INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

IJARIIT

ISSN: 2454-132X
Impact factor: 4.295
(Volume 5, Issue 2)
Available online at: www.ijariit.com

Prediction model for brain tumor patients based on MRI images

Ashutosh Kaushik
ashu7g@gmail.com
S. J. C. Institute of Technology, Chikkaballapur, Karnataka

Ankit Kumar
ankit.k012@hotmail.com
S. J. C. Institute of Technology, Chikkaballapur, Karnataka

Rashmi K. A.
rashmisjcit@gmail.com
S. J. C. Institute of Technology, Chikkaballapur, Karnataka

ABSTRACT

The paper presented here describes the predicted survival time for brain tumor patients using Magnetic Resonance Imaging (MRI). The accuracy is improved using the denoising wavelet transform (DWT) method. For this work BraTS, a dataset is used. MRI images are used to extract the histogram features so that the prediction model can be trained using the machine learning methods. MRI information is damaged due to the noise in MRI imaging. And the 2D wavelet transform was able to improve the results. The SVM with a 10 folds cross-validation helps to achieve the best accuracy by Daubechies 4 level 4 (db4-L4). With the same 10 folds, Daubechies 2 level 1 and 3 produces better results when the age factor is removed. An accuracy of 66.7% is achieved with a 10 % hold validation method in Daubechies 2 level 3.

Keywords— *Denoising wavelet transform, Machine learning, MRI, Histogram, MRI images, Glioma brain tumor*

1. INTRODUCTION

The most dangerous type of brain tumors are Glioma brain tumors and due to which the survival time of patients do not go beyond 2 years. The most used or renowned way to find and locate the size of the brain tumor is MRI or Magnetic Resonance Imaging. In 2017, BraTS introduced a challenge so that anyone can use classification or other approaches like a regression for estimating the overall survival time of brain tumor patients. The file containing the segmentation annotation or MRI modalities is present in each sample. This leads us to three types of patients: the short-termed of which have less than 10months, the mid-term or ones that have 10 to 15 months to survive and lastly the long-term which have more than 15 months to live. There are various types of machine learning methods and also various features that are used to design and develop a prediction model that is based on the classification method.

The pre-trained Alexnet that is trained by Linear Discriminant has helped in achieving the best classification accuracy. When histogram features were used the accuracy didn't exceed 40%. Noisy MRI images were the output of histogram distribution. Due to which this resulted in low accuracy. To be precise a high-quality image is required to achieve accurate medical prediction and its diagnosis. As a result, the DWT method is used for improving the quality of MRI images before we use the histogram features.

2. NOISE IN MRI DATA

The raw MRI data were damaged due to the noise in the MRI System. The coils in the imaging system affected the model of the noise in the MRI images. In single coil system, the noise is presented as Rician distribution and for multi-coils or parallel imaging system as non-central chi-distribution. For denoting the raw frequency-domain measurements or K-space we use:

$$D(x,y) = D_{Re}(x_f, y_f) + iD_{Im}(x_f, y_f) \tag{1}$$

The two components Real and Img are present in the inverse Discrete Fourier Transform of raw data in K-space and described as follow:

$$d(x,y) = Real + i\,Img \tag{2}$$

The Real and Img are defined respectively as:

$$Real = s(x,y)\,cos(\theta(x,y)) + n_{Re}(x,y) \tag{3}$$
$$Img = s(x,y)\,sin(\theta(x,y)) + n_{Im}(x,y) \tag{4}$$

Where the signal is denoted as $s(x,y)$, and $n_{Re}(x,y)$ *and* n_{Im} represents the White Gaussian Noise (WGN) in the real imaginary components respectively. The MRI image magnitude value of $d(x,y)$.

$$|d(x,y)| = \sqrt[2]{(Real)^2 + i(Img)^2} \tag{5}$$

A Rician noise in MRI image is produced due to the magnitude of the imaginary and real part each corrupted by Gaussian noise.

3. RESULTS AND PERFORMANCE EVALUATION

There are 3 steps in the overall prediction system which are as follows: the preprocessing that is used to denoise the MRI images, next is the feature extraction in order to get the histogram features and finally the Machine Learning step to train our prediction model. The patients' age information and the histogram features that were extracted from MRI images were combined to train a prediction model for the overall time for which the brain tumor affected person will survive.

The image is split into 4 subbands, called HH, HL, LH and LL subbands in the first level of decomposition. The performance of the trained prediction model used the accuracy measure which is defined as:

$$ACC = \frac{TP + TN}{P + N} \tag{6}$$

Where TP denotes the number of true positive predictions, and TN denotes the number of true negative predictions; P denotes the number of positive samples, and N denotes the number of negative samples. To improve the accuracy of a prediction model based on denoised data, various mother wavelets transform with different levels, and various ML were employed. The best accuracy before and after implementing DWT are listed in table 1.

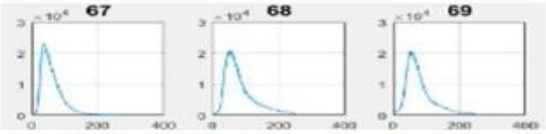


Fig. 1: Histogram of the brain region in MRI images for samples 67, 68, and 69

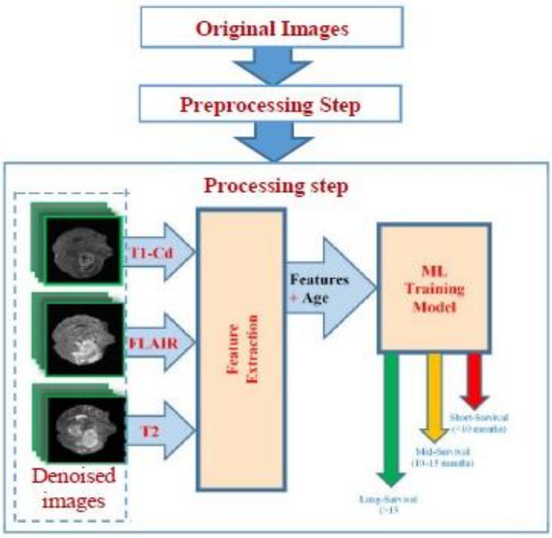
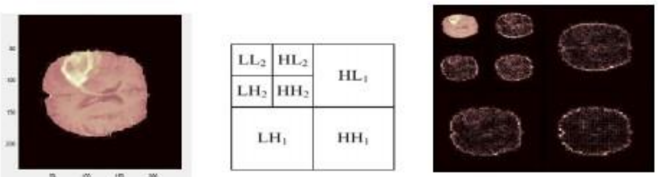


Fig. 2: The overall prediction system



Original noisy MRI image

2-Dimension DWT-Level 2 applied to each slice of each MRI sample

Denoising MRI slice 80 from sample 1 using db 2- Level 2

Fig. 3: Denoising MRI image using 2 level daubechies wavelet transform

Table 1: Accuracy of the prediction model based on histogram features with denoised MRI images by 2D WT

Method	# of Folds	Best ML	Overall accuracy
Noisy +age	10 folds	Simple Tree	41.5%
Noisy +age	10% holdout	KNN	53.3%
Db4-L4+age	10 folds	Linear SVM	44%
Db2-L1, Db2-L4	10 folds	Simple Tree	44.7%
Db2-L3+age	10% holdout	Simple Tree	66.7 %

4. CONCLUSIONS AND EXPANDED FUTURE WORKS

The quality of the MRI images and also the performance of a prediction model is improved by 2D DWT for the survival time of the brain tumor patients. For future, we may work on bringing the 3D DWT instead of 2D DWT. And we can also use both 2D as well as 3D DWT to get better results. Apart from these methods, image and feature fusion techniques can be taken into consideration in order to improve our results.

5. REFERENCES

[1] E. C. Holland, "Progenitor cells and glioma formation," J. Current. Opinion in Neurology, vol. 14, no. 6, pp. 683–688, 2001.

[2] H. Ohgaki and P. Kleihues, "Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglia gliomas," J. of Neuropathology & Experimental Neurology, vol. 64, no. 6, pp. 479–489, June 2005.

[3] Alex Lobera," Imaging in Glioblastoma Multiforme", Updated: Feb 10, 2017, https://emedicine.medscape.com/article/340870-overview#a

[4] BH Menze et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", J. IEEE Transactions on Medical Imaging, vol. 34, no. 10, 2015, pp. 1993-2024, doi: 10.1109/TMI.2014.2377694.

[5] S. Bakas et al., "Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radio mic features", J. Nature Scientific Data, Sept. 5 2017, doi:10.1038/sdata.2017.117.

[6] L. Chato and S. Latifi, "Machine Learning and Deep Learning Techniques to Predict Overall Survival of Brain Tumor Patients using MRI Images", on Proc. 17th IEEE Int. Conf. Bio-Informatics and Bio-Engineering BIBE 2017, Washington DC, pp. 9-14, 23rd-25th Oct. 2017, 2471-7819/17/31.00 ©2017 IEEE, doi: 10.1109/BIBE.2017.00009.

[7] D. Drumheller, "General expressions for Rician density and distribution functions," IEEE Trans. Aerosp. Electron. Syst., vol. 29, no. 2, pp. 580–588, Apr. 1993.

[8] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," Magn. Reson. Med., vol. 34, pp. 910–914, Aug. 1995.