Journal of Scientific & Industrial Research
Vol. 80, August 2021, pp. 707-715

# Segmentation of Natural Images with K-Means and Hierarchical Algorithm based on Mixture of Pearson Distributions

P Chandra Sekhar[1], N Thirupathi Rao[2]*, Debnath Bhattacharyya[3] and Tai-hoon Kim[4]

[1]Department of Computer Science and Engineering, GITAM Deemed to be University, Visakhapatnam 530 045, India

[2]Department of Computer Science and Engineering Vignan's Institute of Information Technology (A), Visakhapatnam 530 049, AP, India

[3]Department of Computer Science and Engineering, Koneru Laksmaiah Education Foundation, Greenfield, Vaddeswaram, Guntur 522 502, India

[4]Glocal Campus, Konkuk University, Chungju-si Chungcheongbuk-do, 27 478, Korea

*Received 03 October 2020; revised 07 May 2021; accepted 22 July 2021*

In this paper, an attempt has been made to analyze the performance of the image segmented algorithms with the addition of the Pearsonian Type III mixture model. By using the Type III Pearsonian system of distributions the image segmentation process was carried out in the current article which is a novel technique. With the help of K-component combination of Pearsonian Type III distribution, it is considered that the whole input images are characterized. The performance parameters PRI (Probabilistic Rand Index), GCE (Global Consistency Error) and VOI (Volume of Interest) for the currently considered model are estimated with the help of EM (Expectation Maximization) algorithm. For analyzing the proposed model's performance, four random images are selected as input for the current model from Berkeley image database. The performance metric parameters PRI, GCE and VOI values given the results as the currently proposed method is providing more précise results for the input images where the regions of the input images selected are with tiles having long upper model and the left skewed images. By the help of image quality measures, the proposed method is performing well for the purpose of retrieving the images with respect to the picture segmenting process which is based on GMM (Gaussian Mixture Model). The current model performance was compared with the other existing models like the k-means hierarchical clustering model and the 3-papRameter regression models.

**Keywords:** Berkeley image database, EM-algorithm, Image quality metrics, Non-symmetric model, Type III Pearsonian distribution

## Introduction

Analyzing and retrieving of data from a picture is highly used in different applications like the remote sensing areas, segmentation areas of images, authentication of an entity and also for processing of videos. For segmentation of images, model based segmenting methods are getting used mostly for accurate results. Several methods are available to analyze and develop the image models and their structures by using various methods. Segmentation of images is also another important method used for the same processes.[1–3] In general, several methods are available in literature also to use these sorts of tasks. The important model of the method was the usage of a mixture of Pearson Type I distribution with the combination of hierarchical algorithms or K-means clustering algorithms. These methods are highly used in some situations — like the intensities of the pixels of feature vectors present on the various regions of the images are the skewed type of representation. These types of models may have the extensive and higher or greater tail with a right skewed mode of working nature. Also, the fact is that if the researchers or working people want to work on these types of images, the intensities of the considered images must have the distributions with right skewed models.[4,5] Hence, in the current article an attempt has been made to analyze the performance of the model by considering the Pearsonian Type III distribution. The images are portraying the right skewed and higher with greater tails and as a result the distribution is chosen.

## Literature Review

There have been several publications on symmetric distribution-based image segmentation with the

*Author for Correspondence
E-mail: nakkathiru@gmail.com

combination of K-Means and clustering algorithms. The details of various works on this particular problem of work are given as follows,

Rao *et al.*[1] considered the image segmentation process with k-means clustering algorithms and also considered the Pearson type VI distribution and the results were calculated. The authors had verified the performance of the considered model with various input images. They had compared the currently considered model for various sets of input images. Pal & Pal[2] discussed in detail about the various techniques available and used for the segmentation of images. They exemplified and presented the advantages and disadvantages of various techniques available. Dae-Hoon *et al.*[3] analyzed the model of the image segmentation process with the usage of k-means algorithm and also considered various performance metrics to calculate the performance of the considered model. Gao *et al.*[4] discussed about various clustering algorithms and compared those clustering algorithms and had given the advantages and disadvantages for the better utilization of these algorithms in various scenarios. Jun *et. al.*[5] discussed the breaking of a video part to two parts like the foreground player and background player. The proposed method was verified with images collected from various sources and is compared with the other set of image datasets and the other video datasets. Shameem *et al.*[6] considered image segmentation as one of the significant tools for image analysis. They had worked with an experimental study of image segmentation using K-Means clustering Algorithm.

Satyanarayana *et al*.[7] considered the image decomposition or segmentation. The authors had worked on three parameter logistic type mixture distributions for decomposing the input images in any model. In their work, they assumed that the three-parameter logistic type probability distribution was followed by the various intensities of the picture regions. The performance metrics are calculated or evaluated by using the EM algorithm. Pillai *et al*.[8] proposed a new technique called Local Diagonal Extrema Number Pattern (LDENP) used for the recognition of face. Here the recognition process was carried out by the decoding of bidirectional features of the face. The problem of dimension reduction was resolved by using the current method. The current model considers the local diagonal pixels rather than the local neighbor pixels. Naik *et al*.[9] proposed a new algorithm called the Teaching Learning-Based Optimization (TLBO) for solving the clustering problems. The algorithm works in two stages, in first stage the algorithm tries to find out the optimal cluster centers and in second stage the algorithm tries to identify the best cluster by using the c-means algorithm. Soundrapandiyan *et al*.[10] discussed about the people identification mechanism by using the SVM model. The difficulties faced in earlier models were resolved by the SVM model. They had also identified the false and true alarms and also classifiers for SVM algorithm are used. Shriranjani *et al*.[11] developed and discussed in detail about the processing of retinal images. For preprocessing of the images, they used the chaotic bat algorithm using the Kapur's Threshold method and for post processing of the images they used the active contour segmentation method. They also discussed the merits and demerits of each method discussed in the article.[12,13]

From the literature review, it is understood that several works were done and discussed about the symmetric image models for identification of problems in retina-based images and also MRI images of black spinal problems. The gaps identified from the previous works are like the authors used the neural networks model, fuzzy C means model, optimized convolutional neural network models etc. for segmenting and processing the images for identifying the content on images. Very few works were present dealing with non-symmetric distribution models with image segmentation process for dealing with non-symmetric functional images. Very few works are reported on segmenting the images using the Type III Person distribution. But no works were reported with the combination of Type III Pearsonian system distribution with EM algorithm and Hierarchical clustering algorithms except the work of Satyanarayana *et al*.[12] who worked on the three-parameter logistic type mixture model with hierarchical clustering work for the segmentation of images. Hence, this had motivated us to design and implement non-symmetric image segmenting method of using the Type III Pearsonian system distribution model using the K-Means and EM algorithm for the segmentation of natural images and analyzing the performance of the method considered. The results obtained by using this model are presented and discussed in the results section. The performance of the current model are compared with the other existing models for better understanding of the current model's performance with respect to the segmenting of the natural images.

**Objectives of the Current Work**

The main objective of the current work was to segment the Non-Symmetric images with Type III Pearsonian distribution.

Many images observed in the areas of medical imaging and security images are laptykurtic images which may not generate better results in symmetric segmenting models. The current Type III Pearsonian distribution was a non-Symmetric model may be verified for non-symmetric functional images for better results. The performance of the current model should be analyzed for the laptykurtic images using Non-Symmetric Type III Pearsonian distribution. The results should be analyzed for the performance metrics like the GCE, VI and PRI.

**Methodology**

The method used in the current study is based on the intensities of the pixels on images. The statistical model with Pearsonian distribution model was chosen and implemented in the current work. Image histograms are used for the assessment of the existence of the number of regions present on images. These estimations and the utilization of the regions of the images are updated by the usage of EM algorithm.[6–8] The model parameters choosing, estimation and initialization was completed with the usage of K-means and a moment method of estimation. By using the Bayes theorem model framework, the segmentation model algorithm is developed and implemented. The performance of the current model considered can be evaluated by calculating various performances measuring metrics like the GCE, PRI and VOI etc. In order to check the performance with some input images, the model is being tested with four types of images like the Water, Hill, the Sea and Boat. The current model utilization was compared with the other models like the Gaussian Mixture model etc. The current model's efficiency was calculated by using various performance metrics like the maximum distance, mean square error, the average difference, image quality index and signals to noise ratio.

The pixel intensities are the values that can clear the knowledge or the data about the image region. The intensity of the pixel are taken as, p = f (r, s) to the considered point as a pixel of (r, s) is an arbitrary variable, due to the reason that a normal point in the regions of an image is identified or observed clearly by using the intensity with various important feature or the factors like humidity, visualization, light and environmental conditions etc.[9–11] For modeling or to identify the image regions to identify the human and animal images, the utilization of non-symmetric Pearson Type III distribution is required.

**Pearson Type III Mixture Distributions**

For modeling or to identify the image regions to identify the human and animal images, the utilization of Pearson Type III distribution is required. The PDF (Probability Distribution Function) of the current model is as follows,

$$f_i(z/a_i,q_i)=\frac{(q_ia_i)^{(q_ia_i+1)}}{e^{q_ia_i}a_i\Gamma(q_ia_i+1)}e^{-q_iz_s}\left(1+\frac{z_s}{a_i}\right)^{q_ia_i},\ -a_i\le z_s<\infty,-\infty\le q_i<\infty \quad (1)$$

where $\Gamma$ is a gamma function, the curves associated with Pearson Type III distribution are shown in Fig.1 as follows,

In general, the images of both human beings and animals are a collection of image regions those can be characterized by using the current distribution. The PDF of the model can be represented as follows,

$$q(y)=\sum_{j=1}^{K}\beta_i f_i\left(\frac{z}{b_i},r_i\right) \quad \text{... (2)}$$

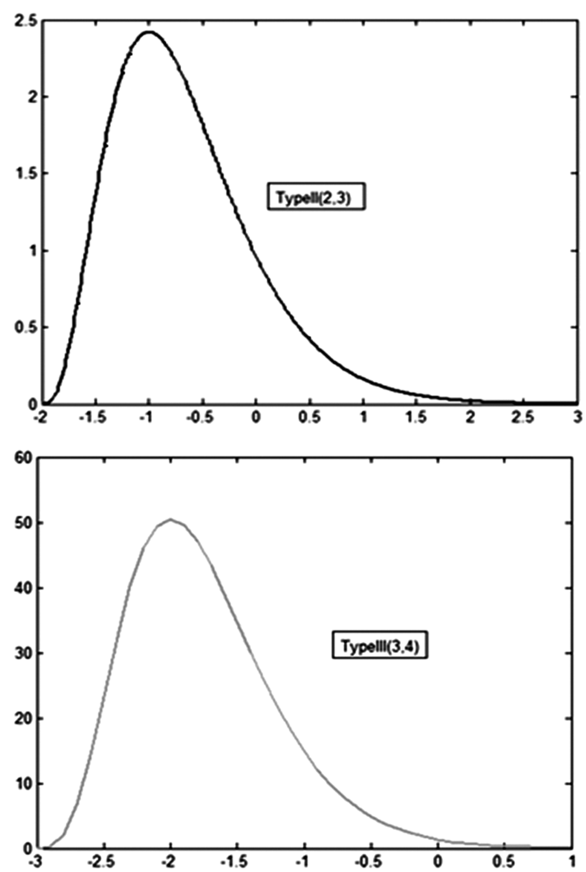


Fig.1 — Frequency curves of Pearson type III distribution

where, K gives the region numbers, $0 \le \beta_i \le 1$ are weights such that $\sum\beta_i = 1$; $\beta_i$ is the weight associated with i$^{th}$ region in the whole image.

**Evaluation of the Parameters in Current Model by EM Algorithm**

The EM (Expectation Maximization) algorithm is used to identify or to calculate the model parameters. The PDF is given by,

$$\log L(\theta) = \sum_{s=1}^{N} \log\left[\sum_{i=1}^{K} \frac{\alpha_i (q_i a_i)^{(q_i a_i+1)}}{e^{q_i a_i} a_i \Gamma(q_i a_i+1)} e^{-q_i z_s}\left(1+\frac{z_s}{a_i}\right)^{q_i a_i}\right] \quad (3)$$

The first step in the EM algorithm was to initiate the parameters $(a_i, q_i; i = 1, 2, ..K)$.

**E-STEP:**

In the current step,

Log likelihood function and its expectation is considered as,

$$Q\left(\theta;\theta^{(l)}\right) = E_{\theta^{(l)}}\left[\log L(\theta)/\bar{z}\right]$$

By using the various discussions, we have

$$Q(\theta;\theta^{(l)}) = \sum_{i=1}^{K}\sum_{s=1}^{N}\left(t_j(z_s,\theta^{(l)})\left(\log p_j(T_s,\mu^{(l)}) + \log\beta_{ij}^{(l)}\right)\right) \quad ...(4)$$

**M-STEP:**

The modernized equation of $\beta_j$ for (K+1)$^{th}$ iteration is

$$\alpha_i^{(l+1)} = \frac{1}{N}\sum_{s=1}^{N} t_i(z_s,\theta^{(l)}) = \frac{1}{N}\sum_{s=1}^{N}\left[\frac{\alpha_i^{(l)} f_i(z_s,\theta^{(l)})}{\sum_{i=1}^{K}\alpha_i^{(l)} f_i(z_s,\theta^{(l)})}\right] \quad ...(5)$$

The updated equation of $a_i$ at $(l+1)^{th}$ iteration is

$$a_i^{(l+1)} = \sum_{s=1}^{N} \frac{t_i(z_s,\theta^{(l)})}{\left[\frac{q_i^{(l)} z_s}{a_i^{(l)}+z_s} + q_i^{(l)}\Gamma\left(q_i^{(l)}a_i^{(l)}+1\right) - q_i^{(l)}\log\left(q_i^{(l)}a_i^{(l)}\left(a_i+z_s\right)\right)\right] t_i(z_s,\theta^{(l)})} \quad ...(6)$$

The updated equation of $q_{i1}$ at $(l+1)^{th}$ iteration is

$$q_i^{(l+1)} = \frac{\sum_{s=1}^{N} q_i\left[a_i^{(l)}\Gamma(q_i^{(l)}a_i^{(l)}+1) + (a_i^{(l)}+z_s) - a_i^{(l)}\log\left(q_i^{(l)}a_i^{(l)}\left(\frac{z_s+a_i^{(l)}}{a_i^{(l)}}\right)\right)\right]}{a_i^{(l)}\sum_{s=1}^{N} t_i(z_s,\theta^{(l)})} - 1 \quad ...(7)$$

## Commencement of Parameters for the Current Considered Model

The EM algorithm is used to identify the parameters of the image by depending on the regions of the images. This mode of values is being identified by various modes and plotting histogram is one of the best methods to analyze such estimations. The mixing parameters $\alpha_i$ and the model parameters $a_i$, $q_i$ is usually considered as known apriori.

**Segmentation Algorithm**

Once the cleansing and refinement process of the parameters of the model are done. The selection and distribution of pixels of the images are given to the images with various segments.[11] The algorithm works as follows,

***Step 1***: At first the whole image is considered and histogram for the whole image is being generated at the first step.

***Step 2***: The initial expected values are achieved through either hierarchical or K-means algorithms.

***Step 3***: In the next step, the maximum effort t was given to achieve the modified expected values of the model.

***Step 4***: Once the refined values are obtained, the assignment of these values to the various regions of the segmented images is done.

The PDF of the current considered model is as follows,

That is $L_j = \max_{j\,\in\,k}\left[\frac{(q_j a_j)^{(q_j a_j+1)}}{e^{q_j a_j} a_j\Gamma(q_j a_j+1)} e^{-q_j z_s}\left(1+\frac{z_s}{a_j}\right)^{q_j a_j}\right]$,

$$-a_i \le z_i < \infty, -\infty < q_j < \infty.$$

## Results and Discussion

In order to check and verify the performance of the current considered model with various algorithms considered here K-means and Hierarchical algorithms with various performance metrics are considered and analyzed.[14,15] The results and the performance are represented for both the model-based algorithms and the results are as follows:

**Initialization of Parameters by K-Means Algorithm**

In order to consider and verify the performance of the model that was considered in the current model are explained in detail. Several images are considered for the segmentation of the input images. Those images are WATER, SEA, BOAT and HILL. The K represents the number of peaks of the image generated histogram. Every image has its own number of peaks depending on the histogram by K-means algorithm. Three peaks were considered for the histogram of the current images. Hence, the value of K was considered as 3. The Pearson Type III distribution mixture was assumed to be followed for the various intensities of the pixels of the images considered here. The input images are considered from the Berkeley image dataset. Those output histograms are observed at Fig. 2 and the various values considered for K estimates and other estimation are given in Tables 1–3 as,

From Table 1, we observe that the images BOAT and SEA have three segments and images WATER and HILL have four segments each.

By using the estimates obtained finally for the parameters considered are used for calculating the PDF for various intensities of the images elected. The segmented and the actual input images are shown in Fig. 3 as,

**Commencement of Model Parameters by Hierarchical Clustering Algorithm**

In the current case, a random four number of images are selected and processed for further process and the results are analyzed. The current algorithm with these input images is processed by using the MATLAB Code and various estimated can be observed at Tables 4–7 as,

By using the estimates obtained finally for the parameters considered are used for calculating the pdf for various intensities of the images elected. The segmented and the actual input images are shown in the following Fig. 4.

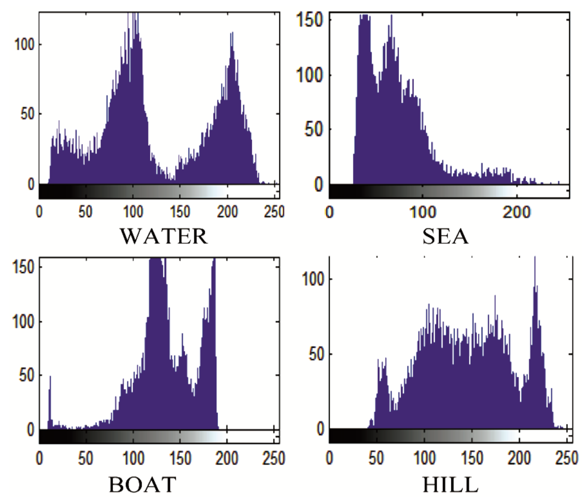


Fig. 2 — Images with histograms

Table. 1 — Values considered for K as estimates initially

| Input | WATER | SEA | BOAT | HILL |
|---|---|---|---|---|
| K-Estimate | 4 | 3 | 3 | 4 |

Table 2 — Image's parameters estimated values regions of image and its number (K =3)

| | Parameters | Initial parameters estimation: Regions of image i | ii | iii | Final parameters estimation with Expectation Maximization Algorithm: Regions of image i | ii | iii |
|---|---|---|---|---|---|---|---|
| BOAT | $\alpha_i$ | 0.325 | 0.325 | 0.325 | 0.3333 | 0.3333 | 0.3333 |
| | $a_i$ | 192.884 | −63.253 | 74.3194 | 1.17E-06 | −7.0873 | 0.4950 |
| | $q_i$ | 0.0600 | 0.0983 | −0.3195 | 853276.3565 | −0.2347 | 0.4373 |
| SEA | $\alpha_i$ | 0.333 | 0.333 | 0.333 | 0.3333 | 0.3333 | 0.3333 |
| | $a_i$ | 107.216 | 92.0780 | 36.660 | 1.4553 | 81.1089 | 0.4307 |
| | $q_i$ | −0.1071 | 0.0007 | −0.0549 | 0.1255 | 0.0024 | 2.1523 |
| WATER | $\alpha_i$ | 0.250 | 0.250 | 0.250 | 0.250 | 0.250 | 0.250 |
| | $a_i$ | 21.692 | −30.122 | 31.145 | −24.648 | −45.189 | −6.5345 |
| | $q_i$ | 0.0566 | 0.0794 | −0.1056 | −0.0329 | 0.2932 | −0.4969 |

Table 3 — HILL Image's parameters estimated values regions of image and its number (K =3)

| Parameters | Initial parameters estimation: Regions of image 1 | 2-Type I | 3 | 4 | Final parameters estimation with Expectation Maximization Algorithm: Regions of image 1 | 2-Type I | 3 | 4 |
|---|---|---|---|---|---|---|---|---|
| $\alpha_i$ | 0.250 | 0.250 | 0.250 | 0.250 | 0.250 | 0.250 | 0.250 | 0.250 |
| $a_i$ | −207.1 | $a_{i1} = −22.48$ | −254.4 | −67.67 | −207.1 | $a_{i1} = −0.174$ | 281.51 | −11.448 |
| $q_i$ | −0.340 | $a_{i2} = 24.3$ | −0.006 | 0.1558 | −0.340 | $a_{i2} = 0.374$ | −0.007 | −0.501 |
| — | — | $m_{i1} = 0.481$ | — | — | — | $m_{i1} = 2.635$ | — | — |
| — | — | $m_{i1} = −0.52$ | — | — | — | $m_{i1} = 0.407$ | — | — |

**Performance Comparison with Previous Models**

The current model and the method are tested by using various set of input images and the results are discussed in the current sections. The results are as follows in the following Table 8,

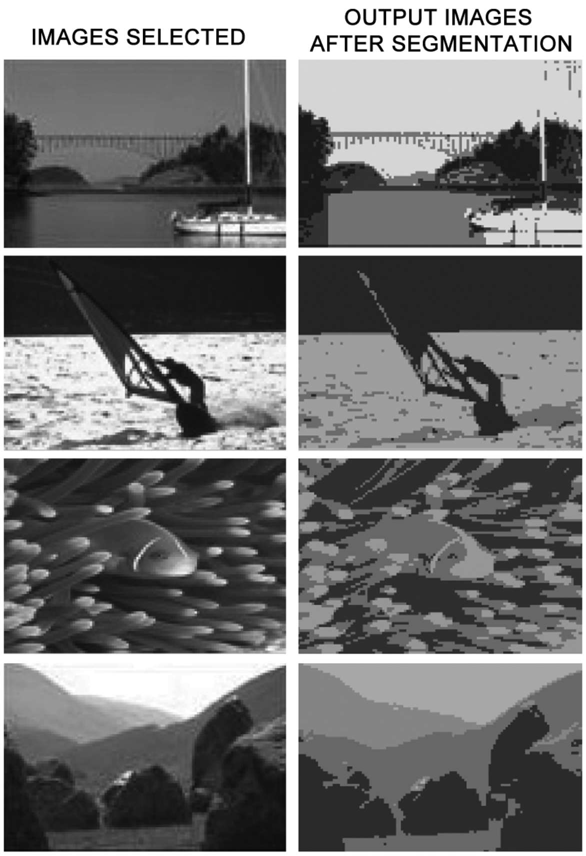


Fig. 3 — Original and segmented images

From the above Table 8, it is clearly understood that the values raised from the PRI for four images are working better or generating better output when compared with the existing methods which were based on Gaussian mixture models. From the results, it is understood that the currently considered model was performing well when compared with the old model.

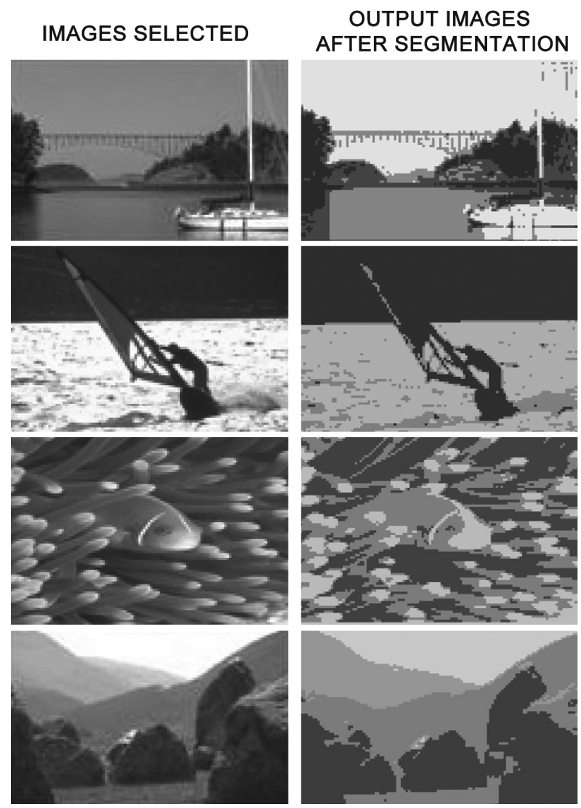


Fig. 4 — Original and segmented images

Table 4 — BOAT image's parameters estimated values regions of image and its number (K =3)

| Parameters | Initial Parameters Estimation | | | Final Parameters Estimation with Expectation Maximization Algorithm | | |
|---|---|---|---|---|---|---|
| | Regions of Image | | | Regions of Image | | |
| | 1 | 2 | 3 | 1 | 2 | 3 |
| $\alpha_i$ | 0.3330 | 0.3330 | 0.3330 | 0.3333 | 0.3333 | 0.3333 |
| $a_i$ | −147.53 | −46.9490 | 151.02 | −0.74935 | 12.0699 | 10.531 |
| $q_i$ | 0.3471 | 0.05542 | 0.0257 | −0.94506 | −0.1868 | 0.1246 |

Table 5 — SEA Image's Parameters Estimated Values Regions of Image and its number (K =3)

| Parameters | Initial Parameters Estimation | | | Final Parameters Estimation with Expectation Maximization Algorithm | | |
|---|---|---|---|---|---|---|
| | Regions of Image | | | Regions of Image | | |
| | 1 | 2 | TYPEI-3 | 1 | 2 | TYPEI-3 |
| $\alpha_i$ | 0.333 | 0.333 | 0.333 | 0.0640 | 0.06401 | 0.8719 |
| $a_i$ | 163.38 | 50.78 | $a_{i1} = -16.54$ | 2.27E-07 | 1.6537 | $a_{i1} =-0.2169$ |
| $q_i$ | 0.074 | −0.093 | $a_{i2} = 59.62$ | 441383 | −0.1738 | $a_{i2} = 267.13$ |
| | — | — | $m_{i1}$=0.2172 | — | — | $m_{i1}$=4.15387 |
| | — | — | $m_{i2}$=0.7828 | — | — | $m_{i2}$= 0.20633 |

Table 6 — WATER Image's parameters estimated values regions of image and its number (K =3)

| Parameters | Initial parameters estimation | | | | Final parameters estimation with Expectation Maximization Algorithm | | | |
|---|---|---|---|---|---|---|---|---|
| | Regions of image | | | | Regions of image | | | |
| | 1 | 2 | TypeI-3 | 4 | 1 | 2 | TypeI-3 | 4 |
| $\alpha_i$ | 0.250 | 0.250 | 0.250 | 0.250 | 0.126 | 0.1260 | 0.6219 | 0.1260 |
| $a_i$ | 6.444 | −38.68 | −33.04 | −33.12 | −28.93 | −1.79 | −0.107 | −33.120 |
| $q_i$ | 0.105 | −0.082 | 12.395 | 0.064 | 1.9409 | 0.1836 | −0.0909 | 0.0643 |
| | — | — | 0.727 | — | — | — | 2.2191 | — |
| | — | — | −0.273 | — | — | — | 0.5478 | — |

Table 7 — HILL image's parameters estimated values regions of image and its number (K =3)

| Parameters | Initial parameters estimation | | | | Final parameters estimation with Expectation Maximization Algorithm | | | |
|---|---|---|---|---|---|---|---|---|
| | Regions of image | | | | Regions of image | | | |
| | 1 | 2 | TypeI-3 | TypeI-4 | 1 | 2 | TypeI-3 | TypeI-4 |
| $\alpha_i$ | 0.250 | 0.250 | 0.250 | 0.250 | 0.250 | 0.25 | 0.250 | 0.250 |
| $a_i$ | 275.01 | −150.3 | $a_{i1}$ =−25.01 | $a_{i1}$ = −29.01 | 3.83E-210 | 1.0 | $a_{i1}$ = −0.067 | $a_{i1}$ = −0.079 |
| $q_i$ | 0.459 | 1.423 | $a_{i2}$ = 25.71 | $a_{i2}$ = 27.87 | 2.61E+209 | 1.0 | $a_{i2}$ = 0.0498 | $a_{i2}$ =−0.082 |
| | — | — | $m_{i1}$=0.493 | $m_{i1}$ =0.509 | — | — | $m_{i1}$= 2.605 | $m_{i1}$= 2.565 |
| | — | — | $m_{i2}$= − 0.51 | $m_{i2}$= −0.49 | — | — | $m_{i2}$= 0.414 | $m_{i2}$= 0.425 |

Table 8 — Performance measures of all three models

| Input pictures | METHOD | Performace measures | | |
|---|---|---|---|---|
| | | PRI | GCE | VOI |
| BOAT | GMM | 0.8813 | 0.6759 | 8.4587 |
| | PTIIID-K | 0.8844 | 0.6826 | 8.1262 |
| | PTIIID-H | 0.9863 | 0.8344 | 5.3877 |
| SEA | GMM | 0.8934 | 0.7815 | 8.4123 |
| | PTIIID-K | 0.9202 | 0.7644 | 8.2244 |
| | PTIIID-H | 0.9307 | 0.6626 | 8.1200 |
| WATER | GMM | 0.0109 | 0.0978 | 6.7759 |
| | PTIIID-K | 0.0112 | 0.0810 | 6.7502 |
| | PTIIID-H | 0.1221 | 0.0855 | 6.7053 |
| HILL | GMM | 0.0056 | 0.0089 | 7.8925 |
| | PTIIID-K | 0.0060 | 0.0075 | 7.4553 |
| | PTIIID-H | 0.0181 | 0.0056 | 7.2654 |

In order to understand the performance of the considered model, the current model values are compared with the other set of models like the GMM with K means and the hierarchical algorithm. The performance metrics are calculated for various input set of images like the HILL, WATER, SEA and BOAT. All these results obtained are shown in Table 9.

The results obtained and shown in the above Table 9 are very meaningful and can be used further for various types of other images too. The quality metrics of all the images considered here in the current methods are working properly and the results are going to meet the standard quality and criteria of the images considered. The methods used here are working properly and the results are more encouraging to continue further for the next further applications on these techniques. The current model results are better than the Gaussian Mixture model which can be considered as the standard model for the same type of applications.

From the above table, it is very clear that all the performance metrics of the images are considered and their performance were understood and analyzed with numerical values. The models have worked. The five performance metrics of the image are average difference, maximum distance, image fidelity, mean square error and the signal to noise ratio had the best output performance for the current considered models when compared with the existing Gaussian Mixture model, 3-parameter distribution model and the

Table 9 — Comparative study of image quality metrics

| IMAGE | Quality Metrics | Existing GMM | Current Considered PTIIID-K | Current Considered PTIIID - H | Existing 3-Parameter K-means | Existing K-means clustering algorithm |
|---|---|---|---|---|---|---|
| | Average Difference | 0.9545 | 0.8259 | 0.8256 | 0.9865 | 106.7574 |
| | Maximum Distance | 1.0000 | 1.0000 | 1.0000 | 0.5715 | 0.250 |
| BOAT/OSTRICH | Image Fidelity | 0.9568 | 0.9605 | 0.9784 | 0.8978 | 0.04863 |
| | Mean Square Error | 0.6584 | 0.5624 | 0.3275 | 0.0592 | 813 |
| | Signal to Noise Ratio | 6.8524 | 10.9467 | 15.785922 | 24.215 | 43.815 |
| | Average Difference | 0.8845 | 0.87542 | 0.84323 | 0.5845 | 153.8644 |
| | Maximum Distance | 1.000 | 1.000 | 1.000 | 0.9814 | 0.245 |
| SEA/WOMAN | Image Fidelity | 0.9999 | 0.9999 | 0.9999 | 0.4928 | 0.0412 |
| | Mean Square Error | 0.8549 | 0.3744 | 0.2932 | 0.0548 | 1264.33 |
| | Signal to Noise Ratio | 5.8964 | 3.7892 | 5.8980 | 5.1878 | 41.5788 |
| | Average Difference | 0.7942 | 0.7256 | 0.0006 | 0.0958 | 89.6244 |
| | Maximum Distance | 1.000 | 1.000 | 1.000 | 0.8914 | 251 |
| WATER/HILLS | Image Fidelity | 0.8973 | 0.9034 | 0.9902 | 0.9856 | 0.0445 |
| | Mean Square Error | 0.4582 | 0.2871 | 0.1581 | 0.3111 | 1023.55 |
| | Signal to Noise Ratio | 5.8597 | 1.29736 | 1.595354 | 2.1897 | 41.5107 |
| | Average Difference | 0.4897 | 0.2006 | 0.1991 | 0.3502 | 130.342 |
| | Maximum Distance | 1.000 | 1.000 | 1.000 | 0.7817 | 247 |
| HILL/EAGLE | Image Fidelity | 0.4589 | 0.5858 | 0.9142 | 0.9978 | 0.0543 |
| | Mean Square Error | 0.8943 | 0.8418 | 0.3133 | 19.245 | 844.3 |
| | Signal to Noise Ratio | 5.4589 | 0.0177 | 0.0126 | 0.9916 | 43.4396 |

hierarchical clustering models. The current considered model is performing better in almost all the performance metrics and can be implemented or used for further more models.

## Conclusions

In the current article, an attempt has been made to analyze the performance of the image segmented algorithms with the addition of the Pearsonian Type III mixture model. In the case of performance, the considered Pearsonian Type III mixture distribution with hierarchical model is showing better performance when compared with the Pearsonian Type III distribution mixture model with K-means and Gaussian mixture models.

**List of Abbreviations**

Some of the abbreviations used in the current article are,
PRI: Probabilistic Rand Index
GCE: Global Consistency Error
VOI: Volume of Interest
GMM: Gaussian Mixture Model
Pdf: Probability distribution function
EM Algorithm: Expectation Maximization Algorithm

**Declarations**

**Availability of data and materials**

The data and the material were available online as a free download for dataset.

**Competing interests:** N/A

**Funding:** Funding was supported by the Konkuk University, Korea.

**Authors' contributions:** The work is a joint research work of all the authors.

**Acknowledgements:** The authors should say thanks for the respective institutes for providing the research facilities to conduct this research.

## References

1 Srinivas Y, Rao K S, Reddy P V G D P, Unsupervised image segmentation using generalized Gaussian distribution with hierarchical clustering, *J Adv Res Comput Eng*, **4(1)** (2010) 43–51.

2 Pal N R & Pal S K, A review on image segmentation techniques, *Pattern Recognit*, **26(9)** (1993) 1277–1294.

3 Wang D & Whangbo T, Automatic diagnostic system for Parkinson’s disease based on deep learning using midbrain magnetic resonance images, *Int J Adv Sci Technol*, **124** (2019)1–20.

4 Gao T, Wang W, Liu W & Yang D, Rock particle image segmentation based on improved normalized cut, *Int J Control and Automation*, **10(4)** (2017) 271–286.

5 Yu J, Rui Y, Tang Y Y & Tao D, High-orderdistance-based multiview stochastic learning in image classification, *IEEE Trans Cybernet*, **44(12)** (2014) 2431–2442.

6 Fatima S & Seshashayee M, Experimental study of image segmentation using K-means clustering algorithm with image quality metrics, *Int J Eng Sci Invent*, **7(10)** (2018) 37–43.

7 Satyanarayana K V, Rao K S & Rao P S, Regional decomposition of images using three parameter logistic type

mixture model with K-Means, *Int J Comput Appl*, **181(32)** (2019)13–20.

8 Pillai A, Soundrapandiyan R, Satapathy S, Satapathy S C, Jung K-H, & Krishnan R, Local diagonal extrema number pattern: A new feature descriptor for face recognition, *Future Gener Comput Syst*, **81** (2018) 297–306. doi: 10.1016/j.future.2017.09.055

9 Naik A, Satapathy S C, Parvathi K, Improvement of initial cluster center of c-means using teaching learning based optimization, *Procedia Technol*, **6** (2012) 428–435.

10 Soundrapandiyan R & Chandra Mouli P V S S R, An approach to adaptive pedestrian detection and classification in infrared images based on human visual mechanism and support vector machine, *Arabian J Sci Eng*, **43** (2018) 3951–3963.

11 Shriranjani D, Tebby S G, Satapathy S C, Dey N & Rajinikanth V, Kapur's entropy and active contour-based segmentation and analysis of retinal optic disc, *Computational Signal Processing and Analysis*, (Springer, Singapore) (2018) 287–295.

12 Geetha R and Mohan J, Hybrid techniques for mri spine image classification, *J Sci Ind Res*, **79(12)** (2020) 829–832.

13 Palak Mahajan, Pawanesh Abrol & Parveen K Lehana, Scene based classification of aerial images using convolutional neural networks, *J Sci Ind Res*, **79(12)** (2020) 1087–1094.

14 Shaik H A & Rajesh V, Image processing based segmentation techniques for spinal cord in MRI, *Indian J Public Health Res Dev*, **9(6)** (2018) 317–323.

15 Lavanya G & Bhattacharyya D, Automatic identification of colloid cyst in brain through MRI/CT scan images, *ICMISC*, 3-5 September 2020.