

# SUPERVISED LEARNING APPROACH FOR BRAIN STROKE CLASSIFICATION USING DEEP LEARNING TECHNIQUES

S.Keerthana<sup>1</sup>, K.Sathiyakumari<sup>2</sup>

<sup>1</sup>Research Scholar, Computer Science, PSGR Krishnammal College for Women, Coimbatore, India

<sup>2</sup>Assistant Professor, Computer Science, PSGR Krishnammal College for Women, Coimbatore, India

## Abstract

This research presents the detection and segmentation of brain stroke using fuzzy c-means clustering and H2O deep learning algorithms. Brain stroke segmentation in magnetic resonance imaging (MRI) has become an evolving research area in the field of a medical imaging system. Brain stroke detection helps in finding the exact size, shape, extraction and location of the stroke. The system is consisting of three stages to detect and segment a brain stroke. An efficient algorithm is proposed for stroke detection based on segmentation and preprocessing techniques. The firstly quality of a scanned image is enhanced and then preprocessing techniques are applied to detect the stroke in the scanned image. In this system film artifacts removal, skull extraction and filtering methods are used to enhance the image. The second stage preprocessed image is segmented using fuzzy c-means clustering to obtain stroke region and edges are detected for accurate prediction of stroke location. After that edge detection operator is applied for boundary extraction and to find the size of the stroke, which helps doctors to make a decision about stroke location, size and etc. Finally, H2O deep learning method is used to classify stroke based on texture features and statistical features. The experimental result shows that the proposed work is performed well in detecting brain stroke efficiently.

**Keywords:** Key MRI, H2O, Deep Learning, Hemorrhage, Ischemic, GRLM, FOH and etc.,

\*\*\*

## 1. INTRODUCTION

Digital image processing is a massive field which can be using various applications for various researchers. It includes detection of the criminal face, figure and palmprint authentication system, in the medical field, object recognition, pattern matching, disease identification etc. Brain stroke detection plays an important role in the medical field. Brain stroke detection is the detection of stroke affected part in the brain along with its shape, boundary and size, so it helpful in the medical field to make the correct decision at corrects time.

Recent techniques accomplished in researches for detection of brain stroke can be broadly classified as

1. Histogram based method for stroke detection.
2. The morphological operation is applied to MRI images of the brain to locate brain stroke.
3. Edge base segmentation and color base segmentation of MRI image.
4. K-mean clustering algorithm.

Brain stroke detection can be done in both gray as well as color image researches in the field are still going on but the remarkable result is not achieved until now. Accurate prediction in brain diagnosis is quite difficult because of diverse shapes, sizes and appearances and location of the stroke. Stroke causes a blood clot in the artery and bleeding inside the brain may lead to die, so accurate detection of brain stroke is an important one in a medical field. In this research work, the detection and classification of brain stroke using various algorithms are carried out.

The work is organized as follows; Section II gives methodologies used in this work. Section III describes proposed algorithm and section IV shown experiment and its result in brief. Finally, section V explains the conclusion and future work.

## 2. INTRODUCTION

MRI images need preprocessing operations because of unorganized nature of the brain tissues that is why preprocessing methods are applied for the diagnosis of the infraction. Data collection is the most important part in every work. In this research, the data are collected from reputed hospital in Coimbatore. For each patient, the MRI scan images are captured with the particular time period. Every patient has 34 images in the specific time interval with various pixel levels. The image has a high resolution of 1105×649 pixels, which means each size is around 81KB. So the image is resized to 256×256 pixels for further proceedings.

### 2.1 Film Artifacts Removal

MRI brain images have film artifacts or labels like patient's name, age, date, time, comment and etc. These labels are removed using the `roifill()` function in Matlab. The following figure (Fig1) shows the original image and film artifacts removed image.

### 2.2 Skull Extraction

The removal of the bony skull surrounding the brain tissue is considered as a challenge to the brain seclusion. This process will allow extending the segmentation of the stroke.

The `bwareopen()`, `imfill()` and `imerode()` Matlab methods and mathematical operations are used to perform the skull removal. The figure (Fig1) shows the skull extracted the image.

### 2.3 Weighted Median filtering

To remove this impulse noise Median filter is used. It is less sensitive compared to outliers as the value of each pixel is obtained by the median of the neighboring pixel. The below figure (Fig1) shown the result of weighted median filter MRI image.

### 2.4 Segmentation using Fuzzy C-means Clustering

Segmentation refers to the process of separate a digital image into several segments. Image segmentation is logically used to locate objects and boundaries in an image. It is a process of assigning a label to every pixel in an image in which pixels with the same label share certain visual characteristics. It subdivides an image into its essential region or objects. The subdivision is carried depends on the problem being solved. In this work the second approach is used. The methodology used to segment brain stroke region is Fuzzy C-Means (FCM) clustering algorithm is used. The following figure (Fig1) illustrates the segmented stroke region and boundary extracted stroke region.

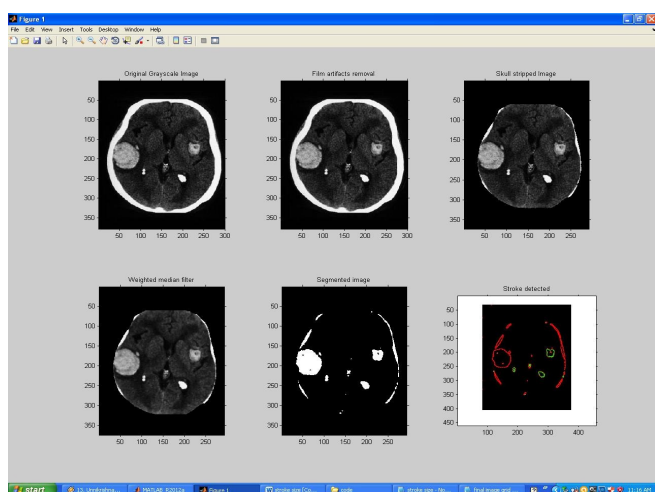


Fig -1: Stroke segmentation

## 3. PROPOSED ALGORITHM

The proposed work includes five different phases. The first and important phase is data collection; the MRI images of different patient's data are collected from reputed hospital lab in Coimbatore area for study purpose. The second and main phase of image processing are pre-processing the input image with various techniques like filtering and enhancement. The third phase is about the segmentation of brain stroke region from the original image using Fuzzy C-Means (FCM) algorithm. In the fourth phase, features are extracted using two different methods like First-Order Histogram and Gray-Level Run Length Matrix (GLRLM). Finally, H2O deep learning model is used to classify brain stroke diseases and shown classification accuracy for help doctors to know about the stroke impacts clearly and display

result image. The above figure (Fig2) illustrates the proposed framework for segmentation of brain stroke.

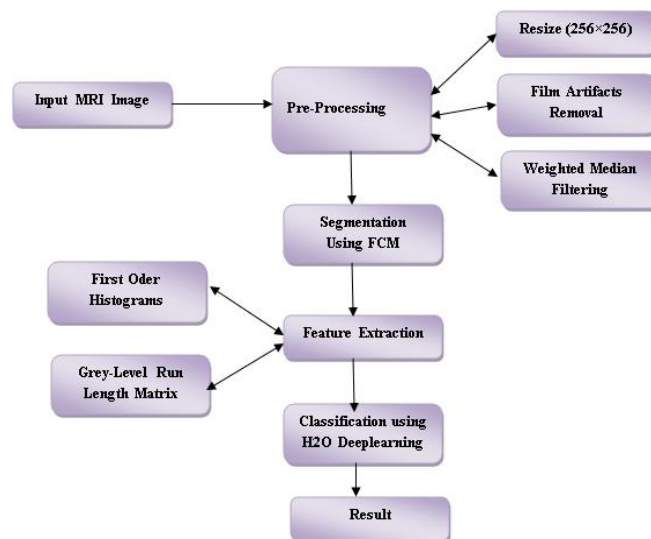


Fig-2: Proposed Framework

### 3.1 Feature Extraction

The feature extraction includes two different methods like First Order Histogram (FOH) features such as Mean, Standard Deviation, Energy, Entropy, Variance, Skewness, and Kurtosis. Another method is Gray level run length matrix (GLRLM) features like Short Run Emphasis (SRE), Long Run Emphasis (LRE), Run Length Nonuniformity (RLN), Low Gray-Level Run Emphasis (LGRE), Gray-Level Nonuniformity (GLN), Run Percentage (RP), High Gray-Level Run Emphasis (HGRE), Short Run High Gray-Level Emphasis (SRHGE), Long Run Low Gray-Level Emphasis (LRLGE), Short Run Low Gray-Level Emphasis (SRLGE), Long Run High Gray-Level Emphasis (LRHGE). Based on these feature vectors dataset is created for classification.

### 3.2 H2O

H2O is fast, scalable, open-source machine learning and deep learning for smarter applications. H2O's platform includes interfaces for Java, R, JSON, Python, Scala, and CffeeScript/ JavaScript, as well as a built-in web interface, Flow.

### 3.3 Deep Learning

Deep learning is known as hierarchical learning, deep structured learning or deep machine learning. It is a branch of machine learning based on a group of algorithms that attempt to model high-level generalizations of data by using the deep graph with multiple processing layers and composed of multiple linear, non-linear transformations methods.

Deep learning is part of machine learning methods based on learning a type of data. An image can be characterized in various ways such as a vector of intensity values per pixel, regions of a particular shape, or in a more abstract way as a

set of edges, etc. Some representations are better than others at shortening the learning task such as face recognition or facial expression recognition. One of the oaths of deep learning is replacing features with efficient algorithms for unsupervised or semi-supervised feature learning and hierarchical feature extraction techniques.

#### Algorithm: Brain Stroke Type Detection (BSTD)

**Procedure** BSTD ( $S, m, n, C, W$ ) where  $S$  is the dataset of  $m$  brain MRI images.  $n$  is the number of features.  $C$  is the set of classifiers to be tested along with their parameter vectors  $W$ .

$D \leftarrow \text{MAKEDATASET}(S, m, n)$

**for all**  $c \in C$  **do**

$A_c \leftarrow \text{CLASSIFIERTESTING}(D, c, cWC)$

**end for**

**return**  $c$  with best accuracy measure  $A_c$

**end procedure**

**Procedure** MAKEDATASET( $S, m, n$ ) where  $S$  is the dataset of  $m$  brain MRI images.  $n$  is the number of features.

Initialize  $D$  to an empty  $n$ -column table

**for**  $i \leftarrow 0, 9$  **do**

$t \leftarrow \text{EXTRACTFEATURES}(i, n)$

Append  $t$  into the end of  $D$

**end for**

**return**  $D$

**end procedure**

**Procedure** FEATURESEXTRACTION ( $i, n$ ) where  $i$  is the brain CT scan.

**If** the intensity of any pixel in  $i$  is  $\geq 250$ , it is set to 0

Apply preprocessing techniques on  $i$

Apply the fuzzy  $c$ -means segmentation step on  $i$

On the resulting image, apply MATLAB's first-order histogram in addition to gray level run length matrix code to compute the element feature vector  $v_i$  of the segmented boundary

Apply the classifier

**return**  $v_i$

**end procedure**

**Procedure** CLASSIFIERTESTING( $D, c, wc$ ) where  $D$  is the dataset of brain MRI scans.  $c$  is a classifier with parameter vector  $wc$ . The testing will be using the 10-fold cross-validation technique.

Randomly partition  $D$  into 10 subsets:  $D_0, \dots, D_9$

**for**  $j \leftarrow 0, 9$  **do**

$T_s \leftarrow D_j$  where  $T_s$  is the testing set.

where  $T_r$  is the training set.

**For** classification method  $c$ , build a model using parameters  $wc$  and  $T_r$  on the R tool

Use the model on  $T_s$  and compute the accuracy measures,  $A_j$ , (See chapter 4 for more details)

**return** the average of all  $A_j$ 's for  $j = 0, \dots, 9$

**end for**

**end procedure**

## 4. EXPERIMENTAL RESULT

The classification algorithms are implemented using a data mining tool called "R". R tool is used by academicians who incessantly provide libraries for new and developing statistical techniques. The H2O deep learning algorithm is implemented as classification task for accurate prediction of brain stroke disease based on different performance evaluation measures. The following figures Fig3 and Fig4 demonstrate the deep learning result and score history for training, validation and testing

```

RGui (64-bit) - [R Console]
File Edit View Misc Packages Windows Help

ACC: 1
Gini: 1

Confusion Matrix for F1-optimal threshold:
      HSM      ISC      Error Rate
HSM   16     0 0.000000 =0/16
ISC    0    22 0.000000 =0/22
Totals 16  22 0.000000 =0/38

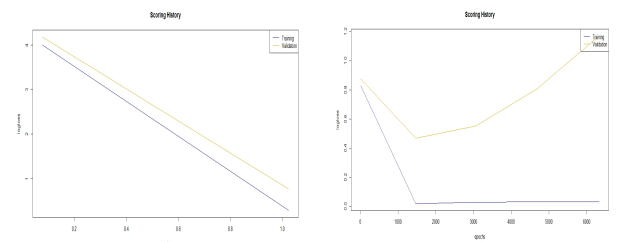
Maximum Metrics: Maximum metrics at their respective thresholds
metric threshold value idc
1 max f1 0.983246 1.000000 22
2 max f2 0.983246 1.000000 22
3 max fpoint5 0.983246 1.000000 22
4 max accuracy 0.983246 1.000000 22
5 max precision 0.999999 1.000000 0
6 max recall 0.983246 1.000000 22
7 max specificity 0.999999 1.000000 0
8 max absolute_MCC 0.983246 1.000000 22
9 max min_per_class_accuracy 0.983246 1.000000 22

Data(Life Table): Extract with 'h2o.gainlift(model, <data>)' or 'h2o.gainlift(model, valid=c('D', xval=c('D','D'))'
> pred <- h2o.predict(m3, test)
|=====| 100%
> pred
      predict      HSM      ISC
1 HSM 3.111532e-01 0.9998008
2 ISC 3.847482e-04 0.9996153
3 ISC 3.429402e-05 0.9998637
4 ISC 4.661410e-04 0.9993338
5 ISC 4.337720e-03 0.9994663
6 ISC 4.703707e-03 0.9995529

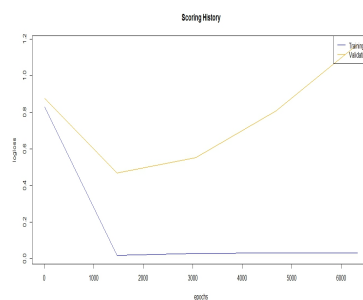
[38 rows x 3 columns]
> test$Accuracy <- pred$predict == test$class
> mean(test$Accuracy)
[1] 0.983246

```

Fig-3: Deep learning result



a) Score history for training      b) Score history for validation



c) Score history for testing

Fig4. Scoring chart for deep learning

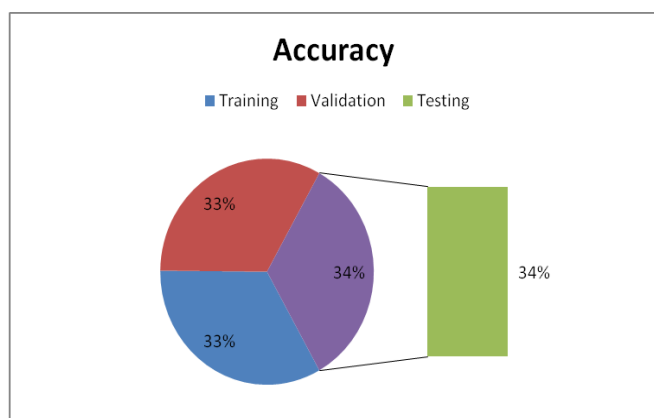
Table1 shows the performance evaluation measures of various deep learning in detail. The following figure (Chart-1) illustrates the chart comparison of deep learning classifier with performance analysis of kappa, sensitivity and specificity. Figures (Chart-2 & Chart-3) demonstrates the Accuracy comparison. In the below table, the evaluation measures like kappa, specificity, sensitivity and accuracy are given for each classifier.

**Table1.** Performance Measures for various Classifiers

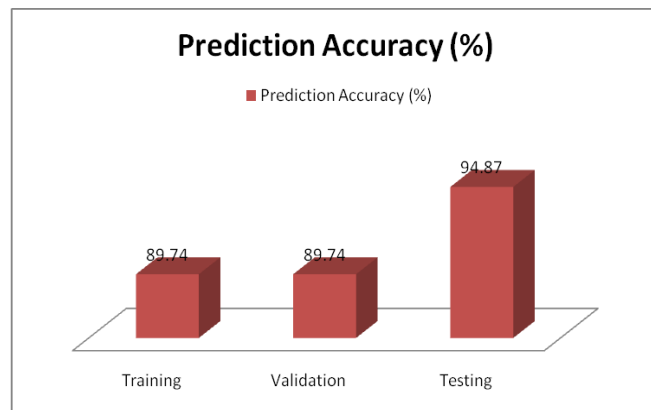
Deep learning / Performance Measure	Training	Validation	Testing
F-score	0.9896	1	0.9974
Precision	1	1	0.9999
Recall	1	1	0.9919
Specificity	1	1	0.9999
Log loss	0.2880	1.1721	0.2723
MSE	0.0621	0.0596	0.5729
Gini	0.9562	0.9375	0.9779
Accuracy	0.9781	0.9687	0.9889
Prediction Accuracy (%)	89.74	89.74	94.87



**Chart -1:** Performance Comparison Chart



**Chart-2:** Accuracy Comparison Chart



**Chart-3:** Prediction Accuracy Comparison Chart

### 5. CONCLUSION AND FUTURE WORK

This research work includes two different phases one is segmentation of stroke area and another one is the classification of stroke types. The procedure for the segmentation of MRI brain image is to detect brain stroke. The brain stroke segmentation process is complicated because it includes various type of process like filtering, enhancement and morphological operations. The Fuzzy C-Means clustering technique is used to segment out the stroke region from input MRI image in three different classes of clustering with different levels of thresholding. The experimental result shows that the proposed segmentation work was able to detect different types of a brain stroke in a different type of position accurately. Finally, the stroke size is calculated based on pixel intensity values to make clearly and effective decisions.

The second phase is about feature extraction and the classification. Feature extraction includes two different methods like First Order Histogram (FOH) features such as Mean, Standard Deviation, Energy, Entropy, Variance, Skewness, and Kurtosis. Another method is Gray level run length matrix (GLRLM) features like Short Run Emphasis (SRE), Long Run Emphasis (LRE), Run Length Nonuniformity (RLN), Low Gray-Level Run Emphasis (LGRE), Gray-Level Nonuniformity (GLN), Run Percentage (RP), High Gray-Level Run Emphasis (HGRE), Short Run High Gray-Level Emphasis (SRHGE), Long Run Low Gray-Level Emphasis (LRLGE), Short Run Low Gray-Level Emphasis (SRLGE), Long Run High Gray-Level Emphasis (LRHGE). Based on these feature values dataset was created. Finally, the H2O deep learning algorithm is used to classify brain stroke types. The experimental result shows that deep learning algorithm performs well in detection brain stroke with above 90% accurately.

In future work, different types of pre-processing techniques and segmentation techniques are used for detecting brain stroke also dataset is increased to find brain stroke subtypes. In this work classification algorithm is limited to deep learning but in future other classification algorithms are used to find stroke types.

**REFERENCES**

- [1]. B. Chanda, D. D. Majumder, Digital Image Processing and Analysis , PHI Learning Private Limited, New Delhi, Isbn 978 -81-203 -4325-2, 2nd Edition, 2011.
- [2]. Issac N. Bankman, "Handbook of medical image processing and analysis", Second edition, Academic press, USA, 2008.
- [3]. O'Brien, J.S., & Sampson, E.L. (1965). Lipid composition of the normal human brain: gray matter, white matter, and myelin. *Journal of lipid research*, 6(4), 537-544.
- [4]. A. Devi, S. P. Rajagopalan, "Brain Stroke Classification Based on Multi-Layer Perceptron Using Watershed Segmentation and Gabor Filter", *Journal of Theoretical and Applied Information Technology*, Vol. 1. 56 No. 2, PP: 410-416, Oct 2013.
- [5]. F. Zhu, "Brain Perfusion Imaging – Performance and Accuracy", Centre for Intelligent System and their Applications, School of Information, University of Edinburgh, 2012.
- [6]. Warburton E. Stroke Management. In Godlee F, ed. *Clinical Evidence*, BMJ Publishing Group Ltd., 2003.
- [7]. Montagu A, Reckless IP, Buchan AM. Stroke: management and prevention. *Medicine*. 2012;(40)9, 490-499.
- [8]. N. Aggarwal and R.K. Agrawal, "First and Second Order Statistics Features for Classification of Magnetic Resonance Brain Images", *Journal of Signal and Information Processing*, Vol. 3, 2012, PP. 146-153.
- [9]. Ma, J., Sheridan, R. P., Liaw, A., Dahl, G. E. & Svetnik, V. Deep neural nets as a method for quantitative structure-activity relationships. *J. Chem. Inf. Model.* 55, 263–274 (2015)
- [10]. Leung, M. K., Xiong, H. Y., Lee, L. J. & Frey, B. J. Deep learning of the tissue regulated splicing code. *Bioinformatics* 30, i121–i129 (2014).