{n}(Xi - Xavg)}{(\sum_{i=1}^{n}(Xi-Xavg)^2)^2} - 3$

Table 1: Statistical Parameter

Ensemble Learning:

In statistics and machine learning, ensemble methods use multiple learning algorithms to obtain better predictive. Statistical ensemble in statistical mechanics, which is usually infinite, a machine learning ensemble consists of only a concrete finite set of alternative models, but typically allows for much more flexible structure to exist among those alternatives.

Supervised learning algorithms are most commonly described as performing the task of searching through a hypothesis space to find a suitable hypothesis that will make good predictions with a particular problem. Even if the hypothesis space contains hypotheses that are very well-suited for a particular problem, it may be very difficult to find a good one. Ensembles combine multiple hypotheses to form a better hypothesis. The term ensemble is usually reserved for methods that generate multiple hypotheses using the same base learner.

Classifier:

In the training phase of the input images taken from a given location to extract input features and the known output will be found by naming the images from the type of hemorrhage. Then the net file can be generated using a train tool for the first time after going through few testing iterations by providing the saved input and output files. Once the input features are calculated and the vector is created, to add the image to train, the output will be defined according to the value that has been received as the output result. Once the input and output files are saved, system can be trained with them. This logic can be used to train the tested images as well. We have using machine learning algorithm and ensemble learning algorithm

Random forest algorithm:

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of over fitting to their training set. This describes a method of building a forest of uncorrelated trees using a CART like procedure, combined with randomized node optimization and bagging. In addition, this combines several ingredients, some previously known and some novel, which form the basis of the modern practice of random forests, in particular:

1. Using out-of-bag error as an estimate of the generalization error.
2. Measuring variable importance through permutation.

Detection

.The Image segmentation technique is actually used to separate the grains components. The field of mathematical morphology contributes to a wide range of operators to image processing. we are only handling binary images. For a binary image, black pixels ("0") are normally taken to represent background regions, while white pixels ("1") denote foreground. The two most basic operations in mathematical morphology are dilation and erosion. These operations can be considered as morphological non-linear filters. The segmented image undergoes a series of morphological operations to detect the exact shape of the object.

Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. Morphological operations are affecting the form, structure or shape of an object, applied on binary images (black & white images – Images with only 2 colors: black and white).

Flowchart:

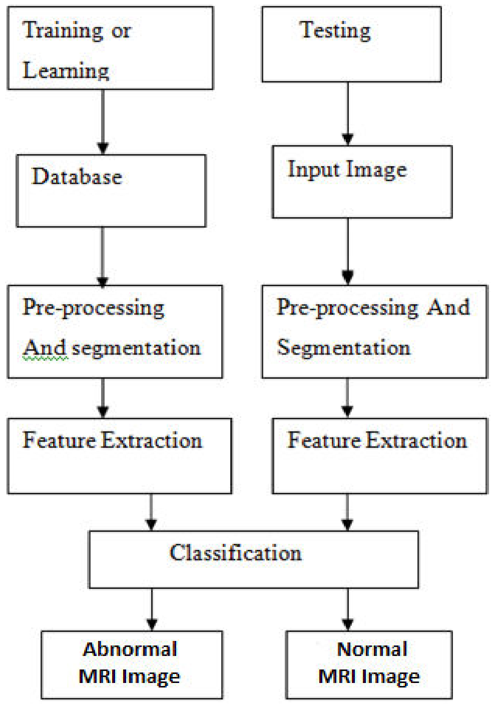


Fig 4.2: Flowchart of Training and Testing

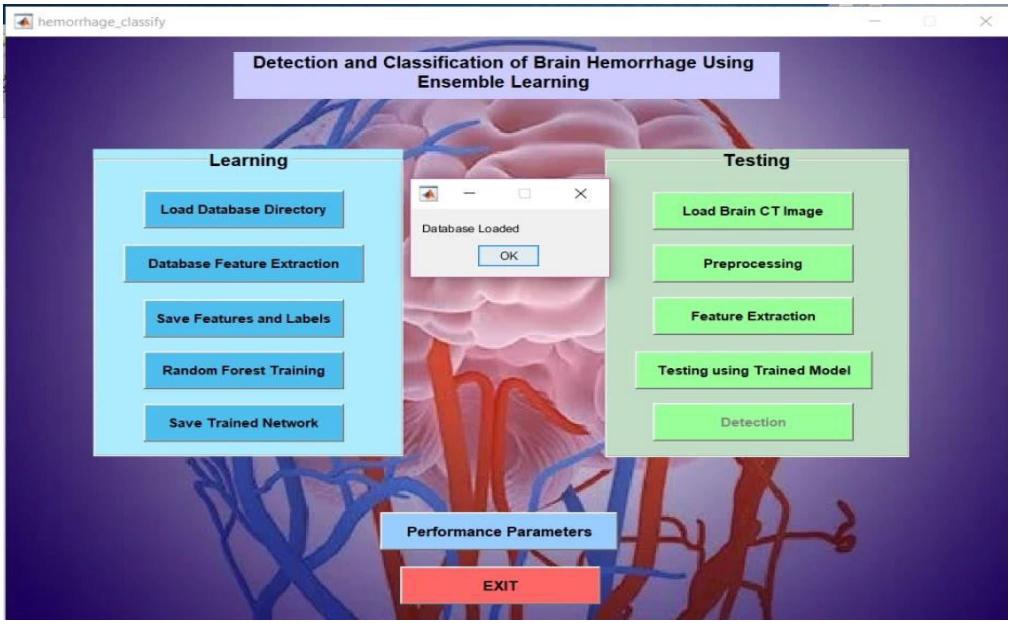
RESULT

The proposed technique has been implemented by using the MATLAB R2016a environment on Core i5, 4GB RAM with processor speed of 1.60GHz.our experiment are focused on brain MRI and CT images(**GUI**). Our database contains hemorrhage, calcification, of different location. . Segmentation phase of the proposed system precisely extracts the calcifies and haemorrhage region from abnormal brain images to certain extent. The output of segmentation becomes the input of the classification phase which classifies the hemorrhage. Effectiveness or correctness of the classifier process is analyzed on the basis of error rate

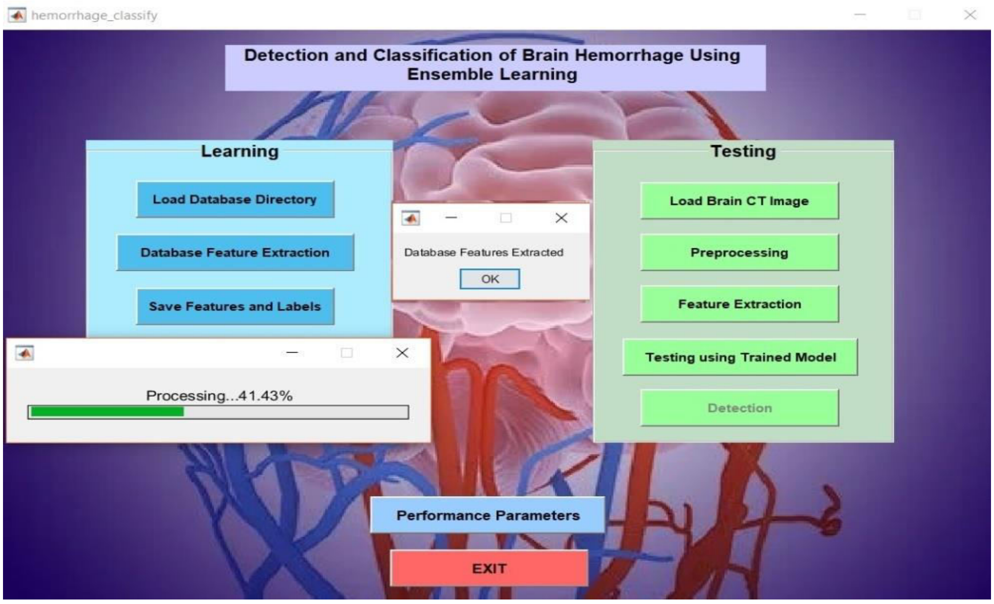
International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 09 | Sept -2020 ISSN: 2582-3930

Training

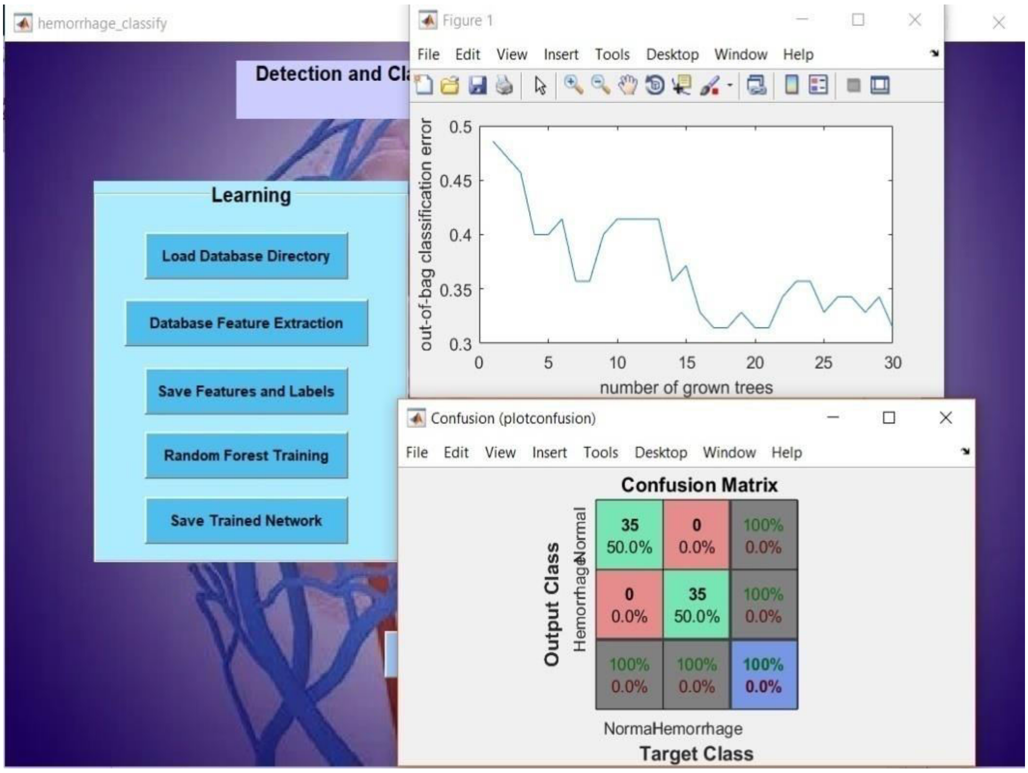
1) **Data load directory:**



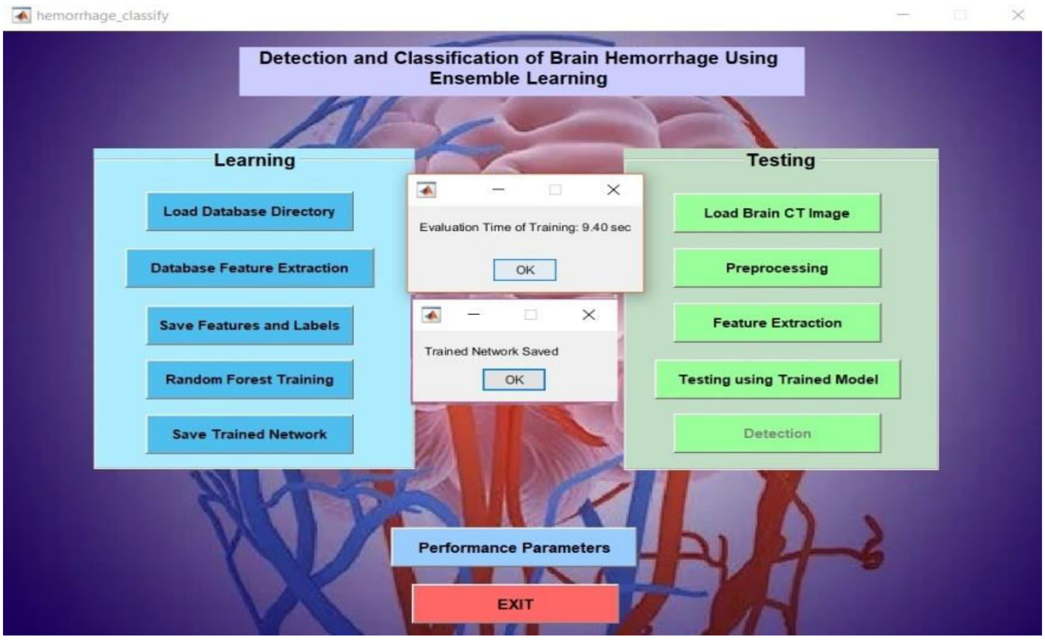
2) **Feature extraction**:



3) Random forest:

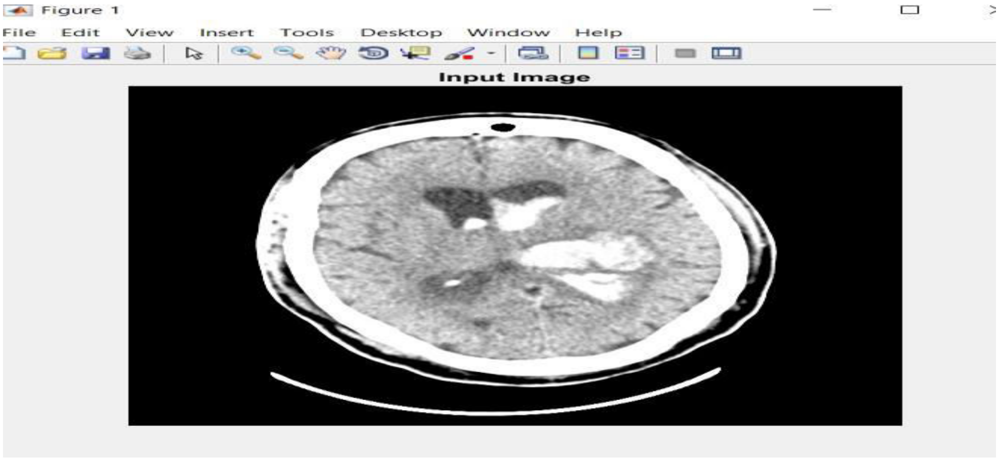


4) Save train model:

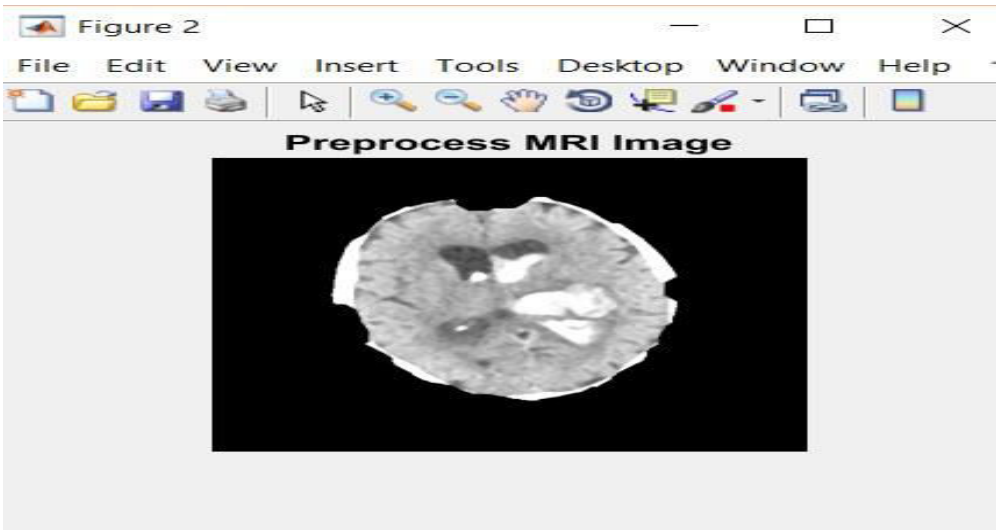


Testing:

1) **Input data :**

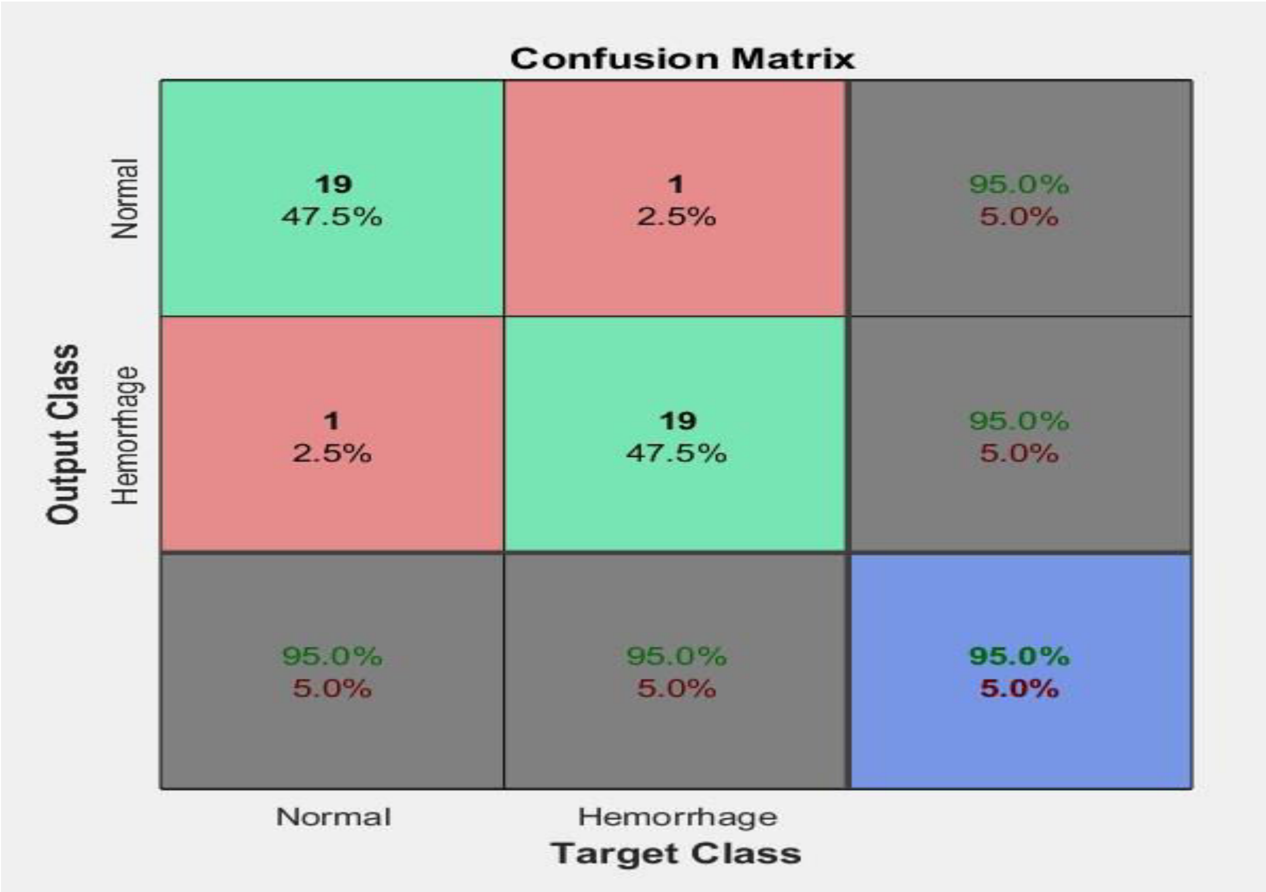


2) **Preprocess**

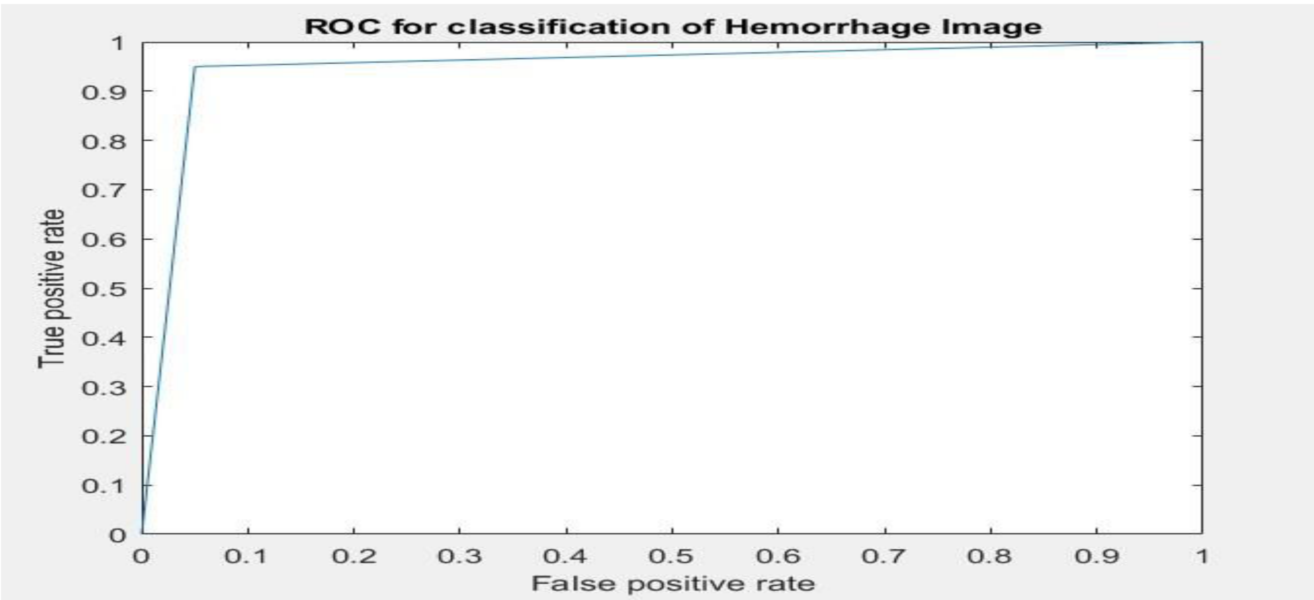


3) **Detection**

Confusion Matrix



ROC



Parameter	Value
TP	19
FN	1
TN	19
FP	1

Table 2: Confusion Matrix Parameter

Parameter	Value
Accuracy	95
Error rate	5
Sensitivity	95
Specificity	95
f-score	95
Positive predictive rate	95
Negative perdicitive rate	95
False positive rate	5
False negative rate	5
Rate of negative prediction	50
Rate of positive prediction	50

Table 3: Statistical Parameter

CONCLUSION

In this paper, we present a method to detect and classify the type of brain hemorrhage from given CT image. This method consists of several stages: image preprocessing, image segmentation, feature extraction and classification. The supervised algorithm is used to segment the image. Various diagnosis techniques for brain hemorrhage were invented. Some of them required high segmentation, noise removal, accuracy, etc. In this project, these problems are overcome by using machine learning. The various features using the computing techniques have been detected with their advantages and limitations and hence it should be provides a good framework for development of emerging medical systems, enabling the better delivery of healthcare with cost effectiveness.

REFERENCES

1.C.Dheeba1, S.Vidhya2 " A Survey on Prediction of Brain Hemorrhage Using Various Techniques ".Indian Journal of Innovations and Developments Vol 5 (6), June, 2016.

2.Rupali Mahajan1, Dr. P. M. Mahajan2 On Diagnosis Of Brain Hemorrhage By Using Artificial Neural Network International Journal of Scientific Research Engineering & Technology (IJSRET), ISSN 2278 – 0882 Volume 5, Issue 6, June 2016

3. Alyaa Hussein Ali, Shahad Imad Abdulsalam, Ihssan Subhi Nema Detection and Segmentation of Hemorrhage Stroke using Textural Analysis on Brain CT Images International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-5 Issue-1, March 2015.

4.M. Chawla, "A method for automatic detection and classification of stroke from brain CT images", 31st Annual International Conference of the IEEE EMBS, Minneapolis, Minnesota, USA, September 2009.

5. Ganesan, R., Radhakrishnan, S. [2009]. Segmentation of Computed Tomography Brain Images Using Genetic Algorithm. International Journal of Soft Computing, 4[4], 157-161

6.Myat Mon Kyaw, [2013]. Computer Aided Detection system For Hemorrhage Contained Region. International Journal of Computational Science and Information Technology (IJCSITY), Vol.1, No.1, February 2013.

7.C. Amutha Devi and Dr. S. P. Rajagopalan [2013]. Brain Stroke Classification Based on Multilayer Perceptron Using Watershed Segment and Gabor Filter. Journal of Theoretical and Applied Information Technology, ISSN: 1992-8645.

8. Nita Kakhandaki, S. B. Kulkarni "A Novel Framework for Detection and Classification of Brain Hemorrhage". International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7 Issue-4, November 2018.

9. Machine Learning For Absolute Beginners: A Plain English Introduction (Second Edition) by Oliver Theobald (Author)

10. Yeu-Sheng Tyan, Ming-Chi Wu,1,2 Chiun-Li Chin,4 Yu-Liang Kuo,2,3 Ming-Sian Lee,4 and Hao-Yan Chang4 "Ischemic Stroke Detection System with a Computer-Aided Diagnostic Ability Using an Unsupervised Feature Perception Enhancement Method" Hindawi Publishing Corporation International

Journal of Biomedical Imaging Volume 2014, Article ID 947539.

11. "https://youtu.be/loNcrMjYh64" For "Random Forest Algorithm".

12.Vincy Davis1, Dr. Satish Devane2 Diagnosis & Classification of Brain Hemorrhage International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391