

# Face Extraction from Image based on K-Means Clustering Algorithms

Yousef Farhang

Faculty of Computer, Khoj Branch, Islamic Azad University, Khoj, Iran

**Abstract**—This paper proposed a new application of K-means clustering algorithm. Due to ease of implementation and application, K-means algorithm can be widely used. However, one of the disadvantages of clustering algorithms is that there is no balance between the clustering algorithm and its applications, and many researchers have paid less attention to clustering algorithm applications. The purpose of this paper is to apply the clustering algorithm application to face extraction. An improved K-means clustering algorithm was proposed in this study. A new method was also proposed for the use of clustering algorithms in image processing. To evaluate the proposed method, two case studies were used, including four standard images and five images selected from LFW standard database. These images were reviewed first by the K-means clustering algorithm and then by the RER-K-means and FE-RER-clustering algorithms. This study showed that the K-means clustering algorithm could extract faces from the image and the proposed algorithm used for this work increased the accuracy rate and, at the same time, reduced the number of iterations, intra cluster distance, and the related processing time.

**Keywords**—K-means; RER-K-means; clustering algorithm; face extraction; edge detection; image clustering

## I. INTRODUCTION

Image segmentation is an important issue in today's world. In image processing, the topic of image extraction, specifically face extraction, has many applications [1]. There are many methods to extract facial images; among them, clustering method has not received adequate attention [2]. Clustering is an important method used in several areas of study such as face mining and knowledge discovery. In the clustering method, a set of objects are divided into subsets in such a way that similar objects are placed within a cluster [3], [4]. Thus, an object is similar to object(s) placed in the same cluster, whereas it is different from those positioned in other clusters in terms of predefined distance or similarity measure. Image clustering is a specific clustering method in which the objects to be clustered are images [5], [6].

This paper addresses a new application of K-means algorithm that is the most popular one among clustering algorithms. This algorithm can be used in many fields, including image mining, audio mining, education, finance, medical image [7], [8], management [9], and image clustering. One problem with clustering algorithms is that researchers have spent a disproportionate amount of time and effort to improve algorithms at the expense of considering additional applications of clustering algorithms. In general, there is an imbalance between clustering algorithms and their applications. In this paper, clustering algorithm is applied to

facial image extraction. The shortages of clustering algorithm that are listed in the second part can be found in the summary section of the Jain article [10].

In 1971, Cormack suggested that clusters should be internally integrative and externally segregated, suggesting a certain degree of uniformity within clusters and heterogeneity between clusters [11]. Thus, many researchers have attempted to operate this description through minimizing within-group disparity [12]. To maximize within-group uniformity, Sebestyen (1962) and MacQueen (1967) separately expanded the K-means style as a strategy for discovering optimal partitions [13], [14]. On the strength of this advancement, K-means has become very popular, earning a place in textbooks on multivariate techniques [15], [16], pattern recognition [17], cluster analysis [18], and image clustering. The clustering applications have been applied to pattern recognition (Anderberg, 1973), information retrieval (Rasmussen, 1992), and image processing (Jain, 1996) [19], [20]. However, despite many studies conducted on K-means clustering algorithms, few researchers have examined the application of this algorithm. In this paper, we examined the application of the K-means clustering algorithm in face extraction.

The rest of this paper is organized as follows. Section 2 reviews the literature in regard to K-means clustering algorithm and clustering algorithm. Section 3 explains the preliminaries used in this study. Section 4 explains the improved K-means algorithm proposed for face extraction. Section 5 reports the experiments carried out on the proposed algorithm and presents the evaluations on the experimental results. Finally, Section 6 concludes the paper.

## II. RELATED WORK

This section briefly reviews the studies previously-conducted on clustering algorithms and their application in image processing. First, the Forgy's method homogeneously allocates each point to one of the K clusters randomly [21]. The centers are then given with the centers of these primary clusters. This style has no theoretical basis. For example, random clusters have no internal homogeneity [22]. Second, Jancey's method [23] assigns a combinatorial point randomly generated in the space of data to each center. A number of these centers may be fully distant from any of the points, except the data set that fills the space, which might lead to the formation of unfilled clusters [24]. In 1967, MacQueen proposed two solutions. The first one is the default choice in the Quick Cluster method of IBM SPSS Statistics [25], which obtains the first K points in X as the centers. An obvious

disadvantage of this technique is its sensitivity to data ordering. The second way selects the centers randomly from among the data points. The basic idea is that by random choice, the selection of points from dense regions is as likely as any other, and these points are suitable to be centers.

Maximum method [26] selects the first center  $c_1$  randomly, and the  $i$ -th ( $i \in \{2, 3, \dots, K\}$ ) center  $c_i$  is selected to be the point that has the most minimum distance to the formerly-chosen centers, that are  $c_1, c_2, \dots, c_{i-1}$  was originally extended as an approximation to the K-center clustering problem. The assignment should include a vector quantization request; Katsavounidis variant obtains the point with the most Euclidean standard as the first center. Al-Daoud's density-based method first regularly partitions the data space into  $M$  decomposed hyper-cubes [27]. Then this randomly selects  $K$   $N_m/N$  points of hypercube  $m$  ( $m \in \{1, 2, \dots, M\}$ ) to create a number of  $K$  centers where  $N_m$  is the number of points in hypercube  $m$ . Bradley and Fayyad's method [28] begins by randomly partitioning the data set into  $J$  subsets. These subsets are clustered by k-means initialized through the MacQueen's second way producing  $J$  sets of intermediate centers, each with  $K$  parts. These center sets are united into a superset which is then clustered through k-means  $J$  times, each time initialized by a diverse center set. Parts of the center set that give the least SSE are then taken as the final centers.

Pizzuti advanced upon Al-Daoud's density-based method using a solution grid way [29]. This method starts with a 2D hypercube and iteratively divides these as the number of points they accept to expand. The k-means++ method [30] interpolates between maximin method and the MacQueen's second way. It opts the first center randomly, and the  $i$ -th ( $i \in \{2, 3, \dots, K\}$ ) center is  $x \in X$ , where  $md(x)$  denotes the minimum distance from a point  $x$  to the previously chosen centers. The PCA-Part method applies a divisive hierarchical way based on PCA (Principal Component Analysis) [31]. In this way, starting from a cluster that contains all data sets, the method iteratively chooses the cluster with the most SSE and divides it into two sub-clusters by a hyper-plane that passes with the center of cluster and is orthogonal to the way of the fundamental eigenvector of the covariance matrix. This process is repeated until  $K$  clusters are taken. The centers are then given through the centers of these clusters. Lu et al.'s method applies a two-phase pyramidal method [32]. The attributes of each point are first coded as integers. These parts of integers are considered to be at stage 0 of the pyramid. In the bottom-up stage, starting from stage 0, adjacent data points at stage  $k$  ( $k \in \{0, 1, \dots\}$ ) are averaged to take weighted points at step  $k + 1$  until at least  $20 K$  points are taken. Onoda's method [33] first computes  $K$  Independent Components (ICs) [34] of  $X$  and then opts the  $i$ -th ( $i \in \{1, 2, \dots, K\}$ ) center as the point that has the least cosine distance.

As the clustering algorithm is easy to implement, it can be widely used. One of the applications of clustering algorithm is in image processing, which was used for the first time in 1996 by Jane [35]. Jain used clustering for image retrieval by color and shape. In 1994, Brandt used fuzzy clustering in medical images such as MRI images [36]. In 1999, Lucchese employed the K-means clustering algorithm in image

segmentation [37]. At the same time, Ray and Turi applied K-means clustering to image segmentation. They proposed intra and inter clusters that could help to find the minimum distance in the cluster centers [38]. In 2002, Clausi proposed a K-means iterative fisher that was applied to image texture segmentation [39]. Chuang, in 2006, employed fuzzy c-means clustering with spatial information for image segmentation, which became a powerful method for noisy image segmentation [40]. Additionally, Cai, in 2007, and Yang, in 2009, used fuzzy c-means clustering for image segmentation [41], [42]. In 2009, Wang proposed adaptive spatial information-theoretic clustering to be used in image segmentation [43]. In 2010, Yu [44] and Das [45] applied pixel clustering to image segmentation. Simultaneously, Juang employed K-means clustering for segmentation in MRI brain images [46]. In 2011, Huang proposed X in which weight was selected in W-K-means clustering algorithm for color image segmentation.

This study uses K-means clustering algorithm for facial image extraction, which is explained in the next sections. The literature shows that the results of each proposed algorithm should be compared to those of other algorithms in terms of four factors: accuracy rate, intra cluster distance, number of iteration, and the duration of process. However, in most studies, the comparison has been made in terms of only one or two factors. Nevertheless, the results obtained from the algorithm proposed in the present paper are compared to those of two other algorithms in terms of all four factors mentioned above.

### III. PRELIMINARIES

In this section, three important items, namely, K-means clustering algorithm, image segmentation, and image feature extraction are described briefly to make them more clarified.

#### A. K-Means Clustering Algorithm

The goal of data clustering, also known as cluster analysis, is to discover the standard grouping of a set of patterns, points, or objects. Cluster analysis is defined as a statistical classification approach used to determine whether the individuals of a population fall into different groups through making quantitative comparisons of manifold characteristics. The aim is to develop a clustering algorithm that will find the normal groupings in the data of unlabeled objects. Clustering or cluster analysis is a technique of assigning a set of objects into clusters where all the objects in the cluster are considered to be similar based on common features. Clustering is an unsupervised learning-based method of statistical data analysis, which is used in many fields, including data mining, image analysis, pattern recognition, and image segmentation [47].

The most popular algorithm among clustering algorithms is the K-means one that is a rather easy but well-known algorithm for grouping objects [48]. For this reason, this algorithm is considered as the equivalent of clustering algorithms. The word "K-Means" was first used by James MacQueen in 1967, though the idea originated with Hugo Steinhaus in 1956. A standard algorithm was first proposed by Stuart Lloyd in 1982 as a method for pulse-code modulation

[49]. The major advantages of the K-means clustering algorithm are its simplicity and high speed, which allow it to run on big datasets [50]. The classical K-means clustering algorithm is aimed to detect a set C of K clusters  $C_j$  with cluster mean  $c_j$  to reduce the sum of squared errors [51]. This is typically described as follows:

$$E = \sum_{j=1}^K \sum_{x_i \in C_j} \|c_j - x_i\|^2 \quad (1)$$

Where, E is sum of the square error (SSE) of objects with cluster means for K cluster. It is also a distance metric between a data point  $x_i$  and a cluster mean  $c_j$ . For instance, the Euclidean distance is defined as:

$$\|x - y\| = \sqrt{\sum_{i=1}^V |x_i - y_i|^2} \quad (2)$$

The mean of cluster  $C_i$  is defined by the following vector:

$$c_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i \quad (3)$$

K-means clustering algorithms are fully described in Table 1.

TABLE I. THE K-MEANS CLUSTERING ALGORITHM

No.	Description of Algorithm Programming
1	Choose k center of cluster randomly
2	Assign initial values for cluster means $c_1$ to $c_k$
3	Repeat (Repeat the following steps until the cluster no longer changes)
4	for $i=1$ to $n$ do
5	Assign each data point $x_i$ to cluster $C_j$ where $\ c_j - x_i\ $ is the minimum
6	end for
7	for $j=1$ to $K$ do
8	Recalculate cluster mean $c_j$ of cluster $C_j$
9	end for
10	until convergence
11	return C

As can be seen in Table 1, the K-means clustering algorithm has four main steps: first, the initial cluster centers are selected randomly. Second, in the overall loop, the main steps of the algorithm are performed to achieve stability. Algorithm stability is determined as the constant sum of distances from cluster centers in the next step. Third, inside the outer loop, there is a main loop in which the core computation algorithm is run. In this loop, the first data row is performed calculated to the last data. This means that the distance for each row is calculated from its primary centers, which have less distance. The lines are placed in its cluster, and this work is performed for all rows. Fourth, after finishing the main loop, another calculation is performed in which new centers are calculated for each cluster and the new centers replace the initial centers. Then, the condition of stable solution algorithm is considered; if the answer of algorithm is not stable, the whole outer loop is run again.

K-means clustering algorithm is a greedy algorithm, which can only converge a local minimum, even though recent studies have exposed the enormous possibility that K-means can converge the overall optimum when clusters are well detached [52], [53]. The K-means begins with a primary partition with K clusters and allocates patterns to clusters so as to decrease the squared error. The key stages of standard K-means algorithm are as follow [54], [55]:

- 1) Select an initial partition with K clusters; repeat stages b and c until membership of cluster stabilizes.
- 2) Create a new partition through assigning each pattern to its closest cluster center.
- 3) Compute new cluster centers.

The problem is that the application of clustering algorithm and K-means algorithm has not been adequately studied. In the view of the researchers, clustering algorithms are applied algorithms; therefore, they can be used in different fields of study. For example, K-means clustering algorithm has the potential to be used for face extraction in image segmentation. This paper applies the K-means clustering algorithms to the face extraction.

### B. RER-K-Means Clustering Algorithm

The K-means clustering algorithm is one of the most practical algorithms that many researchers have used it and have improved. One of the improvements of the clustering algorithm is the reduction of error rate, which is achieved by the Reduction Error Rate in K-means (RER-K-means) clustering algorithm proposed by [56]. This algorithm reduces the number of errors and also increases stability. Standard K-means clustering algorithms are not stable; sometimes they do and sometimes do not get the correct answer. In general, the RER-K-means clustering algorithm reduces the number of errors and increases the stability of the algorithm. In this study, this algorithm is used for extraction of face from images.

### C. Image Segmentation

In this section, the image segmentation is briefly described. Image segmentation is the process of partitioning a digital image into multiple segments in the computer vision that is composed of sets of pixels. The objective of segmentation is the transformation of images into a model that is simple to understand and analyze. Image segmentation is normally used to locate objects and boundaries in the images. Generally, it is the process of assigning a label to each pixel of an image such that pixels through the same label share certain visual characteristics. Every pixel in an area is alike with respect to some characteristics or computed properties such as color, texture, or intensity. Nearby areas are significantly different with respect to the characteristic of similarity. The image segmentation result is a segment set that collectively covers the entire image, or a contour set extracted from the image that is seen as edge detection.

There are different methods to segment images, including compression-based methods, histogram-based methods, region-growing methods, split-and-merge methods, partial differential equation-based methods, graph partitioning

methods, and clustering methods. The last one (i.e., clustering method) is used in this research.

#### D. Image Feature Extraction

Feature extraction is a special form of dimensionality reduction in image processing. If the input data is too large for processing, the input data will be transformed into a reduced representative set of features. Transforming the input data into the set of features is named feature extraction. Whenever the features are carefully selected, the features set is expected to extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the total size input. The feature extraction module can be used to extract features in a format supported with machine learning algorithms from datasets consisting of formats, e.g., text and image.

Feature extraction can be used in image processing that involves using algorithms to find and isolate various desired portions or features of a video stream or digital image. It is an important characteristic in optical character recognition. If the feature extraction is used in image processing, it is called image feature extraction. Image feature extraction is an operation to extract various image features for identifying or interpreting meaningful physical objects from images. Features are classified into three types: spectral features (e.g., color, tone, ratio, and spectral index), geometric features (e.g., edges and lineaments), and textural features (e.g., pattern, homogeneity, and spatial frequency). There are different types of image feature extraction, which include face extraction. In this paper, the face extraction is investigated as explained in the next sections.

### IV. RESEARCH METHOD

In this paper, the K-means clustering algorithm is used to extract the face from an image. The FE-RER-K-Means algorithm improves the initial part in the K-means clustering algorithm, which is used to extract the face. In this section, all steps of the proposed algorithm are fully explained.

#### A. Initialization

In this study, initial value is randomly selected the data set is applied as cluster centers are selected randomly in the initial stage. In this study, MATLAB software is employed for simulation and implementation. Datasets in this study are color or grayscale images that are to be processed. First, an image is applied to MATLAB. For this purpose, the following program is used.

```
a=imread('Name of image .Format of image');
```

If the image was in color, the answer to the above formula is a 3D array. This array changes for case of use to a two-dimensional array with the following program:

```
x=rgb2gray(a);
```

The x array is two-dimensional and is a dataset for this study, which is clustered until the face can be extracted from the image. Then, initial values are found for clusters centers.

The overall goal of the proposed algorithm is to find the initial domain of each cluster. If this domain is an appropriate

choice, the clustering result will be appropriate. To find an appropriate primary domain for each cluster, the K-means clustering algorithm is run several times. For this operation, the initial cluster centers are selected randomly. After finding a suitable domain for each cluster within the range of each domain (in the domain [first member in the domain of each cluster, final member in the domain of each cluster]), an initial cluster center is selected randomly. Then, dataset is clustered based on the following proposed algorithm. If the algorithm is better able to determine the appropriate domain, the data set will be better clustered, and if the data set is better clustered, face extraction is performed better.

#### B. Choosing the Number of Clusters

Determination of the number of clusters in this study requires a focus on images. The size of the face in the image is very important when determining the number of clusters. This study uses four images, all including a face. If the size of the face in the image is greater than 50 percent of the image, (or the image size for segmentation was more than 50 percent of the image), it is better to use two clusters for clustering. Moreover, if the size of the face in the image is greater than 33 percent of the image, it is better to use three clusters for clustering. The general cases are as follows:

If 50% of the image (or 1/2 image) must be segmented  $n=2$   
 $\Rightarrow$  2 clusters should be considered.

If 33% of the image (or 1/3 image) must be segmented  $n=3$   
 $\Rightarrow$  3 clusters should be considered.

If 25% of the image (or 1/4 image) must be segmented  $n=4$   
 $\Rightarrow$  4 clusters should be considered.

If 20% of the image (or 1/5 image) must be segmented  $n=5$   
 $\Rightarrow$  5 clusters should be considered.

If (100/n)% of the image (or 1/n image) must be segmented  $\Rightarrow$  n clusters should be considered.

In this study, four images were used to evaluate the proposed algorithm. As the size of face in all of them was greater than 33 percent, three clusters were used for clustering. Table 2 depicts the way the number of clusters is chosen.

TABLE II. THE WAY TO CHOOSE THE NUMBER OF CLUSTERS IN THE FOUR IMAGES

Image No.	Size of Image	Size of Dataset	Clusters Size	Face Size	Face percent	No. of Cluster
1	96* 96	96	8,53,35	53	55%	2 or 3 cluster
2	512*512	512	104,221,187	221	43%	3 cluster
3	96*96	96	11,49,36	49	51%	2 or 3 cluster
4	225*225	225	87,96,42	96	42%	3 cluster
All Images						3 cluster

Additionally, in the second case study (image database), five images (with the same size of 250\*250) are selected randomly from database. In this case study, the number of clusters is set to three. As can be seen in Table 2, the number of clusters is chosen as the face size. In some images, two of the three clusters were selected. However, for better evaluation of the proposed algorithm, all the three clusters were chosen.

### C. Using Edge Detection Approach for Extraction

A set of mathematical methods for identifying points in a digital image is edge detection in which the image brightness changes sharply has discontinuities. Image brightness changes sharply has points that are typically organized into a set of curved line segments termed edges. Edge detection is a basic tool in image processing, computer vision, machine vision, feature detection, and feature extraction. Applying an edge detection algorithm to an image may considerably decrease the amount of data to be processed and may thus filter out information that may be regarded as less relevant, while preserving the main structural features of an image. An edge might, for instance, be the border between a block of red and a block of orange. In contrast, a line can be a small number of pixels of a distinct color on an otherwise unchanging background and the line can be extracted by a ridge detector. Therefore, there may be one edge on each side of the line for a line.

To show edge detection, changes in the light intensity of the image are used. When the image is segmented, it can be converted into an array, where the intensity of the image color ranges from 0 to 255. In general, 0 is black and 255 is white, and the other colors that range between these two numbers are considered. To illustrate an edge, consider the following example that contains six pixels. It can be seen that there is a too great difference between the third and fourth pixels. Such change in light intensity is indicative of an edge.

2	11	19	150	155	165

In this study, the following formula is used to calculate edge detection. If the previous pixel is considered  $a_1$  and the next one is considered  $a_2$ , then  $|a_1 - a_2|$  should be greater than 15 if the corresponding pixels are to be considered an edge. The value of  $|a_1 - a_2|$  should be greater than 15 pixels since after many experiments, a distance greater than 15 is deemed suitable for face extraction of image.

```
if abs(a1-a2)>15
{
g(Number of row, Number of column)=255;
}
```

Calculating for edge detection can be done after clustering. The first image clustered with clustering algorithm and then the cluster is located the face, in this cluster done to calculate edge detection to be extracted face from image.

### D. Problem Formulation

The purpose of this paper is to extract the face from facial images using the K-means clustering algorithms. The method used in this study is as follows: first, an image should be selected to be clustered by clustering algorithms. The image must be carefully chosen to ensure that it includes a face. In the next step, the clustering algorithm converts digital images into an array and the array is then used for clustering. An array is two dimensional and if the color image is divided into three dimensions, images are first converted to grayscale and then converted to an array using MATLAB software. Each color has its own numeric identity, with numbers ranging from 0 to 255. Similarly, an array contains numbers from 0 to 255, which are divided into groups using the clustering algorithms. During clustering, a cluster is a face because every face has its own color, and when converting face to an array, the facial numbers are close, which leads these images to being clustered together.

The number of clusters is then specified, which in this paper has been set to three. Next, the initial cluster centers are selected randomly, which in the K-means clustering algorithm means that cluster centers are selected randomly from the total array. Then, K-means clustering algorithm is applied and the array is divided into three clusters. This process continues until the algorithm becomes stable. Finally, three clusters consist of the three obtained arrays, which an array of face image is converted to image that is the answer. The framework of the facial image extraction applied in this study is shown in Fig. 1.

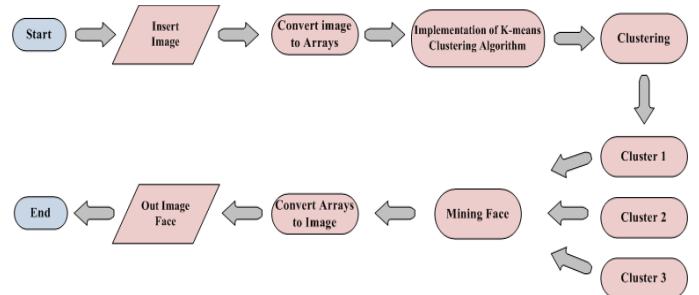


Fig. 1. Framework of facial image extraction using based K-means clustering algorithm.

In this paper, in addition to the use of K-means clustering algorithms for facial extraction as mentioned above, the proposed algorithm improves the K-means clustering algorithm, as described in the next section. Overall, the proposed algorithm's framework is similar to the framework presented in Fig. 1; however, there is a difference; the proposed algorithm chooses the initial values for cluster centers, making it a better method.

### E. Proposed Algorithm; Face Extraction with RER-K-Means Algorithm (FE-RER-K-Means)

In this section, an algorithm is proposed to improve the efficiency of RER-K-means clustering algorithms for face extraction. The name of the proposed algorithm is Face Extraction with RER-K-means algorithm (FE-RER-K-means

algorithm). The proposed algorithm is fully described in Table 3. The FE-RER-K-means clustering algorithm is composed of five parts. In the first part, image (dataset) is applied to MATLAB software. In the second part, the FE-RER-K-means algorithm finds the best domain for clusters. In the third part, the algorithm clusters image (dataset). In the fourth part, the proposed algorithm finds face clusters and to finding outs of program. In the fifth part, the algorithm converts the face cluster to an image (face mining).

TABLE III. THE FE-RER-K-MEANS CLUSTERING ALGORITHM

<p>Input: Image (Dataset), Number of Cluster, Size of Image                  Number of Iterations for Finding Domain: 10 Times                  Number of Iterations: 50 Times                  Number of Runs: 20 Times                  Output: Clusters, Face Image, Iteration, Related Processing Time, Accuracy Rate</p>
<p>..... Part 1: Applying of Image .....</p> <p>Step 1: Initially, the target image is applied to the MATLAB software. The image must meet the clustering conditions. In addition, the image itself should be a face.</p> <p>Step 2: The image is converted to an array.</p> <p>Step 3: Image analysis in terms of color or grayscale; if the image is in color, the image is three-dimensional.</p> <p>Step 4: The three-dimensional array is converted to two-dimensional.</p> <p>..... Part 2: Finding of Domain for Clusters .....</p> <p>Step 5: The two-dimensional array is ready for clustering. In this part, the problem is that no domain is the best for clusters; a problem that should be solved.</p> <p>Step 6: To find the best domain, K-means clustering algorithm is run several times on the dataset (in this study the algorithm is run 10 times).</p> <p>Step 7: Thus, the initial cluster centers are selected in the entire domain of dataset randomly. The number of rows of the dataset is found prior to random selection of the desired numbers of rows as cluster centers. The selected attributes of the random rows are assumed to be initial cluster centers.</p> <p>Step 8: The number of iterations required to find the ideal domain (for the purpose of this study) is set to 10. All main processes are placed into this loop.</p> <p>Step 9: A loop is made for the first to the last dataset in which all the main instructions can be placed.</p> <p>Step 10: The distances of cluster centers, which have been previously considered from all members of the dataset, are calculated. To calculate the distance, the coordinates of the cluster center in one array and attributes of a row as dataset in another array are placed, and then the distance between these two arrays is calculated using the following formula. This operation is carried out for all cluster centers in one step.</p> $D(j) = \sum_{i=1}^m \sqrt{\sum_{k=1}^n [A(i, k) - A(j, k)]^2}$ <p>Step 11: The distance of all cluster centers from one of the datasets is calculated separately and the minimum distance is taken into consideration. Members of datasets are then placed in the cluster with the minimum distance.</p> <p>Step 12: This step is the end of the internal loop. It means that steps 9 to 11 are run until the termination condition of the inter loop occurs.</p> <p>Step 13: The means of any cluster should be determined separately. Then, at the end of any step, the determined means are considered as cluster centers for the next step.</p> <p>Step 14: This step is the end of the outer loop. It means steps 8 to 13 are run</p>

until the termination condition of the outer loop occurs.

Step 15: Initial clustering in this episode is over and the domain is determined for each cluster.

..... Part 3: Clustering of Image .....

Step 16: After selecting the domain for each cluster, initial cluster centers for each cluster are selected randomly and separately.

Step 17: The number of iterations is fixed at 20 for all datasets in this study. All main processes are placed into this loop.

Step 18: A loop is created from the first to the last dataset in which all the main instructions can be placed. This loop is the internal loop.

Step 19: The distances of cluster centers, which have been previously considered from all members of the dataset, are calculated. To calculate the distance, the coordinates of the cluster center in one array and attributes of a row as dataset in another array are placed, and then the distance between these two arrays is calculated using the formula presented in step10. This operation is carried out for all cluster centers in one step.

Step 20: The distances of all cluster centers from one of the datasets are calculated separately and the minimum distance is taken into consideration. At this time, members of datasets are placed in the cluster with the minimum distance.

Step 21: Some variables are defined to represent summation of distances between cluster center and its members. The number of defined variables should be equal to the number of clusters. For instance, if there are three clusters, three variables, namely s<sub>1</sub>, s<sub>2</sub>, and s<sub>3</sub> are defined in which s<sub>i</sub> is summation of distances between i<sup>th</sup> cluster center to its member (i=1, 2, 3).

Step 22: This step is the end of the internal loop. It means that steps 18 to 21 are run until the termination condition of the internal loop occurs.

Step23: In this step, the number of iterations, accuracy rate, and related processing time (s) are calculated.

Step 24: Variable S which is intra cluster distance is defined as summation of s<sub>1</sub>, s<sub>2</sub>, s<sub>3</sub>, and so on. The convergence of S indicates that the algorithm has been stabilized.

Step 25: The means of any cluster should be determined separately. Then, at the end of any step, the determined means are considered as cluster centers for the next step.

Step 26: This step is the end of outer loop. It means steps 17 to 25 are run until the termination condition of the outer loop occurs.

..... Part 4: Finding the Face Cluster .....

Step 27: After clustering the dataset (image), a face cluster should be found between clusters.

Step 28: To find a face cluster, the largest cluster is selected since a face in the selected image has maximum domain.

..... Part 5: Face Mining .....

Step 29: After finding the cluster image, the face should be extracted. This is because in database clustering, all attributes are discovered and included in all columns of dataset. In general, the figure should be extracted from the respective cluster.

Step 30: The edge detection method is used for face mining. First, an array is created with the same size as the face cluster, in which all pixels are placed at 0 that represents the color black.

Step 31: A loop is created for the first to the last face cluster size; this loop is the outer loop.

Step 32: A loop is created for the first to the last number of attributes; this loop is the internal loop.

Step 33: The edge detection is calculated. If the previous pixel is considered

$a_1$  and the next pixel is considered  $a_2$ , the calculation is  $|a_1 - a_2|$ .  
 Step 34: If  $|a_1 - a_2| > 15$ , it is considered as an edge and  $g$  (Number of row, Number of column) = 255. The number 255 is represented in white color.  
 Step 35: This step is the end of the internal loop. It means that steps 32 to 34 are run until the termination condition of the internal loop occurs.  
 Step 36: This step is the end of the outer loop. It means steps 31 to 35 are run until the termination condition of the outer loop occurs.  
 Step 37: The array obtained in 34<sup>th</sup> step (array  $g$ ) is converted to an image. This image is the final answer.  
 Step33: The edge detection is calculated.If the previous pixel is considered  $a_1$  and the next pixel is considered  $a_2$ , the calculation is  $|a_1 - a_2|$ .  
 Step34: If  $|a_1 - a_2| > 15$ , it is considered an edge and  $g$  (Number of row, Number of column) = 255. The number 255 is represented in white color.  
 Step35: This step is the end of the inter loop. It means that steps 32 to 34 are run until the end condition of the inter loop.  
 Step36: This step is the end of the outer loop. It means steps 31 to 35 are run until the end condition of the outer loop.  
 Step37: The array obtained in 34<sup>th</sup> stage (array  $g$ ) is converted to an image. This image is the final answer.

In Table 3, the proposed algorithm is fully and clearly explained. In the next section, the results of the implementation of the K-means clustering algorithm, RER-K-means clustering algorithm, and the FE-RER-K-means algorithm are presented. The results obtained from the three algorithms are compared to each other and the advantages and disadvantages of the three algorithms are described.

### V. EXPERIMENTAL RESULTS AND EVALUATION

This section falls into two parts: experimental results and a comparison between them. First, in the part experimental results, the results obtained by the K-means clustering algorithm and the proposed algorithm are presented. Second, the results reported in the previous part are evaluated. Also, two case studies (i.e., standard images and LFW image database) are investigated for evaluation of the proposed algorithm.





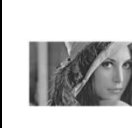






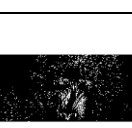
#### A. Experimental Results for Standard Images

In this section, K-means clustering and the proposed algorithms are used to extract a face from an image. First, using the MATLAB software, the K-means clustering algorithm and the proposed algorithm described in the previous sections have been implemented, and their databases use images depicted in Table 4.

In Table 4, the image database used in this paper is presented. In this section, four images are used each of which contains a face. In the first step, an original image is applied to the clustering algorithm. In the second step, after processing, a clustered image is obtained, which includes an array. In the third step, this array is processed and the face is extracted. To

perform this step, the color method in grayscale images is used since each pixel converted into an array takes on a number between 0 and 255. Everyone has the same face color and when the pixel is converted to an array, the numbers come closer to each other. Finally, face extraction is achieved when the difference between two adjacent pixels is calculated.

TABLE IV. THE STANDARD IMAGE CONTAINING THE ORIGINAL IMAGE, IMAGE CLUSTERING AND FACE MINING

Name of image	Original Image	Image clustering	Face Mining
Person			
Lenna			
Monkey			
Mandrill			

As suggested in Table 4, after extracting the facial images using K-means clustering, RER-K-means, and the FE-RER-K-means algorithms, factors can be examined in all of these algorithms. In Table 5, four facial extraction images are considered using K-means clustering and two others algorithms and these three algorithms are evaluated in regard to nine criteria. Factors evaluated using these three algorithms include four factors for the number of iterations (average and standard deviation), one factor for accuracy rate, and four factors for the related processing times (average and standard deviation). Additionally, Table 5 expresses image size and the number of clusters for all images.

As can be seen in Table 5, in all factors studied, the FE-RER-K-means algorithm is better than the other algorithms. In the next section, the other case study is checked for evaluation proposed algorithm.



TABLE V. THE NUMBER OF ITERATIONS, NUMBER OF CLUSTERS, ACCURACY RATE, AND RELATED PROCESSING TIMES IN THE THREE ALGORITHMS TO EXTRACT FACES FROM FOUR STANDARD IMAGES

Name of image	Name of Algorithm	Number of iteration		Intra Cluster Distance		Related processing time (s)		Accuracy rate (%)
		Average	Std. Dev.	Average	Std. Dev.	Average	Std. Dev.	
Person	K-means	8.25	2.65	35900	1381	0.38	0.18	85
	RER-K-means	5.30	1.62	34948	621	0.80	0.14	90
	FE-RER-K-means	3.35	0.74	34192	306	0.68	0.11	95
Lenna	K-means	8.20	1.98	138166	1341	6.17	0.33	80
	RER-K-means	6.80	1.54	137716	617	2.45	0.30	85
	FE-RER-K-means	4.95	1.39	137678	507	2.26	0.16	95
Monkey	K-means	7.25	1.68	51161	444	0.34	0.17	65
	RER-K-means	5.55	1.60	50877	435	0.59	0.05	70
	FE-RER-K-means	4.55	1.14	50612	327	0.53	0.02	85
Mandrill	K-means	9.60	2.30	107302	1910	0.60	0.40	75
	RER-K-means	9.35	2.23	105213	983	1.43	0.24	80
	FE-RER-K-means	7.80	1.88	104362	270	1.21	0.06	95

*B. Experimental Results for LFW Standard Images Database*

In this section, the LFW (Labeled Faces in the Wild) standard image database used for testing and evolution of the proposed algorithm is presented. LFW is a database of face photographs designed by the LFW University of Massachusetts, Amherst for studying the problem of unconstrained face recognition [57]. This database contains 5749 subjects among which 5 images were selected randomly to be used in this study.

All algorithms were run 20 times for five images. In this stage shows the result obtained from the proposed algorithm (FE-RER-K-means algorithm) for image3, which was run 20

times. As can be seen, run No. 3 has error and its clustering is not good. Apart from that, the rest of the steps were implemented correctly.

After extracting the facial images using the three algorithms, as suggested in the previous table, factors can be examined in all of the algorithms. In Table 6, the results obtained regarding five images from LFW standard image database are presented. Factors and algorithms are the same as those in the first case study.

As can be seen in Table 6, in all factors studied, the FE-RER-K-means algorithm is more successful than the other algorithms. In the next section, all these factors are fully evaluated.

TABLE VI. THE NUMBER OF ITERATIONS, NUMBER OF CLUSTERS, ACCURACY RATE, AND RELATED PROCESSING TIMES IN THE THREE ALGORITHMS TO EXTRACT FACES FROM LFW STANDARD IMAGE DATABASE

Name of image	Name of Algorithm	Number of iteration		Intra Cluster Distance		Related processing time (s)		Accuracy rate (%)
		Average	Std. Dev.	Average	Std. Dev.	Average	Std. Dev.	
Image1	K-means	9.60	1.66	127919	407	3.64	0.17	95
	RER-K-means	9.50	1.53	127672	199	1.61	0.05	95
	FE-RER-K-means	6.45	1.14	127579	109	1.53	0.02	95
Image2	K-means	7.45	1.53	152056	755	3.66	0.19	95
	RER-K-means	7.10	1.37	151865	595	1.60	0.04	95
	FE-RER-K-means	4.45	1.09	151612	266	1.54	0.01	95
Image3	K-means	6.60	1.14	152043	2702	3.72	0.04	70
	RER-K-means	5.20	0.83	151453	2353	1.58	0.03	75
	FE-RER-K-means	4.30	0.73	150225	520	1.47	0.02	95
Image4	K-means	6.25	0.96	146407	3652	3.76	0.08	90
	RER-K-means	5.55	0.75	145788	2659	1.57	0.05	95
	FE-RER-K-means	4.75	0.71	145788	2659	1.52	0.02	95
Image5	K-means	6.55	1.39	161276	3773	3.74	0.05	85
	RER-K-means	6.45	1.31	161164	3479	1.55	0.03	85
	FE-RER-K-means	5.95	0.94	159912	2828	1.52	0.02	95



### C. Evaluation Results for Standard Images

In this section, the results obtained from the three algorithms are evaluated using four standard images. To extract faces from images, K-means clustering algorithm, RER-K-means Clustering algorithm, and FE-RER-K-means algorithm are employed. In this study, MATLAB software is used for evaluation. In addition, the evaluation and system implementation used processor (Intel(R) core(TM) i3-2310M CPU @ 2.10GHz), with an installed memory of (4.00 GB) and system type (64-bit operating system). To evaluate the performance of the three algorithms in this part, five factors are used, and a comparison chart for all five factors is drawn. All of these graphs use data presented in Table 5, and this data is used for all four images.

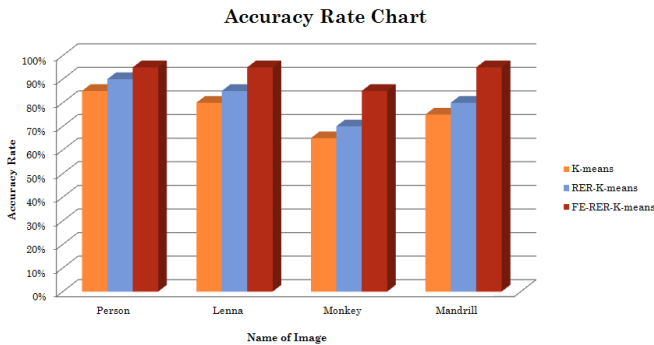


Fig. 2. Diagram of accuracy rate in three algorithms for face extraction from four images.

First, the accuracy rate of the K-means, RER-K-means, and FE-RER-K-means algorithms is discussed. As shown in Fig. 2, the FE-RER-K-means algorithm has better performance than the K-means algorithm in all four images, which indicates that this algorithm appears to deliver more answers. This efficiency is higher in the fourth image, meaning that the proposed algorithm outperforms the others on these images. Fig. 2 shows the accuracy rate in the K-means and the proposed algorithms. In general, the FE-RER-K-means algorithm has fewer errors than the RER-K-means and it has fewer errors than K-means clustering algorithm in all four images.

Second, the K-means, RER-K-means, and FE-RER-K-means algorithms are compared in regard to the average number of iterations. Fig. 3 compares the three algorithms based on the average number of iterations. This diagram is seen as the benchmark for all images, and shows that the proposed algorithm is more successful than the K-means and RER-K-means algorithms. The first image is a personal photograph, and the difference is large, indicating that the proposed algorithm works best on images that contain personal photos.

Third, three algorithms are compared using the standard deviation of the number of iterations. In Fig. 4, four images are presented and they are compared using this factor in both algorithms. In all images, it can be seen that the proposed algorithm outperforms the others. In the first image, the proposed algorithm substantially reduces the standard deviation.

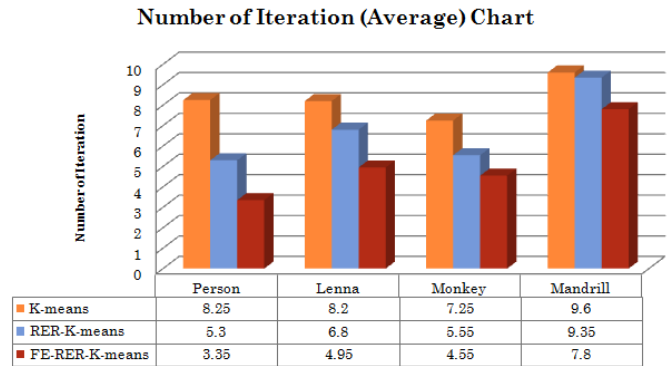


Fig. 3. Diagram of the average number of iterations in three algorithms for face extraction from four images.

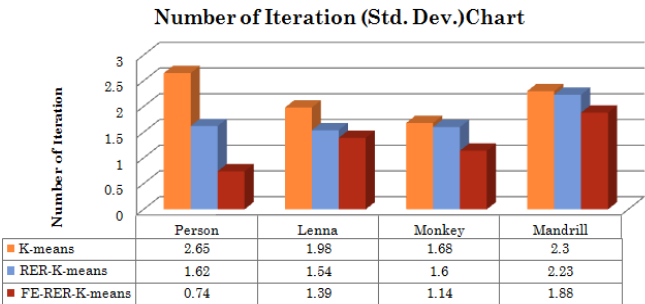


Fig. 4. Diagram of the standard deviation of the number of iterations in three algorithms for face extraction from four images.

Fourth, the three algorithms are evaluated based on processing time. In Fig. 5, the proposed algorithm is compared with K-means algorithm in terms of the average related processing times in all four images.

Finally, the three algorithms are examined using the standard deviation of related processing times. Fig. 6 makes a comparison between the proposed algorithm and K-means and RER-K-means algorithms in regard to all four images. It can be seen that the proposed algorithm is more successful than others in all four image processing times and has less standard deviation in terms of related processing times. In the third and fourth images, the proposed algorithm reduced substantially the standard deviation.

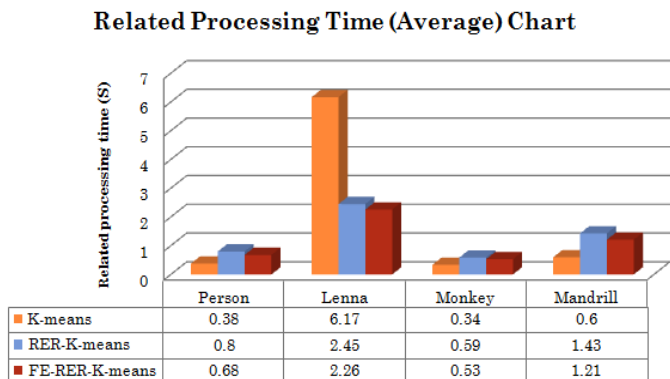


Fig. 5. Diagram of the average related processing times in three algorithms for face extraction from four images.

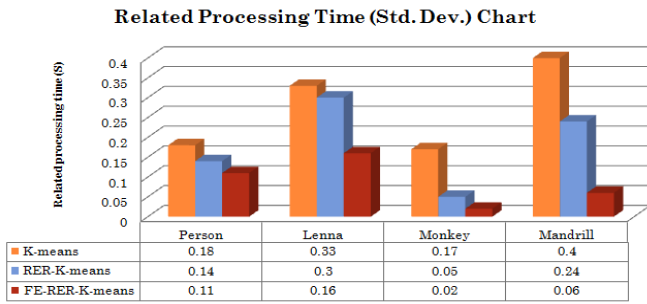


Fig. 6. Diagram of standard deviation of related processing time in three algorithms for face extraction from four images.

In this section, the results reported in the previous section are evaluated and discussed with regard to four standard images. Results obtained through the use of five factors (accuracy rate, average number of iterations, standard deviation of the number of iterations, average of related processing time, and standard deviation of related processing time) have been evaluated. In the five evaluations, the performance of the FE-RER-K-means algorithm is better than RER-K-means and K-means algorithms, suggesting that the proposed algorithm is an improved version of the K-means algorithm.

**D. Evaluation Results for LFW Standard Images Database**

In this section, the results obtained through implementation of the three algorithms are evaluated using LFW standard image database. To extract faces from images, K-means clustering, RER-K-means clustering, and the proposed algorithms are used. To evaluate the performance of the three algorithms in this part, five factors are employed, and a comparison chart for all five factors can be drawn.

First, the accuracy rate of the K-means clustering algorithm and the two other algorithms is discussed. As shown in Fig. 7, the FE-RER-K-means algorithm has better performance than the other algorithms in case of all five images, which indicates that this algorithm appears to deliver more answers. This efficiency is higher in the third image, indicating that the proposed algorithm gives a better answer in all images. The diagram depicted in Fig. 7 shows the percentage of the accuracy rate. Generally, the proposed algorithm has fewer errors than the K-means clustering algorithm in all five images.

Second, the K-means, RER-K-means, and the proposed algorithms are compared in terms of the average number of iterations (see Fig. 8). The diagram is seen as the benchmark for all images, and it can be seen that the FE-RER-K-means algorithm outperforms the K-means and RER-K-means algorithm.

Third, three algorithms are compared regarding the standard deviation of the number of iterations. In Fig. 9, five images are seen and they are compared using this factor in three algorithms. In all images, it can be seen that the FE-RER-K-means algorithm has better performance and in the first image, the proposed method substantially reduces the standard deviation.

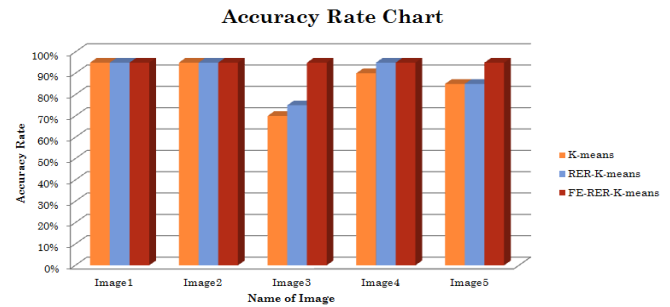


Fig. 7. Diagram of accuracy rate in three algorithms for face extraction from five images in LFW Database.

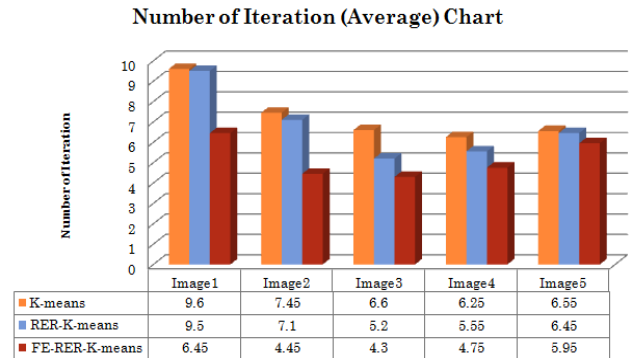


Fig. 8. Diagram of the average number of iterations in three algorithms for face extraction from five images in LFW database.

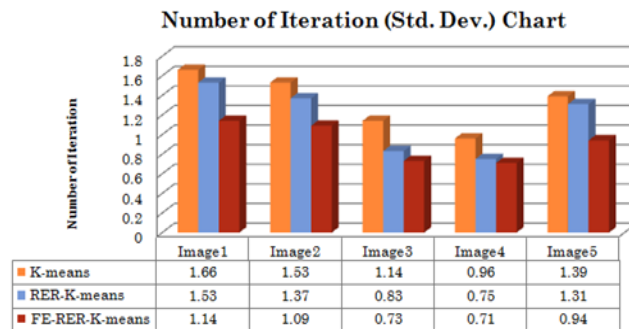


Fig. 9. Diagram of the standard deviation of the number of iterations in three algorithms for face extraction from five images in LFW database.

Fourth, the three algorithms are evaluated in terms of the processing time. In Fig. 10, the proposed algorithm is compared to the K-means and RER-K-means in terms of the average related processing times in all five images. In case of all images, it can be seen that the proposed algorithm has a better performance and requires less processing time.

Finally, the three algorithms are examined in terms of the standard deviation of related processing times. In Fig. 11, the proposed algorithms are compared with K-means algorithm in case of all five images. It can be seen that the proposed algorithm outperforms the others in all five image processing times and has less standard deviation in related processing times. In the first and second images, the proposed algorithm delivers a substantially reduced standard deviation.

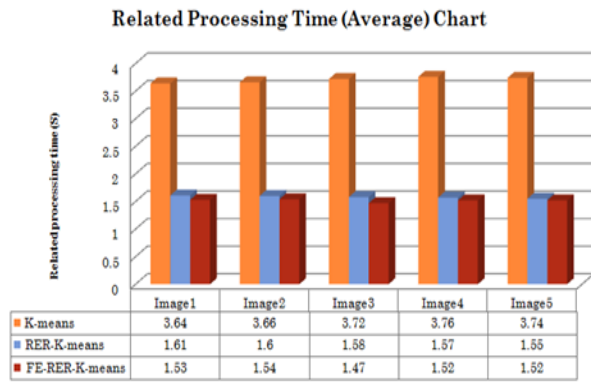


Fig. 10. Diagram of the average related processing times in three algorithms for face extraction from five images in LFW database.

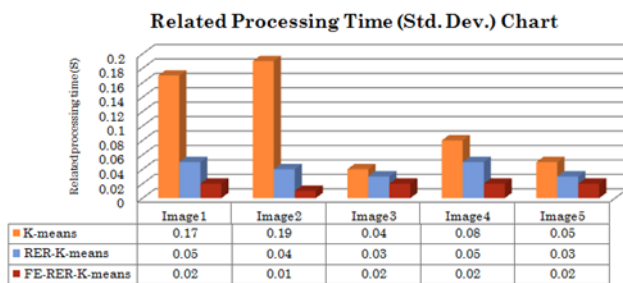


Fig. 11. Diagram of standard deviation of related processing time in three algorithms for face extraction from five images in LFW database.

In this section, the results obtained in the previous section were evaluated and discussed with regard to five images in LFW standard database. Results were obtained using five factors (i.e., accuracy rate, average number of iterations, standard deviation of the number of iterations, average of related processing time, and standard deviation of related processing time). The evaluations demonstrated that the performance of the proposed algorithm (FE-RER-K-means) was better than the *K*-means clustering algorithm in terms of all five factors mentioned above, suggesting that the proposed algorithm is an improved version of the *K*-means algorithm.

## VI. CONCLUSION

This paper focused on the application of clustering algorithms; particularly, which clustering algorithm was the best in terms of extracting faces from images. It also noted that one of the problems with clustering algorithms was that researchers had made more effort to improve the existing algorithms and less effort on the applications of algorithms. In general, there is an imbalance between the application of the algorithm and the improvement of algorithms. To solve this problem, this paper used *K*-means and RER-*K*-means algorithms to extract images and proposed an improved algorithm. Then, results of these three algorithms were reviewed based on 13 factors (average number of iterations, standard deviation of the number of iterations, best of intra cluster distance, worst of intra cluster distance, average of intra cluster distance, standard deviation of intra cluster distance, average of related processing times, standard deviation of related processing times, and accuracy rate). It was shown that the proposed algorithm (FE-RER-*K*-means

algorithm) outperformed the others in terms of the all factors. To summarize, this article attempted to find a balance between the applications of clustering algorithms and the improvement of clustering algorithms.

In this study, a method was proposed to solve one of the problems of clustering algorithms, i.e., the imbalance of the clustering algorithms. In this paper, for the first time, the *K*-means clustering algorithm for face extraction was used. Additionally, an innovative, improved clustering algorithm was proposed to extract the face. In general, the purpose of this paper was face extraction through *K*-means clustering algorithm, which in combination with the proposed improved algorithm caused a reduction in processing times, number of iterations, and intra cluster distance, and an increase in the accuracy rate. In future studies, other problems of clustering can be addressed using clustering algorithm. Also, the proposed algorithm can be evaluated with other criteria. Finally, the database can consider medical images such as those used in radiology, mammography, and in cancer patients.

## REFERENCES

- [1] M. Castrillón, *et al.*, "ENCARA2: Real-time detection of multiple faces at different resolutions in video streams," *Journal of Visual Communication and Image Representation*, vol. 18, pp. 130-140, 2007.
- [2] S. Young, *et al.*, "Hierarchical spatiotemporal feature extraction using recurrent online clustering," *Pattern Recognition Letters*, vol. 37, pp. 115-123, 2014.
- [3] J. M. Duarte, *et al.*, "A Constraint Acquisition Method for Data Clustering," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, ed: Springer, 2013, pp. 108-116.
- [4] O. M. Jafar and R. Sivakumar, "A Comparative Study of Hard and Fuzzy Data Clustering Algorithms with Cluster Validity Indices," ed: Elsevier Publications, 2013.
- [5] S. N. Sulaiman and N. A. M. Isa, "Adaptive fuzzy-*K*-means clustering algorithm for image segmentation," *Consumer Electronics, IEEE Transactions on*, vol. 56, pp. 2661-2668, 2010.
- [6] P. Thakur and C. Lingam, "Generalized Spatial Kernel based Fuzzy C-Means Clustering Algorithm for Image Segmentation," *International Journal*, 2013.
- [7] H. H. A. Afzali, *et al.*, "A model-based evaluation of collaborative care in management of patients with type 2 diabetes in Australia: an initial report," *Australian Health Review*, vol. 36, pp. 258-263, 2012.
- [8] E. Hemphill and C. T. Kulik, "Segmenting a general practitioner market to improve recruitment outcomes," *Australian Health Review*, vol. 35, pp. 117-123, 2011.
- [9] Y. Farhang, *et al.*, "Granular Computing Based Data Mining in the Views of Rough Set and Fuzzy Set," in *Informatics Engineering and Information Science*, ed: Springer, 2011, pp. 624-629.
- [10] A. K. Jain, "Data clustering: 50 years beyond *K*-means," *Pattern Recognition Letters*, vol. 31, pp. 651-666, 2010.
- [11] R. M. Cormack, "A review of classification," *Journal of the Royal Statistical Society. Series A (General)*, pp. 321-367, 1971.
- [12] D. Steinley, "K-means clustering: A half-century synthesis," *British Journal of Mathematical and Statistical Psychology*, vol. 59, pp. 1-34, 2006.
- [13] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1967, p. 14.
- [14] G. S. Sebestyen, "Decision-making processes in pattern recognition (ACM monograph series)," 1962.
- [15] R. A. Johnson and D. W. Wichern, *Applied multivariate statistical analysis* vol. 5: Prentice hall Upper Saddle River, NJ, 2002.

- [16] J. M. Lattin, *et al.*, *Analyzing multivariate data*: Thomson Brooks/Cole Pacific Grove, CA, 2003.
- [17] R. O. Duda, *et al.*, "Unsupervised learning and clustering," *Pattern classification*, p. 571, 2001.
- [18] A. Gordon, "Classification. 1999," *Chapman&Hall, CRC, Boca Raton, FL*, 1999.
- [19] F. Bayat, *et al.*, "A non-parametric heuristic algorithm for convex and non-convex data clustering based on equipotential surfaces," *Expert Systems with Applications*, vol. 37, pp. 3318-3325, 2010.
- [20] E. Rasmussen, "Clustering algorithms," in *Information retrieval*, B. F. William and B.-Y. Ricardo, Eds., ed: Prentice-Hall, Inc., 1992, pp. 419-442.
- [21] E. W. Forgy, "Cluster analysis of multivariate data: efficiency versus interpretability of classifications," *Biometrics*, vol. 21, pp. 768-769, 1965.
- [22] M. R. Anderberg, "Cluster analysis for applications," DTIC Document 1973.
- [23] R. Jancey, "Multidimensional group analysis," *Australian Journal of Botany*, vol. 14, pp. 127-130, 1966.
- [24] M. EmreCelebi, *et al.*, "A comparative study of efficient initialization methods for the k-means clustering algorithm," *Expert Systems with Applications*, 2012.
- [25] M. J. Norusis, *IBM SPSS statistics 19 statistical procedures companion*: Prentice Hall, 2012.
- [26] I. Katsavounidis, *et al.*, "A new initialization technique for generalized Lloyd iteration," *Signal Processing Letters, IEEE*, vol. 1, pp. 144-146, 1994.
- [27] M. d. B. Al-Daoud and S. A. Roberts, "New methods for the initialisation of clusters," *Pattern Recognition Letters*, vol. 17, pp. 451-455, 1996.
- [28] P. S. Bradley and U. M. Fayyad, "Refining Initial Points for K-Means Clustering," in *ICML*, 1998, pp. 91-99.
- [29] C. Pizzuti, *et al.*, "A divisive initialisation method for clustering algorithms," in *Principles of Data Mining and Knowledge Discovery*, ed: Springer, 1999, pp. 484-491.
- [30] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 2007, pp. 1027-1035.
- [31] T. Su and J. G. Dy, "In search of deterministic methods for initializing K-means and Gaussian mixture clustering," *Intelligent Data Analysis*, vol. 11, pp. 319-338, 2007.
- [32] J. Lu, *et al.*, "Hierarchical initialization approach for K-Means clustering," *Pattern Recognition Letters*, vol. 29, pp. 787-795, 2008.
- [33] M. Meilă, "The uniqueness of a good optimum for k-means," in *Proceedings of the 23rd international conference on Machine learning*, 2006, pp. 625-632.
- [34] W.-L. Hung, *et al.*, "Weight selection in k-means algorithm with an application in color image segmentation," *Computers & Mathematics with Applications*, vol. 62, pp. 668-676, 2011.
- [35] A. K. Jain and A. Vailaya, "Image retrieval using color and shape," *Pattern Recognition*, vol. 29, pp. 1233-1244, 1996.
- [36] M. E. Brandt, *et al.*, "Estimation of CSF, white and gray matter volumes in hydrocephalic children using fuzzy clustering of MR images," *Computerized Medical Imaging and Graphics*, vol. 18, pp. 25-34, 1994.
- [37] L. Lucchese and S. Mitra, "Unsupervised segmentation of color images based on k-means clustering in the chromaticity plane," in *Content-Based Access of Image and Video Libraries, 1999.(CBAIVL'99) Proceedings. IEEE Workshop on*, 1999, pp. 74-78.
- [38] S. Ray and R. H. Turi, "Determination of number of clusters in k-means clustering and application in colour image segmentation."
- [39] D. A. Clausi, "K-means Iterative Fisher (KIF) unsupervised clustering algorithm applied to image texture segmentation," *Pattern Recognition*, vol. 35, pp. 1959-1972, 2002.
- [40] K.-S. Chuang, *et al.*, "Fuzzy c-means clustering with spatial information for image segmentation," *Computerized Medical Imaging and Graphics*, vol. 30, pp. 9-15, 2006.
- [41] W. Cai, *et al.*, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Pattern Recognition*, vol. 40, pp. 825-838, 2007.
- [42] Z. Yang, *et al.*, "Robust fuzzy clustering-based image segmentation," *Applied Soft Computing*, vol. 9, pp. 80-84, 2009.
- [43] Z. M. Wang, *et al.*, "Adaptive spatial information-theoretic clustering for image segmentation," *Pattern Recognition*, vol. 42, pp. 2029-2044, 2009.
- [44] Z. Yu, *et al.*, "An adaptive unsupervised approach toward pixel clustering and color image segmentation," *Pattern Recognition*, vol. 43, pp. 1889-1906, 2010.
- [45] S. Das and S. Sil, "Kernel-induced fuzzy clustering of image pixels with an improved differential evolution algorithm," *Information Sciences*, vol. 180, pp. 1237-1256, 2010.
- [46] H. Juang and M.-N. Wu, "MRI brain lesion image detection based on color-converted K-means clustering segmentation," *Measurement*, vol. 43, pp. 941-949, 2010.
- [47] D. Chang, *et al.*, "A genetic clustering algorithm using a message-based similarity measure," *Expert Systems with Applications*, vol. 39, pp. 2194-2202, 2012.
- [48] C. Ju and C. Xu, "A New Collaborative Recommendation Approach Based on Users Clustering Using Artificial Bee Colony Algorithm," *The Scientific World Journal*, vol. 2013, 2013.
- [49] S. Lloyd, "Least squares quantization in PCM," *Information Theory, IEEE Transactions on*, vol. 28, pp. 129-137, 1982.
- [50] Q. Wu, *et al.*, "Follow the Leader": A Centrality Guided Clustering and Its Application to Social Network Analysis," *The Scientific World Journal*, vol. 2013, 2013.
- [51] M. Chau, *et al.*, "Uncertain data mining: a new research direction," in *Proceedings of the Workshop on the Sciences of the Artificial, Hualien, Taiwan*, 2005, pp. 199-204.
- [52] H. Jiawei and M. Kamber, "Data mining: concepts and techniques," *San Francisco, CA, itd: Morgan Kaufmann*, vol. 5, 2001.
- [53] H. Steinhaus, "Sur la division des corp materiels en parties," *Bull. Acad. Polon. Sci*, vol. 1, pp. 801-804, 1956.
- [54] Y. Farhang, "Development of the Meta-Heuristic of PSOGA with K-means Algorithm," *IJCSNS*, 17(6), 29, 2017.
- [55] A. K. Jain and R. C. Dubes, *Algorithms for clustering data*: Prentice-Hall, Inc., 1988.
- [56] Y. Farhang and S. M. H. Shamsuddin, "A Novel Method for Reduction of Error Rates in K-Means Clustering Algorithm," *Life Science Journal*, vol. 11, 2014.
- [57] G. B. Huang, *et al.*, "Labeled Faces in the Wild: A Database for Studying," *month*, 2007.