Bio-medical Image Retrieval using Various Statistical Methods

Vinutha N, Sandeep S, P Deepa Shenoy, Venugopal K R

Abstract: In recent, the healthcare sectors rely more on imaging technologies for early detection and diagnosis of the disease. But, the abundant images obtained from these imaging technologies have complex disease patterns associated with them and thus an expert requires more time to analyze and arrive at the decision. Hence, the image retrieval techniques have a significant role to assist the experts by retrieving the most similar images existing in the database and also help them to compare a new scan of the patient with the top matched images and arrive at the quick decision during the diagnosis of a patient. So, we have performed our studies on the two-dimensional structural Magnetic Resonance Imaging of the Open Access Series of Imaging Studies dataset. The collected images are preprocessed and categorized into different groups based on the ventricular region of the brain. After the categorization, we employ second and higher-order statistical approaches to extract the textural features. Then the computed textural features of the images existing in the dataset are compared with the textural features of a query image to retrieve the top matched images using similarity distance as the metric. Then the image retrieval performances of the proposed hybrid based statistical methods are measured. The obtained results shows that the combined features of Gray Level Co-occurrence Matrix and Law's Texture Energy Measure attains the highest precision across the categorized groups of a dataset and it achieves 80% precision for Group1, Group2 images and 60% precision for Group3 images.

Keywords : Alzheimer's Disease, Content-based Image Retrieval, Magnetic Resonance Imaging, Statistical Methods, Textural Features, Ventricle.

I. INTRODUCTION

Alzheimer's Disease is a common form of Dementia in older adults [1]. It is characterized by the deposition of proteins in the different regions of the brain. Thus, it results in the shrinkage and enlargement of different brain regions and these are termed as the significant biomarkers. In the recent past, the healthcare sectors rely more on a diverse range of biomedical images such as Fluorodeoxyglucose Positron Emission Tomography (FDG-PET), Structural Magnetic Resonance Imaging (sMRI) and Computed Tomography (CT) to identify the significant biomarkers. But MRI, a non-invasive technique that has been widely employed in neuroimaging analysis because it acquires the high-resolution images in sagittal, coronal and axial planes and it also shows different intensity levels between the different tissues of the brain.

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Vinutha N *, University Visvesvaraya College of Engineering, Bengaluru, India, E-mail: vinutha1v@gmail.com.

Sandeep S, Practo Technologies Private Limited, Bengaluru, India.

P DeepaShenoy, University Visvesvaraya College of Engineering, Bengaluru, India.

Venugopal K R, Bangalore University, Bengaluru, India.

Due to the various advantages associated with it, the researchers and doctors prefer MRI images to diagnose the disorders and differentiate the different stages of disease. Therefore it helps to monitor the disease status of a person by visualizing the pathologies of the affected brain. The various imaging techniques, demographic information, clinical scores anddiagnostic information are formed as the standardized datasets. These datasets are freely available to the doctors, researchers and scientific communities and are considered as an important source of information that is essential to develop and test the efficient techniques required for the Computer-Aided Diagnosis (CAD) [2], [3]. And the large collection of images has led to an increase in the size of medical datasets. Thus, the researchers also focus on the development of image retrieval techniques that are necessary to provide efficient search and quick access to the most similar images existing in the database. Retrieval of similar images aids the doctor to improve the diagnosis and also to provide an effective treatment. The Content-based Image Retrieval (CBIR) is an image retrieval technique that is utilized to index and retrieve the images based on the visual content such as color, shape and texture [4]. But the textural features define the characteristics of an image concerning various properties such as coarseness, entropy, uniformity, randomness, etc. So it plays an important role to identify the discriminative patterns of an image and thereby it helps to distinguish an abnormal person from the healthy subjects. The different methods such as Transform based approach,Model-based approachand Statistical based approach are available for textural feature extraction. The Model based approach commonly employs Autoregressive, Markov Random Field (MRF) [5] model to derive the descriptors.But a large computation has to be performed, if a higher-order neighborhood is preferred to extract the textural features. The Transform based approach [6], [7] utilizes the filter bank to derive the features from either the frequency or spatial domain. Some of the transform-based approaches utilized to derive the textural features are Gabor Transform, Fourier Transform and Wavelet Transform. Another method utilized for textural feature extraction is the Statistical based approach, which involves a simple computation to obtain the different textural properties of an image and the simple calculation is performed on the image pixels by using the standardized mathematical equation. The statistical methods are categorized into different groups based upon the number of pixels utilized in the computation of the features and the different groups are First-order statistics, Second-order statistics and Higher-order statistics. The calculation of First-order statistics is based on the utilization of the single pixel value in the textural feature extraction, while Second-order statistics obtains the textural features by utilizing the two neighboring pixel values and the Higher-order statistics is calculated based on the spatial relationship of more than two pixels.

Therefore, in our study we extract the textural features from sMRI images of Open Access Series of Imaging Studies (OASIS) dataset [8] by employing the second and higher order statistical methods such as Gray Level Co-occurrence Matrix (GLCM), Neighborhood Gray Tone Difference Matrix (NGTDM)and Law's Texture Energy Measure. In the next step, a similarity distance is computed between the textural features of the query image and images existing in the database. Finally, we index and retrieve the top matched images from the database which are similar to the given query image.

Motivation: The people affected with the disorders are exponentially increasing day by day. As a result, many of them visit to the hospitals regularly to monitor disease status and to obtain a better diagnosis. This has resulted in the massive accumulation of data in the health care sectors. Therefore it is most essential for the researchers to develop efficient techniques for Computer-Aided Diagnosis (CAD) and also to retrieve the most similar images from the database for a given query image.

Objective: The main objective of our proposed work is to improve the precision of the retrieved images from the database matching to the given query image.

Contribution: The images are categorized into different groups based on the ventricular region of the brain. Categorization is followed by the extraction of textural features using both the second and higher-order statistical approaches. Then the textural features are utilized to retrieve similar images from the database.

Organization: The Literature Survey is discussed in Section 2. The proposed image retrieval framework is explained in Section 3. Experiments and Results are presented in Section 4 and the conclusion of the paper is written in Section 5.

II. LITERATURE SURVEY

In this section, we discuss the different pattern based image retrieval techniques applied to the bio-medical images by extracting the local and global descriptors as the features from the images.

Kumar et al., [9] extracted the Zernike Moments (ZM) as the features from MRI and CT images of the OASIS and Lung database. ZM features are the global descriptors that extract both the shape and texture of the image and the feature extraction method is invariant to the rotation and scale of the images. By using these features it helps to retrieve the relevant images from the database, even though the imagesare either rotated or have different scaled versions. The retrieval performance of the proposed method using ZM as the features shows better results than other local descriptors based methods such as Local Binary Pattern (LBP), Uniform Local Binary Pattern (LBP), etc.

In paper [10] the authors have proposed Local Mesh Vector Co-occurrence pattern-based image retrieval and tested its performance on the OASIS database. The preprocessing steps applied on these MRI images are: Resizing of Images to 256×256 dimension, grayscale conversion, noise removal by Gaussian filter, an Adaptive Histogram Equalization is applied for image enhancement. Then the authors extracted pattern maps from these preprocessed images using Local Vector Pattern (LVP) and Local Mesh Pattern (LMP) techniques. Further, the GLCM features are computed and combined from these two pattern maps and then the combined features are used to retrieve the similar images from the database.

The authors of paper [11] have improvisedthe Local Ternary Pattern (LTP) technique and applied to the standardized bio-medical databases. In the first step, the three local mesh patterns are obtained from the images. Then the local mesh pattern extraction is followed with the extraction of a ternary pattern of all the collected images from the database. Later a threshold value is set to convert the obtained ternary pattern of an image into two Local Binary Patterns (LBP). Then in the last step, a product of a binomial weight with two LBP's generates the unique ternary patterns and finally, the histograms are obtained as the features from these patterns. Further, by following the above steps the histograms are also generated from the other mesh patterns of the image. At the last, all the histograms of different ternary patterns are fused to perform image retrieval.

Table-I: Image Retrieval Techniques Applied on Bio-Medical and Natural dataset

Author	Dataset	Technique	Type of Descriptor
Kumar et al.,[9] (2017)	Emphysema CT database and OASIS-MRI	Zernike Moments based framework	Global Descriptor
Jenitta et al., [10] (2017)	OASIS-MRI	Local Mesh Vector Co-occurrence pattern	Local Descriptor
Verma et al., [12] (2017)	Texture and natural database	Local Neighborhood Difference Pattern	Local Descriptor
Murala et al., [13] (2015)	Biomedical dataset, natural and texture database	Spherical Symmetric 3D Local Ternary Pattern	Local Descriptor
Galshetwar et al., [14](2018)	VIA/I-ELCAP (CT images) MESSIDOR (Retinal images), and OASIS-MRI databases	Local Directional Mask Maximum Edge Patterns	Local Descriptor

Verma et al., [12] extracted the local descriptors using two different methods such as Local Neighborhood Difference Pattern (LNDP) and conventional Local Binary Pattern (LBP) techniques. A 3×3 matrix is considered to extract the features using LNDP. Then a binary pattern for each pixel is obtained by calculating the difference between the other two neighborhood pixels, which are located either in a horizontal or vertical direction of a matrix. Then the histograms obtained for these pattern maps are utilized to perform the image retrieval.

In paper [13] Murala et al., proposed the Spherical Symmetric 3D Local Ternary Pattern (SS-3DLTP) based image retrieval technique and tested it on natural, biomedical and texture databases. They considered the three-dimensional images from the above specified databases and converted them into three 2D planes by applying Gaussian Filter Bank with a suitable standard deviation.

Then the Spherical Symmetric patterns are obtained from the 2D planes in five spherical directions and it is followed with the generation of LTP patterns. Further, SS-3D LTP's are converted into lower LTP and upper LTP and then generated the histogram for these two patterns. The proposed method outperforms the other transform-based and spatial domain techniques in terms of Average Retrieval Rate (ARR).

The authors of paper [14] have developed the Local Directional Mask Maximum Edge Patterns (LDMaMEP) method. It is suitable to extract edge pattern-based features from both the grayscale and color images. In the first step, an image is converted to three different 2D planes. If a color image is used as an input then the 2D planes are obtained by using RGB colors and for a grayscale, it is obtained by Gaussian filter bank. Then, in the next step five directional images are obtained by traversing the three different planes in vertical, horizontal, anti-diagonal and diagonal directions. Then these directional images are convolved with the eight directional masks to produce edge vectors and then these edge vectors are used as features to improve the retrieval performance.

III. METHODOLOGY

In this section, the proposed framework is represented in Fig.1. and we describe it in detail in the following subsections: It begins with the data collection from OASIS dataset followed by the preprocessing and categorization of images into different groups based on the ventricular region of the brain. Then the textural features are extracted from these MRI images using second and higher order statistical approaches. Later, the extracted textural features are utilized in image matching and retrieval.

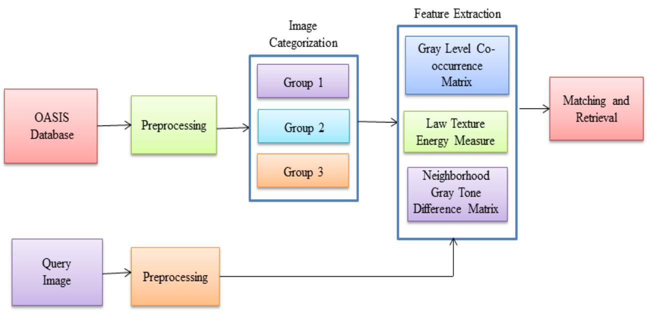

Fig. 1. Bio-medical Image Retrieval using Various Statistical Methods

A. Data Collection

We obtained the two dimensional MRI images of demented and non-demented subjects from the OASIS dataset. The collected data consists of cross-sectional images of 436 subjects with 268 females and 168 males having their age group between 18-96 years. The dataset also provides the clinical diagnostic information of the older adults of being diagnosed as mild/moderate AD or as non-demented subjects. The OASIS dataset is freely available to all the researchers and scientific communities. So it allows them to perform the experimentation on biomedical images using efficient techniques and therefore it improves the CAD and as well as image retrieval performance.

B. Preprocessing

The collected images are pre-processedto obtain the images with similar intensityand therefore it aids in the feature extraction and retrieval of top matched images. Fig.2. shows two brain images having different intensity.So, if we subject these images into feature extraction step, then it results in the extraction of distinct features. Hence, the grey levels are adjusted for the collected images and for our dataset; we found the following values to be most effective. If the grey level varies between 0.41-1, then we adjust it to 0.75. Else if, the grey level varies between 0.25-0.31, then we adjust it to 0.50. Else, the grey level varies between 0.1-0.25, then we adjust it to 0.25.

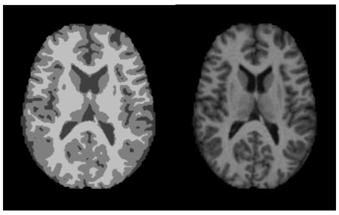

Fig. 2.Images of OASIS Dataset with Different Intensity

C. Grouping the Images

The preprocessed MRI images are categorized into different groups based on the ventricular region of the brain because the size of the ventricle increases with AD progression. Hence,it is considered to be an important criterion to group the images at different stages of AD. So to perform the categorization, we initially begin with the selection of the pre-defined area in such a way that it encloses only the ventricle and the remaining area of the brain is excluded and it is followed with the determination of a specific value of a pixel from the dataset images to filter out other parts of the brain excluding the ventricle. In our research we found pixel values less than or equal to 27 to be an appropriate value for the same.Later, we count the number of pixels present in the ventricular region of the brain and set the threshold of pixel count to categorize the images into different groups.

D. Feature Extraction

The categorization of images into different groups is followed by the feature extraction using the Grey Level Co-Occurrence Matrix (GLCM) [15], Law's Texture Energy Measure[16] and Neighborhood Gray Tone Difference Matrix (NGTDM)[17].

- **Grey Level Co-Occurrence Matrix (GLCM)**

GLCM is a second-order statistical approach utilized to extract the different textural features such as Entropy, Contrast, Inverse Difference Moment, Energy and Correlation using equations 1, 2, 3, 4 and 5 respectively. The feature calculation by GLCM is based on analyzing the co-occurrence pixels pairs that are separated with distance 'd'from each other.

Entropy: Entropy is a measure of randomness in an image. A random distribution of pixel grey levels would lead to high entropy.

$Entropy = \sum_a \sum_b P(a,b) \log P(a,b)$(1)

Contrast: Contrast is a measure of Grey level variation of pixels in an image. If the difference between two neighboring pixels is large for an image. Then it indicates that an image has high contrast.

$$Contrast = \sum_a \sum_b (a-b)^2 P(a,b) \quad (2)$$

Inverse Difference Moment: Inverse Difference Moment is a measure of local homogeneity in an image. If an image comprises of repeated grey level values then it is said that to possess a homogeneous texture. Homogeneity also determines whether the image is non-textured or textured based on the range of values obtained by its calculation using equation 3.

$$IDF = \sum_a \sum_b \frac{1}{1+(a-b)^2} P(a,b) \quad (3)$$

Energy: It is also known as Uniformity or angular second moment. It identifies the repeated pixel pairs in an image as a measure of uniformity. If an image has number of repeated pixel pairs, then it results in high homogeneity and high energy.

$$Energy = \sum_a \sum_b \{P(a,b)\}^2 \quad (4)$$

Correlation: Correlation provides information about the pixel and its neighbor. If an obtained value is +1, then it represents the pixels are positively correlated, -1 represents the pixels are negatively correlated and 0 indicates that pixels are uncorrelated.

$$Correlation = \sum_a \sum_b \frac{(a-\mu_1)(b-\mu_2)P(a,b)}{\sqrt{\sigma_1^2 \sigma_2^2}} \quad (5)$$

- **Law's Texture Energy Measure**

Law's Texture Energy Measure performs the convolutional operation on an image using the filter masks and extracts the different textural features. A filter mask is defined as a two-dimensional array formed by the combination of two different 1D convolutional kernels. The filter masks with 3×3 as its matrix size are formed by using 1D convolutional kernels of vector length 3 and the filter masks with 5×5 as its matrix size are formed by using 1D convolutional kernels of vector length 5. For the formation of 5×5 filter masks, five different1D convolutional kernels are used and these kernels are represented as S5, W5, R5, E5, and L5 respectively. Here S5 is used to detect the spot, W5 identifies the wave, R5 detects the ripple, E5 finds the edges and L5 obtains the center-weighted local average. The combinations of these 1D convolutional kernels have resulted in the formation of twenty five filter masks as shown in Table-II. Finally, these masks are convolved on an image to generate energy maps.

$$Spot\ S5 = [-1 \quad 0 \quad 2 \quad 0 \quad -1]$$

$$Wave\ W5 = [\,-1 \quad 2 \quad 0 \quad -2 \quad 1\,]$$

$$Ripple\ R5 = [1 \quad -4 \quad 6 \quad -4 \quad 1]$$

$$Edge\ E5 = [\,-1 \quad -2 \quad 0 \quad 2 \quad 1\,]$$

$$Level\ L5 = [1 \quad 4 \quad 6 \quad 4 \quad 1\,]$$

The formation of 5×5 filter mask using $S_5 \times W_5$ convolutional kernels is illustrated below:

$$\begin{bmatrix} 1 & -2 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & 4 & 0 & -4 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 2 & -1 \end{bmatrix}$$

Table-II : 2D Mask formed by the combination of 1D kernels

	S5	W5	R5	E5	L5
S5	S5S5	S5W5	S5R5	S5E5	S5L5
W5	W5S5	W5W5	W5R5	W5E5	W5L5
R5	R5S5	R5W5	R5R5	R5E5	R5L5
E5	E5S5	E5W5	E5R5	E5E5	E5L5
L5	L5S5	L5W5	L5R5	L5E5	L5L5

- **Neighborhood Gray Tone Difference Matrix**

Let 'i' be the gray tone value of the pixel at (a, b). Then we compute the average gray tone '$\overline{G_i}$' over the neighborhood which is centered at but the gray tone at (a, b) is excluded during the computation.

$$\overline{G_i} = \frac{1}{w-1}\left[\sum_{x=-d}^{d} \sum_{y=-d}^{d} f(a+x, b+y)\right] \quad (6)$$

where (a, b) ≠ (0, 0)

whereW= $(2\mathbf{d}+\mathbf{1})^2$ (7)

Where'd' denotes the size of the neighborhoodand 'W' denotes the sum total of pixels in the neighborhood, x and y represent the neighborhood pixels of a and b.

For example if d=1 then it results in the 3×3 neighborhood having W=9 pixels and if d=2 then it results in a 5×5 matrix with W=25pixels. Then for each gray tone in the NGTDM S(i) is calculated as shown in equation 8.

$S(i)=\sum|i-\overline{G_i}|$ for $i \in N_i$ if $N_i \neq 0$ (8)

Else S(i) = 0.

'N_i' represents the set of pixels of Gray Tone 'i' present in the neighborhood.

$$\begin{matrix} 1 & 2 & 5 & 3 & 1 \\ 4 & 0 & 1 & 2 & 3 \\ 3 & 1 & 4 & 5 & 0 \\ 0 & 2 & 4 & 4 & 5 \\ 2 & 0 & 4 & 3 & 1 \end{matrix}$$

Fig .3. : A 5×5 matrix with different Grey levels in a sample image

Table-III:Neighborhood Gray Tone Difference Matrix

i	N_i	p_i	S(i)
0	1	0.111	2.625
1	2	0.222	3
2	2	0.222	1
3	0	0	0
4	3	0.333	3.5
5	1	0.111	2.125

Consider a 5×5 matrix in a sample image having different grey levels as shown in Fig.3. If we choose'd=1' as a distance, then it results in a 3×3 neighborhood. The neighborhood which can be centered at only these pixels are indicated with red color and the remaining pixels which are at the periphery of the matrix are indicated with black color.

For example, let us consider two pixels having gray tone 'i=1' is present in the neighborhood. So we calculate S(i) using equation 8.

$$\mathbf{S(1)} = \left|\mathbf{1} - \frac{\mathbf{22}}{\mathbf{8}}\right| + \left|\mathbf{1} - \frac{\mathbf{18}}{\mathbf{8}}\right| = \mathbf{3}$$

Similarly, we calculate for S (0) = 2.625; S (2) = 1; S (3) = 0; S (4) = 3.5; S (5) = 2.125;

The grey level probability 'p_i' is calculated by taking the ratio of set of pixels having gray tone 'i' and the sum total of pixels 'N_{PV}' with valid neighbors located at a distance 1.

$$p_i = \frac{N_i}{N_{PV}} \tag{9}$$

$$p_1 = \frac{2}{9} = 0.222$$

Similarly, we calculate for p0 = 0.111; p2 = 0.111; p3 = 0; p4 =0.333; p5 =0.111;

The obtained values for S(i) and p_i for theconsidered gray tone 'i' are tabulated in the Table-III.Further, we calculate five different NGTDM features such as coarseness, contrast, busyness, complexity and strength using equations 10, 11, 12, 13 and 14 respectively.

Coarseness: Coarseness is an inverse measure of the rate of change of spatial intensity in an image. Rate of change of spatial intensity in an image can be defined as "A uniform texture has small difference between the average of pixel intensities of the neighborhood and the intensity value of pixels. The sum totals of such differences are computed for all the pixel values to indicate the rate of change of spatial intensity". The basic pattern or primitives are large in coarse texture. So the texture possesses uniform intensity in image. Thus, the higher value of coarseness results in a more uniform texture and smaller the gray tone differences in an image. The coarseness 'F_{cos}' is calculated as shown in the equation 10.

$$F_{cos} = \left[\epsilon + \sum_{i=0}^{G_{hi}} p_i\, s(i)\right]^{-1} \tag{10}$$

Where 'ϵ' represents a small number that avoids the 'F_{cos}' becoming infinite.

Contrast: A contrast depends on the frequency of change in spatial intensity and the dynamic range of grey levels in an image. The contrast 'F_{con}' is termed as the product of dynamic range of gray levels and the frequency of change in spatial intensity. If both these factors are high then the intensity difference of neighboring pixels is also high. So it results in an image with high contrast.

$$F_{con} = \left[\frac{1}{N_{gl}(N_{gl}-1)} \sum_{i=0}^{G_{hi}} \sum_{j=0}^{G_{hi}} p_i p_j (i-j)^2\right] \left[\frac{1}{n^2} \sum_{i=0}^{G_{hi}} S(i)\right] \tag{11}$$

Where 'Ngl' represents the sum total of grey levels present in an image.

$$N_{gl} = \sum_{i=0}^{G_{hi}} U_i, \text{Where} U_i = \begin{Bmatrix} 1, & if\ p_i \neq 0 \\ & \text{else } 0; \end{Bmatrix} \tag{12}$$

Busyness: A rapid change in the grey levels between pixels and its neighborhood is a measure of busyness.The level of busyness mainly depends on the frequency of change in spatial intensity.

$$F_{bus} = \frac{\left[\sum_{i=0}^{G_{hi}} p_i\, s(i)\right]}{\left[\sum_{i=0}^{G_{hi}} \sum_{i=0}^{G_{hi}} ip_i - jp_j\right]} \tag{13}$$

Where $p_i \neq 0,\ p_j \neq 0$

Complexity: Complexity is a measure of visual information in a texture. If non-uniformity is high then it leads to high visual information content and high complexity. This is mainly due to the presence of more primitives or patterns in the texture and also because of these patterns having different average intensities present in it. The complexity is partly correlated to contrast and busyness.

$$F_{com} = \sum_{i=0}^{G_{hi}} \sum_{j=0}^{G_{hi}} \left\{\frac{|i-j|}{n^2(p_i+p_j)}\right\} \{p_i S(i) + p_j S(j)\} \tag{14}$$

Strength: Strength measures the image primitives. If the primitives in an image are clearly visible, attractive and easy to define, then it leads to high strength. The contrast and coarseness may be correlated to the texture strength. The texture strength is calculated using equation 15. Here the numerator represents the intensity difference between the primitives and denominator represents the texture primitive size.

$$F_{str} = \frac{\left[\sum_{i=0}^{G_{hi}} \sum_{j=0}^{G_{hi}} (pi+pj)(i-j)^2\right]}{\left[\epsilon + \sum_{i=0}^{G_{hi}} S(i)\right]} \tag{15}$$

Where $p_i \neq 0,\ p_j \neq 0$

- ***Similarity Measurement***

The similarity measurement is a distance measure computed between the feature vector of the database and feature vector of the query image as shown in equation 16.

$$D = \sum_{i=1}^{p} \frac{f_{diz} - f_{qi}}{1 + f_{diz} + f_{qi}} \tag{16}$$

Where f_{diz} - Feature of zth image in the database.
f_{qi} - Feature of a query image.
D - Similarity Distance metric between the feature vector of the database and a query image.

- ***Metrics for Performance analysis***

The Precision and Error are the different metrics utilized to analyze the image retrieval performance of the top matched images from the database for a given query image. The Precision and Error is calculated using equation 17 and 18.

$$\text{Precision} = \frac{Relevant\ Images\ Retrieved}{Total\ no\ of\ Images\ Retrieved} \times 100 \tag{17}$$

$$\text{Error} = \frac{Irrelevant\ Images\ Retrieved}{Total\ no\ of\ Images\ Present\ in\ the\ Database} \times 100 \tag{18}$$

IV. EXPERIMENTS AND RESULTS

The cross-sectional MRI images are collected from the OASIS dataset and they arepreprocessed and categorized into different groups based on the total number of pixels present in the ventricular region of the brain and also performed experimentation to adjust a suitable threshold range to each group. So we carried out our implementation by selecting the ventricular region of the brain using the rectangular shape. In the second step, we calculate the total number of pixels having gray scale value less than or equal to 27. Then, in the last step, we categorize the images into three groups based on the pixel count in the ventricular region of the brain. In the first group, the number of pixel count ranges between 1-500, while for the second group the number of pixel count ranges between 501-1000 and for the last group it ranges between 1001 to 1500.

After the data collection, preprocessing and categorization of images, they are subjected to the feature extraction step using the second and higher-order statistical approaches. The GLCM is employed as a second-order statistical method to extract different textural features such as Entropy, Contrast, Inverse Difference Moment, Angular Second Moment and Correlation. Further, the NGTDM and Law's Texture Energy Measure are employed as the higher-order statistical approaches for the extraction of textural features. The obtained NGTDM features for each image are coarseness, contrast, busyness, complexity, strength and the obtained 25 energy maps from Law's Texture Energy Measure are S5S5, W5S5, R5S5, E5S5, L5S5, S5W5, W5W5, R5W5, E5W5, L5W5, S5R5, W5R5, R5R5, E5R5, L5R5, S5E5, W5E5, R5E5, E5E5, L5E5, S5L5, W5L5, R5L5, E5L5 and L5L5 respectively. But, we choose 9 non-repetitive energy maps (E5S5, L5E5, S5R5, L5R5, R5R5, S5S5, E5R5 E5E5 and L5S5) among the obtained 25 energy maps and evaluated the retrieval performance.

The retrieval performance is measured using precision and error as the metrics for the different textural features extracted from GLCM, NGTDM and Law's Texture Energy Measure and also by using the combined features from second and higher order statistical methods.Then, the obtained results are plotted in the Fig.4. and Fig.5. respectively and these figuresshows the high precision and low error for the combination five features from GLCM and nine feature maps from Law's Texture Energy Measure.

The proposed hybrid based statistical methods achieves 80% precision for Group1 and Group2 images and 60% precision for Group 3 images and are represented in the Fig.6. , Fig.7. and Fig.8. respectively..

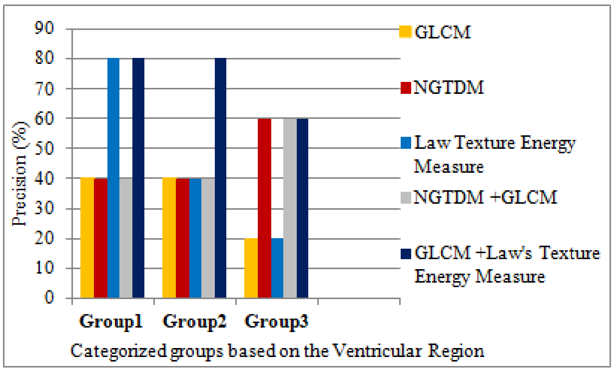

Fig.4. The Precision Result of the Categorized Groups using Second and Higher Order Statistical Approach

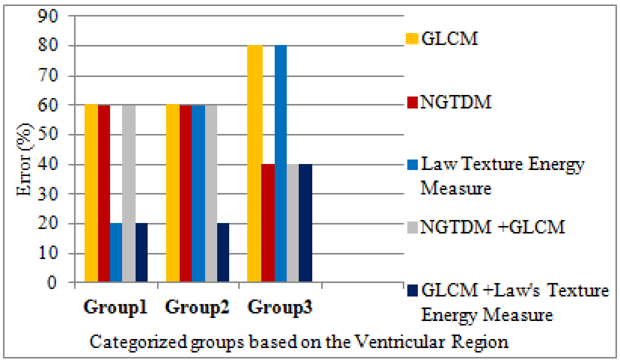

Fig.5. The Error (%) of the Categorized Groups using Second and Higher Order Statistical Approach

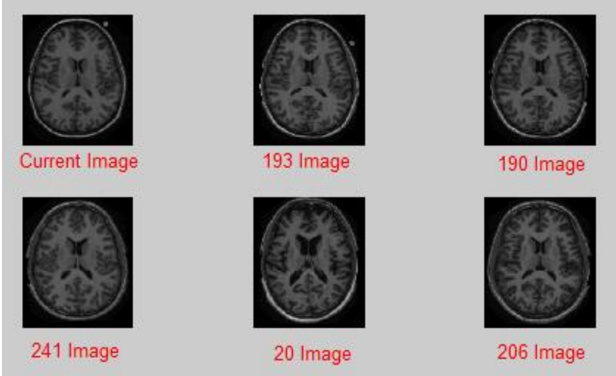

Fig.6. TheImage Retrieval Result of Group1 Images by the Fusion of Features from Law's Texture Energy and GLCM

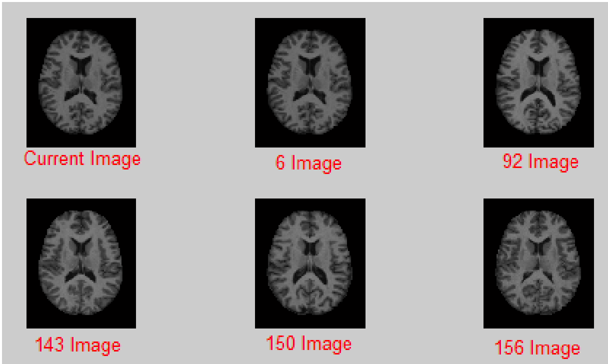

Fig.7. The Image Retrieval Result of Group2 Images by the Fusion of Features from Law's Texture Energy and GLCM

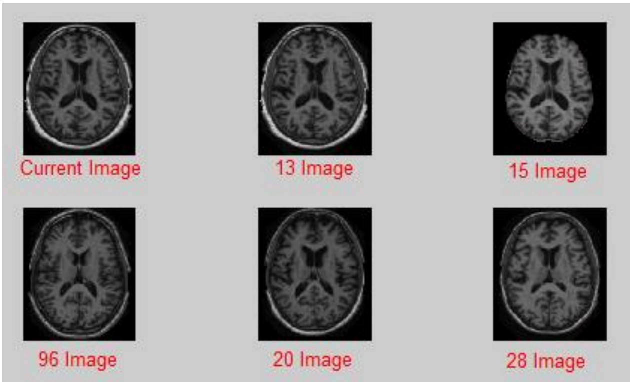

Fig.8. The Image Retrieval Result of Group3 Images by the Fusion of Features from Law's Texture Energy and GLCM

V. CONCLUSIONS

The precision result obtained by the fusion of features from GLCM and Law's Texture Energy Measure outperforms the fused features from other statistical methods across all the three different categorized groups and therefore, this approach can be utilized to improve the Computer Aided Diagnosis of AD and also assist the doctors to arrive at the quick decision. In future, we extend our work by considering the Region of Interest and three dimensional images of other standardized medical datasets.

ACKNOWLEDGMENT

We thank the curators of OASIS dataset for providing the dataset

REFERENCES

1. Alzheimer's, Association. "2015 Alzheimer's Disease Facts and Figures." *Alzheimer's and Dementia: The Journal of the Alzheimer's Association*, vol.11, no. 3, pp. 332, 2015.
2. Vinutha N, P. DeepaShenoy, Venugopal K. R., "Efficient Morphometric Techniques in Disease Detection: Survey and Tools." *American Journal of Neuroscience, Science publication*,ISSN: 2524-2237, Vol. 7, no.2, pp.19-44, 2016.
3. Sandhya Joshi, V Simha, P DeepaShenoy, Venugopal K R, L M Patnaik, "Classification and Treatment of Different Stages of Alzheimer's Disease using Various Machine Learning Methods." *International Journal of Bioinformatics Research*, vol. 2, no. 1, pp. 44–52, 2010.
4. Liu, Guang-Hai, Jing-Yu Yang, and ZuoYong Li. "Content-based Image Retrieval using Computational Visual Attention Model." *Pattern Recognition*, vol. 48, no. 8, pp. 2554-2566, 2015.
5. Hung, Chih-Cheng, Enmin Song, and YihuaLan. "Texture Features and Image Texture Models." *In Image Texture Analysis, Springer*, pp. 15-50. 2019.
6. Ghasemzadeh, Ardalan, SaeedSarbazi Azad, and ElhamEsmaeili. "Breast Cancer Detection based on Gabor-Wavelet transform and machine learning methods." *International Journal of Machine Learning and Cybernetics*,vol.10, no. 7, pp. 1603-1612, 2019.
7. Alaei, Fahimeh, AlirezaAlaei, Umapada Pal, and Michael Blumenstein. "A Comparative Study of Different Texture Features for Document Image Retrieval." *Expert Systems with Applications*,vol.121,pp.97-114, 2019.
8. Open Access Series of Imaging Studies. Accessed: June 20, 2019. [Online]. Available: https://www.oasis-brains.org/.
9. Kumar, Y., Aggarwal, A., Tiwari, S. and Singh, K., "An Efficient and Robust Approach for Biomedical Image Retrieval using Zernike Moments." *Biomedical Signal Processing and Control*, vol.39, pp.459-473, 2018.
10. Jenitta A and R Samson Ravindran. "Image Retrieval Based on Local Mesh Vector Co-occurrence Pattern for Medical Diagnosis from MRI Brain Images." *Journal of Medical Systems*, vol. 41, no. 10, pp.1-10, 2017.
11. Deep G, L Kaur, and S Gupta, "Local Mesh Ternary Patterns: A New Descriptor for MRI and CT Biomedical Image Indexing and Retrieval." *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, vol.6, no. 2, pp.155-169, 2018.
12. Verma, Manisha, and Balasubramanian Raman. "Local neighborhood difference pattern: A New Feature Descriptor for Natural and Texture Image Retrieval." *Multimedia Tools and Applications*, vol. 77, no. 10 pp.11843-11866, 2018.
13. MuralaSubrahmanyam, and Q M Jonathan Wu. "Spherical Symmetric 3D Local Ternary Patterns for Natural, Texture and Biomedical Image Indexing and Retrieval." *Neurocomputing*, vol.149, pp.1502-1514, 2015.
14. Galshetwar G M, L M Waghmare, A B Gonde, and S Murala. "Multi-Dimensional Multi-Directional Mask Maximum Edge Pattern for Bio-Medical Image Retrieval." *International Journal of Multimedia Information Retrieval*, pp. 1-9, 2018.
15. DhruvBhawna, Neetu Mittal, and MeghaModi. "Study of Haralick's and GLCM Texture Analysis on 3D Medical Images." *International Journal of Neuroscience*, pp.1-29, 2018.
16. Kamal Khurram, RehanQayyum, SenthanMathavan and TayyabZafar. "Wood Defects Classification using Laws Texture Energy Measures and Supervised Learning Approach." *Advanced Engineering Informatics*, vol.34, pp. 125-135, 2017.
17. Mahmood, Usman, AdityaApte, Joseph O. Deasy, C. Ross Schmidtlein, and AmitaShukla-Dave. "Investigating the Robustness Neighborhood Grey Tone Difference Matrix (NGTDM) and Grey Level Co-Occurrence Matrix (GLCM) RadiomicFeatures on Clinical Computed Tomography Systems using Anthropomorphic Phantoms: Evidence from a multivendor study." *Journal of Computer Assisted Tomography*,vol.41, no. 6, pp. 995-1001, 2017.

AUTHOR'S PROFILE

Vinutha Nobtained her Bachelor of Engineering in Biotechnology, VTU, India. She received her Master of Engineering in Bioinformatics from University Visvesvaraya College of Engineering, Bangalore University, Bangalore, India and currently she is full time Research Scholar, pursuing Phd in the Department of Computer Science and Engineering at University Visvesvaraya College of Engineering, India. She is IEEE student member and her current research interests include Machine learning, Medical Image Processing and Bioinformatics.

Sandeep Sobtained his Bachelor of Engineering in Computer Science and Engineering at University Visvesvaraya College of Engineering, India.. He is currently working in Practo Technologies Private Limited, Bengaluru, India. Hisresearch interests include Medical Image Processing and Satellite Image Processing.

P Deepa Shenoy is currently working as Professor in the Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bangalore, India. She did her doctorate in the area of Data Mining from Bangalore University in the year 2005. Her areas of research include Data Mining, Soft Computing, Biometrics, Bioinformatics, Image Processing and Social Media Analysis. She has published more than 150 papers in refereed International Conferences and Journals.

Venugopal K R iscurrently working as Vice Chancellor, Bangalore University, Bangalore. He obtained his Bachelor of Engineering from University Visvesvaraya College of Engineering. He received his Master's degree in Computer Science and Automation from Indian Institute of Science Bangalore. He was awarded Ph.D in Economics from Bangalore University and Ph.D in Computer Science from Indian Institute of Technology, Madras. He has a distinguished academic career and has degrees in Electronics, Economics, Law, Business Finance, Public Relations, Communications, Industrial Relations, Computer Science and Journalism. He has authored and edited 64 books on Computer Science and Economics, which include Petrodollar and the World Economy, C Aptitude, Mastering C, Microprocessor Programming, Mastering C++ and Digital Circuits and Systems etc., He has filed 101 patents. During his three decades of service at UVCE he has over 700 research papers to his credit. His research interests include Computer Networks, Wireless Sensor Networks, Parallel and Distributed Systems, Digital Signal Processing and Data Mining. He is a Fellow of IEEE and ACM Distinguished Educator.