Ischemic Brain Stroke Detection from MRI Image using Logistic Regression Classifier

Conference Paper · January 2021

DOI: 10.1109/ICREST51555.2021.9331090

5 authors, including:

Md.Shabuj Hossain
Pabna University of Science and Technology
1 PUBLICATION **3** CITATIONS
SEE PROFILE

Liton Chandra Paul
Pabna University of Science and Technology
122 PUBLICATIONS **1,050** CITATIONS
SEE PROFILE

Rezaul Azim
University of Chittagong
156 PUBLICATIONS **2,615** CITATIONS
SEE PROFILE

Abdulla Al Suman
Macquarie University
32 PUBLICATIONS **321** CITATIONS
SEE PROFILE

Ischemic Brain Stroke Detection from MRI Image using Logistic Regression Classifier

Md.Shabuj Hossain, Subir Saha
Dept. of CSE, Pabna University of Science and Technology, Pabna-6600, Bangladesh
shabujpust@gmail.com, ss.buet@gmail.com

Liton Chandra Paul
Dept. of EECE, Pabna University of Science and Technology, Pabna-6600, Bangladesh
litonpaulete@gmail.com

Rezaul Azim,
Dept. of Physics, University of Chittagong, Chittagong-4331, Bangladesh
rezaulazim@yahoo.com

Abdulla Al Suman
School of Engineering and Information Technology, UNSW, Canberra, ACT 2600, Australia
Abdulla.Suman@student.unsw.edu.au

***Abstract*— In this paper an efficient model for ischemic brain stroke detection from magnetic resonance imaging (MRI) using machine learning approach namely logistic regression classifier is proposed. The MRI images are pre-processed to reduce noise and converted into gray images. Then the stroke portions of the MRI gray images are segmented by using hue, saturation, and value (HSV) color threshold and the segmented stroke images are converted into binary images to reduce computational complexity. The stroke features namely mean hue, standard deviation, mean variance and area of affected lesion i.e. stroke portion have been extracted. Finally, logistic regression classifier is used to identify the classes of test image. The proposed model shows an accuracy of 96%, sensitivity of 92.3% and specificity of 100% for testing datasets.**

Index Terms—Brain stroke detection; MRI; Logistic regression classifier; HSV color space; Machine learning;

I. INTRODUCTION

Brain stroke occurs when blood circulation to the brain is either reduced or interrupted. When this happens, the brain does not get sufficient oxygen or others necessary ingredients and then the brain cell starts to die. It is also known as brain attack or cerebrovascular accident (CVA).Brain stroke is one of the leading causes of death all over the world. Worldwide approximately 15 million people suffer each year and 5.8 million people die from it. One in six persons would have stroke in their entire life [1]. There are mainly three types [2-3] of brain strokes namely ischemic strokes, hemorrhagic strokes and transient ischemic attack(TIA) which is also known as a warning or mini-stroke. Ischemic strokes are occurred because of arteries blocking or narrowing. Ischemic stroke is the one of the most case of stroke and almost 87% of strokes are ischemic strokes [3]. For this reason, treatment is needed immediately for restoring an adequate flow of blood to brain. Hemorrhagic strokes are happened because of blood leaking into the brain. In this case, control of bleedingis needed to decrease extra pressure from brain as soon as possible. Transient ischemic attack is an indicator of small stroke, which means the current of blood cannot reach properly to a particular part of brain for a while. For detecting stroke, mainly two types of invasive screening test like computed tomography (CT) or magnetic resonance imaging (MRI) is used.

Researchers all over the world are working for developing advanced method to predict and detect human ischemic brain stroke. For example, Laxmi et al., proposed a voxel based lesion segmentation through SVM classifier for effective brain stroke detection. They have compared with random forest algorithm. Their accuracy, Sensitivity and Specificity are 88%, 95%, and 66%, respectively [4]. In [5], Alhawaimil focused on segmentation of brain strokes image. He only discussed how to segment brain strokes part of MRI image. S. Keerthana et al., proposed a method on brain stroke segmentation using fuzzy C-means clustering. Their method focused on detection and extraction of brain strokes from different patients' MRI images [6].The author of [7] proposed a method on analyzing training information from random forests for improved image segmentation. He used random forest classifier and discussed about segmentation. Hossain et al., discussed about ulcer detection in image converted from video footage by new methods and tried to improve classification problem by using logistic regression classifier [8]. Unsupervised and supervised methods for lesion segmentation are presented in [9]. Rajini et al., discussed on computer aided detection of ischemic stroke using segmentation and texture features. They worked on CT images for strokes detection and used different classification algorithm for prediction [10]. A discussion on Ischemic stroke detection system using an unsupervised feature perception method has been presented in [11]. The paper [12] presents a segmentation and extraction of brain injury lesions from MRI images. Their objective is to create a segmentation program to extract an injured area in the brain using minimal user contribution. KFCM algorithm has been using for brain stroke detection in [13]. Authors have been implemented Kernelized fuzzy C-means clustering and adaptive threshold for stroke segmentation and precise detection. They obtained an accuracy of 93%.Rajendra et al., presented an ischemic stroke detection using higher order spectra features in brain MRI images. Their developed technique effectively detected the stroke lesion [14]. Using expectation maximization and random forest classifier, segmentation and classification of brain stoke have been presented by the authors of [15]. They implemented random forest classifier which provided an accuracy of 93.4%.

Although many research works have already been done on the detection of human brain stroke, till now many researchers of different parts of the world are doing research on the same topic in order to make precise prediction and more accurate detection in smarter as well as easier way. In this paper, we are intended to develop a model for brain stroke detection using magnetic resonance imaging. The proposed model shows an accuracy of 96%, sensitivity of 92.3% and specificity of 100%.The rest of the paper has been organized as follows: the proposed model has been described step by step with necessary diagrams and data tables in section II. In section III, the results and performance of the model has been analysed briefly and also compared with the performance of some similar works published by other researchers. Finally concluding remarks have been pointed out in section IV.

II. PROPOSED MODEL

An enhanced model for human brain stroke detection from MRI images using digital image processing has been presented and discussed in detail in this section. The flow chart of the proposed model is shown in Fig.1. The proposed detection process has been subdivided into seven consecutive steps. These steps are named as MRI images acquisition, MRI images pre-processing, segmentation using HSV color threshold, conversion of HSV color threshold images to binary images, features extraction from binary images and logistic regression classifier.

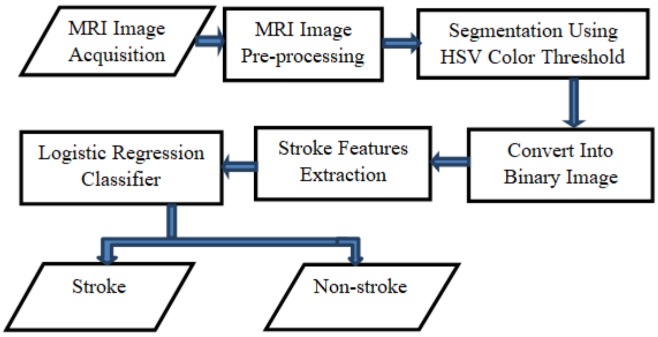
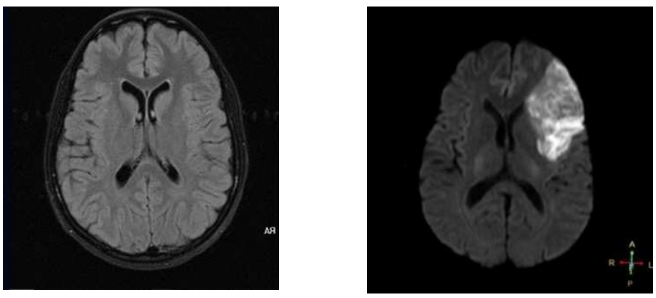


Fig.1. Block diagram of proposed brain stroke detection model

A. MRI Image Acquisition

Image acquisition means capturing photographic image of an object or physical scene which is the first step of any digital image processing system. A Physician may take the brain image by using either magnetic resonance imaging (MRI) or computed tomography (CT) scan.MRI is a type of medical imaging process that uses strong magnetic fields and radio waves to produce detailed images of organs and tissues inside human body. CT scan is another medical imaging process which can also produce image of organs, bones and tissues inside body but it uses a combination of X-ray measurement taken from different angles to create the image.



(a) Non-stroke image (b) Stroke image

Fig.2. MRI images of human brain: (a) Non-stroke image, (b) Stroke image

In this work, MRI is opted for image acquisition because MRI may detect more subtle change in the content of brain tissue than CT scan. Samples of a MRI non-stroke image and a MRI stroke image of human brain are shown in Fig.2.

B. MRI Image Pre-processing

At this stage, the MRI images are pre-processed by using digital image processing techniques. Firstly, noise has been reduced from the MRI images. Then, the RGB images are converted to gray images.

C. Segmentation using HSV Color Threshold

After pre-processing, HSV color threshold is used to make segmentation the portions of brain stroke of the MRI stroke gray images. It helps to remove the unnecessary portion from the brain stroke MRI gray images that fall within a specified color range. It also helps to detect the objects of dominant color values. In this case, unnecessary portion means the non-stroke portion of MRI image and the segmented portion indicates the portion of the brain stroke. HSV color space is the indicator of hue, saturation and brightness at a specified range which detects and segments the stroke portions from MRI stroke images. A segmented portion of brain stroke and non strokes by using HSV color space is shown in Fig. 3 and Fig. 4 respectively. The histograms of hue, saturation and brightness of the segmented stoke portion of brain are shown in Fig. 5.

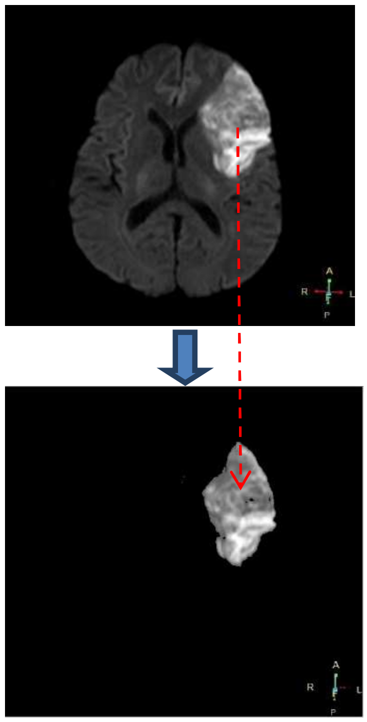


Fig.3. Segmented portion of brain stroke by HSV color threshold

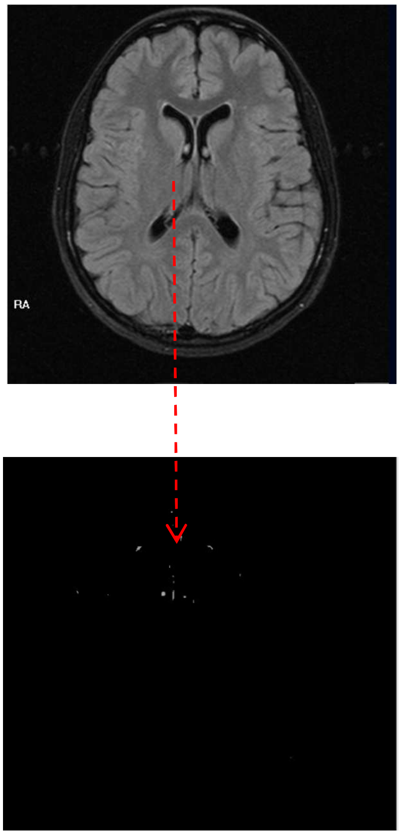


Fig.4. Segmented portion of non-stroke by HSV color threshold

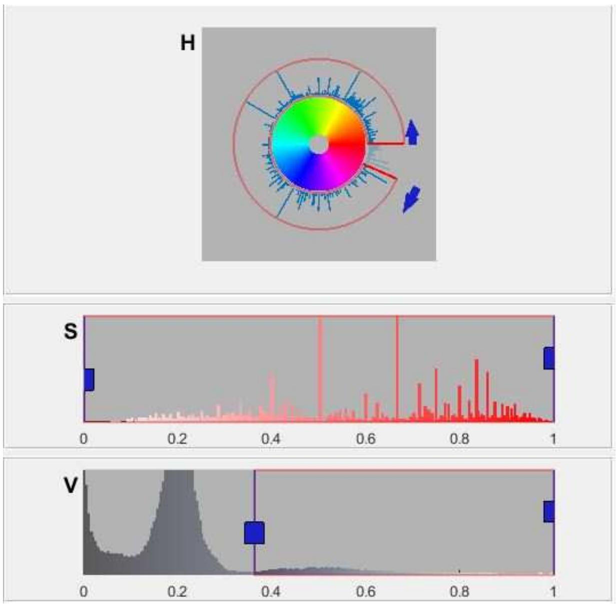


Fig. 5. HSV color threshold histograms

D. Conversion of HSV Color Threshold Image to Binary Image

For further calculation and features extraction, the segmented HSV color threshold images have been converted into binary image as shown in Fig. 6. Binary image is such an image which consists of pixels that can have one of exactly two colors, usually either black or white. The white and black pixels of the image are represented by 1 and 0 respectively. The binary image helps to reduce the volume of calculation.

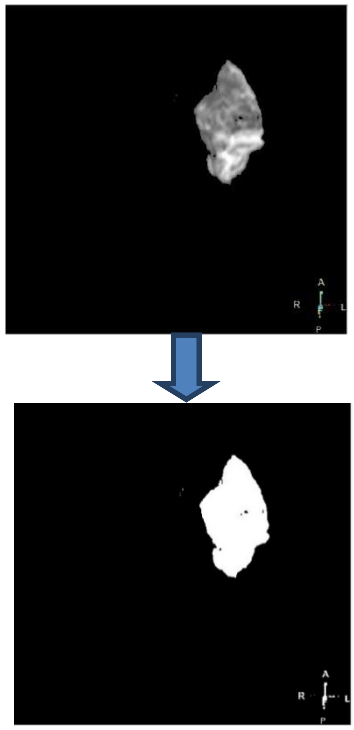


Fig. 6. Conversion of HSV color threshold image to binary image

E. Features Extraction from Binary Image

In image processing, features extraction is a process to select some features or attributes which helps to detect something specified. It reduces the dimension of data by building derived values i.e. features intended to be non-redundant and informative. In this work, features are extracted from MRI images to detect brain strokes. The stroke features namely mean hue, standard deviation, mean variance and area of stroke portion of the MRI images have been calculated. Here, the mean hue indicates the mean of dominant wavelengths of color pixels of stroke portions which differentiates the stroke portion from the other unwanted colors. The standard deviation is used to determine the amount of variation or dispersion level of a set of pixels. The less amount variation of pixels indicates that the pixels are close to stroke portion of an image which helps to detect the brain stroke easily. Here, mean variance is used to measure how far a set of random pixels are spread out from their average pixel value. It is used to determine a value which indicates how it is near about stroke pixels. Finally, the area of affected part by stroke is the most important part for the detection of brain stroke. It indicates how much area is affected by stroke which also assists to know the level of stroke that means whether is it ischemic or mini stroke. For this reason, the affected area has been calculated from MRI segmented images. The MRI images that contain any stroke portion will have a particular numerical value. However, MRI images without any strokes portion will contain near about zero numerical value. Extracted features' values of MRI stroke images and non stroke images of some training images are given in following Table I.

TABLE I EXTRACTED FEATURES FOR SOME TRAINING IMAGES

Case	Mean Hue	Standard Deviation	Mean Variance	Stroke Area
STROKE	0.025207860900	0.156755939801	0.022621070	10073.6
	0.061892149400	0.240959563600	0.047404517	67624.2
	0.070899056201	0.256656151400	0.053516539	85108.3
	0.069648014400	0.2545540283302	0.051865019	7758.7
	0.0363254738702	0.1870987274800	0.029752326	39370.2
	0.0325190355300	0.1773757923500	0.028800885	1670.2
	0.0514371179600	0.2208890568203	0.042641305	3988.3
	0.0394231417200	0.1946016657100	0.032138241	1769.6
	0.6336657801418	0.4818076422545	0.160160938	28615.2
	0.0657808857808	0.2479006168891	0.049979470	4272.5
	0.0681576624791	0.2520167336815	0.057877785	18288.7
	0.0685683139534	0.2527201493030	0.054388524	5674.8
	0.0383844104950	0.1921232157818	0.032956853	5041.5
	0.1127934921190	0.3163419816326	0.087302236	10084.3
	0.0734082830857	0.2608054966064	0.055326388	72307.1
	0.0300630735658	0.1707609005998	0.026099200	40267.8
	0.0438108330173	0.2046739942631	0.034278911	55501.6
	0.0408719135802	0.1979934348925	0.033480005	47709.7
	0.0319866588517	0.1759645205922	0.026605620	45886.3
	0.0220671929502	0.1469025663772	0.018681895	3619.12
	0.0727658071352	0.2597536092134	0.062608997	5013
NON STROKE	0	0	0	0
	0.0000131578947	0.0036273693182	0.00001.314	2
	0	0	0	0
	0	0	0	0
	0	0	0	0
	0	0	0	0
	0.0006117724867	0.0247264680291	0.000608005	530
	0.0000537240537	0.0073294725219	0.000053204	44.50
	0	0	0	0
	0	0	0	0
	0	0	0	0
	0.0003544305539	0.0188229894799	0.000348726	418.87
	0	0	0	0
	0.0001804092713	0.0134304401943	0.000177275	170.12
	0.0005863628748	0.0242079994004	0.000577817	43.25
	0	0	0	0
	0.0000402762953	0.0063462489686	0.000040236	8
	0.001181957570	0.034359835090	0.001168880	38.50
	0	0	0	0

F. Logistic Regression Classifier

A machine learning technique namely logistic regression classifier has been used to classify the image as either brain stroke or non-stroke. Logistic regression classifier is a classification algorithm which uses complex cost function and it works with sigmoid function. It can be used for binomial, ordinal or multinomial purposes. Logistic regression classifier can be used for different pattern recognition and different field of statistics. The aforementioned four features of measurement variables have been extracted for the classification. In this work, 5-fold cross validation technique has been used to get the result. After classification of the data, brain stroke and non-strokes images are identified properly.

III. EXPERIMENTS AND RESULT ANALYSIS

A. Dataset Description

Total 65 MRI images have been used in this work. Among 65 MRI images, 40 MRI images are used for the training purpose of model and 25 MRI images are used for the testing purpose. The type of images is .jpg. All the MRI images for both training dataset and testing dataset are collected from different online resources [16-18]. We would like to express our gratitude to owners of the resources.

B. Performance Analysis

A well-known machine learning classifier namely logistic regression classifier has been used in this work to classify stroke and non-stroke cases. The performance of logistic regression classifier for the MRI images is measured by accuracy, specificity and sensitivity which are extensively used to evaluate performances of any classification technique. Another important thing needs to mention that 5-fold cross validation is appliedto obtain the decision or result.The result is obtained from confusion matrix. Approximately 61.53% of the total MRI images are used as the training images. An accuracy of 97.7%, sensitivity of 100% and specificity of 95% are obtained for the training dataset. Then the rest of approximately 38.47% of total MRI images are used as test dataset to validate the result of our proposed brain stroke detection model. For the test dataset, the proposed human ischemic brain stroke detection model shows an accuracy of 96%, sensitivity of 92.3% and specificity of 100%.To ensure better accuracy, sensitivity and specificity, ten times of calculation result of confusion matrix on different images datasets with different numbers of fold and validation have been taken. Then the average value has taken form all ten observations which helps us to find improved results of the proposed system. The formula of accuracy, sensitivity and specificity are given below:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (3)$$

where TP = True positive, TN= True negative, FP= False positive and FN= False negative.

Here, the accuracy indicates how well the model detects both stroke and non-stroke MRI images. It also means the overall accuracy of the proposed model. The sensitivity indicates how well the model detects brain ischemic strokes while specificity indicates how well the model detects non-strokes. The results of logistic regression classifier obtained from the validation phase are summarized in the Table II. A comparison of performance of the proposed human brain ischemic stroke detection model in terms of accuracy, sensitivity and specificity using logistic regression classifier with some published works has been shown in Table III. The proposed model using logistic regression classifier shows very good performance compared to some reference works.

TABLE II ESTIMATED PERFORMANCE PARAMETERS OF THE PROPOSED MODEL USING LOGISTIC REGRESSION CLASSIFIER

Parameters	Values
Accuracy	96 ±0.40%
Sensitivity	92.3 ± 0.50%
Specificity	100%

TABLE III COMPARISON WITH SIMILAR WORKS

Parameters	SVM [4]	Random Forest [15]	Proposed Model (Logistic Regression)
Accuracy	88%	93.4%	96%
Sensitivity	95%	92.5%	92.3%
Specificity	66%	95.8%	100%

IV. CONCLUSION

In this work, an efficient model for human ischemic brain stroke detection from MRI images using machine learning algorithm namely logistic regression classifier is proposed. HSV color space is applied to detect and make segmentation the stroke portion of the MRI images. Hue, saturation and brightness of the segmented stoke portion are calculated. Then the segmented MRI images have been converted into binary images to reduce calculation size and extract four features from the segmented stroke portion. The four stroke features namely mean hue, standard deviation, mean variance and area of stroke portion of the MRI images have been extracted from the segmented stroke portion. Finally, the analysis of brain stroke features have been prepared for the logistic regression classifier that deals with the best demonstration of strokes and non-strokes images. For the test dataset, the proposed human ischemic brain stroke detection model shows an accuracy of 96%, sensitivity of 92.3% and specificity of 100%. In order to enhance the performance of the proposed systema 5-fold cross validation is applied to obtain the decision or result. The estimated performance parameters of the proposed detection system are very encouraging compared to some other published works.

REFERENCES

[1] www.world-stroke.org (Accessed on November 1, 2020)

[2] www.medicalnewstoday.com (Accessed on November 1, 2020)

[3] www.cdc.gov/stroke/types_of_stroke.htm (Accessed on November 1, 2020)

[4] R. P. Lakshmi, M. S. Babu, and V. Vijayalakshmi, “Voxel based lesion segmentation through SVM classifier for effective brain stroke detection,” Proceedings of International Conference on Wireless Communications, Signal Processing and Networking,22-24 March 2017, Chennai, India, pp. 1064-1067.

[5] A. R. Alhawaimil, “Segmentation of Brain Strokes image,” International Journal of Advanced Research in Computer and Communication Engineering, vol. 4, no. 9, pp.375-378,2015.

[6] S. Keerthana and K. Sathiyakumari, “Brain Strokes Prediction using fuzzy C-means Clustering” International Journal of Computer Applications, vol. 154, no. 4 , pp. 26-30, 2016.

[7] D.Mahapatra, “Analyzing training information from random forests for improved image segmentation,” IEEE Transactions on Image Processing, vol. 23, no. 4,pp. 1504 - 1512,2014.

[8] A. A. Mamun and M. S. Hossain, “Ulcer Detection in Image Converted from Video Footage of Wireless Capsule Endoscopy,” Proceedings of 1st International Conference on Advances in Science, Engineering and Robotics Technology, 3-5 May 2019, Dhaka, Bangladesh, pp. 1-4.

[9] T.Jerman, A.Galimzianova, F.Pernuš, B.Likar, and Z.Spiclin, “Combining unsupervised and supervised methods for lesion segmentation,” BrainLes 2015, Lecture Notes in Computer Science, vol. 9556, Springer, Cham, vol. 9, pp. 45–56, 2015.

[10] N. Rajini and R.Bhavani, “Computer aided detection of ischemic stroke using segmentation and texture features,” Measurement, vol. 46, no. 6, pp. 1865–1874, 2013.

[11] Y.-S. Tyan, M.-C. Wu, C.-L. Chin, Y.-L. Kuo, M.-S. Lee, and H.-Y. Chang,“Ischemic stroke detection system with a computer-aided diagnostic ability using an unsupervised feature perception enhancement method,” International journal of biomedical imaging, vol. 2014, pp. 1-12, 2014.

[12] R. Noureddine, K. Tarhini, and S. Saleh, “Segmentation and extraction of brain injury lesions from MRI images: Matlab implementation,” Proceedings of International Conference on Advances in Biomedical Engineering, 16-18 September 2015, Beirut, Lebanon, pp. 45-48.

[13] V. Vijayalakshmi, M. Sunil Babu, and R. P. Lakshmi, “KFCM Algorithm for Effective Brain Stroke Detection through SVM Classifier,” Proceedings of IEEE International Conference on System, Computation, Automation and Networking, 6-7 July 2018, Pondicherry, India, pp. 1-6.

[14] U. R.Acharya, K. M. Meiburger, O. Faust, J. E.W.Koh, S. L. Oh, E. J. Ciaccio, A.Subudhi, V. Jahmunah, and S.Sabut, “Automatic detection of ischemic stroke using higher order spectra features in brain MRI images,” Cognitive Systems Research, vol. 58,pp. 134-142,2019.

[15] A. Subudhi, M. Dash, and S. Sabut, “Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier,” Biocybernetics and Biomedical Engineering, vol. 40, no. 1, pp. 277-289, 2020.

[16] www.kaggle.com (Accessed on November 1, 2020)

[17] www.radiopaedia.org (Accessed on November 1, 2020)

[18] www.medscape.com (Accessed on November 1, 2020)