RESEARCHERID
THOMSON REUTERS

[Lanka* *et al.*, 7(2): February, 2018]
IC™ Value: 3.00

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

MULTIPLE SCLEROSIS DETECTION BASED ON MULTIVARIATE ASYMMETRIC DISTRIBUTION

Shanthi Raju Lanka [*1], Srinivas Yarramalle [2]

[*1, 2] Department of IT, GITAM Institute of Technology, GITAM, Visakhapatnam, 530045, India

DOI: 10.5281/zenodo.1165644

ABSTRACT

Multiple Sclerosis is one of the predominant occurring brain diseases during the last decade. Many automatic segmentation models have been widely investigated with objective to identify the disease at the early stage and thereby helping the patients. In this research, a Multivariate Asymmetric Gaussian Mixture Model is considered for segmenting the brain MRI images and to identify the Multiple Sclerosis disease from the T1 weighted, T2 weighted and photon density images. The earlier works on GMM failed to converge to arrive at optimal value and thus becomes time consuming and also could not able to achieve the expected results. To overcome these disadvantages, this research study aims at proposing a model based on Multivariate Asymmetric Gaussian Mixture Model. The images are compressed by using Non Negative factorization method and the decomposed images are considered for further evaluation. The results of the proposed methods are tested on brain web images and in most of the cases; the recognition accuracy is underlined at above 85%. This shows that the performance of the proposed method is far better than the existing methods and also this methodology is very much time efficient.

KEYWORDS: Asymmetric Gaussian Mixture Model, Multivariate Distribution, Multiple Sclerosis, performance evaluation, T1 and T2 weighted.

I. INTRODUCTION

Multiple Sclerosis is one of the most recurrent diseases witnessed during last decade and this disease if affected, causes disability in young adults. The recent medical statistics confer that women are mostly affected by this disease. The patients suffering with multiple sclerosis forego the immunity system and the central nervous system also gets attacked and damages the myelin sheath which leads to sclerosis. Basing on the location of multiple sclerosis, the corresponding area in the brain gets damaged, resulting into deformities such as vision, Neuro deficiency, gait and paralysis also occurs in advanced stages.

Therefore, it is necessary to identify the disease at early stage, which enables to treat the patient successfully. However, due to the abnormal data accumulated in the brain, analyzing manually is a relatively difficult task and hence automatic segmentation techniques are to be developed. Recently, much emphasis is given towards the identification and treatment of disease using both parametric and non-parametric approaches [1] [2] [3] [4]. Among the non-parametric models, most of the researchers have utilized the concept of edge detection, region of interest identification, SVM based classification, SVD, GGLCM [5] [6] [7] [8] [9] [10]. However, the brain is a complicated organ, which involves different issues like gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). Most of the brain related diseases are subjected to occur within these tissues, as such making it next to impossible for the proper identification of the disease.

Alternative techniques based on parametric models are also suggested in the literature. Most of these works are based on Gaussian Mixture Model, Constrained Gaussian Mixture Model, Bivariate Gaussian Mixture Model, etc., [11] [12] [13] [14] [15]. However, the brain image is asymmetric in nature and therefore affective results can be assumed only by considering the asymmetric nature of brain distribution. Also, the disease will be accumulated to a particular region in the brain resulting into difficulties during the identification phase. Also, as the brain is a combination of several neurons, in order to identify and manage the disease more exactly within a minimum time stamp, one need to consider compression algorithms.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Features play a vital role during the identification of disease. Since these diseases are mostly concealed under brain tissues, it is necessary to consider more features such that effective analysis can be carried out. Hence, in this article an attempt is made in this direction to develop a model by considering multiple features. To compress the image data and identify the disease more robustly Multivariate Skew Distribution together with Non-Negative factorization is considered. The rest of the paper is organized as follows. In section 2 of the paper, Multivariate Skew Symmetric Model is presented. Section 3 of the paper deals with the feature set. In section 4 data set considered is presented. Section 5 of the paper deals with image compression based on Non-Negative factorization and section 6 of the paper highlights the FCM Clustering algorithm. The methodology is explained in Section 7. The experimental results derived thereof are highlighted in section 8 of the paper. Section 9 of the paper concludes the paper.

II. MULTIVARIATE SKEW SYMMETRIC GAUSSIAN MIXTURE MODEL (MSSGMM):

Most of the issues in the brain are mostly non homogenous and are asymmetric in nature. To analyze the disease, one need to consider distributions that can understand the asymmetry. Skew Symmetric Gaussian Mixture Model is one among such distributions.

The probability density function of the skew normal distribution is given by

$$f_{\mu,\ \sigma,\ \lambda}(x) = \frac{2}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right).\Phi\left(\lambda\frac{x-\mu}{\sigma}\right) \quad \ldots(1)$$

Where $\mu \in R, \sigma > 0$ and $\lambda \in R$ represents the location, scale and shape parameters respectively. Where ϕ and Φ denote the probability density function and the cumulative density function of the standard normal distribution. The maximum and minimum intensity of pixels within the image regions are denoted by 'a' and 'b'. Truncating the data between these limits helps to minimize the image space. Using this concept, Truncating equation (1) between these limits 'a' and 'b' we have

$$F_{\mu,\ \sigma,\ \lambda}(x)\int_a^b = F_{\mu,\sigma,\lambda}(b) - F_{\mu,\sigma,\lambda}(a) \quad \ldots(2)$$

$$F_{\mu,\ \sigma,\ \lambda}(a) = \int_{-\infty}^{a} F_{\mu,\sigma,\lambda}(x)dx \quad \ldots(3)$$

$$F_{\mu,\ \sigma,\ \lambda}(b) = \int_{-\infty}^{b} F_{\mu,\sigma,\lambda}(x)dx \quad \ldots(4)$$

$f\ \mu, \sigma,\ \lambda(x)$ is as given in equation (1)

$$Q = \int_{-\infty}^{b}\frac{2}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)\bar{\varphi}\left(\frac{\lambda(x-\mu)}{\sigma}\right)dx - \int_{-\infty}^{a}\frac{2}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)\bar{\varphi}\left(\frac{\lambda(x-\mu)}{\sigma}\right)dx \quad \ldots(5)$$

III. FEATURE SET CONSIDERED:

Features play a vital role in the identification of the disease. In order to extract the relevant images more accurately, one need to consider multiple features rather than a solitary feature. In order to fulfill this aspect, in the present article we have considered the features based on shape, texture and color. These features play a consolidating role in the identification of the disease.

IV. IMAGE COMPRESSION BASED ON NON-NEGATIVE FACTORIZATION (NNF):

NNF is considered for effective compression of the images considered. Each of the images is first pre processed to eliminate noise and the processed images are considered for enhancement. Each of the missing data is considered and the corresponding pixels within the image dataset are mapped against the most recurrent existing pixels such that the missing data is eliminated. This image matrix is considered and the concept of NNF is applied against this data so that the corresponding resulted image is free from redundancy.

V. DATASET CONSIDERED:

In order to present the proposed work, we have considered the MRI Medical images from American Medical Society and Brain web images. Each of these images is subjected to compression based on NNF and the resultant preprocessed image is considered for further analysis.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

VI. FUZZY CLUSTERING ALGORITHMS:

To classify the disease more accurately from the heterogeneous group of images, one need to cluster the data, FCM is considered in this article. The main reason behind the choice of FCM is that medical image pixels are vogue in nature and each pixel can assume multiple characteristics such as damaged undamaged and partially damaged. FCM has the capability to interpret this variable nature of the pixel and help towards proper ratification. It is based on the calculation of fuzzy weights based on the reciprocal weights. The cluster centers are calculated using the Gaussian weights and it is a recursive process where it uses large initial prototypes for estimating the cluster centers, identifying the eliminating pixels and finally merging the pixels into the clusters basing on the weights.

The step wise algorithmic representation is presented below

1. The basic idea behind the FCM is estimating the total mean square error by considering the minimum weights between the pixels. The calculation for the weights is as follows

$$J(Wij, z(j)) = \sigma (j=1,J)\, \sigma (j=1,J)\, (Wij)\| x(i)- z(j)\|2 \quad \ldots(6)$$

$\sigma (j=1,J) (Wij) = 1$ for each i

$$Wij = (1/(Dij)2)1/(p-1) / \sigma (j=1,J)\, (1/(Dij)2)1/(p-1)\ ,\ p > 1 \quad \ldots(7)$$

The FCM considers every pixel within the cluster to be in the range of fuzzy truth value which lies between 0 and 1 and is computed using the formula given in equation (7) presented above. Using the above algorithm, one can assign a feature vector to a cluster according to the maximum weight of the feature vector over all clusters.

VII. METHODOLOGY:

In order to experiment with the proposed methodology, we have considered the dataset consisting of around 2500 images of which 50 images are considered for testing purpose and 50 images are considered for training purpose. Each of the images is preprocessed and compressed. Each image is normalized into a fixed size of 150x150. Each image from the training set is considered and the corresponding PDF's are generated and stored against a database. Each of the images from the training data is considered and again the PDF's are generated. The PDF are mapped based on the criteria of maximum likelihood estimate.

Each image is considered and clustered so as to identify the damaged pixels and the undamaged pixels. In order to cluster the data, Fuzzy C-Means (FCM) algorithm is considered. Each image pixel is fit into the corresponding PDF equation (1) proposed in section II of the paper. For each of the medical images, the corresponding PDF images are generated. Using the concepts of Maximum Likelihood criteria and comparing the resultant images using KL-Divergence.

VIII. EXPERIMENTATION:

The experimentation is conducted in Dot Net Framework. Initial training set considered for the identification of the disease is presented below.

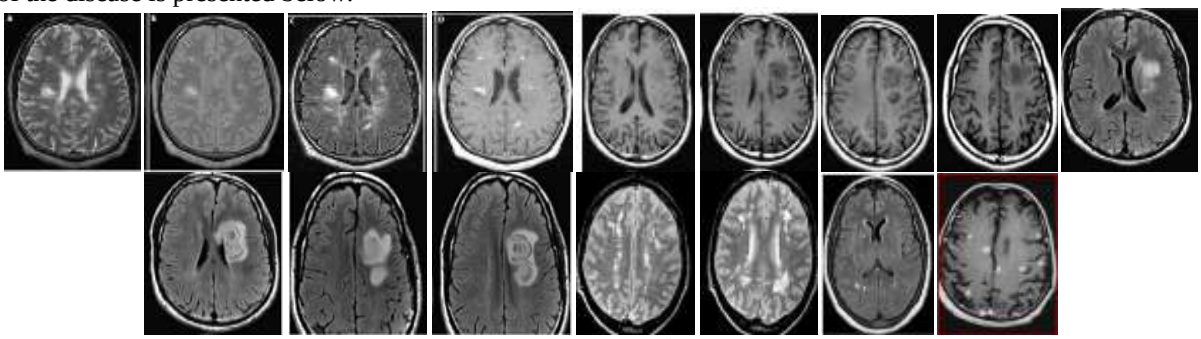
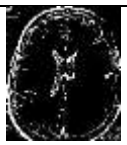
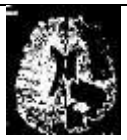


Fig. 1. Testing Images considered for experimentation

The results obtained are presented below. The results are tested on 8 images and the results are compared with that of methodologies presented using GMM and the results are evaluated using benchmark quality metrics such as Average Difference, Maximum Distance, Image Fidelity, Mean Squared error and Signal to noise ratio. The standard limits and the standard criteria are also presented in the following Table 1 of the article.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Table 1. Showcasing the identification of diseased images

Image	Quality Metric	GMM	SGMM with Fuzzy C-Mean	Standard Limits	Standard Criteria
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.573 0.422 0.416 0.04 17.41	0.8451 0.945 0.9756 9.3E-07 108.42	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.37 0. 221 0.336 0 2404 14.45	0.49 0.931 0.9046 3.6E-06 102.5	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.456 0.345 0.44 0.22 19.88	0.6721 0.911 0.9366 2.43E-06 104.27	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.231 0. 224 0.212 0.24 21.42	0.7731 0.9001 0.8835 4.46E-06 101.634	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.342 0.317 0.391 0.2514 3.241	0.6957 0.815 0.985 4.62E-07 111.482	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.21 0.21 0.2134 0.06 13.43	0.3653 0.892 0.787 0.145 49.22	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.3232 0.123 0.233 0.01 11.11	0.322 0.212 0.897 0.4345 27.267	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.314 0.241 0.293 0.18 21.214	0.338 0.249 0.683 0.197 78.19	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible

IX. CONCLUSION:

From the above Table.1 it can be clearly seen that the diseased images are identified most significantly by the proposed methodology. The derived results are also computed against image quality metrics like Average Difference, Maximum Distance, Image Fidelity, Mean Squared error and Signal to noise ratio and the results obtained against this metrics clearly showcase that the present methodology has the ability to identify the diseased images from the set of given images.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

X. REFERENCES

[1] Ch Murali Krishna, Y.Srinivas, Unsupervised Color Image Segmentation Using Bivariate Features and Dimensionality Reduction Methodology, International Journal of Modern Computer Science and Applications (IJMCSA), Volume No.3, Issue No.1, January, 2015
[2] L.Shanthi Raju, Kumar Vasantha, Y.Srinivas, Content Based Image Retrievals Based on Generalization of GMM, International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 3 (6), 2012.
[3] Ashraf A. Aly, Safaai Bin Deris, Nazar Zaki, RESEARCH REVIEW FOR DIGITAL IMAGE SEGMENTATION TECHNIQUES, International Journal of Computer Science & Information Technology (IJCSIT) Vol 3, No 5, Oct 2011.
[4] Jia Li, Surajit Ray, Bruce G. Lindsay, A Nonparametric Statistical Approach to Clustering via Mode Identification,Journal of Machine Learning Research, Vol 8 (2007) 1687-1723.
[5] Sonam Saluja, Aradhana Kumari Singh, Sonu Agrawal, A Study of Edge-Detection Methods, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 1, January 2013.
[6] T.Johncy Rani, S.Suja Priyadharsini, Region of Interest Tracking In Video Sequences, International Journal of Computer Applications, Volume 3 – No.7, June 2010.
[7] Benoît Magnin, Lilia Mesrob, Serge Kinkingnéhun, Mélanie Pélégrini-Issac, Olivier Colliot, Marie Sarazin, Bruno Dubois, Stéphane Lehéricy, Habib Benali, Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI, Neuroradiology, Springer, September 2008.
[8] Nidhal K. El Abbadi, Adil Mohamad and Mohammed Abdul-Hameed, IMAGE ENCRYPTION BASED ON SINGULAR VALUE DECOMPOSITION, Journal of Computer Science, Vol 10 (7): 1222-1230, 2014.
[9] P. Mohanaiah, P. Sathyanarayana, L. GuruKumar, Image Texture Feature Extraction Using GLCM Approach, International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013.
[10] PratikChavada, Narendra Patel, KanuPatel, Region of Interest Based Image Compression, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 1, January 2014.
[11] Prasad Reddy P.V.G.D, Srinivas Rao. K, Srinivas Yarramalle, Unsupervised Image Segmentation Method based on Finite Generalized Gaussian Distribution with EM & K-Means Algorithm, IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.4, April 2007.
[12] Oren Freifeld, Hayit Greenspan, and Jacob Goldberger, Multiple Sclerosis Lesion Detection Using Constrained GMM and Curve Evolution,Hindawi Publishing Corporation, International Journal of Biomedical Imaging, Volume 2009, Article ID 715124, 13 pages
[13] Nagesh Vadaparthi, Srinivas Yarramalle, Suresh Varma Penumatsa, An Improved Brain Mr Image Segmentation using Truncated Skew Gaussian Mixture, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 7, 2015.
[14] Hayit Greenspan, Amit Ruf, and Jacob Goldberger, Constrained Gaussian Mixture Model Framework for Automatic Segmentation of MR Brain Images, IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 25, NO. 9, SEPTEMBER 2006.
[15] M Vamsi Krishna, Dr. P. V. G. D. Prasad Reddy, Dr. Y. Srinivas, Dr.S.S.Nayak, Bivariate Gaussian Mixture Model Based Segmentation for Effective Identification of Sclerosis in Brain MRI Images, International Journal of Engineering and Technical Research (IJETR), Volume-3, Issue-1, January 2015
[16] Alireza Karimian, Simin Jafari, A New Method to Segment the Multiple Sclerosis Lesions on Brain Magnetic Resonance Images, Journal of Medical Signals & Sensors, Vol 5, No 4 (2015).

CITE AN ARTICLE

Lanka , S. R., & Yarramalle, S. (n.d.). MULTIPLE SCLEROSIS DETECTION BASED ON MULTIVARIATE ASYMMETRIC DISTRIBUTION. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, 7*(2), 83-87.