An Efficient Model For Brain Stroke Detection Using Machine Learning Techniques

SINDHUJA.V MYTHILY.J SHOBANA PRIYA.H
Dr. K.VALARMATHI, M.E., Ph.D
Panimalar Engineering College,Department of Computer Science And Engineering,Chennai-600123

***Abstract*— One of the most popular applications of Artificial Intelligence that has seen an immense growth in the digital era is Machine Learning Techniques where the system studies and improves its performance through progressive learning without any explicit programming. Machine learning is widely used in numerous applications one of them being medical analysis. Feature extraction and Image classification are considered to be the most popularly used approaches done using machine learning process. In this paper, we have proposed an efficient model for detecting brain stroke by preprocessing CT/MRI scan images andfurther segmenting it for feature extraction. Preprocessing of the images are done using Medical Image Fusion that generates a high quality fused image with spectral and spatial information. The fused image is trained with SVM classifier to classify whether the image is benign or malignant. The efficiency of the proposed model is evaluated which produces 80.48% sensitivity, 99.9% specificity,and 99.69% accuracy. The performance of the model is more accurate when compared with traditional medical analysis as it uses a fused image having higher quality.**

Keywords—***Brain tumor, Classification, Accuracy, SVM,K-Means Clustering, Fusion, Medical Image Analysis*.**

I. Introduction

The growth of Artificial Intelligence has given birth to many new technologies out of which the popular ones are being Machine Learning Approaches and Deep Learning Techniques. Machine learning techniques are widely used in numerous applications such as medical image analysis [1], robot path planning [2], flood detection in a particular city or area [3] and land cover classification [4]. Machine learning technique is a process of learning a specific task without any human intervention and improving the performance only by the continuous learning process. The learning is of two types: supervised learning [5], where the labeling is given for the features of the training dataset and unsupervised learning [6], where no labels are given and the system needs to label the features of the dataset. Feature extraction is a vital process in all the machine learning approaches. The extracted features could then be used for various other approaches like classification or regression. Some of the classifiers used predominantly in machine learning techniques are SVM classifiers [7], Decision Trees [8], Naive Bayes [9] and Logistic Regression [10]. When larger datasets are being used in an application then the use of Artificial Neural Networks (ANNs) is preferred for feature extraction as it produces more accurate results [11]. Use of ANNs is widely called Deep Learning Approaches as the neural network learns each and every layer very deeply and uses the output of a layer as the input of the next layer. ANNs serves as a classifier resembling the function of a biological neuron having numerous layers connected to each other through weights [12]. In image processing, the number of pixels (picture elements) depends upon the input image. The image is the replica of reality which provides as much information as possible about an object. The arrangement of neurons forming layers and the connection patterns formed within and between each layer is called Network Structures. ANNs are information processing structures that can solve any problem through learned examples rather than pre-specified algorithms [13].

Machine Learning Approaches finds unique importance in the field of medical image analysis. Images pertaining to the measurement of various parts of the human body based on different scales like microscopic or macroscopic are coined as biomedical images. These biomedical images are generated through various instruments such as Ultrasound machines and CT scan images. In this paper, we propose a technique ofidentifying and detecting brain stroke by segmenting tumor cells from CT/MRI scan images. The proposed model makes use of Medical Image Fusion method to fuse the scanned imaged so as to generate a high quality image. This further could be used for training the classifier and detecting whether the cell is benign or malignant. Rest of the paper consists of some related works proposed by various researchers, the methodologies and algorithms used in the proposed technique and experimental results

II. Related Work

Numerous research works are proposed by various researchers for segmenting or classifying a specific object or a certain area from an image. A semi-automatic segmentation method proposed by Rui Lu et al. estimated the volume of tumor cells in liver [14]. The boundaries of the tumor cells were localized from a CT image. Though the process consumed quite a lot of time for computation it claimed to be very efficient in finding the volume of the tumor cells by segmenting it into slices. In [15], Kostas Haris introduced a hybrid image segmentation using a morphological algorithm for watersheds that combined edge and region-based techniques. The technique found to be effective as it reduced the number of false edge detection. Emission of noise served to a part for indirectly increasing the processing time for the computation. Fanman Meng [16], designed an efficient and robust supervised image co-segmentation model comprising of a strategy called color reward and an active contour model. The model was evaluated on numerous images from a database and could efficiently pair the common objects with minimal error rate. A braintumor as detected by Jason J. Corso in [17] by integrating segmentation with Bayesian model. Weighted aggregation algorithm was used for the detection of tumor cells. The performance evaluation was done on a larger dataset where stochastic models were used extracting the features. The model could be enhanced by providing an accurate boundary of the tumor cells. Simultaneous image segmentation and bias correction were performed by Kaihua Zhang [18],where minimization of energy was performed by an efficient Bayesian Learning Approach. The technique stated that the experimental results were performed on a real dataset of images and produced an accurate intensity of homogeneity. Md. Badrul et al., [19] classified lung cancer images from CT scan using MLFFNN technique yielding a good accuracy of 96.67%. FFNN approach used in [20], described the classification technique producing an accuracy of 92%. Rajesh Kumar Tripathy [21] combined SVM with Least Square and provided an accuracy of 95.34%. Persi et al [22] used the Particle Swarm Optimization technique for predicting heart disease yielding 92.2% accuracy.

III. PROPOSED WORK

Numerous feature extraction methods and classification techniques are used widely to identify and detect the location of the tumor cells in a human body. Various classifiers such as SVM, K-means clustering and Decision Trees that are used widely for the image segmentation applications have been already discussed. In this paper, we majorly focus on the comparative analysis of various existing algorithms and the performance evaluation is done by using an application-oriented model. The enhanced and efficient model designed in the current paper first extracts the sequence of images from the scanner database. The images are then pre-processed where all the features of the images are extracted during neural networks. The preprocessing is initially done by Medical Image Analysis where all the images are processed in such a way that it generates new fused images that have high quality when compared to the original images. These fused images are easier for training and classifying the classifier as it contains more spatial and spectral information. The validation of the dataset is done frame by frame for a particular time period. As the proposed model is an iterative process, the model tends to produce a more accurate result when compared with other existing classifiers. The various steps involved in the entire model is discussed as follows:

Image Acquisition: CT/MRI images obtained from patients suffering from brain stroke are collected from Database and hospitals. The file formats collected from the database are of .jpeg, .tift, .png file formats and the file dimensions consist of rows and columns.

Image Pre-processing: The aim of image pre-processing is to improve the quality of the image and to enhance the features of an image for further processing. Subsequent to the image acquisition stage, the images pass through the pre-processing stage. There will be a change in the output image for the given input image. The occurrence of this change is due to the reduction of noise and/or enhancement of contrast. Image pre-processing is needed to vary itslightning condition. Pre-processing involves Color Normalization, Noise Reduction, Edge Detection, andHistogram Equalization. Many filters are used to reduce the effect of noise on the image. Median Filter is used for Color Normalization and Noise Reduction. Gabor Filter is used for Edge Enhancement and Histogram Equalization. Image Enhancement is used to improve the perception of information of images. It also modifies the attributes of an image and makes it suitable for a task. Highest Peak Signal to Noise Ratio (PSNR) is calculated for the entire data set. The term peak signal-to-noise ratio (PSNR) is an expression for relating the ratio between the maximum possible value(power) of a signal and the power of distorting noise that affects the quality of its representation.

Feature Extraction: Transforming the input data into a set of features is called Feature Extraction. It is a method of capturing visual content of the image. Features are extracted from the segmented image which includes area, perimeter, equivalent diameter, irregularity index, mean, standard deviation and entropy. Extracted Features are used in a Neural classifier to train the model in such a way that it could recognize a particular class as normal or abnormal. The classifier will assign the unknown object to the correct class depending on the extracted features.

Segmentation of Tumor: In this process, the homogeneous regions are obtained from the input image. The regions of interest in an image are found by using the process of segmentation. It reduces the number of pixels of an image to make it easier for the next step of feature extraction and classification. Segmentation is tougher in CT scan images and when combined with a huge amount of data because of the extra dimensions that need to be considered for the neighborhood calculation. Fig. 1 depicts the flow chart of the entire classification process.

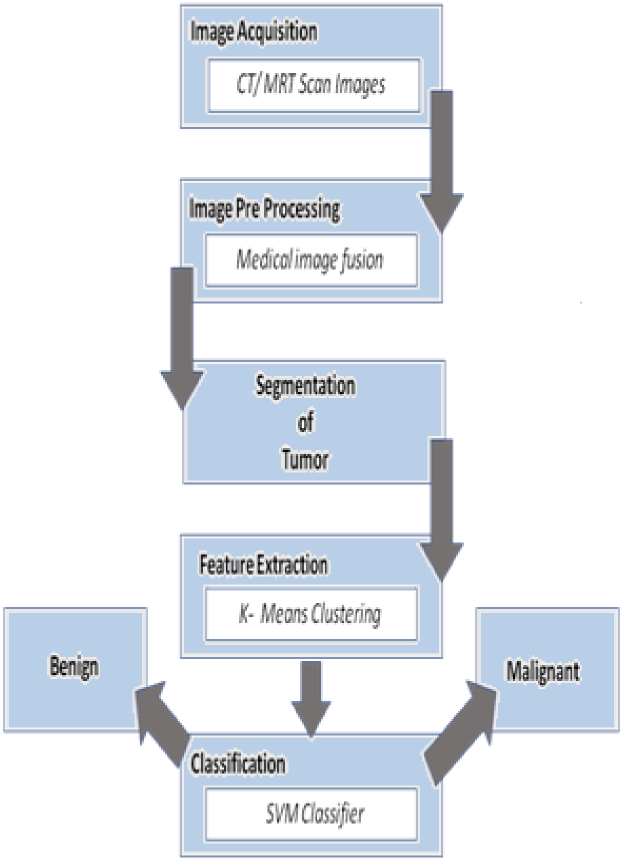

Figure. 1 Workflow of Classification of images

Figure. 2 represents the entire architecture of how the classification is done. The Input image is given from the scanner database and is first converted into Grayscale. Thethe Discrete Cosine Transform is applied to the image when the contrast of the image is enhanced in such a way that even the small features of the images are perfectly visible. Then the features areextracted for further study and then finally Medical Image Fusion is performed on the image so as to obtain the segmented image. Once the fused images are obtained it is then trained using an SVM classifier that is responsible for extracting the features of the image. The feature extraction is done using K-Means clustering algorithm where all the features are extracted for further classification.

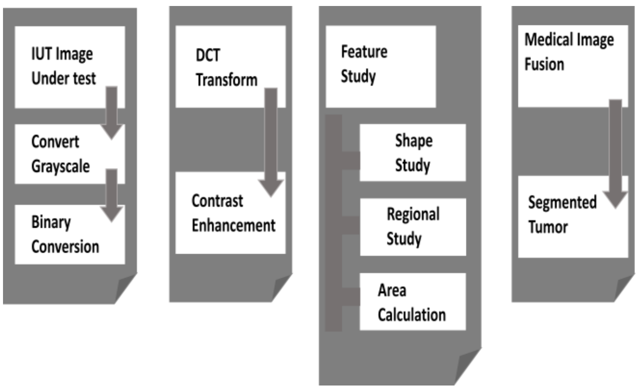

Figure. 2 Architecture Diagram

IV. EXPERIMENTAL RESULTS

The experiments are performed on MATLAB R2017b version. The computations are performed using image Classification Learner Toolbox that is readily available in MATLAB. First, the input image from the scanner database is given to the tool. Figure.3 is a sample input image obtained from a scanner database. When this image is passed the pre-processing takes place where all the features are extracted for further segmentation.

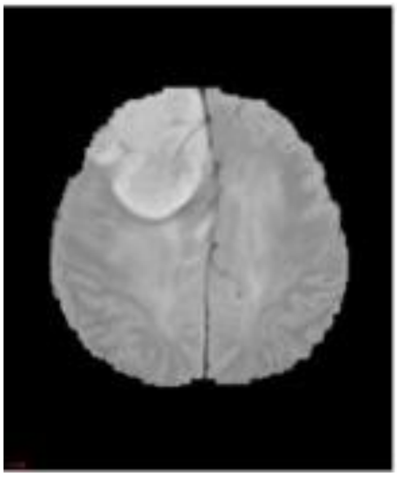

Figure. 3 Input Image

As the next step, the RGB components are extracted from the image and then converted to Grayscale image. Figure.3 shows the output image when the Grayscale filter is applied to the image.

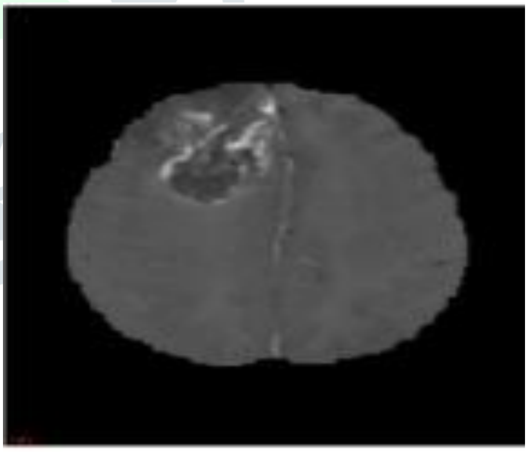

Figure. 4 Filter Output

Both Figure.3 and Figure.4 are compared with each other to find the difference between the features and then Medical Image fusion is applied to the images to identify the tumor. The features are then trained with SVM classifier where some images are kept for training and the rest are kept for testing of images. The classifier learns the features and the successfully classifies when any new image is given to it. It identifies where the image is benign or malignant. Figure. 5 shows the identification of the tumor cell from the image and Figure.6 show the segmented tumor from the image. The accuracy of the SVM classifier is observed with 99.69% whereas 99.9% specificity and 80.48% sensitivity. The segmentation of the tumor from the exact image gives us various parameters for measuring such as its intensity, volume,and size. This helps in diagnosing and treating the tumor cells more effectively.

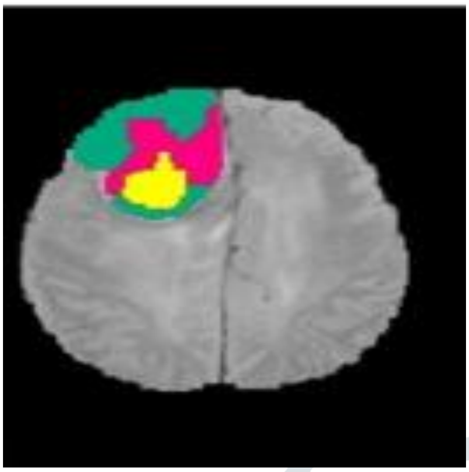

Figure. 5 Image Segmentation

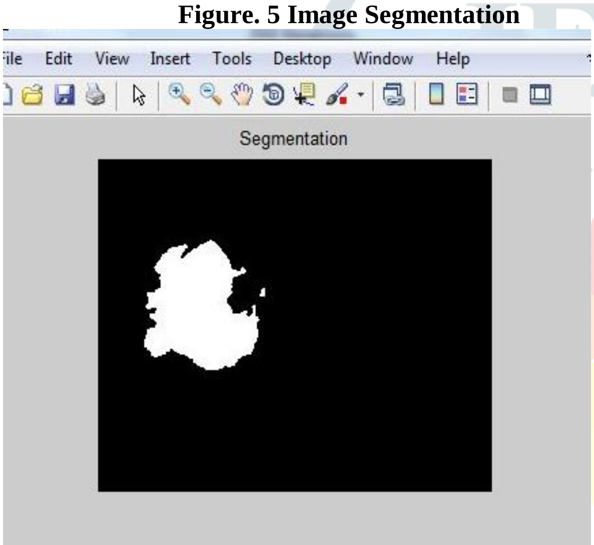

Figure. 6 Segmented Output

V. Conclusion

Brain strokes are considered as one of the most upcoming health issues where more patients are likely to suffer. The detection of these brain tumors at the initial stages could minimize the amount of risk in the medical field. In this paper, we have proposed an efficient method of diagnosing brain tumors at an earlier stage by using Medical Image Analysis. The images obtained from CT/MRI scanners are used for fusion in such a way that generates high image quality. These images are then used for further classification. SVM classifier is used for classification of the tumors and has obtained an accuracy level of about 99.69%. The future work can include the use of deep learning techniques where it could reduce the computational time to a greater extent.

References

[1] Rahman, M. M., Bhattacharya, P., & Desai, B. C. (2007). A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. IEEE Transactions on Information Technology in Biomedicine, 11(1), 58-69.

[2] Morales, M., Tapia, L., Pearce, R., Rodriguez, S., & Amato, N. M. (2004). A machine learning approach for feature-sensitive motion planning. In Algorithmic Foundations of Robotics VI(pp. 361-376). Springer, Berlin, Heidelberg.

[3] Ireland, G., Volpi, M., & Petropoulos, G. P. (2015). Examining the capability of supervised machine learning classifiers in extracting flooded areas from Landsat TM imagery: a case study from a Mediterranean flood. Remote sensing, 7(3), 3372-3399.

[4] Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. International Journal of remote sensing, 23(4), 725-749.

[5] Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160, 3-24.

[6] Cheriyadat, A. M. (2014). Unsupervised feature learning for aerial scene classification. IEEE Transactions on Geoscience and Remote Sensing, 52(1), 439-451.

[7] Bauer, S., Nolte, L. P., & Reyes, M. (2011, September). Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 354-361). Springer, Berlin, Heidelberg.

[8] Liu, K., Tong, M., Xie, S., & Zeng, Z. (2014, August). Fusing decision trees based on genetic programming for classification of microarray datasets. In International Conference on Intelligent Computing (pp. 126-134). Springer, Cham.

[9] Dumitru, D. (2009). Prediction of recurrent events in breast cancer using the Naive Bayesian classification. Annals of the University of Craiova-Mathematics and Computer Science Series, 36(2), 92-96.

[10] Chou, Y. H., Tiu, C. M., Hung, G. S., Wu, S. C., Chang, T. Y., & Chiang, H. K. (2001). Stepwise logistic regression analysis of tumor contour features for breast ultrasound diagnosis. Ultrasound in medicine & biology, 27(11), 1493-1498.

[11] Sarhan, A. M. (2009). Cancer classification based on microarray gene expression data using DCT and ANN. Journal of Theoretical & Applied Information Technology, 6(2).

[12] Steven K. Rogers, Dennis W. Ruck, Matthew Kabrisky, Artificial Neural Netwroks for early detection and diagnosis of cancer, Elsevier Scientific Publishers, pp. 79-83, 1994.

[13] Md. Badrul Alam Miah, Mohammad Abu Tousuf, Detection of Lung Cancer from CT Image Using Image Processing and Neural Network, IEEE, In Proceedings of 2nd Int'l Conference on Elelctrical Engineering and Information & Communication Technology, 2015

[14] Lu, R., Marziliano, P., & Thng, C. H. (2006, January). Liver tumor volume estimation by semi-automatic segmentation method. In Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the(pp. 3296-3299). IEEE.

[15] Haris, K., Efstratiadis, S. N., Maglaveras, N., & Katsaggelos, A. K. (1998).Hybrid image segmentation using

watersheds and fast region merging. IEEETransactions on image processing, 7(12), 1684-1699.

[16] Meng, F., Li, H., Liu, G., & Ngan, K. N. (2013). Image cosegmentation by incorporating color reward strategy and active contour model. IEEE transactions on cybernetics, 43(2), 725-737.

[17] Corso, J. J., Sharon, E., Dube, S., El-Saden, S., Sinha, U., & Yuille, A. (2008). Efficient multilevel brain tumor segmentation with integrated bayesian model classification. IEEE transactions on medical imaging, 27(5), 629-640.

[18] Zhang, K., Liu, Q., Song, H., & Li, X. (2015). A variational approach to simultaneous image segmentation and bias correction. IEEE Transactions on Cybernetics, 45(8), 1426-1437.

[19] Md. Badrul Alam Miah, Mohammad Abu Tousuf, Detection of Lung Cancer from CT Image Using Image Processing and Neural Network, IEEE, In Proceedings of 2nd Int'l Conference on Elelctrical Engineering and Information & Communication Technology, 2015

[20] Shubhangi Khobragade, Aditya Tiwari, C.Y.Patil , Vikram Narke, Automatic Detection of Major Lung Diseases using Chest radiographs and Classification by Feed- forward Artificial Neural Networks, In Proceedings of 1st IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems, pp.1-5, 2016

[21] Rajesh Kumar Tripathy, Sailendra Mahanta Subhankar Paul, Artificial Intelligence –based Classification of Breast Cancer using Cellular images RSC Advances , Issue 18, Page 8939 to 9411, 2014

[22] Persi Pamela I, Gayathri P, N. Jaisankar, A Fuzzy optimization Techniques for the Prediction of Coronary Heart Disease Using Decision Tree, International Journal of Engineering and Technology, Vol 5, No 3,pp. 2506-2513, 2013.