International Journal on Emerging Technologies 11(3): 1185-1193(2020)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

Image Compression using Clustering Techniques for Bio Medical Applications

G. Vimala Kumari[1], M. Sunil Prakash[2], P. Surya Prasad [3], B. Srinivas[4] and B. Lavanya[5]
[1]Assistant Professor, Department of ECE, MVGR College of Engineering (A) (Andhra Pradesh), India.
[2]Professor, Department of ECE, MVGR College of Engineering (A),Vizianagaram (Andhra Pradesh), India.
[3]Associate Professor, Department of ECE, MVGR College of Engineering(A), Vizianagaram (Andhra Pradesh), India.
[4, 5]Assistant Professor, Department of ECE, MVGR College of Engineering (A), Vizianagaram (Andhra Pradesh), India.

(Corresponding author: G Vimala Kumari)
(Received 17 April 2020, Revised 12 June 2020, Accepted 24 June 2020)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Demand for communication of multimedia data through the telecommunications network and Internet is growing exponentially. Significant portion of the multimedia data consists of images and they occupy the major portion of the communication bandwidth. The digitization of medical image information is a prime concern for the medical community. In this scenario, compression is an essential component for creating file sizes of manageable and transmittable dimensions. Image compression is nothing but decreasing the size of bytes of the file it can allows the user to store a huge amount of information within a static memory. This paper discusses an innovative proposed compression process to do lossy image compression by using clustering techniques such as K-Means, Fuzzy C-Means and Density Based Spatial Clustering of Applications with Noise (DBSCAN). The main objective is to compress the image where it is compressed to a reasonable size as well as preserving the quality of an image. The method segments the image using clustering techniques. The segmented image is encoded with various encoding techniques like Bit-Plane slicing, LZW coding, and Huffman coding. The enactment of proposed hybrid techniques are compared in terms of enactment parameters such as PSNR, MSE, SSIM and compression Ratio, and the results shows that the proposed technique is useful in Bio medical applications.

Keywords: Image Compression, K-means, Fuzzy c-means, DBSCAN, Lempel-Ziv-Welch (LZW), Peak Signal to Noise Ratio.

I. INTRODUCTION

Now a days, processing of medical images could be developing and necessary field. Every year, terabytes of medicative image information is produced over progressive modalities such as magnetic resonance imaging (MRI), ultrasound (US), computed tomography (CT), X-rays and plenty of more modern techniques of medical imaging. The digitization of medical image information is of huge concentration to the medicative region which might result in the implementation of e-health, telemedicine, and telematics. Thereby, to decrease transmission time and storage costs, effective image data compression systems without degradation of image quality is needed by removing the redundancy and irrelevancy of the image data. Image compression can be attained using many methods and is classified into two categories, one with transform technique and the other without transform technique. With transformed techniques i.e., spatial domain image is transformed to frequency domain with transformation techniques. The transform techniques like Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) are considered as promising tools for image compression for the past two decades. But due to their time consuming nature, high cost, and requirement of more hardware, traditional non-transformed techniques are developed. These techniques include vector quantization, segmentation using clustering techniques, multi-level image thresholding, and back propagation neural network. So, compression in spatial domain is implemented in this work using segmentation based clustering techniques with less complexity. Image segmentation is a process of optimizing similar regions in an image which results in effective clustering of image, hence better image compression is achieved by running cascaded encoding techniques on cluster image. Performance of the encoding techniques depends upon the number of similar clusters and probability occurrence of same cluster centroids in an image respectively. So, its performance depends upon effective clustering technique. An image can be viewed as a geographical dataset that has large quantity of information which when sort out becomes intelligible. The key objective of image segmentation is to process the image in such a way that the image becomes easier to analyse. An image could be represented as a 2-D function g(x, y), where x and y are the co-ordinates of the pixels. The value of g(x, y) gives the values of the intensity of Image present at (x, y). The actual image may not be useful for study. So, image segmentation is often needed for analysing an image where a part of the image must be studied. Image segmentation can be defined as the grouping of pixels in an image into various groups that exhibit homogeneity. The homogeneity of the pixel may be in the form of the intensity of the pixel, density of the pixels or texture. Different clusters should not intersect with one another. Image Segmentation (IMS) has a dynamic part in Image analysis. Many methods for this problem have emerged over the years. One of the important methods to segment the image is clustering technique. The cluster analysis is to dividing image information set into a number of separate collections or clusters. The paper aims to segment the image for better image compression using clustering techniques like K-means, Fuzzy c-means (FCM) and Density Based Spatial Clustering of Applications with Noise (DBSCAN). K-means clustering is one of the widespread methods, since its ease and computational effectiveness. FCM procedure has extra tractability for the Images to go to numerous modules with changing amount of membership. Density based clustering methods are used to find temporal connectivity and colour comparison of the pixels, which is used to determine clusters of arbitrary shape leading to the dividing of pixels and further segregating the noise points.

So in this paper, cluster based IMS is suggested for capable and effective results of clusters in better image compression, which assists to increase the Bit rate for broadcast as well as keep up an considerable quantity of quality or reliability of the image. IMS is a procedure of improving alike community in an image which effects in effective clustering of image, hence improved image compression is attained by successively cascaded Bit-Plane slicing, LZW coding, and Huffman coding on cluster image. Performance of the Bit-Plane slicing, LZW coding, and Huffman coding depends upon the number of similar clusters and probability occurrence of same cluster centroids in an image respectively. So, its performance depends upon effective clustering technique.

II. LITERATURE SURVEY

Image compression is achieved by appropriate image thresholding and these thresholds are obtained with a principal of moment preserving and were proposed by Chen and Wen [1]. The proposed method achieved a high compression ratio with better reconstructed image quality. Kaur *et al*., proposed an image compression method which consumes less time and follows a strategy where thresholds are optimized with optimization techniques for which objective function is distortion [2]. Birge–Massart thresholding is inbuilt thresholding technique which is used for image compression and obtained results are compared with the uni-modal thresholding in terms of reconstructed image quality and compression ratio [3]. In [4] Electrocardiography (ECG) signals are compressed by transforming the signal with the help of discrete wavelet transform. Another kind of image compression where image to be compressed is transformed to frequency domain with the help of bandlet and required bandlet coefficients are obtained with type II Fuzzy thresholding and results are compared with the ordinary thresholding [5]. Prashant and Ioana proposed a non-uniform thresholding and observed the effects of thresholding on reconstructed image quality [6]. Tony and Zhou proposed a technique for noise removal and image compression in wavelet domain thresholding which is based on Partial Differential Equation (PDE) and it takes the advantage of variations in framework [7]. Image compression can also be performed with Multistage Lattice Vector Quantization (MLVQ) and by thresholding of DWT coefficients. Proposed combination tries to minimize the quantization error and its computational complexity is less compared to ordinary VQ [8]. Kavehet et al. proposed a 2-D discrete wavelet transform based image thresholding by optimal thresholding the DWT coefficients with Particle Swarm Optimization (PSO) for image compression. They did three level decomposition of DWT and 62.5% of thresholds are assigned and optimized for the approximation coefficients and the remaining 37.5% equally assigned to horizontal, vertical, and diagonal coefficients [9]. They compared the results with the Set Partition in Hierarchical Tree (SPHIT), Chrysafis, JPEG and JPEG-2000 and proved better in Peak Signal to Noise Ratio (PSNR) and Bits per Pixel (BPP).Image is compressed with the help of Multistage Lattice vector quantization (MLVQ) and threshold the DWT coefficients. The Proposed method always tries to reduce the quantization error and its computational time is less as compared to normal VQ [10]. Recently, to improve the performance of the compression methods several revolutionary algorithms have been developed for the purpose of codebook design with the technique of Vector Quantization. G.V. Kumari *et al*., Flower Pollination Algorithm (FPA) is proposed for codebook design to compress the medical images. FPA is advantageous as it is simple and flexible in terms of a number of parameters. FPA has just one key parameter along with a scaling factor, which makes it easier to execute [11]. Fred, A. Lenin, *et al*. highlighted the application of bat optimization algorithm in medical image compression. The bat optimization algorithm is employed for the optimum codebook design in Vector Quantization (VQ) algorithm. The performance of the BAT-VQ compression scheme was compared with the Classical VQ, Contextual Vector Quantization (CVQ) and JPEG lossless schemes for the abdomen CT images [12]. Aslam *et al.,* proposes an analysis on various methods of optimization with requisition to image processing. Traditional techniques in general are unsuccessful to resolve such large scale problems particularly with nonlinear objective functions. Some optimization techniques frequently fail to resolve optimization problems that have many local optima. To prevail over these problems, there is a need to extend further powerful optimization techniques. The modern techniques are used to solve linear, nonlinear, differential and non-differential optimization problems [13]. Standard stacked denoising autoencoder compression sampling (SDA-CS) approach improves compression sampling process by extracting features functional for disease detection from samples recorded using microscope [14]. Man Hoang et al., proposed a new layered image compression framework with encoder-decoder matched semantic segmentation (EDMS). And then, followed by the semantic segmentation, a special convolution neural network is used to enhance the inaccurate semantic segment. As a result, the accurate semantic segment can be obtained in the decoder without requiring extra bits. The experimental results show that the proposed EDMS framework can get up to 35.31% BD-rate reduction over the HEVC-based (BPG) codec, 5% bitrate and 24% encoding time saving compare to the state-of-the-art semantic-based image codec [15]. Chen *et al*., proposed a Learning based Facial Image Compression (LFIC) framework with a novel Regionally Adaptive Pooling (RAP) module whose parameters can be automatically optimized according to gradient feedback from an integrated hybrid semantic fidelity metric, including a successfully exploration to apply Generative

Adversarial Network (GAN) as metric directly in image compression scheme. The experimental results verify the framework's efficiency by demonstrating performance improvement of 71.41%, 48.28% and 52.67% Bit Rate Saving separately over JPEG2000, WebP and neural network-based codecs under the same face verification accuracy distortion metric [16]. Data transmission of images captured by the smart underwater objects is very challenging due to the nature of underwater environment and necessitates an efficient image transmission strategy for IoUT. Krishnaraj, N., *et al.*, modeled and implemented a discrete wavelet transform (DWT) based deep learning model for image compression in IoUT. For achieving effective compression with better reconstruction image quality, convolution neural network (CNN) is used at the encoding as well as decoding side [17]. An adaptive neuro fuzzy inference system (ANFIS) is developed for segmentation based on the automatic seed point selection range. The pixels intensity of the proposed algorithm is not dependent on the tumor type [18-20]. Li, Mu, *et al.,* presented a content-weighted encoder-decoder model, where the channel-wise multi-valued quantization is deployed for the discretization of the encoder features, and an importance map subnet is introduced to generate the importance masks for spatially varying code pruning. Consequently, the summation of importance masks can serve as an upper bound of the length of bit stream [21]. For image compression, clustering based segmentation is a technique to extract the objects from the background of a scene which benefits for analysis and to examine the image information. In this work, image segmentation is suggested for capable and effective results of clusters in better image compression. Meaningful/useful clusters are possible with clustering techniques, which lead to better image segmentation, results to objective of image compression. The obtained results are compared with individual clustering techniques such as K-Means, Fuzzy C-Means and Density Based Spatial Clustering of Applications with Noise (DBSCAN).

III. IMAGE SEGMENTATION USING CLUSTERING TECHNIQUES

Clustering is considered as the most important unsupervised learning problem, as it deals with finding a structure in a collection of unlabelled data. A loose definition of clustering could be the process of organizing objects into groups whose members are similar in some way. A cluster is therefore a collection of objects which are similar between them and are dissimilar to the objects belonging to the clusters. In this case we easily identify the 3 clusters into which the data can be divided; the similarity criterion is distance: two or more objects belong to the same cluster if they are close. Another kind of clustering is conceptual clustering: two or more objects belong to the same cluster if this one defines a concept common to all that objects. In other words, objects are grouped according to their fit to descriptive concepts, not according to simple similarity measures. In this paper, three different clustering techniques are used for segmenting an image. The three clustering techniques are

A. K-Means clustering; B. Fuzzy C-Means clustering; C. DBSCAN clustering

A. K-Means clustering

K-means clustering is one amongst the essential clustering techniques. In K-Means segmentation, we should mention the number of clusters that need to be present in the segmented image. Let us consider that the number of clusters to be present in the image should be Q. Now, the image will be divided into Q parts and a centroid for each of the parts will be calculated. Initially, the distance for the data points is calculated. There are various kinds of distances that can be calculated. The output segmented image will depend on the type of distance being considered. Euclidean distance has been considered in this paper. The nearest data points to the centroid will be clubbed together to form a cluster. The centroid is calculated again after a cluster is formed. This process is iterated until the distance of the centroid is the least that can be formed from the data points.

Take an image of dimensions, R x S and the image is required to be clustered into Q clusters. Let $f(R, S)$ be input pixels and c_Q be the centroids. Then the algorithm for K-Means clustering can be given by,

— Initiate by Q Clusters and c_Q centroids.

— For every data point i.e., pixel calculate the distance from the centroid using the Euclidean distance.

$$D = \|f(R, S) - c_Q\|$$

— Allot the pixels to the nearest centroid depending on the distance D.

— After every pixel is assigned to one of the clusters, new centroids are calculated.

$$c_Q = \frac{1}{Q} \sum_{y \in c_Q} \sum_{x \in c_Q} f(R, S)$$

— Iterate the process until the centroids from the previous and the present iteration are the same.

Here, 30 Clusters are considered for segmenting the images.

B. Fuzzy C-Means clustering

Fuzzy C-Means is a process where a pixel can be present in single and many clusters. In Fuzzy-C means segmentation, it's required to take a defined number of clusters like the K-Means clustering. But unlike the K-Means clustering where a point be appropriate to only one cluster, in Fuzzy-C means a single information point can fit in to various clusters at a time. A coefficient ε is maintained such that if a centroid of a cluster has a coefficient value less than ε then the cluster is finalized.

Let $Y = \{y_1, y_2, y_3, \ldots, y_n\}$ be the usual of information points and $U = \{u_1, u_2, u_3, \ldots, u_n\}$ be the set of centers. The algorithm for the fuzzy c segmentation can be given by,

— Randomly select 'Q' cluster centres.

— Compute the fuzzy affiliation 'μ_{ij}' using:

$$\mu_{ij} = \frac{1}{\sum_{l=1}^{Q} \left(\frac{d_{ij}}{d_{ik}}\right)^{\left(\frac{2}{m-1}\right)}}.$$

— Calculate the fuzzy centresuj using:

$$u_j = \left(\sum_{i=1}^{n} \left(\mu_{ij}\right)^m x_i\right) \Big/ \left(\sum_{i=1}^{n} \left(\mu_{ij}\right)^m\right), \forall j = 1,2, \ldots, Q$$

— Recap step 2 and 3 up until the least possible 'J' value is attained or$\|V^{(k+1)} - V^{(k)}\| < \varepsilon$ where,

Where 'k' is the repetition number, 'ε' is the conclusion principle among [0, 1], '$V = \left(\mu_{ij}\right)_{nxQ}$' is the fuzzy involvement matrix and 'J' is the impartial function. Thirty clusters are considered for segmenting the image using the fuzzy c-Means clustering also.

C. DBSCAN clustering

DBSCAN is a clustering method based on density of pixels present in the image. DBSCAN identifies regions that have large density and divided by sections having lower mass. Unlike Fuzzy-C means and K-Means methods, we need not mention the number of clusters to be present in the image. The method works by clustering the image based on the difference of the intensity of the node pixel, distance of the pixel from the node and the number of information points to be present within the region of the node to be called as a cluster. If the data point or the pixel is present within the limits of the values, then a cluster is made. In this algorithm, a cluster can be formed within a cluster. The algorithm can be given as,

Let Y= (y1, y2, y3, … , ,yn) be the set of information points. DBSCAN needs two parameters: ε (epsilon) and the least no. of points required to made a cluster (m).

— Start with a pixel that has not been visited yet.

— Obtain the neighbourhood of the pixel with ε (The points that lie within a distance of ε are called as neighbourhood).

— If there are enough points in the neighbourhood then this can be called as a cluster else this is called as noise.

— If a point 'p' belongs to a cluster 'c' then the neighbourhood of the point within ε belongs to the cluster 'c'. This step is iterated until all the points in the cluster are determined.

— After the formation of a cluster, a new data point is visited, and the respective cluster formation takes place.

— This process iterates until all the data points have been visited and marked.

For the pixel to be said in a neighbourhood, the pixel value should lie within the threshold (Th). In the algorithm, we have considered, 40 to be the minimum number of points and 1000 to be the value of epsilon.

II. PROPOSED MODEL FOR IMAGE COMPRESSION

An image is compressed is obtained with the combination of segmentation and an encoding technique. The image is segmented using clustering methods such as K-means, Fuzzy c-means and DBSCAN. In the process of segmentation, some data is lost itself which leads to better image compression.

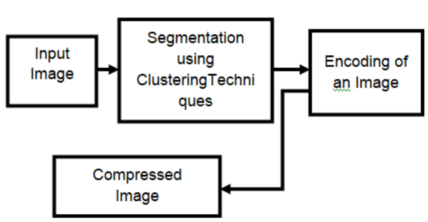


Fig. 1. Block Diagram of Proposed Model.

After the segmentation process, image compression is achieved using different encoding techniques such as Bit plane slicing, Huffman coding and LZW coding and the performance parameters such as PSNR, MSE, SSIM are calculated and compared with one another to determine the better combination among these methods. The block diagram of proposed model for image compression is shown in Fig. 1.

IV. RESULTS AND DISCUSSION

In this paper, the enactment of the recommended model is evaluated on three images of brain containing different tumors namely, Glioma, Astrocytoma and Neuroectodermal. These images are compressed using different clustering techniques and compared using performance parameters. In this paper, the above mentioned algorithms are implemented using MATLAB 16a for compression of images. Three MRI images of brain are considered and segmented using K-means, Fuzzy C-Means and DBSCAN. After the segmentation process, image compression is achieved using different encoding techniques such as Bit plane slicing, Huffman coding and LZW coding. The performance parameters such as Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE), Structural Similarity Index (SSIM), Compression Ratio (CR) are calculated and compared with one another to determine the better among the methods.

Mean Square Error (MSE): The MSE is the cumulative squared error among the compressed and the real image. The methods for the calculation of MSE is given as

$$\frac{1}{MN}\sum_{y=1}^{M}\sum_{x=1}^{N}\left[I(x,y)-I'(x,y)\right]^2$$

Where, I(x,y) is the original image, I'(x,y) is the decompressed image, and M, N are the dimensions of the image. A lower value for MSE means smaller error.

Peak Signal to Noise Ratio (PSNR)

PSNR is usually expressed in terms of the logarithmic decibel scale (dB). The PSNR is most commonly used as a measure of quality of restoration in image compression. The formula for calculation of PSNR is given as

$$\mathrm{PSNR} = 20\log_{10}\left(\frac{255}{\sqrt{\mathrm{MSE}}}\right)$$

Usual values for the PSNR in lossy image compression are among 30 and 50 dB, where higher is better. PSNR is computed by calculating the pixel difference among the real and compressed image.

Compression Ratio (CR)

The size of the real and compressed images is A and B respectively then, the compression ratio is determined as the ratio of the size of the original to the size of the compressed image.

$$C = \frac{A}{B}$$

where, C is compression ratio.

Structural Similarity Index (SSIM): The Structural Similarity (SSIM) Index quality valuation is depends on the calculation of 3 terms, namely the luminance term, the contrast term and the structural term. The total index is a multiplicative of the 3 terms. SSIM of y and $\bar{y}$ is calculated as

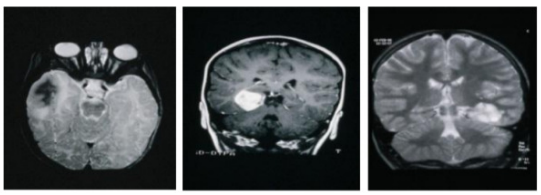
$$SSIM(y,\bar{y}) = \left[L(I,\tilde{I})\right]^{\varepsilon_1}\left[C(I,\tilde{I})\right]^{\varepsilon_2}\left[S(I,\tilde{I})\right]^{\varepsilon_3}$$

Where, α, β and γ are the adaptable parameters, which gives the comparative significance of the 3 terms and are equivalent to one in this paper for easy calculation of SSIM.

$$SSIM = \frac{\left(2\mu_I \mu_{\tilde{I}} + C1\right)\left(2\sigma_{I\tilde{I}} + C2\right)}{\left(\mu_I{}^2 + \mu_{\tilde{I}}{}^2 - C1\right)\left(\sigma_I^2 + \sigma_{\tilde{I}}^2 - C2\right)}$$

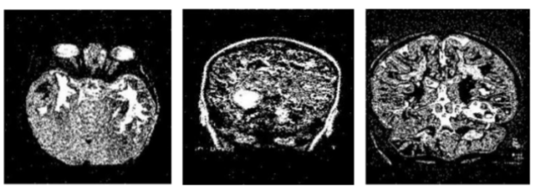
Where μI and μĨ are the mean values of the real image I and reconstructed image Ĩ, σI and σĨ are the standard deviation of real and reconstructed image, σIĨ is the cross-correlation and C1& C2 are constants which are equivalent to 0.065.

The three MRI images of brain Glioma, Astrocytoma and Neuroectodermal are shown in Fig. 2. The results after compression of the three images with different clustering techniques K-Means, Fuzzy C-Means and, DBSCAN with encoding techniques are shown in below figures. Fig. 3-5 shows decompressed images of K-Means clustering with encoding techniques Bit plane slicing, Huffman coding and LZW coding respectively.



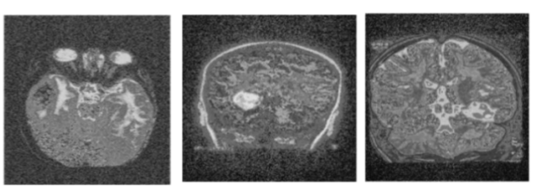
(a) (b) (c)

Fig. 2. Input Brain Images: (a) Glioma (b) Astrocytoma and (c) Neuro ectodermal.



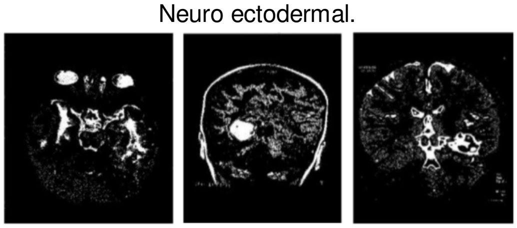
(a) (b) (c)

Fig. 3. Decompressed Image of K-means clustering with Bit plane slicing: (a) Glioma(b) Astrocytoma and (c) Neuroectodermal.



(a) (b) (c)

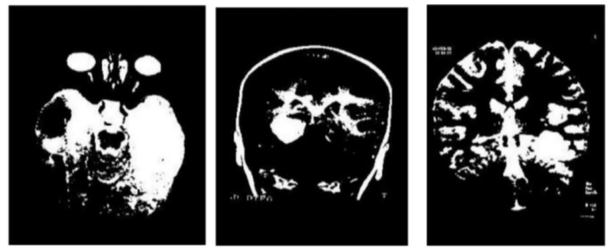
Fig. 4. Decompressed Image of K-means clustering with Huffman coding: (a) Glioma(b) Astrocytoma and (c) Neuro ectodermal.



(a) (b) (c)

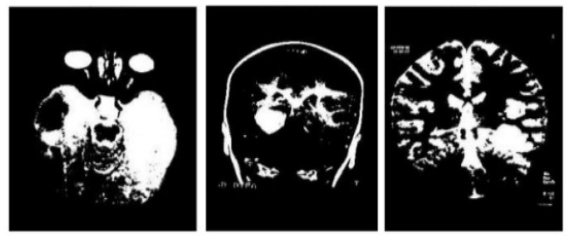
Fig. 5. Decompressed Image of K-means clustering with LZW coding: (a) Glioma(b) Astrocytoma and (c) Neuroectodermal.

Figs. 6-8 shows decompressed images of Fuzzy C-Means clustering with encoding techniques Bit plane slicing, Huffman coding and LZW coding respectively.



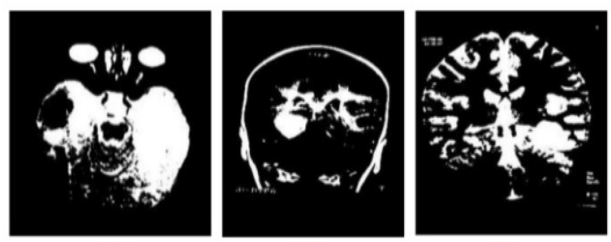
(a) (b) (c)

Fig. 6. Decompressed Images of Fuzzy C-Means with Bit plane slicing: (a) Glioma(b) Astrocytoma and (c) Neuro ectodermal



(a) (b) (c)

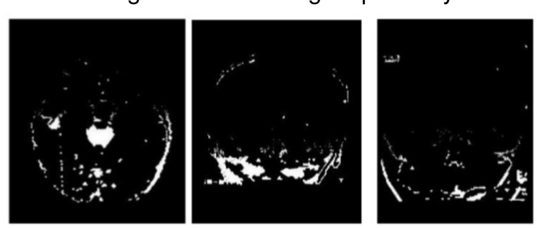
Fig. 7. Decompressed Images of Fuzzy C-Means with Huffman coding: (a) Glioma(b) Astrocytoma and (c) Neuro ectodermal.



(a) (b) (c)

Fig. 8. Decompressed Images of Fuzzy C-Means with LZW: (a) Glioma(b) Astrocytoma and (c) Neuro ectodermal.

Figs. 9 -11 shows decompressed images of DBSCAN clustering with encoding techniques Bit plane slicing, Huffman coding and LZW coding respectively.



(a) (b) (c)

Fig. 9. Decompressed Images of DBSCAN with Bit Plane Slicing: (a) Glioma(b) Astrocytoma and (c) Neuroectodermal.

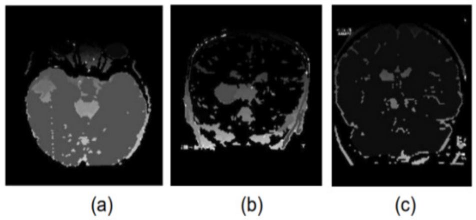


Fig. 10. Decompressed Images of DBSCAN with Huffman coding: (a) Glioma(b) Astrocytoma and (c) Neuroectodermal.

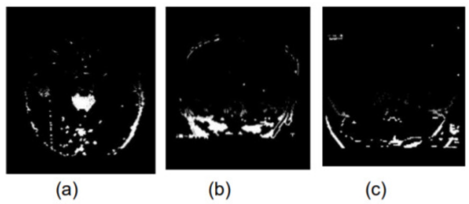


Fig. 11. Decompressed Images of DBSCAN with LZW: (a) Glioma(b) Astrocytoma and (c) Neuroectodermal.

From Fig. 3-11, it is observed that recreated Image quality of the clustering techniques with Huffman encoding is better as compared to other encoding techniques. The compression parameters are compared for all clustering techniques and with state of the art techniques Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) in the below Tables 1-3 for three brain images Glioma, Astrocytoma and Neuroectodermal.

From Tables it is observed that, Huffman encoding technique gives better result in terms of PSNR, MSE and, SSIM when compared to another encoding techniques with all clustering techniques.

From results it is observed that DBSCAN clustering technique gives better result in CR when compared to other techniques with all encoding techniques. From results it is also found that clustering based approach provides good compression results when compared to state of art techniques in all performance parameters.

Table 1: Performance parameters of clustering Techniques for compression on Glioma brain Image.

Performance Parameter	K-Means with LZW	K-Means with Bit-plane	K-Means with Huffman	DCT	DWT
Compression ratio	5.57	4.53	5.865	4.12	4.03
Output memory size (KB)	40.7	50.1	38.7	89.34	60.24
MSE	82.82	53.28	40.95	98.35	89.586
PSNR (in dB)	28.98	30.899	32.04	27.34	28.02
SSIM	0.855	0.8932	0.9240	0.689	0.756
	Fuzzy C-Means with LZW	Fuzzy C-Means with Bit-plane	Fuzzy C-Means with Huffman	DCT	DWT
Compression ratio	5.820	8.37	8.345	5.34	4.89
Output memory size (KB)	39.0	27.1	27.2	78.95	64.37
MSE	79.39	73.85	73.87	119	95
PSNR (in dB)	29.167	29.48	29.48	26.98	27.89
SSIM	0.858	0.870	0.869	0.634	0.735
	DBSCAN with LZW	DBSCAN with Bit-plane	DBSCAN with Huffman	DCT	DWT
Compression ratio	10.607	16.569	14.367	5.65	4.634
Output memory size (KB)	21.4	13.7	15.8	67.23	57.89
MSE	86.08	76.36	75.8	134	98.35
PSNR (in dB)	28.82	29.335	29.36	25.76	26.856
SSIM	0.851	0.874	0.9074	0.649	0.727

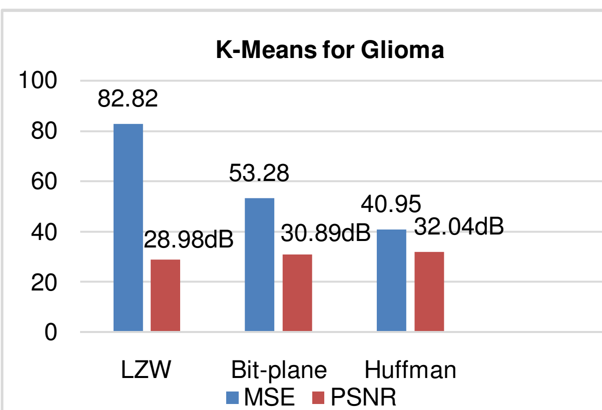
Table 2: Performance parameters of clustering Techniques for compression on Astrocytoma brain image.

Performance Parameter	K-Means with LZW	K-Means with Bit-plane	K-Means with Huffman	DCT	DWT
Compression ratio	6.012	8.189	6.917	4.56	5.03
Output memory size (KB)	47.4	34.8	41.2	83.76	59.434
MSE	73.43	62.26	31.68	89.67	87.235
PSNR (in dB)	29.50	30.222	33.15	26.14	27.96
SSIM	0.795	0.814	0.834	0.656	0.712
	Fuzzy C-Means with LZW	Fuzzy C-Means with Bit-plane	Fuzzy C-Means with Huffman	DCT	DWT
Compression ratio	9.531	13.507	13.443	5.65	4.60
Output memory size (KB)	29.9	21.1	21.2	75.34	63.35
MSE	78.4	72.62	72.62	123	106
PSNR (in dB)	29.22	29.55	29.55	24.58	25.83
SSIM	0.793	0.8015	0.801	0.656	0.732
	DBSCAN with LZW	DBSCAN with Bit-plane	DBSCAN with Huffman	DCT	DWT
Compression ratio	12.66	20.503	17.484	5.34	4.867
Output memory size (KB)	22.5	13.9	16.3	65.23	54.89
MSE	77.53	69.17	67.17	128	96.42
PSNR (in dB)	29.27	29.76	29.89	23.56	24.345
SSIM	0.794	0.807	0.819	0.678	0.719

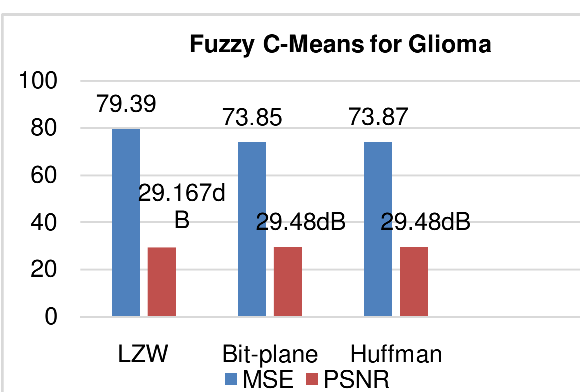
Table 3: Performance parameters of clustering Techniques for compression on Neuroectodermal brain image.

Performance Parameter	K-Means with LZW	K-Means with Bit-plane	K-Means with Huffman	DCT	DWT
Compression ratio	6.943	9.7027	9.020	4.45	5.78
Output memory size (KB)	51.7	37	39.8	84.34	67.24
MSE	82.05	60.16	42.08	102.35	92.586
PSNR (in dB)	29.02	30.512	31.92	26.97	27.08
SSIM	0.829	0.854	0.876	0.623	0.753
	Fuzzy C-Means with LZW	Fuzzy C-Means with Bit-plane	Fuzzy C-Means with Huffman	DCT	DWT
Compression ratio	11.360	16.244	16.098	5.79	4.64
Output memory size (KB)	31.6	22.1	22.3	75.74	68.12
MSE	77.49	71.41	71.19	109	96
PSNR (in dB)	29.27	29.627	29.64	25.98	26.89
SSIM	0.8424	0.8517	0.85155	0.681	0.749
	Fuzzy C-Means with LZW	Fuzzy C-Means with Bit-plane	Fuzzy C-Means with Huffman	DCT	DWT
Compression ratio	15.276	24.256	24.93055	5.34	4.653
Output memory size (KB)	23..5	14.8	14.4	66.23	59.89
MSE	87.05	75.33	71.61	132	94.35
PSNR (in dB)	28.76	29.39	29.615	23.13	25.329
SSIM	0.825	0.846	0.89778	0.660	0.703

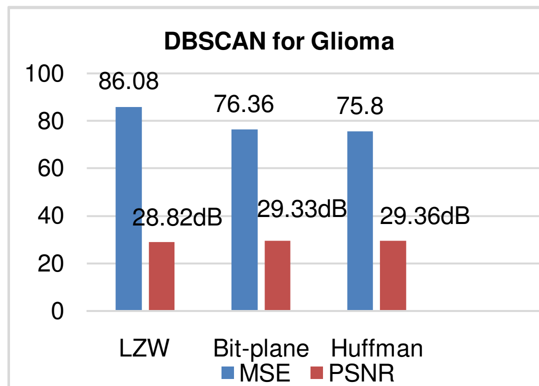
The Performance parameters PSNR and MSE of three clustering techniques for image compression on MRI brain images Glioma, Astrocytoma and Neuroectodermal are plotted using bar graphs shown in below Fig. 12-14. From bar graphs it is observed that Huffman encoding technique gives better PSNR and lower MSE for all clustering techniques when compared to other encoding techniques Bit plane slicing and LZW coding.



(a)

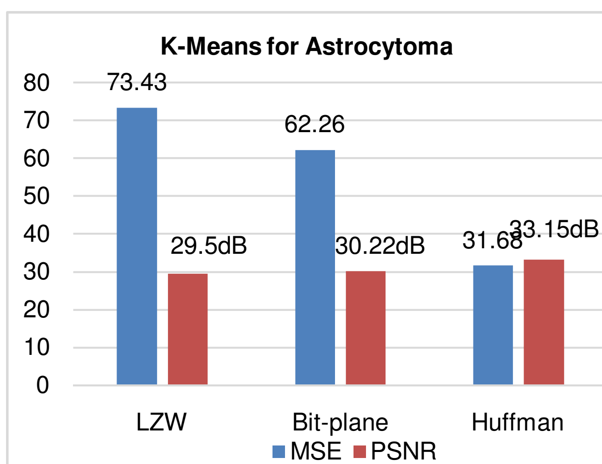


(b)

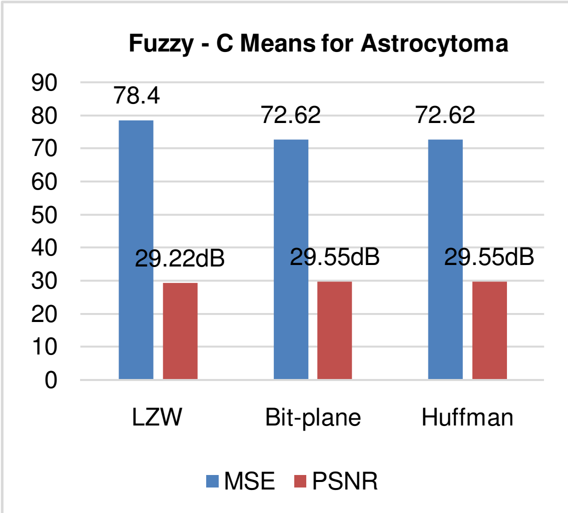


(c)

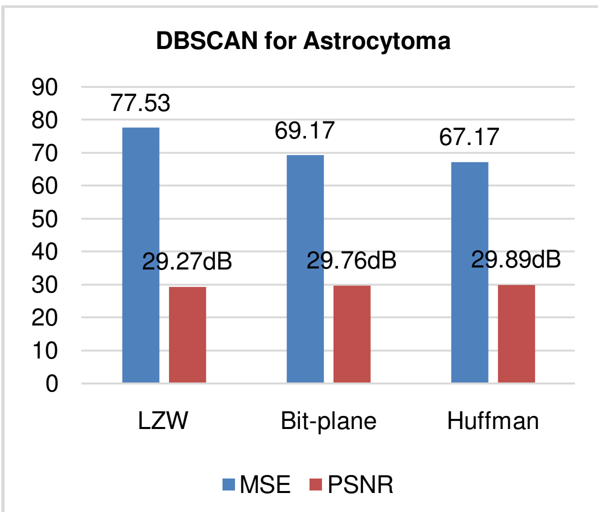
Fig. 12. Performance parameters of clustering techniques on MRI Glioma brain Image: (a) K-Means (b) Fuzzy C-Means (c) DBSCAN.



(a)

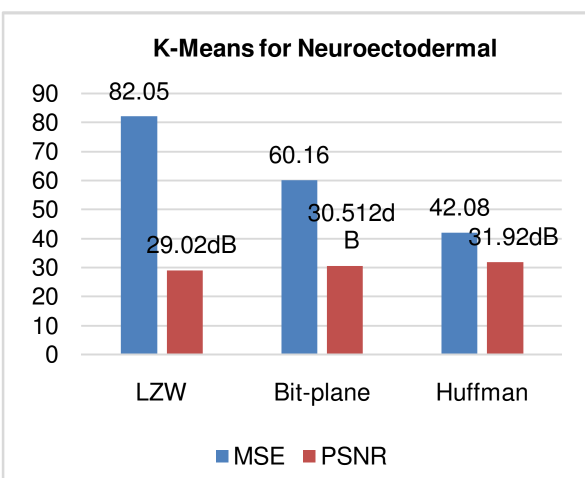


(b)

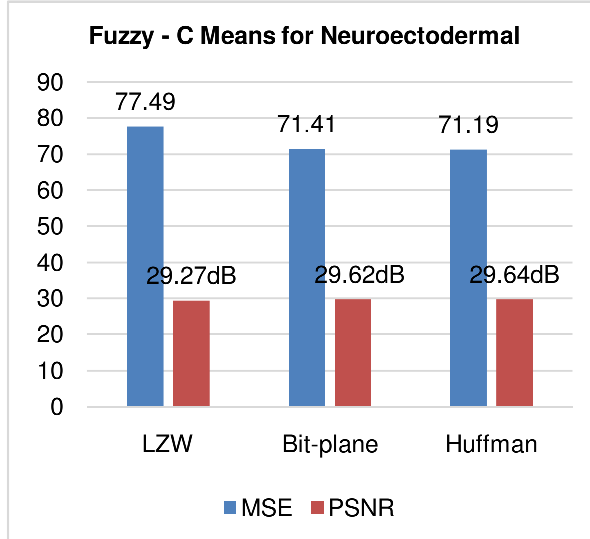


(c)

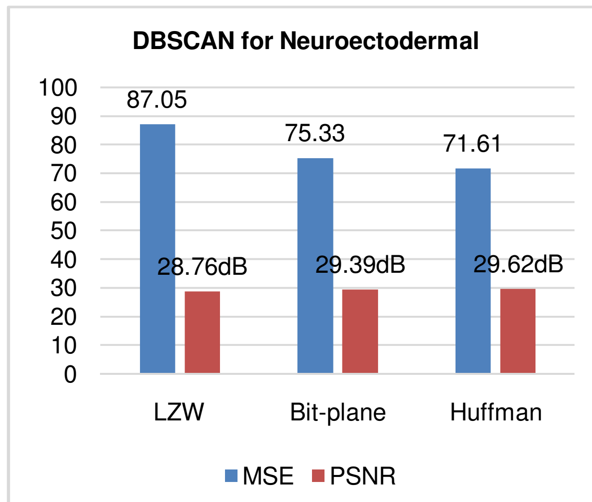
Fig. 13. Performance parameters of clustering techniques on MRI Astrocytoma brain Image: (a) K-Means (b) Fuzzy C-Means (c) DBSCAN



(a)



(b)



(c)

Fig. 14. Performance parameters of clustering techniques on MRI Neuroectodermal brain Image: (a) K-Means (b) Fuzzy C-Means (c) DBSCAN.

V. CONCLUSION

In this paper, image compression is obtained by using clustering methods like K-Means, Fuzzy C-Means and DBSCAN. The clustering techniques were used for segmentation of images for effective compression. The segmentation results are encoded using Bit plane slicing, Huffman coding and LZW coding for efficient transmission and storage. By observing the decompressed images, the quality of images obtained by Huffman encoding technique is high than other two techniques for all three clustering techniques. The proposed method is validated on different MRI brain images Glioma, Astrocytoma and Neuroectodermal and compared in terms of performance parameters like CR, PSNR, MSE and SSIM. The outcomes display that the Huffman encoding method gives better result in terms of PSNR, MSE and SSIM when compared to other encoding techniques. Higher compression ratio is obtained for Huffman encoding with DBSCAN clustering technique.

FUTURE SCOPE

The proposed model can be further extended to some domain-based applications such as finger print recognition, retina identification, and object detection etc., for large image database. Classifier algorithms can also use the outcomes of the project, as segmentation is already performed. The computation time required can be further decreased during the segmentation process with other algorithms. This area can be further developed by using other algorithms like mean shift and agglomerative hierarchical clustering algorithms.

Conflict of Interest. The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest in the subject matter or materials discussed in this manuscript.

ACKNOWLEDGEMENT

I would like to take this great opportunity to thank Dr. K.V.L. Raju, Correspondent, MANSAS Educational trust and Principal, MVGR College of Engineering (A), Chintalavalasa, Viziangaram, for his constant encouragement and support during my independent study. I should also thank the Faculty of Electronics and Communication Engineering for giving me the opportunity and also thank all those who gave me the valuable advices that allowed me to make this independent study a success.

REFERENCES

[1]. Yang, C. K., & Tsai, W. H. (1998). Color image compression using quantization, thresholding, and edge detection techniques all based on the moment-preserving principle. *Pattern Recognition Letters, 19*(2), 205-215.

[2]. Kaur, L., Gupta, S., Chauhan, R. C., & Saxena, S. C. (2007). Medical ultrasound image compression using joint optimization of thresholding quantization and best-basis selection of wavelet packets. *Digital Signal Processing, 17*(1), 189-198.

[3]. Sidhik, S. (2015). Comparative study of Birge–Massart strategy and unimodalthresholding for image compression using wavelet transform. Optik, *126*(24), 5952-5955.

[4]. Mohammadpour, T. I., & Mollaei, M. R. K. (2009). ECG compression with thresholding of 2-D wavelet transform coefficients and run length coding. *European Journal of Scientific Research, 27*(2), 248-257.

[5]. Rajeswari, R., & Rajesh, R. (2012). Type-2 fuzzy thresholded band let transform for image compression. Procedia engineering, *38*, 385-390.

[6]. Sansgiry, P., & Mihaila, I. (2003). Selective Thresholding in Wavelet Image Compression. In Wavelets and Signal Processing, Birkhäuser, Boston, MA, 377-381.

[7]. Chan, T. F., & Zhou, H. M. (2007). Total variation wavelet thresholding. *Journal of Scientific Computing, 32*(2), 315-341.

[8]. Salleh, M. F. M., & Soraghan, J. (2007). A new multistage lattice vector quantization with adaptive subbandthresholding for image compression. *EURASIP Journal on Advances in Signal Processing,* 1-11.

[9]. Ahmadi, K., Javaid, A. Y., & Salari, E. (2015). An efficient compression scheme based on adaptive thresholding in wavelet domain using particle swarm optimization. signal processing: image communication, *32*, 33-39.

[10]. Salleh. M. F. M and Soraghan. J. (2007). A New Multistage Lattice Vector Quantization with Adaptive Sub-band Thresholding for Image Compression. *EURASIP Journal on Advances in Signal Processing.*

[11]. Kumari, G.V., Rao, G.S. and Rao, B.P. (2018). Flower pollination-based K-means algorithm for medical image compression. Int. J. Advanced Intelligence Paradigms.

[12]. Fred, A. L., Kumar, S. N., Kumar, H. A., & Abisha, W. (2019). Bat Optimization Based Vector Quantization Algorithm for Medical Image Compression. In Nature Inspired Optimization Techniques for Image Processing Application, Springer, 29-54.

[13]. Aslam, Y., & Santhi, N. (2020). A Comprehensive Survey on Optimization Techniques in Image Processing. Materials Today: *Proceedings, 24*, 1758-1765.

[14]. Pattanaik, P. A. (2020). A Stacked Denoising Autoencoder Compression Sampling Method for Compressing Microscopic Images. In Smart Healthcare Analytics in IoT Enabled Environment, Springer, Cham,191-199.

[15]. Man Hoang, T., Zhou, J., & Fan, Y. (2020). Image Compression With Encoder-Decoder Matched Semantic Segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 160-161.

[16]. Chen, Z., & He, T. (2019). Learning based facial image compression with semantic fidelity metric. *Neurocomputing, 338*, 16-25.

[17]. Krishnaraj, N., Elhoseny, M., Thenmozhi, M., Selim, M. M., & Shankar, K. (2019). Deep learning model for real-time image compression in Internet of Underwater Things (IoUT). Journal of Real-Time Image Processing, 1-15.

[18]. P. Shanthakumar, P. Ganeshkumar. (2015). Performance analysis of classifier for brain tumor detection and diagnosis. *Computers and Electrical Engineering.*

[19]. Srinivas, B., & Rao, G. S. (2019). Performance evaluation of fuzzy C means segmentation and support vector machine classification for MRI brain tumor. In Soft computing for problem solving, Springer, Singapore, 355-367.

[20]. Srinivas, B., & Rao, G. S. (2018). Unsupervised learning algorithms for MRI brain tumor segmentation. In *2018 Conference on Signal Processing And Communication Engineering Systems (SPACES), IEEE,* 181-184.

[21]. Li, M., Zuo, W., Gu, S., You, J., & Zhang, D. (2020). Learning content-weighted deep image compression. IEEE Transactions on Pattern Analysis and Machine Intelligence.

How to cite this article: Kumari, G. V., Prakash, M. S., Prasad, P. S., Srinivas, B. and Lavanya, B. (2020). Image Compression using Clustering Techniques for Bio Medical Applications. *International Journal on Emerging Technologies*, *11*(3): 1185–1193.