

Application of Fisher Score and mRMR Techniques for Feature Selection in Compressed Medical Images

Vamsidhar Enireddy

Associate Professor, Department of CSE,
MVR College of Engineering, Vijayawada, A.P
enireddy.vamsidhar@gmail.com

Dr.D.V.V.S.PhaniKumar

Professor, Department of CSE,
Geethanjali Institute of Tehnology & Science, Nellore, A,P
phanikumard@yahoo.com

Dr. Gunna Kishore

Professor, Department of CSE,
MVR College of Engineering, Vijayawada, A.P
kishorebrbm@gmail.com

Abstract: In nowadays there is a large increase in the digital medical images and different medical imaging equipments are available for diagnoses, medical professionals are increasingly relying on computer aided techniques for both indexing these images and retrieving similar images from large repositories. To develop systems which are computationally less intensive without compromising on the accuracy from the high dimensional feature space is always challenging. In this paper an investigation is made on the retrieval of compressed medical images. Images are compressed using the visually lossless compression technique. Shape and texture features are extracted and best features are selected using the fisher technique and mRMR. Using these selected features RNN with BPTT was utilized for classification of the compressed images.

Keywords: Image Compression, Image retrieval, Haar Wavelet, Gabor filter, Feature Selection, Recurrent Neural Network.

Introduction

The advancement of technology in the medical field has seen a large increase in the digital images which are used for diagnosis and therapy. To maintain and access these huge image repositories are difficult to manage. The images are accessed using the identification number or study characteristics [1]. Content-based image retrieval (CBIR) is a technique which retrieves images from a large scale databases according to user interests by using the visual contents of the images. It has been a fast growing research area. The previous techniques utilized textual annotations to retrieve the image which have many drawbacks. The early text-based image retrieval method surveys can be found in [2, 3]. Content-based image retrieval uses the features such as color [7], shape [8], texture [9] and spatial layout of an image to represent and index the image. During 1990, large interest is shown in this technique ranging from government [4], industry [5, 6] to universities [10]. Retrieving of medical images is also one of the important applications of CBIR [11]. The main goal of the medical information systems is to give the required information at the right time to improve the quality and efficiency of the care process [12]. CBIR has played a greater role in providing the required information for clinical decision making system. The medical images are now growing larger in number and to store these images and transmission is also a challenging factor. The image needs to be transferred from one place to another in a case where a specialist may not be available. The similar images have to be retrieved from the database so that the professional can take a decision and guide for the further treatment.

Feature selection searches and extracts a subset of “relevant features” so that effectiveness of classification accuracy can be optimized. By applying feature selection, only a compact subset of highly relevant features is retained, to simplify classification process and make it more effective [71].

To store and transmit the data efficiently we need to reduce the image data which is done by the method of compression. It removes the irrelevant and redundant data. It compress the image without degrading the image [13]. Hence by image compression space is saved and hence more images can be stored and also the same time the transmission of the image is also easy as required bandwidth reduces and it can be downloaded swiftly. Image compression is of two types lossy and lossless. Medical images, clip arts, technical drawings uses lossless

methods[14]. Natural images can utilize the lossy compression methods as minor loss of information is much acceptable. If the lossy compression can produce unnoticeable differences then these are called as visually lossless [15]. Spatial, temporal and spectral are the different types of redundancies [16]. spatial redundancies are focused in this research as these are foremost type of medical imaging data used [17].

LITERATURE REVIEW

Yang et al., [18] presented a boosting framework for distance metric learning that aims to preserve both visual and semantic similarities. The boosting framework first learns a binary representation using side information. Syam et al., [19] proposed a CBIR that using Medical images for retrieval and the feature extraction is used along with color, shape and texture feature extraction to extract the query image from the database medical images. The proposed CBIR technique is evaluated by querying different medical images and the retrieval efficiency is evaluated in the retrieval results. Han et al., [20] proposed a novel Back Propagation -based image retrieval (BPBIR) system for image retrieval. Kumar & Kumaraswamy [21] proposed to implement a novel feature selection mechanism using Discrete Sine Transforms (DST) with Information Gain for feature reduction.

Peter et al., [61] developed a feature selection approach by projecting the image into sub-space based on Fisher's Linear Discriminant. The proposed method was effective in separating classes in a low-dimensional subspace. Chengjun & Harry [62] developed an enhanced Gabor Fisher Classifier (GFC) for face recognition. Shiguang et al., [63] employed extended fisher face for face recognition. This was done to overcome the problem of inability to apply fisher face where single image is available for training. Ling et al., [64] used a set of fisher scores for calculating through partial derivative analysis of the parameters estimated in each HMM. These fisher scores are further combined with traditional features such as log likelihood and appearance based features to form feature vectors that exploit the strengths of both local and holistic features of human face. Xiao-Ming et al., [65] developed an extended fisher face with 3D Morphable model. Dinakardas et al., [66] developed Inverse Fisher Discriminant Analysis which combined PCA plus LDA techniques that has two- phrase framework to deal with high dimensional space and singular cases. Zhuang et al., [67] Inverse fisher criterion was augmented by adding a constraint in PCA procedure so that singularity phenomenon will not occur. Nithya & Menaka [68] proposed Rough set theory (RST) based feature selection technique for removing irrelevant features and producing high accuracy for post processing data.

Huda et al., [69] proposed two filter and wrapper hybrid approaches for feature selection techniques by combining the filter's feature ranking score in the wrapper stage. The first approach hybridizes a Mutual Information (MI) based Maximum Relevance (MR) filter ranking heuristic with an Artificial Neural Network (ANN) based wrapper approach where Artificial Neural Network Input Gain Measurement Approximation (ANNIGMA) has been combined with MR (MR-ANNIGMA) to guide the search process in the wrapper. The second hybrid combines an improved version of MI based (Maximum Relevance and Minimum Redundancy; MaxRel-MinRed) filter ranking heuristic with the wrapper heuristic ANNIGMA (MaxRel-MinRed-ANNIGMA). Unler et al., [70] presented a hybrid filter-wrapper feature subset selection algorithm based on Particle Swarm Optimization (PSO) for support vector machine (SVM) classification. This hybrid algorithm, called maximum relevance minimum redundancy PSO (mr2PSO), uses the MI available from the filter model to weigh the bit selection probabilities in the discrete PSO.

Compressed medical image retrieval has not been studied extensively in literature. In this paper, investigation has been carried out for classification of compressed medical images.

METHODOLOGY

In this work, medical images were compressed using Haar. The compressed images were segmented and features extracted using Gabor filter for texture, Sobel edge detector for edge features and fusion of Gabor and edge features using concatenation. Features are selected using Fisher Score and mRMR feature selection methods. Recurrent Neural Network (RNN) is used for classification of the compressed images. Figure 1 shows the proposed framework to retrieve medical images. The techniques used are detailed in subsequent subsections.

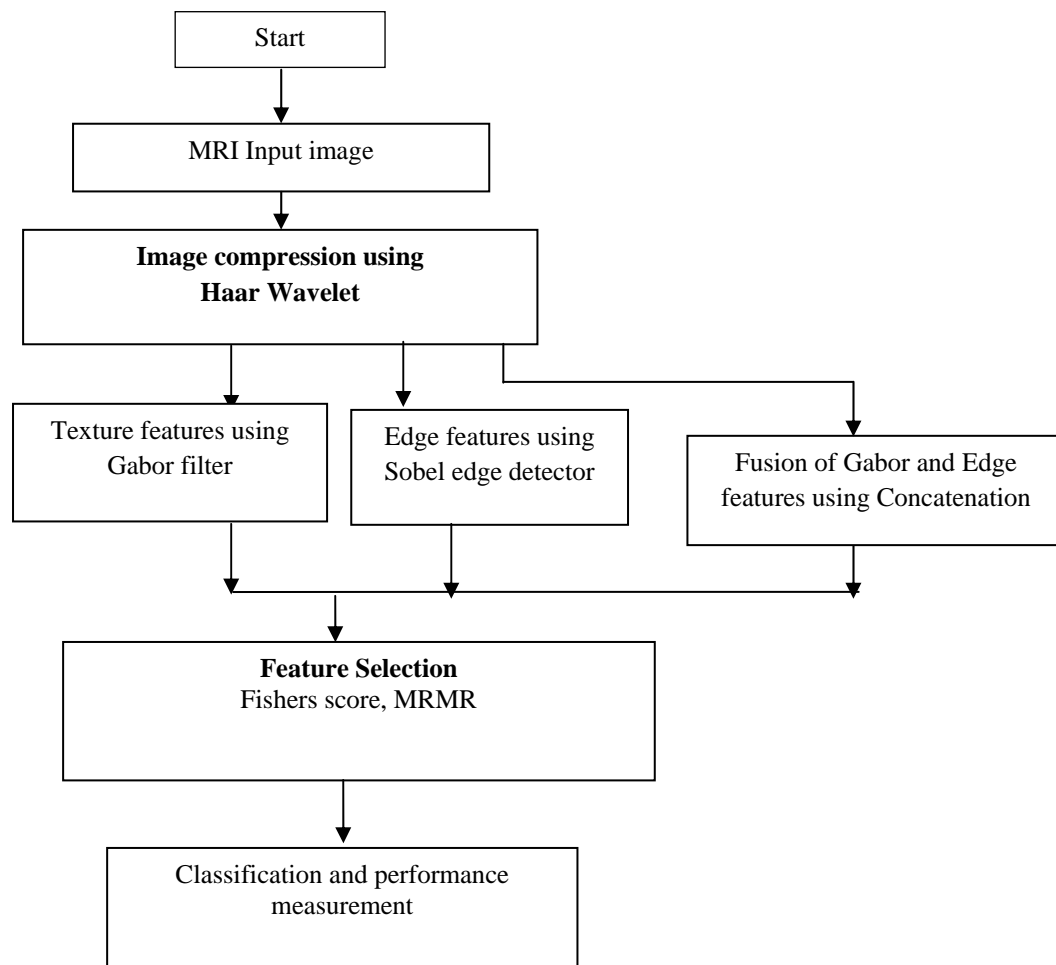


Figure 1: Flowchart of Proposed Methodology

2.1 Image Compression

Haar wavelet is the simplest wavelet introduced by Hungarian Mathematician Alfred Haar [22]. The use of the wavelet was done in 1981 by Jean Morlet a Geophysicist. Morlet and the physicist Alex Grossman invented the term wavelet in 1984 [23]. Haar wavelet is orthogonal wavelet, compact and dyadic [24, 25]. The Wavelet transforms the image from the space domain to a local frequency domain [22, 26]. It requires low computing requirements. The operation of the Haar wavelet operates on the adjacent elements and calculates the sum and the differences of the data.

The mother wavelet function of Haar $\psi(t)$ can be given as:

$$\psi(t) = \begin{cases} 1 & t \in (0, 1/2) \\ -1 & t \in (1/2, 1) \\ 0 & t \notin (0, 1) \end{cases}$$

And the scaling function can $\phi(t)$ can be described as

$$\phi(t) = \begin{cases} 1 & t \in (0, 1) \\ 0 & t \notin (0, 1) \end{cases}$$

To decompose the signal Hilbert transform was used and the signal was decomposed into two components average (approximation) and detail (fluctuation).

A signal having 2^n sample values, the first average sub signal $a^1 = (a_1, a_2, a_3, \dots, a_{N/2})$ for a signal length of N is given as: $a_n = \frac{x_{2n-1} + x_{2n}}{\sqrt{2}}, n=1, 2, 3, \dots, N/2$. The first detail sub signal $d^1 = (d_1, d_2, \dots, d_{N/2})$. A resultant matrix M of first level is formed and shown below

A represents approximation area, H -horizontal area, V - vertical area and D -diagonal area [27].

$$M = \begin{bmatrix} t_{11} & t_{12} & \vdots & t_{13} & t_{14} \\ t_{21} & t_{22} & \vdots & t_{23} & t_{24} \\ \dots & \dots & \vdots & \dots & \dots \\ t_{31} & t_{32} & \vdots & t_{33} & t_{34} \\ t_{41} & t_{42} & \vdots & t_{43} & t_{44} \end{bmatrix}$$

$$A = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \quad H = \begin{pmatrix} t_{13} & t_{14} \\ t_{23} & t_{24} \end{pmatrix}$$

$$V = \begin{pmatrix} t_{31} & t_{32} \\ t_{41} & t_{42} \end{pmatrix} \quad D = \begin{pmatrix} t_{43} & t_{44} \\ t_{53} & t_{54} \end{pmatrix}$$

2.2 Feature Extraction

Texture: Gabor filters [28,29,30] is most widely used for texture feature extraction as it performs better and correspond to the characteristics of the illustration for the edge detection[31,32]. The images are rotated in certain directions to measure the texture and information may be lost during the process of rotation of image, shifts or scale changes [33]

Gabor filtered output of the image is obtained by the convolution of the image with Gabor function for each of the orientation spatial frequency (scale) orientation (Fig. 2). Given an image $F(x, y)$, we filter this image with $Gab(x, y, W, \theta, \sigma_x, \sigma_y)$ [59]

$$FGab(x, y, W, \theta, \sigma_x, \sigma_y) = \sum_k \sum_l F(x-k, y-l) * Gab(x, y, W, \theta, \sigma_x, \sigma_y)$$

The magnitudes of the Gabor filters responses are represented by three moments

$$\mu(W, \theta, \sigma_x, \sigma_y) = \frac{1}{XY} \sum_{x=1}^X \sum_{y=1}^Y FGab(x, y, W, \theta, \sigma_x, \sigma_y)$$

$$std(W, \theta, \sigma_x, \sigma_y) =$$

$$= \sqrt{\sum_{x=1}^X \sum_{y=1}^Y \|FGab(x, y, W, \theta, \sigma_x, \sigma_y) - \mu(W, \theta, \sigma_x, \sigma_y)\|^2}$$

$$Skew = \frac{1}{XY} X \sum_{x=1}^X \sum_{y=1}^Y \left(\frac{FGab(x, y, W, \theta, \sigma_x, \sigma_y) - \mu(W, \theta, \sigma_x, \sigma_y)}{std(W, \theta, \sigma_x, \sigma_y)} \right)^3$$

The feature vector is constructed using $\mu(W, \theta, \sigma_x, \sigma_y)$, $std(W, \theta, \sigma_x, \sigma_y)$ and Skew as feature components.

Sobel Edge Detector: To Identify and locate sharp discontinuities in an image, due to sudden pixel intensity change, characterizing object boundaries within a scene is called edge detection [35]. To perform the Edge detection it applied the Sobel method. The gradient in X-direction and Y- direction is calculated using two 3X3 size masks [34].

The array is given below

-1	0	+1
-2	0	+2
-1	0	+1

Gx

+1	+2	+1
0	0	0
-1	-2	-1

Gy

Figure 2: Convolution kernels of Sobel detector

The horizontal and vertical gradient (Gx and Gy), are calculated and combined together to find the absolute magnitude of the gradient. The absolute magnitude is given by

$$|G| = \sqrt{G_x^2 + G_y^2}$$

The shape descriptors [59] are used to extract the features from the objects and described here are computes a no of circularity, discontinuity angle irregularity, aspect ratio, length irregularity, sharpness, complexity, right-angleness, directedness. Those are translation, rotation (except angle), and scale invariant shape descriptors. It can also use the edges to extract the image contours. Shape information can be obtained from the contour. From individual contour a set of shape features can be extracted and stored. Figure 3 represents these shapes.

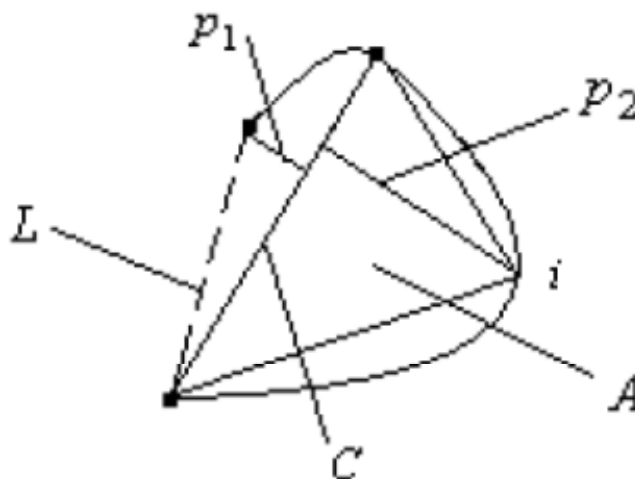


Figure3. Shape and measures used to compute features

$$1) \text{ Circularity } cir = \frac{4PA}{P^2}$$

$$2) \text{ Aspect Ratio } ar = \frac{p1 + p2}{C}$$

$$3) \text{ Discontinuity Angle Irregularity } dar = \sqrt{\frac{\sum |\theta_i - \theta_{i+1}|}{2\pi(n-2)}}$$

$$4) \text{ Length Irregularity } lir = \frac{\sum |L_i - L_{i+1}|}{K}$$

$$5) \text{ Complexity } com = \frac{-3}{n}$$

$$6) \text{ Right-Angleness } ra = r/n$$

$$7) \text{ Sharpness } sh = \sum \frac{\max(0.1 - (\frac{2|\theta - \pi|}{\pi})^2)}{n}$$

$$8) \text{ Directedness } dir = \frac{M}{\sum P_i}$$

Where: n - number of sides of polygon, A - area of polygon, P - perimeter of polygon,

C-length of longest boundary chord, p1, p2 - greatest perpendicular distances from longest chord to boundary, in each half-space either side of line through longest chord,

θ_i - discontinuity angle between (i-1)-th and i-th boundary segment, r – number of discontinuity angles = a right-angle within a specified tolerance, and M - total length of straight-line segments parallel to mode direction of straight-line segments within a specified tolerance.

2.3 Feature Selection

In the real world classification problems require supervised learning techniques are required where the underlying class probabilities and class-conditional probabilities are unknown, and each instance is associated with a class label [36]. Unknown relevant features are a priori. An irrelevant feature will not affect the target. Redundant feature does not add anything new to the target [37]. Redundant features can be removed so that running time of algorithm can also be reduced. It gives a better sight into the real world classification problem [38, 39].

The following lists those that are conceptually different and cover a range of definitions [36].

- a) Idealized: Necessary and sufficient minimally sized feature subset must be find out [40].
- b) Classical: Selecting a subset of D features from a set of S features, $D < S$, such that the value of a standard function is optimized over all subsets of size M [41].
- c) Improve Prediction accuracy: The selected features should improve the prediction accuracy [39].
- d) Approximating original class distribution: the selected features class distribution should be close to the original class distribution given all feature values [39].

The main criteria of the feature selection is classification accuracy should not decrease and the selected features class distribution should be close to the class distribution of all the features. Feature selection methods searches for best features using some evaluation function.

The feature selection method available in the literature are: filter-based, wrapper based and embedded methods [42]. The features are ranked as pre-processing step prior to learning algorithm and high ranking scores are selected in Filter- based methods. In Wrapper based methods features are selected using the by giving the scores to the features and these scores are given using a learning algorithm. In Embedded methods feature selection is combined with the learning algorithm. In this paper we used Fisher Score and mRMR (minimum redundancy – maximum relevance)

Fisher Score:

The Fisher Score [43] is a method to find the subset of most relevant features for classification. In order to obtain the best features it utilizes the discriminative methods, and generative statistical models. The features are selected using the principle that distance between data points lying in same class should be small and distance between data points lying in different classes should be large. Let us consider in general that the given data set is $\{(x_i, y_i)\}_{i=1}^n$ where the $x_i \in R^d$ and $y_i \in \{1,2,3,\dots,c\}$. The aim is to find a subset of m most important informative features. The data matrix is represented as $X=[x_1, x_2, \dots, x_n] \in R^{d \times n}$ -1 is a vector of all ones and 0 is vector of all zeros. I is the Identity matrix. So given f features it have to choose a subset of $m < f$ features. Consider the input data matrix is $A \in R^{d \times n}$ reduces to $B \in R^{m \times n}$ m features are selected. The Fisher Score can be computed as [44]:

$$F(Z) = \text{tr}\{(\overline{S_b})(\overline{S_t} + \gamma I)^{-1}\}$$

Where γ –positive regularization parameter, $\overline{S_b}$ –between- class scatter matrix, $\overline{S_t}$ - total scatter matrix

$$\overline{S_b} = \sum_{k=1}^c n_k (\overline{\mu_k} - \overline{\mu})(\overline{\mu_k} - \overline{\mu})^T$$

$$\overline{S_t} = \sum_{i=1}^n (z_i - \overline{\mu})(z_i - \overline{\mu})^T$$

Where $\overline{\mu_k}$ and n_k are mean vector and size of K-th class. The overall mean vector of reduced data is given by $\overline{\mu} = \sum_{k=1}^c n_k \overline{\mu_k}$. let μ_k^j and σ_k^j be the mean and standard deviation of k-th class related to the j-th feature μ^j and σ^j represent the mean and standard deviation of the set of the j-th feature. The Fisher score is given by: $F(x^j) = \frac{\sum_{k=1}^c n_k (\mu_k^j - \mu^j)^2}{(\sigma^j)^2}$ where $(\sigma^j)^2 = \sum_{k=1}^c n_k (\sigma_k^j)^2$

The features with top scores are selected by the Fisher Score.

The generalized Fisher Score Algorithm:

Input : C and m;

Output : V and

Initialize $V = \frac{1}{n} \mathbf{1}_n \mathbf{1}_c^T$ *and* $t = 1$;

Find the most violated constraint p^1 ,

and set $\Omega_1 = \{p^1\}$;

repeat

Initialize $\lambda = \frac{1}{t}$;

repeat

Solve V *using*

$$V = \left(\frac{1}{\gamma} \sum_{t=1}^{|\Omega_t|} \lambda_t \sum_{j=1}^d p_j^t K_j + I \right)^{-1} H$$

under the current λ ;

Solve for V *using gradient descent as in*

$$\nabla_{vg}(\lambda, V) = -\frac{1}{2\gamma} \text{tr} \left(V^T \sum_{j=1}^d p_j^t K_j V \right)$$

until converge

Fine the most violated constraint p^{t+1} *and set*

$$\Omega_{t+1} = \Omega_t \cup p^{t+1};$$

$t = t + 1$;

until converge

Maximum Relevance and Minimum Redundancy (mRMR)

The mRMR [45] method is used for feature selection. Features are selected by ranking them based on its relevance to target and redundancy of features. The best features are selected such that it has minimum redundancy within the features and maximum relevance to target variable [46].

Mutual information (MI) is a measure of general interdependence between random variables [47]. It is used to quantify both relevance and redundancy. The MI is defined as [46]

$$I(X, Y) = \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)}$$

Where X, Y are vectors, joint probabilistic density is represented by $p(x, y)$ marginal probabilistic densities are given by $p(x)$ and $p(y)$

Given the N data points drawn from the joint probability distribution (x_i, y_i) , $i = 1, \dots, N$, Gaussian kernel estimator was used to estimate the joint and marginal densities [48,49]

$$P(x, y) = \frac{1}{M} \sum \frac{1}{2\pi h^2} e^{-\frac{1}{2h^2}((x-x_i)^2 + (y-y_i)^2)}$$

And the $P(x)$, $P(y)$ can be defined as

$$P(x) = \frac{1}{M} \sum \frac{1}{\sqrt{2\pi}h} e^{-\frac{1}{2h^2}(x-x_i)^2}$$

And

$$P(y) = \frac{1}{M} \sum \frac{1}{\sqrt{2\pi h^2}} e^{-\frac{1}{2h^2}(y-y_i)^2}$$

the width of the kernels are controlled by using h.

Let us denote the whole feature set as β . β_s Denotes the already selected feature set which contains n features. β_t Represent to be selected feature set which contains m features. The relevance S of the feature f in β_t having the target c is calculated by $S = I(f, c)$ and redundancy R of the feature f in β_t of all features in β_s is calculated by

$$R = \frac{1}{m} \sum_{f_i \in \beta_s} I(f, f_i)$$

The features f_j are obtained from the β_t with maximum relevance and minimum redundancy, S and R are combined with the mRMR function:

$$\max_{f_j \in \beta_s} [I(f_j, c) - \frac{1}{m} \sum_{f_i \in \beta_s} I(f_j, f_i)] \quad (j = 1, 2, \dots, n)$$

The feature set is obtained by mRMR method

$$F = \{f'_1, f'_2, \dots, f'_h, \dots, f'_N\}$$

With the N features $N = (m+n)$. Smaller index h values are obtained when the features extracted are better being the index indicating the importance of the feature.

Classification

The selected feature set is used for the classification of the compressed image.

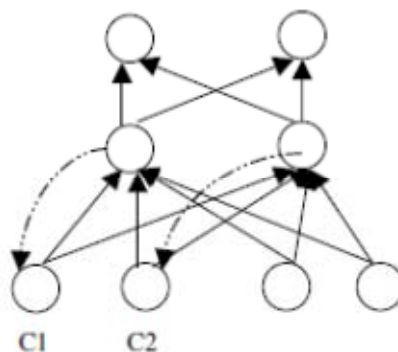
The extracted features are classified using Recurrent Neural Network (RNN). A recurrent net is a neural network with feedback connections [50, 51]. In literature many RNN are defined such as Hopfield, Elman, Boltzmann machine, BAM, and recurrent back propagation nets [52].

These are applied in many fields including classification trajectory problems, language processing, filtering and so on.

Recurrent neural networks architectures can be in many forms with two common features:

1. Incorporating some form of Multi-Layer Perceptron as a sub-system.
2. Exploiting Multi-Layer Perceptron's powerful non-linear mapping capabilities, plus some form of memory.

RNN architectures include multilayer feed forward networks with input, hidden and output layers. Each node has input from other nodes. Feedback to the node is possible. In simple RNNs (Figure 4) some nodes form part of a feed forward structure, others provide sequential context and receive feedback from other nodes. Context units C1 and C2 weights are processed similar to those for input units. Time delayed feedback is received from second layer to the context unit that is shown in Fig 4.



Elman [53] introduced feedback from hidden layers to the input layers context portion. This method emphasizes input values sequence. Jordan recurrent neural networks [54] use output layer feedback to input

layer context nodes emphasizing output values sequences. . A learning algorithm, Error Back-Propagation through Time (BPTT) is used for RNN and derived from classic BP to be used on FFNN [55, 56]. Compared to the other learning algorithms, implementing BPTT is relatively easy. The basic of the BPTT algorithm is transformation of RNN “unfolded” to the FFNN comprising more layers, which can be applied to classic BP with minor changes. By the unfolding of RNN, every newly produced layer of FFNN signifies one recursion in time of RNN considered by T - depth of unfolding of RNN.

In the prompt units of time updation of the activation states occur. The initial step consists of recounting method of network that consists in replicating times (folds) the RNN obtaining an equivalent FFNN and at all times, each replicated connection shares its value [57]. Using the Backpropagation algorithm this FFNN is trained .It is BPTT algorithm.

The output time of each neuron of back propagation is defined as

$$y_i(t) = f(x_i(t))$$

$$x_i(t) = \sum_{j \in H} y_j(t)w_{ij} + \sum_{j \in I} x_j^{in}(t)w_{ij} + \sum_{j \in M} y_j(t - \tau_{ij})w_{ij}$$

where f denotes neuron activation function, input and hidden neuron indexes are represented by I and H. M is the indexes of neurons which contains Information about the previous stages given by the M which is the indexes of neurons τ_{ij} is an integer representing displacement in recurrent connections through time [58].

Since the top 50 features obtained from Fisher Score and mRMR methods and for classification the Partial Recurrent Neural Network is used in which the size of the input layer was fixed at 50 with one hidden layer consisting of 20% of the neurons used in the input layer.

RESULTS AND DISCUSSION

The recurrent neural network is evaluated using a dataset containing 4 different class of image with 750 images. The experiments are run 10 fold cross validation. The Haar Wavelet is used for image compression. Gabor filters are used to extract texture features. Sobel edge detector is used for edge features .Feature selection is achieved through Fisher Score and mRMR methods. The top 50 ranked features of Fisher Score and mRMR are utilized for classifying the medical images.

The design parameters of the proposed recurrent neural network are given in Table 2.

Table 1: Parameters Used in the Proposed Model

Number of Hidden Layer	1
Number of Neurons in hidden layer	10
Activation function of hidden layer	Tanh
Learning Rule	BPTT
Step Size	0.01
Momentum	0.5

Sample medical images used in the experimental setup are shown in Figure 5.

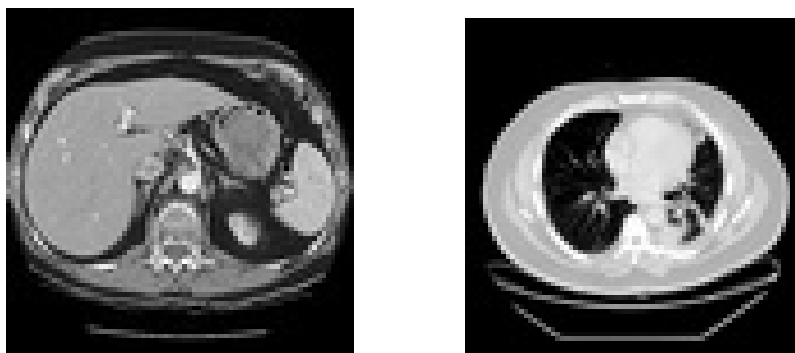


Figure 5: Images used in the investigation.

The classification accuracy is evaluated using RNN [60].and the figure shows the classification accuracy of the RNN using the Fisher Score and mRMR feature selection methods. It also shows the different methods used in the feature extraction techniques.

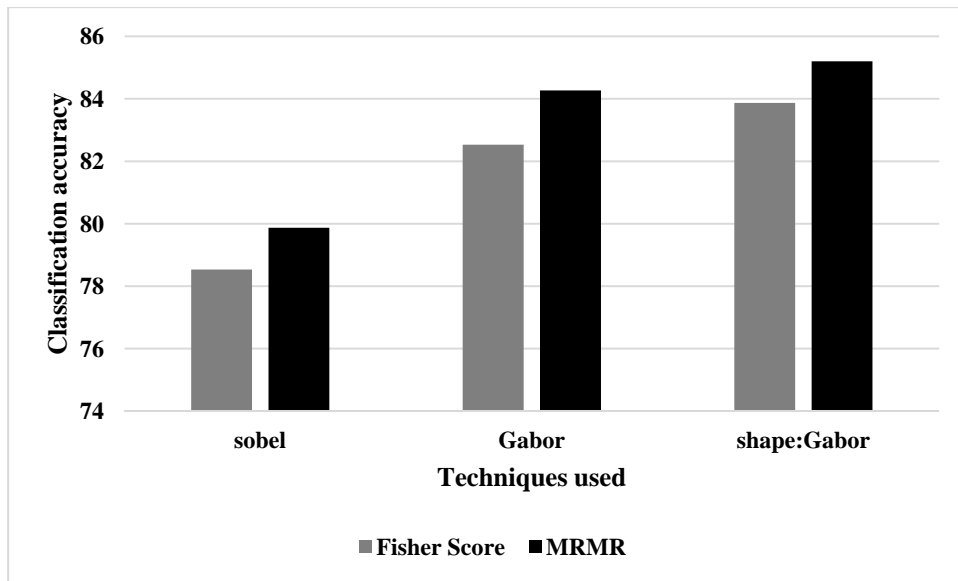


Figure 6 Classification accuracy

From the figure 6, it can be observed that the Fisher score, the sobel method improves classification accuracy by 4.96% and 6.57% when compared with Gabor and shape: Gabor methods. Similarly MRMR, the sobel method increased classification accuracy by 5.36% and 6.45% when compared with Gabor and shape: Gabor methods.

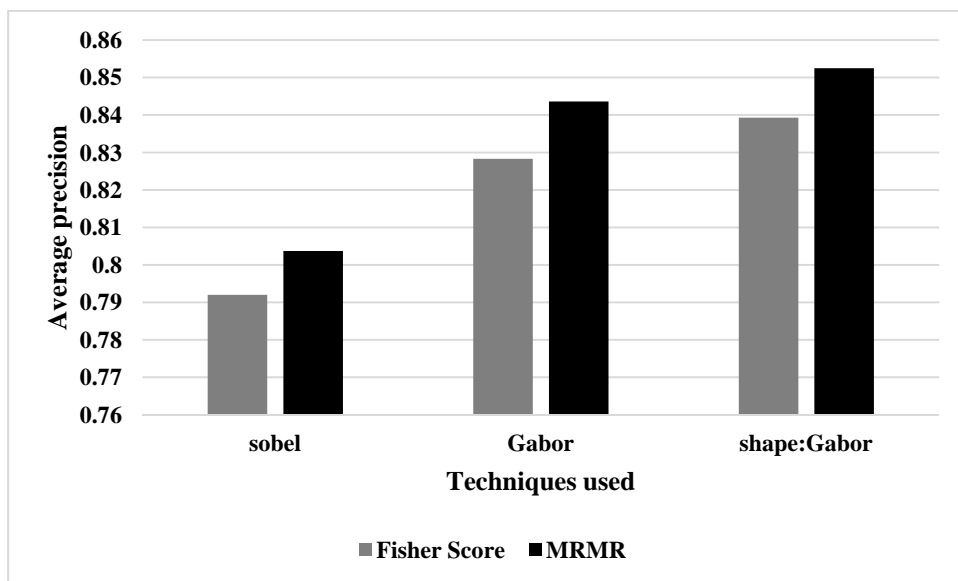


Figure 7 Average Precision

From the figure 7, it can be observed that the Fisher score, the sobel method improves average precision by 4.48% and 5.79% when compared with Gabor and shape: Gabor methods. Similarly MRMR, the sobel method increased average precision by 4.84% and 5.89% when compared with Gabor and shape: Gabor methods.

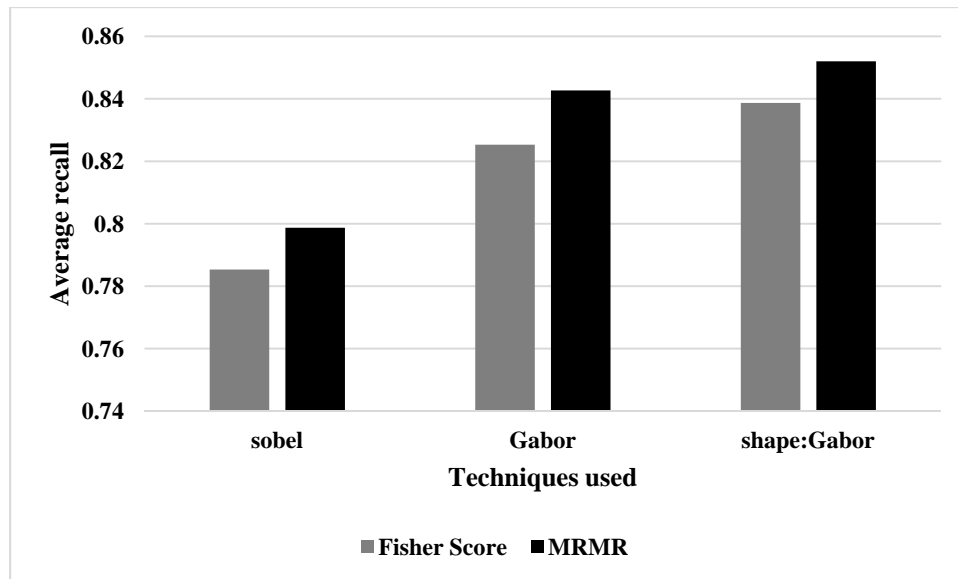


Figure 8 Average Recall

From the figure 8, it can be observed that the Fisher score, the sobel method improves average recall by 4.96% and 6.57% when compared with Gabor and shape: Gabor methods. Similarly MRMR, the sobel method increased average recall by 5.36% and 6.45% when compared with Gabor and shape: Gabor methods.

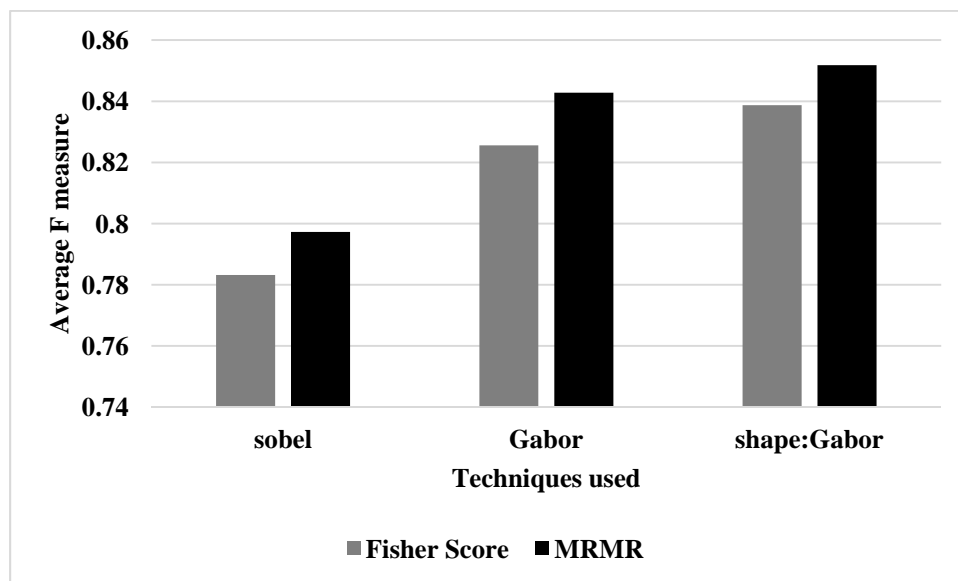


Figure 9 Average F Measure

From the figure 9, it can be observed that the Fisher score, the sobel method improves average f measure by 5.27% and 6.84% when compared with Gabor and shape: Gabor methods. Similarly MRMR, the sobel method increased average f measure by 5.54% and 6.60% when compared with Gabor and shape: Gabor methods.

CONCLUSION

The proposed method integrates image retrieval to retrieve diagnostic cases similar to the query medical image and image compression techniques to minimize the bandwidth utilization. Haar Wavelet is used for obtaining the visually lossless image compression. Shape and texture features are extracted from the medical compressed medical images using Gabor transforms and Sobel edge detector. 50 top ranked features selected through Fisher Score and mRMR Techniques with ranking are used for classification. The mRMR Feature selection method has performed better than the Fisher Score method. Using the selected features from these methods the mRMR features have shown better performance in classification of compressed medical images using the Partial Recurrent neural Network. Simulation results achieved are satisfactory.

REFERENCES

- [1] T.M. Lehmann, M.O. Güld, C. Thies, B. Fischer, M. Keyzers, D. Kohnen, H. Schubert, B.B. Wein, Content-based image retrieval in medical applications for picture archiving and communication systems, in: Proceedings of the SPIE Conference on Medical Imaging, vol. 5033, San Diego, CA, USA, 2003
- [2] S. K. Chang, and A. Hsu, "Image information systems: where do we go from here?" IEEE Trans. On Knowledge and Data Engineering, Vol.5, No.5, pp. 431-442, Oct.1992.
- [3] H. Tamura, and N.Yokoya, "Image database systems: A survey, " Pattern Recognition, Vol.17, No.1.pp. 29-43, 1984.
- [4] R. Jain,A.Pentland, and D.Petkovic, "NSF-ARPA workshop on visual information management systems",Cambridge,MA,June1995
- [5] Zobeir Raisi, Farahnaz Mohanna, Mehdi Rezaei "Applying Content-Based Image Retrieval Techniques to Provide New Services for Tourism Industry" Int. J. Advanced Networking and Applications Volume: 6 Issue: 2 Pages: 2222-2232 (2014) ISSN : 0975-0290.
- [6] James Dowe, Content-based retrieval in multimedia imagining" in Proc. SPIE Storage and Retrieval for Image and Video Databases.
- [7] M.J. Swain, D.H. Ballard, "color indexing" Int. J Computer Vision, Vol. 7, no.(1), pp. 11-32, 1991.
- [8] H. Tamura, S. Mori, T. Yamawaki, 'Textural Features Corresponding to Visual Perception', IEEE Transactions on Systems Man and Cybernetics, Vol 8, pp. 460-473, 1978
- [9] Y. Liu, D. Zhang, G. Lu, W.-Y. Ma, 'A survey of content-based image retrieval with high-level semantics', Pattern Recognition , Vol. 40, (1), pp. 262-282, 2007.
- [10] W.Y.Ma andB.S.Manjunath, "Netra: A toolbox for navigating large image databases" in Proc. IEEE Int Conf. on Image Proc. 1997.
- [11] Samuel, G., Armato III, et al.: Lung image database consortium – Developing a resource for the medical imaging research community, in Radiology . 232, 739-748 (2004)
- [12] Winter, A and Haux, R. A three-level graph-based model for the management of hospital informationsystems. Methods Information Med. 1995; 34: 378–396
- [13] Dezhgoshka, K., Sylla, A. K., &Ngouyassa, E. (1994, May). Lossless and lossy image compression algorithms for on-board processing in spacecrafts. InAerospace and Electronics Conference, 1994. NAECON 1994., Proceedings of the IEEE 1994 National (pp. 416-423). IEEE.
- [14] Srikanth, R., &Ramakrishnan, A. G. (2005). Contextual encoding in uniform and adaptive mesh-based lossless compression of MR images. Medical Imaging, IEEE Transactions on, 24(9), 1199-1206.
- [15] ME, S. S., Vijayakumar, V. R., &Anuja, R. (2012). A Survey on Various Compression Methods for Medical Images. International Journal of Intelligent Systems and Applications (IJISA), 4(3), 13.
- [16] Xiong, Z., Wu, X., Cheng, S., &Hua, J. (2003). Lossy-to-lossless compression of medical volumetric data using three-dimensional integer wavelet transforms.Medical Imaging, IEEE Transactions on, 22(3), 459-470.
- [17] Zukoski, M. J., Boulton, T., &Lyriboz, T. (2006). A novel approach to medical image compression. International journal of bioinformatics research and applications, 2(1), 89-103.
- [18] Yang, L., Jin, R., Mummert, L., Sukthankar, R., Goode, A., Zheng, B., &Satyanarayanan, M. (2010). A boosting framework for visually-preserving distance metric learning and its application to medical image retrieval. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32(1), 30-44.
- [19] Syam, B., Victor, J. S. R., &Rao, Y. S. (2013, March). Efficient similarity measure via Genetic algorithm for content based medical image retrieval with extensive features. In Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on (pp. 704-711). IEEE.
- [20] Han, J. H., Huang, D. S., Lok, T. M., &Lyu, M. R. (2005, July). A novel image retrieval system based on BP neural network. In Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on (Vol. 4, pp. 2561-2564). IEEE.
- [21] Kumar, M. S., &Kumaraswamy, Y. S. (2012, March). An improved support vector machine kernel for medical image retrieval system. In Pattern Recognition, Informatics and Medical Engineering (PRIME), 2012 International Conference on (pp. 257-260). IEEE.
- [22] Haar A.: ZurTheorie der orthogonalenFunktionensysteme. MathematischeAnnalen, 69, 331–371.1973.
- [23] Chun-Lin, Liu "A Tutorial of the Wavelet Transform" February 23, 2010
- [24] Castleman KR. Digital image processing. Englewood Cliffs: Prentice-Hall; 1996.
- [25] Moharir PS. Pattern recognition transforms. New York: Wiley; 1992.
- [26] Kamrul Hasan Talukder , Koichi Harada , Haar Wavelet Based Approach for Image Compression and Quality Assessment of Compressed Image, in IAENG International Journal of Applied Mathematics, pp 49-56, 2007.
- [27] E.vamsidhar and Kiran Kumar Reddi, An Data Mining Approach for Compressed Medical Image Retrieval,"International journal of computer applications(IJCA),Volume 52-No.August 2012.
- [28] D.M. Squire, W. Müller, H. Müller, T. Pun, Content-based query of image databases: in-spirations from text retrieval, Pattern Recognition Letters, vol. 21, 2000, pp. 1193–1198 (Selected papers from the 11th Scandinavian Conference on Image Analysis, SCIA'99).
- [29] W. Ma, B. Manjunath, Texture features and learning similarity, in: Proceedings of the 1996 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'96), San Francisco, California, 1996, pp. 425–430.
- [30] S. Santini, R. Jain, Gabor space and the development of preattentive similarity, in: Proceedings of the 13th International Conference on Pattern Recognition (ICPR'96), IEEE, Vienna, Austria, 1996, pp. 40.
- [31] Daugman, J.G. An information theoretic view of analog representation in striate cortex. Computat. Neurosci.1990; 2: 9–18
- [32] Daugman, J.G. High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. Machine Intel. 1993; 15: 1148–1161.
- [33] Milanese, R and Cherbuliez, M. A rotation, translation and scale-invariant approach to content- based image retrieval. J. Visual Commun. Image Represent. 1999; 10: 186–196
- [34] Rezai-Rad, G. and M. Aghababaei (2006). Comparison of SUSAN and Sobel Edge Detection in MRI Images for Feature Extraction. in Information and Communication Technologies. ICTTA '06. 2nd. pp 1103 - 1106.
- [35] Pal, N. R., & Pal, S. K. (1993). A review on image segmentation techniques. Pattern recognition, 26(9), 1277-1294
- [36] M.Dash,H.Liu "Feature Selection for Classification" in Intelligent Data Analysis I (1997) 131-156.
- [37] John, G.H., Kohavi, R. and Pfleger, K., Irrelevant features and the subset selection problem. In: Proceedings of the Eleventh International Conference on Machine Learning, 121-129, 1994.
- [38] Kohavi, R. and Sommetfield, D., Feature subset selection using the wrapper method: Overfitting and dynamic search space topology. In: Proceedings of First International Conference on Knowledge Discovery and Data Mining, Morgan Kaufmann, 192-197, 1995.
- [39] Koller, D. and Sahami, M., Toward optimal feature selection. In: Proceedings of International Conference on Machine Learning, 1996.
- [40] Kira, K. and Rendell, L.A., The feature selection problem: Traditional methods and a new algorithm. In: Proceedings of Ninth National Conference on Artificial Intelligence, 129-134, 1992.
- [41] Narendra, P.M. and Fukunaga, K., A branch and bound algorithm for feature selection. IEEE Transactions on Computers, C- 26(9):917-922, September 1977.

- [42] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003.
- [43] P. E. H. R. O. Duda and D. G. Stork. *Pattern Classification*. Wiley-Interscience Publication, 2001.
- [44] Quanquan Gu, Zhenhui Li and Jiawei Han “Generalized Fisher Score for Feature Selection”
- [45] Peng H, Long F, Ding C (2005) Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell* 27(8): 1226–1238. doi:10.1109/TPAMI.2005.159
- [46] Yudong Cai et.,al “Prediction of lysine ubiquitination with mRMR feature selectionand analysis” *Amino Acids* DOI 10.1007/s00726-011-0835-0.
- [47] Tourassi, G. D., Frederick, E. D., Markey, M. K., & Floyd Jr, C. E. (2001). Application of the mutual information criterion for feature selection in computer-aided diagnosis. *Medical Physics*, 28(12), 2394-2402
- [48] Beirlant J, Dudewicz E, Gyorfi L, Meulen Evd (1997) Nonparametric entropy estimation: an overview. *Int J Math Stat Sci* 6(1):17–39
- [49] Qiu P, Gentles AJ, Plevritis SK (2009) Fast calculation of pairwise mutual information for gene regulatory network reconstruction. *Comput Methods Programs Biomed* 94(2):177–180. doi:10.1016/j.cmpb.2008.11.003.
- [50] Pal, S. K., &Mitra, S. (1992). Multilayer perceptron, fuzzy sets, and classification. *IEEE Transactions on Neural Networks*, 3(5), 683-697.
- [51] Fausett, L.: *Fundamentals of Neural Networks; Architectures, Algorithms and Applications*, Prentice-Hall, Inc. New Jersey, 07632, 1994.
- [52] Hecht-Nielsen, R. (1989, June). Theory of the backpropagation neural network. In *Neural Networks, 1989. IJCNN., International Joint Conference on* (pp. 593-605). IEEE.
- [53] Elman, J. L.: *Finding Structure in Time*, *Cognitive Science*, 14, 179–211, 1990.
- [54] Jordan M (1986) Attractor dynamics and parallelism in a connectionist sequence machine. In: *Proc of Ninth Annual conference of Cognitive Science Society*. Lawrence Earlbaum, New York, pp 531–546
- [55] Martens, J., &Sutskever, I. (2011). Learning recurrent neural networks with Hessian-free optimization. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)* (pp. 1033-1040).
- [56] Bodén, M. (2002). A guide to recurrent neural networks and backpropagation. The Dallas project, SICS technical report.
- [57] I. J. Koscak., (2010). *StochasticWeight Selection in Backpropagation Through Time*
- [58] Grau, I., Nápoles, G., Bonet, I., &García, M. M. (2013). Backpropagation through Time Algorithm for Training Recurrent Neural Networks using Variable Length Instances. *Computación y Sistemas*, 17(1), 15-24.
- [59] Choras, R. S. (2007). Image feature extraction techniques and their applications for CBIR and biometrics systems. *International journal of biology and biomedical engineering*, 1(1), 6-16.
- [60] E. Vamsidhar, Kiran Kumar Reddi “ Compressed Medical Image Retrieval Using Neural Networks” *International Journal of Applied Engineering Research* ISSN 0973-4562 Volume 10, Number 3 (2015) pp. 8953-8965
- [61] Peter, N. B., Jooao, P. H., & David, J. K. (1997). EigenfacevsFisherfaceRecognition usingclass specific linear projection.*IEEE Transactions on pattern Analysis and machineIntelligence*, 19(7), 711-720.
- [62] Chengjun, L., & Harry, W. (2002). Gabor feature based classification using the Enhanced Fisher Linear Discriminant Model for face Recognition. *IEEE Transaction on image processing*, 11(4), 467-476.
- [63] Shinguang, S. Bo, C., Wen, G., & Debin, Z. (2002). Extended fisherface for face recognition from a single example image per person. FRJDL institute of computing technology .Chinese Academy of science.Department of Computer Science.Harbin Institute of Technology.
- [64] Ling C, Hong M, Ara V. N, (2005). Face recognition based on multi-class mappingofFisher scores. *Pattern Recognition*, 30, 799-811
- [65] Xiao-Ming B, Bao-Cal Y, Qin S, and Yan-Feng S, (2005). Face Recognition Using Extended Fisherface with 3D Morphable Model.*Proceedings of the fourth International Conference on Machine learning and Cybernetics*, Guangahon, 4481-4486.
- [66] Dinakardas, C.N, Perumal S. and Nisha G, (2013). Multimodal identification system in monozygotic twins.*International Journal of Image Processing (IJIP)*, 7(1), 72-80.
- [67] Zhuang, X.-S, Dai, D.-Q, and Yuen, P.C, (2005). *Face Recognition by Inverse Fisher Discriminant Feature*. Springer-Verlag Berlin Heidelberg, 92-98.
- [68] Nithya, P., & Menaka, T. (2013). Feature Selection for Post Processing In High Dimensional Data. *representations*, 2(9).
- [69] Huda, S., Yearwood, J., & Stranieri, A. (2011, January). Hybrid wrapper-filter approaches for input feature selection using maximum relevance-minimum redundancy and artificial neural network input gain measurement approximation (ANNIGMA). In *Proceedings of the Thirty-Fourth Australasian Computer Science Conference-Volume 113* (pp. 43-52). Australian Computer Society, Inc..
- [70] Unler, A., Murat, A., & Chinnam, R. B. (2011). mr 2 PSO: a maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification. *Information Sciences*, 181(20), 4625-4641.
- [71] Lu, L., Liu, M., Ye, X., Yu, S., & Huang, H. (2011, October). Coarse-to-fine classification via parametric and nonparametric models for computer-aided diagnosis. In *Proceedings of the 20th ACM international conference on Information and knowledge management* (pp. 2509-2512). ACM.