Segmentation of image based on k-means and modified subtractive clustering

Simon Tongbram1, Benjamin A. Shimray2, Loitongbam Surajkumar Singh3
1,3ECE Department, National Institute of Technology Manipur, India
2EE Department, National Institute of Technology Manipur, India

ABSTRACT

Image segmentation has widespread applications in medical science, for example, classification of different tissues, identification of tumors, estimation of tumor size, surgery planning, and atlas matching. Clustering is a widely implemented unsupervised technique used for image segmentation mainly because of its simplicity and fast computation. However, the quality and efficiency of clustering-based segmentation is highly depended on the initial value of the cluster centroid. In this paper, a new hybrid segmentation approach based on k-means clustering and modified subtractive clustering is proposed. K-means clustering is a very efficient and powerful algorithm but it requires initialization of cluster centroid. And, the consistency of the clustering outcomes of k-means algorithm depends on the initial selection of the cluster center. To overcome this drawback, a modified subtractive clustering algorithm based on distance relations between cluster centers and data points is proposed which finds a more accurate cluster centers compared to the conventional subtractive clustering. These cluster centroids obtained from the modified subtractive clustering are used in k-means algorithm for segmentation of the image. The proposed method is compared with other existing conventional segmentation methods by using several synthetic and real images and experimental finding validates the superiority of the proposed method.

Keywords:
Clustering techniques
Image segmentation
K-means
Modified subtractive clustering
Subtractive clustering

This is an open access article under the CC BY-SA license.

Corresponding Author:
Simon Tongbram
Department of Electronics and Communication Engineering
National Institute of Technology Manipur
Langol Rd, Lamphelpat, Imphal, Manipur 795004, India
Email: simontongbram@gmail.com

1. INTRODUCTION

Image segmentation is a process that is used for partitioning an image into different regions containing similar characteristics in terms of color, intensity, or texture so that; the image can be easily understood for analysis [1]. Techniques such as object recognition, computer vision, tracking, and image analysis are primarily focused on image segmentation [2]. The main aim of image segmentation is to partition the image into different segments or regions related to a certain specific application and also to distinguish similar segments within the image, whether they are discrete or associated with specific objects. Several methods and algorithms [3] have been introduced for the process of image segmentation. The highly sophisticated medical imaging techniques of today, such as magnetic resonance imaging (MRI) and computerized tomography (CT) scans, create complex images which are very difficult to study and manually interpret. The requirement for the introduction of novel and efficient image processing methods or techniques is growing as a result of these complexities, leading to the creation of different processes which are robust.

Journal homepage: http://ijeecs.iaescore.com
and highly efficient at the same time. Applications of the image segmentation process consist of filtration of images, medical applications such as the location of tumors, measuring tissue volumes, diagnosis, computer-guided surgery and highly advanced applications such as facial recognition systems, and fingerprint recognition systems. Various techniques of the image segmentation process can be classified as edge-based [4], [5], threshold-based [6], region-based [7]-[9], watershed-based [10], and clustering-based segmentation.

Clustering is a widely used technique for image segmentation due to its simplicity and effectiveness [11]-[13]. In image processing, clustering is a mechanism by which indistinguishable image primitive are classified and they are combined to form clusters. The criteria for the similarity between the pixels are being introduced which helps in the formation of clusters from the identical pixels. And the criteria are also based on the type of applications. The combination of the identical pixels to form clusters is based on the objective to increase the similarity of the pixels belonging to the same cluster and decrease the similarity among different clusters. A modified k-means clustering algorithm is developed in [14], in which the initial cluster centroid is determined in a more systematic manner rather than choosing randomly as in the case of the conventional k-means method. Experimental results confirm that the developed algorithm lowers the magnitude of mean square error without affecting the time of execution and also, this process also produces much more accurate results in the case of dense datasets as compared to sparse datasets. Outlier rejection fuzzy c-means (ORFCM) algorithm, which is an enhanced version of traditional fuzzy c-means (FCM), is introduced in [15].

The algorithm is less sensitive to outliers due to the introduction of the exponent function in the expression of its membership function. The experimental evidence suggests that the developed method is more efficient in segmenting images with small intra-cluster and large inter-cluster variance. A segmentation algorithm for detecting the tumor of the brain in MRI images is introduced in [16], which uses fuzzy c-means and k-means technique. In this segmentation process, there are four stages consisting of the pre-processing stage, the segmentation process, the feature extraction process, and finally, the reasoning stage. The first stage is performed by the use of a filter which helps in increasing the image’s quality and in the case of the segmentation stage, k-means and fuzzy c-means are implemented. And the segmented image is utilized for the feature extraction process for research and analysis purposes. A more advanced clustering algorithm that is based on intuitionistic fuzzy c-means (IFCM) is suggested in [17] and this method is performed by acknowledging the local spatial information of the image.

The proposed method retains maximum image details and is more noise-resistant and, furthermore, the method does not involve adjusting the parameters. The k-means algorithm has some limitations, which are too much dependent on the initial parameter’s value, and to overcome this limitation, a new improved k-means algorithm is introduced in [18]. An automatic initialization process for the determination of the initial parameters is suggested by utilizing the histogram data of the image. This particular process improves the quality of the segmented image. An adaptive k-means clustering algorithm-based image segmentation method that does not require initialization of ‘k’ value is presented in [19]. The information of the luminance components of the image is used to select an adaptive value of ‘k’. This method increases the segmentation efficiency and also reduces the time of execution. Furthermore, the method reduces the influence of the background on the image and results in a more accurate segmentation. A dynamic k-means clustering algorithm is introduced in [20], to determine the number of clusters dynamically based on the calculated threshold value taken as the centroid of the k-means. Two data points will belong to the same group if the Euclidean distance between them is the same or lower than the selected threshold and if the distance is higher than the threshold, and hence creates a new cluster. A new method for determining the number of clusters for the k-means algorithm for high dimensional data is presented in [21]. The method is being experimented with 18 sets of normal as well as non-normal high dimensional data leading to the conclusion that the calculated number of clusters is far more accurate than other methods.

2. MATERIALS
2.1. K-means clustering algorithm

K-means clustering algorithm [22], [23] is widely used due to its uncomplicated process and fast computation. The image is being partitioned into clusters of data i.e., ‘k’ groups of data. A cluster’s centroid is represented by a point such that the sum of distances from all the other data points present in the cluster to the point is the least. The algorithm is used to determine the centroids for all the clusters.

To understand the algorithm let us take an image with dimension m*n and the objective is to cluster the image into ‘k’ groups of data. Let us take an input pixels y (m, n) which is to be clustered, and considering the cluster centroid as ‘c_i’. The algorithm is as follows: i) The cluster’s center and ‘k’ are initialized; ii) With the following relationship, the Euclidean distance represented by ‘d’ is determined for every image’s pixel;

Segmentation of image based on k-means and modified subtractive clustering (Simon Tongbram)
\[ d = \| y(m, n) - c_k \| \] (1)

iii) Based on the value of ‘d’ obtained, assign the pixels to the closest center; iv) With the following relation, calculate the new position of the centroid;

\[ c_k = \frac{1}{k} \sum_{n \in c_k} \sum_{m \in c_k} y(m, n) \] (2)

v) The process will be repeated till the tolerance value is achieved;

Although K-means algorithm is easy and straightforward to implement, one of the drawbacks of this particular method is that the forthcoming results for the clustering process will depend entirely on the value of the first centroid chosen. So, the segmentation quality and the algorithm efficiency rely on the selection of the initial centroid, the number of clusters chosen, and the extent of iterations to be performed.

2.2. Subtractive clustering algorithm

The subtractive clustering (SC) algorithm [24, 25] takes into account the density of neighbouring data points in a given feature space. The principle of the algorithm is to determine the data points with the highest density. The centre of the cluster will be the data point that has the highest potential value. Then, for the next centroid, the point with the next highest potential value is identified as the next centroid by the algorithm by excluding certain data points which are within a specified radius and this is repeated until a predefined requirement has been fulfilled. Consider a collection ‘Y’ that has ‘n’ number of data points such that \( Y = \{ y_1, y_2, y_3 \ldots y_n \} \) in an M dimensional space. And, we assume that each data points in the set can become the centroid of the cluster.

The centre of the cluster will be the data point that has the highest potential value. Then, for the initial centroid, the number of clusters chosen and is a positive constant. The data point’s potential is expressed as (3).

\[ \rho(y_i) = \sum_{j=1}^{n} e^{-4(y_j - y_i)^2} \] (3)

After deriving the potential value of all the data points, then, the data point having the highest potential is chosen as the first centroid. Considering \( y_1^* \) to be the first cluster center and \( \rho_1^* \) be the potential then, the updated potential of all the data points can be obtained from the following relation:

\[ \rho(y_i) = \rho(y_i) - \rho_1^* e^{-4(y_j - y_i)^2} \] (4)

where, \( r_b \) is a positive constant. Some potential from every data point is then reduced from the first centroid. There would be considerably less potential near the first centroid, so choosing the same centroid as the next centroid is not possible. The constant \( r_b \) is essentially the radius that determines the neighborhood with measurable potential reductions. We have selected \( r_b \) to be slightly larger than \( r_b \) to prevent the close distribution of cluster centers. As shown in (4), the potential value of all the data points is obtained, and pixel with the highest value is chosen as the next center. In general, the relation to update the potential value of each data point is given as (5).

\[ \rho(y_i) = \rho(y_i) - \rho_k^* e^{-4(y_j - y_k^*)^2} \] (5)

Where, \( y_k^* \) is the kth cluster center and potential value \( \rho_k^* \). The process of acquisition and revision of new cluster centres repeats itself until a given predefined condition is satisfied.

\[ \rho_k^* < \tau \rho_i^* \]

Where \( \tau \) is a small fraction. If \( \tau \) is very large, very few data points will be recognized as centroid and if \( \tau \) is very small, too many centroids will be created. The choice of the threshold is crucial for the results. The key issue with the SC algorithm is that it presumed the clustering center to be \( y_1^* \) obtained by the i\(^{th}\) iteration. \( y_1^* \) could, however, it may only present a rough cluster center which is not indicative of the actual cluster center. So, a modified distance measurement-based subtractive algorithm is proposed to identify the actual cluster centroid.
3. PROPOSED METHOD

The paper proposes a hybrid segmentation approach based on k-means and a modified SC algorithm. The initial assumption of cluster centres has a significant impact on the performance of k-means clustering. Therefore, a novel modified subtractive clustering method is employed to obtain a more accurate initial cluster centre, which helps to overcome the constraint encountered by k-means clustering. The algorithm of the modified methodology is explained in the following subsection.

3.1. Modified SC algorithm

It is an enhanced version of the traditional subtractive clustering method. The algorithm takes into account the distance relation between various data points and the center of clusters thereby reducing the computational time of the SC algorithm eventually delivers a more accurate cluster centre. To understand the modified algorithm, let us assume \( y_i^* \) represents the cluster centroids which are derived from SC algorithm in which \( i = 1, 2, 3, \ldots, C \). The algorithm defines a relation between the cluster centroid and the data points as given in (6).

\[
d_{ij} = e^{-\frac{4\sigma y_j-y_i^2}{r_d^2}} \tag{6}
\]

Where, \( i=1,2, \ldots, C; j=1,2, \ldots, n \). \( d_{ij} \) represents the distance relation between \( y_j \) and \( y_i^* \) and high value of \( d_{ij} \) indicates data \( y_j \) approximates the \( i \)th cluster center. Let \( L_k \) represent the number of data points surrounding \( y_k^* \) in which \( d_{ij} \geq \varepsilon \), in which \( \varepsilon \) is a constant. The new \( i \)th cluster centroid is defined by the given relation:

\[
m_i = \frac{\sum_{k=1}^{L_i} y_k^*}{L_i} \tag{7}
\]

where, \( y_k^* \) represents the data points in which \( d_{ij} \geq \varepsilon \). In (7), \( m_i \) represents the mean of the data neighbouring the cluster \( i \) and this relation delivers a more accurate centroid if \( y_i^* \) is determined by the SC algorithm and the potential is updated by:

\[
\rho(y_k^*) = \rho(y_k^*) - \frac{4\sigma y_k^* - m_i \rho^2}{r_b} \tag{8}
\]

3.2. Proposed algorithm

The algorithm of the proposed segmentation approach which includes k-means and modified subtractive is given as shown:

- Step 1: Initialization of parameters of the modified subtractive algorithm such as the number of clusters and the parameters \( r_a \) and \( r_b \)
- Step 2: As shown in (3) is used to determine the potential value.
- Step 3: The actual cluster center \( m_i \) is found out according to shown in (6) and (7).
- Step 4: As shown in (8) is used to update the potential of the remaining data points
- Step 5: Steps 3 and 4 are repeated for all the clusters
- Step 6: Initialize the parameters of the k-means algorithm
- Step 7: As shown in (1) is then used to determine the distance of each centroid from all the pixels
- Step 8: The pixel which has a minimum distance with respect to the centroid is then assigned to the cluster of the respective centroid
- Step 9: The position of the center is recalculated by using shown in (2).
- Step 10: The steps 7-9 are repeated until it satisfies the condition of tolerance or error value
- Step 11: The clusters are reshaped into the image

4. RESULTS AND DISCUSSION

The proposed hybrid method is compared with existing conventional clustering-based methods of segmentation namely, k-means, fuzzy c-means (FCM), intuitionistic fuzzy c-means (IFCM), and subtractive clustering which are implemented in MATLAB to verify the efficacy and superiority of the technique introduced. In the modified SC algorithm, the parameters \( r_a \) and \( r_b \) are set to be 0.5 and 0.75. As the number of clusters needs to be initialized for all the methods and in order to maintain a uniformity the value is given as 4 for all the methods. And, also for the quantitative performance analysis of the various methods, we have calculated mean square error (MSE) and peak signal to noise ratio (PSNR) for all the segmentation
approaches. Less value of MSE signifies the better quality of the output image. Small PSNR value signifies that the segmented image has an inferior quality and more prone to errors.

For experimentation, we have taken a synthetic image and real images obtained from Berkeley database [26] to test the performance of the various segmentation methods. Figure 1 shows the segmentation result of the synthetic image while Figure 2 shows the result for the real images taken from Berkeley database. The experiment was conducted on multiple grayscale images from the database and the results demonstrate superior or comparable performance with the existing conventional methods. Some experimental results have been presented in Figure 2. Visually, we can observe that all the methods perform well in the case of the synthetic image; however, the proposed method performs better for the real images. The superiority of the proposed methods is due to the accurate selection of the initial cluster centroid using the modified SC algorithm. Tables 1 and 2 presents the PSNR and MSE values of the various segmentation methods. Experimented values suggest that the proposed method provides better PSNR and MSE values compared to the other methods. The proposed method achieves a maximum PSNR of 62.35dB for the synthetic image and 45.52dB for the real image and a minimum MSE value of 1.07 for the synthetic image and 6.25 for the real image. Figures 3 and 4 show the graphical representation of the PSNR and MSE comparison of the various segmentation methods.

![Figure 1](image1.png)

**Figure 1. Image segmentation results of synthetic image: (a) original image, (b) K-means, (c) FCM, (d) IFCM, (e) subtractive and (f) proposed method**

| Table 1. PSNR comparison of different segmentation methods |
|--------------------------------------------------------|
| Synthetic image | Real image 1 | Real image 2 | Real image 3 | Real image 4 | Real image 5 | Real image 6 |
|-----------------|--------------|--------------|--------------|--------------|--------------|--------------|
| K-means         | 59.22        | 38.75        | 33.18        | 29.12        | 29.95        | 30.78        | 28.16        |
| FCM             | 58.25        | 40.74        | 32.05        | 30.75        | 29.47        | 32.42        | 29.12        |
| IFCM            | 59.15        | 41.89        | 32.82        | 29.72        | 30.12        | 26.11        | 28.72        |
| Subtractive     | 57.12        | 41.02        | 31.62        | 25.71        | 30.25        | 29.12        | 29.81        |
| Proposed        | 62.35        | 45.52        | 34.85        | 31.82        | 30.15        | 32.76        | 30.12        |

| Table 2. MSE comparison of different segmentation methods |
|--------------------------------------------------------|
| Synthetic image | Real image 1 | Real image 2 | Real image 3 | Real image 4 | Real image 5 | Real image 6 |
|-----------------|--------------|--------------|--------------|--------------|--------------|--------------|
| K-means         | 4.29         | 12.54        | 9.44         | 10.12        | 15.25        | 11.62        | 22.51        |
| FCM             | 5.53         | 10.11        | 9.65         | 11.14        | 16.32        | 12.81        | 24.15        |
| IFCM            | 4.15         | 8.12         | 8.12         | 10.52        | 16.73        | 11.25        | 19.72        |
| Subtractive     | 3.35         | 9.58         | 8.14         | 9.36         | 14.92        | 12.36        | 20.42        |
| Proposed        | 1.07         | 7.56         | 6.25         | 8.21         | 13.81        | 10.38        | 18.13        |
Segmentation of image based on k-means and modified subtractive clustering (Simon Tongbram)

Figure 2. Image segmentation results of real images: (a) original image, (b) K-means, (c) FCM, (d) IFCM, (e) subtractive and (f) proposed method

Figure 3. PSNR plot of different methods

Figure 4. MSE plot of different methods
5. CONCLUSION

A new method for image segmentation is proposed in this paper which is based on k-means and modified subtractive clustering methods. Segmentation is considered an important step in image processing for dividing the image into many regions so that each region represents a significant part of the image. The proposed method uses an enhanced subtractive algorithm which is based on distance relation between the data points and center of the clusters and delivers a more accurate center than the conventional subtractive clustering method. And, the centroids from the modified SC algorithm are used in the algorithm of k-means for segmenting the image. Moreover, for the purpose of evaluating the performance of the method, it is compared with other clustering-based image segmentation methods by using certain evaluation indices, MSE, and PSNR values and experimental results validate the superiority of the proposed method. While good performance indices have resulted from the proposed system, some aspects still remain to be improved. Certain efficient optimization methods can be incorporated with the method proposed in this paper which will define the optimal value of the prior initialization used in the algorithm and will further improve its performance.

REFERENCES

[1] S. A. Raut, M. Raghuvanshi, R. Dharaskar, and A. Raut, “Image segmentation-A state-of-art survey for prediction,” 2009 International Conference on Advanced Computer Control, 2009, pp. 420-424, doi: 10.1109/ICACC.2009.78.

[2] V. Guliashki, H. Toshev, and C. Korsemov, “Survey of evolutionary algorithms used in multiobjective optimization,” Problems of Engineering Cybernetics and Robotics, pp. 42–54, 2009.

[3] S. Saini and K. Arora, “A study analysis on the different image segmentation techniques,” International Journal of Information & Computation Technology, vol. 4, no. 14, pp. 1445-1452, 2014.

[4] N. Senthil Kumar and R. Rajesh, ”A comparative study on the different image segmentation techniques,” IJCSNS International Journal of Computer Science and Network Security, vol. 6, no. 9A, pp. 211-215, 2006.

[5] W. Luo, “Comparison for edge detection of colony images,” ICSSS International Journal of Computer Science and Information Technologies, vol. 5, no. 3, pp. 3289-3293, 2014.

[6] S. Snith and P. Soille, “Image segmentation by using threshold techniques”, Journal of Computing, vol.2, no. 5, pp.83-86, 2010.

[7] S. C. Zhu and A. Yuille, “Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 9, pp. 884-900, Sept. 1996, doi:10.1109/34.537343.

[8] H. G. Kaganani and Z. Beiji, “Region-based segmentation versus edge detection,” 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009, pp. 1217-1221, doi: 10.1109/IHH-MSP.2009.13.

[9] I. Karouzi, R. Fablet, J. Boucher, and J. Augustin, “Variational region-based segmentation using multiple texture statistics,” in IEEE Transactions on Image Processing, vol. 19, no. 12, pp. 3146-3156, Dec. 2010, doi: 10.1109/TIP.2010.2071129.

[10] L. Vincent and P. Soille, “Watersheds in digital spaces: an efficient algorithm based on immersion simulations,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 6, pp. 583-598, June 1991, doi: 10.1109/34.87344.

[11] S. Tongbram, B. A. Shimray, and L. S. Singh, “Clustering-based image segmentation techniques: A review,” Journal of Advanced Research in Dynamical and Control Systems, vol. 12, no. 7, pp. 701–707, 2020, doi: 10.5373/JARDCS/V12SP7/2020160.

[12] D. Naik and P. Shah, “A review on image segmentation clustering algorithms,” International Journal of Computer Science and Information Technologies, vol. 5, no. 3, pp. 3289–3293, 2014.

[13] B. Sathy and R. Manavalan, “Image segmentation by clustering methods: Performance analysis,” International Journal of Computer Applications, vol. 29, no. 11, pp. 27-32, 2011, doi: 10.5120/3688-5127.

[14] P. Purohit and R. Joshi, “A new efficient approach towards k-means clustering algorithm,” International Journal of Computer Applications, vol. 65, no. 11, pp. 7-10, 2013, doi: 10.5120/10966-6097.

[15] F. U. Siddiqui, N. A. M. Isa, and A. Yahya, “Outlier rejection fuzzy c-means (ORFCM) algorithm for image segmentation,” Turkish Journal of Electrical Engineering & Computer Sciences, vol. 21, no. 6, pp. 1801-1819, 2013, doi: 10.3906/elk-1111-29.

[16] A. Jose, S. Ravi and M. Sambath, “Brain tumor segmentation using K-means Clustering and fuzzy C-means algorithm and its area calculation,” International Journal of Innovative Research in Computer and Communication Engineering, vol. 2, no. 3, pp. 3496-3501, 2014.

[17] H. Verma, R. K. Agrawal, and A. Sharan, “An improved intuitionistic fuzzy c-means clustering algorithm incorporating local information for brain image segmentation,” Applied Soft Computing, vol. 46, pp. 543-557, 2016, doi: 10.1016/j.asoc.2015.12.022.

[18] Z. Khan, J. Ji, X. Fan, and P. Shi, “An improved K-means clustering algorithm based on an adaptive initial parameter estimation procedure for image segmentation,” International Journal of Innovative Computing, Information and Control, vol. 13, no. 5, pp. 1509–1526, 2017.
Segmentation of image based on k-means and modified subtractive clustering (Simon Tongbram)

BIOGRAPHIES OF AUTHORS

Simon Tongbram, completed B.E. in Electronics & Communication and M.Tech in Digital communication and Networking and started his teaching carrier in 2013. He is currently working as a Lecturer in EE dept. of National Institute of Technology Manipur, Imphal, India and also pursuing Ph.D. in the dept. of ECE of NIT Manipur. His area of interest is Image processing, data mining and optimization methods.

Dr. Benjamin A Shimray, received his Ph.D. and M.Tech from the National Institute of Technology Manipur (2018), Imphal and The National Institute of Engineering, Mysore (2010) respectively. Since 2011 he has been working as faculty member of department of Electrical Engineering, National Institute of Technology Manipur. He is also a member of various professional organization such as IEEE, IEI and IAENG. His area of interest is application of Soft Computing techniques to renewable energy planning and management, control system applications, and cyber vulnerabilities and system resilience.

Dr. Loitongbam Surajkumar Singh, after completing his Bachelors in Engineering and Masters in Technology in the field of Electronics and Communication Engineering started his teaching and research career as an Assistant Professor in Electronics and Communication Engineering Department of National Institute of Technology Manipur, Imphal, India from the year 2011. After obtaining a PhD degree from the same department, he is now actively carrying out research and projects in many Embedded Systems and VLSI related areas.

[19] X. Zheng, Q. Lei, R. Yao, Y. Gong, and Q. Yin, “Image segmentation based on adaptive K-means algorithm,” EURASIP Journal on Image and Video Processing, vol. 68, 2018, doi: 10.1186/s13640-018-0309-3.
[20] Md. Z. Hossain, Md. N. Akhtar, R.B. Ahmad and M. Rahman, “A dynamic K-means clustering for data mining,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 13, no. 2, pp. 521-526, 2019, doi: 10.11591/ijeecs.v13.i2.pp521-526.
[21] M. C. Alibhutto and N. I. Mahat, “Distance based k-means clustering algorithm for determining number of clusters for high dimensional data,” Decision Science Letters, vol. 9, no. 1, pp. 51-58, 2020, doi: 10.5267/j.dsl.2019.8.002.
[22] M. E. Celebi, H. A. Kingravi, and P. A. Vela, “A comparative study of efficient initialization methods for the k-means clustering algorithm,” Expert Systems with Applications, vol. 40, no. 1, pp. 200–210, 2013, doi: 10.1016/j.eswa.2012.07.021.
[23] M. Alhawarat and M. Hegazi, “Revisiting k-means and topic modeling, a comparison study to cluster arabic documents,” in IEEE Access, vol. 6, pp. 42740-42749, 2018, doi: 10.1109/ACCESS.2018.2852648.
[24] S. L. Chiu, “Fuzzy model identification based on cluster estimation,” Journal of intelligent and fuzzy systems, vol. 2, no. 3, pp. 267-278, 1994, doi: 10.3233/IFS-1994-2306.
[25] J. Y. Chen, Z. Qin and J. Jia “A weighted mean subtractive clustering algorithm,” Information Technology Journal, vol. 7, no. 2, pp. 356-360, 2008, doi: 10.3923/itj.2008.356.360.
[26] Martin D. and Fowlkes C., “The Berkeley segmentation database and benchmark,” Computer Science Department, Berkeley University, 2001.