Computational Visual Media
https://doi.org/10.1007/s41095-019-0151-2 Vol. 5, No. 4, December 2019, 347–361
Research Article

Evaluation of modified adaptive *k*-means segmentation algorithm

Taye Girma Debelee[1,2] (✉), Friedhelm Schwenker[1], Samuel Rahimeto[2], and Dereje Yohannes[2]

Abstract Segmentation is the act of partitioning an image into different regions by creating boundaries between regions. *k*-means image segmentation is the simplest prevalent approach. However, the segmentation quality is contingent on the initial parameters (the cluster centers and their number). In this paper, a convolution-based modified adaptive *k*-means (MAKM) approach is proposed and evaluated using images collected from different sources (MATLAB, Berkeley image database, VOC2012, BGH, MIAS, and MRI). The evaluation shows that the proposed algorithm is superior to *k*-means++, fuzzy *c*-means, histogram-based *k*-means, and subtractive *k*-means algorithms in terms of image segmentation quality (*Q*-value), computational cost, and RMSE. The proposed algorithm was also compared to state-of-the-art learning-based methods in terms of IoU and MIoU; it achieved a higher MIoU value.

Keywords clustering; modified adaptive *k*-means (MAKM); segmentation; *Q*-value

1 Introduction

1.1 Overview

Segmentation is the act of partitioning an image into different regions by creating boundaries that keep regions apart. It is one of the most used steps in zoning pixels of an image [1]. After segmentation, pixels belonging to the same partition have higher similarity values, but higher dissimilarity with pixels in other partitions. Segmentation is a technique used in many fields including health care, image processing, traffic image, pattern recognition, etc. According to the review in Ref. [1], image segmentation techniques can be categorized into two types: layered-based segmentation and block-based segmentation. In layered-based segmentation, the image is divided into layers such as background, foreground, and mask layers. Reconstruction of the final image is decided using the mask layer [2]. This method is not widely applicable to medical image segmentation. Block-based segmentation divides the image into unequal blocks using attributes such as color, histogram, pixels, wavelet coefficients, texture, and gradient [1, 2]. Block-based segmentation can be further grouped into methods based on discontinuity or similarity in the image. It can also be further grouped into three categories: region-based, edge- or boundary-based, and hybrid techniques [1, 2].

1.2 Edge-based segmentation

The discontinuous nature of pixels characterizes all algorithms in the edge-based segmentation family [2]. In this type of image segmentation, images are segmented into partitions based on unanticipated changes in gray intensity in the image. In most cases, edge-based segmentation techniques can identify corners, edges, points, and lines in the image. However, pixel miscategorization errors are the main limitation of the edge-based segmentation category. The edge detection technique is an example of this class of segmentation method [2].

1.3 Region-based segmentation

Edge-based segmentation techniques use the discontinuous nature of pixels. However, region-based techniques use similarity of pixels in the image. Edges, lines, and points are attributes that decide the

1 Institute of Neural Information Processing, Ulm University, 89081Ulm, Germany. E-mail: T. G. Debelee, taye.debelee@uni-ulm.de (✉); F. Schwenker, friedhelm.schwenker@uni-ulm.de.
2 Addis Ababa Science and Technology University, Addis Ababa, 120611, Ethiopia. E-mail: S. Rahimeto, samuelrahimeto@gmail.com; D. Yohannes, derejey@yahoo.com.
Manuscript received: 2019-04-01; accepted: 2019-06-29

effectiveness of region-based techniques. Algorithms like clustering, splitting and merging, normalized cuts, region growing, and thresholding belong to the region-based segmentation family [1, 2]. Our main interest is in clustering algorithms. Schwenker and Trentin [3] presented traditional machine learning as supervised and unsupervised learning: supervised learning associates every observation of the samples with a target label whereas this is not the case in unsupervised learning. Clustering algorithms are very important, especially for unlabeled larger dataset classification [3]; they belong to the unsupervised category. However, there is another machine learning approach, partially supervised machine learning, which lies between unsupervised and supervised machine learning. A detailed review is given in Ref. [3].

1.4 Learning-based segmentation

Deep learning models are well known in object detection, feature extraction, and classification. In addition, semantic image segmentation or image labeling is also an area in which deep learning has been applied. Semantic segmentation is a technique in which semantic labels (like "cat" or "bike") are assigned to every pixel in the image [4]. The most common models that have been applied to semantic image segmentation include FCN-8s [5], DeepLab [4], DeepLab-Msc [4], MSRA-CFM [6], TTI-Zoomout-16 [7], DeepLab-CRF [4], DeepLab-MSc-CRF [4], DeepLab-CRF-7x7 [4], DeepLab-MSc-LargeFOV [4], DeepLab-MSc-CRF-LargeFOV [4], and Front-End Modules [8]. Consecutive application of pooling in deep convolutional neural networks (DCNNs) reduces feature resolution and allows DCNNs to learn abstract representations of objects [9].

2 Related work

2.1 *k*-means segmentation

There has been much research on image segmentation for different application areas, using various techniques from conventional and learning-based methods. Among many segmentation algorithms, k-means is one of the simplest for generating a region of interest [10–12]. It has a time complexity of $O(n)$ for n samples [13]. However, it is sensitive to outliers and initialization parameters [14]. As a result, it gives different clustering results with different cluster numbers and initial centroid values [12]. Much research has considered how to initialize the centers for k-means with the intention of maximizing the efficiency of the algorithm. In the k-means clustering algorithm, each pixel belongs to only one cluster and center, so it is a hard clustering algorithm [10, 11]. Some recent works in clustering methods of segmentation and deep learning based segmentation are addressed in Sections 2.2 and 2.3 respectively.

2.2 Clustering methods for segmentation

In Ref. [15], adaptive k-means clustering is introduced to ameliorate the performance of k-means. Here, the initialization parameters remain consistent for several iterations. However, the initial seed point is computed simply by taking the usual mean of all data values in the input image, making it a simple post-processing operation for good quality image segmentation.

A first attempt to ameliorate the deficiencies of k-means clustering with respect to outliers occurred three decades ago. Bezdek [16] came up with a new algorithm named fuzzy c-means (FCM) in 1981. This algorithm is a membership-based soft clustering algorithm.

Faußer and Schwenker [17] proposed an algorithm that divides the samples into subsets to perform clustering in parallel, and merges the output repeatedly. In their proposed approach they used many kernel-based FCM clustering algorithms. Two datasets (the Breast Cancer database from the UCI repository and Enron Emails) were used to evaluate their algorithm. The experimental analysis proved that the algorithm has high accuracy and works well for large real-life datasets. Benaichouche et al. [18] brought in a region-based image segmentation algorithm using enhanced spatial fuzzy FCM. Lei et al. [19] explained that traditional FCM is susceptible to noise, and describes improvements based on the addition of local spatial information. This solves the robustness problem but greatly increases the computational complexity. First, they used morphological reconstruction to smooth images to enhance robustness and then applied FCM. They also modified FCM by using faster membership filtering instead of the slower distance computation between pixels within local spatial neighborhoods and their cluster centers. The gray-level histogram of the morphologically reconstructed image is used for clustering. The median filter is employed to avoid

noise from the fuzzy membership matrix generated using the histogram. The paper demonstrated that the proposed algorithm is faster and more efficient when compared to FCM and other types of modification.

Arthur and Vassilvitskii [20] introduced a new algorithm called k-means++ that improves upon the initial selection of centroids. The selection for initial clusters is started by selecting one initial center randomly. The other cluster centers are then selected to satisfy specific probabilities determined by "D^2 weighting". The probabilities are defined based on the squared distance of each point to the already chosen centers. The paper claims that k-means++ outperforms the original k-means method in achieving lower intra-cluster separation and in speed. The number of clusters is still chosen by the user. But the algorithm is faster and more effective and even provided as a library in MATLAB.

Zhang et al. [21] used Mahalanobis distance instead of Euclidean distance to allocate every data point to the nearest cluster. Using their new clustering algorithm, PCM clustering, they got better segmentation results. However, their algorithm also has high computational cost and the challenge of initializing parameters.

Purohit and Joshi [22] presented a new approach to improve k-means with aim of reducing the mean square error of the final cluster and attaining minimum computation time. Yedla et al. [23] also introduced an enhanced k-means clustering algorithm with better initial centers. They achieved an effective way to associate data points with appropriate clusters with reduced computation time compared to standard k-means.

Dhanachandra et al. [12] initialized k-means clustering using a subtractive clustering approach which attempted to find optimal centers based on data point density. The first center is chosen to have the highest density value in the data points. After selecting the first center, the potential of the data points near this center decreases. The algorithm then tries to find other centers based on the potential value until the potential of all grid points falls below some threshold. The algorithm is effective at finding centers but the computational complexity increases exponentially as the number of data points increases. The standard k-means algorithm is then initialized with these centers. Since the aim of the paper was the segmentation of medical images, which suffer from poor contrast, a partial spatial starching contrast enhancement technique was applied. After segmentation, filtering is applied to avoid unwanted regions and noise. The paper attempted to illustrate the out-performance of subtractive clustering based k-means over normal k-means. However, it failed to compare it to other methods. Subtractive clustering has a higher computational time than other clustering methods, which is the main drawback of this technique.

Küçükkülahlı et al. [24] tried to initialize k-means by finding both k and centroid locations. First, they used histogram values to determine peaks and pits. Then, by calculating the distances between adjacent peaks and pits, a vertical sweep is done to find the highest peak within some threshold distance. Horizontal sweeping is followed to group peaks that are close to each other, replacing them with a representative by calculating the mean distance and choosing peaks which are above the mean. Once k and centroids have been obtained, the standard k-means method is used for clustering. Even though the approach is dependent on human involvement for assigning the threshold for the vertical sweep, the algorithm automated the k-means method.

2.3 Deep learning in image segmentation

Recently a number of deep learning models have shown astounding results in semantic segmentation [4, 25, 26].

According to Ref. [26], deep learning has shown its success in handwritten digit recognition, speech recognition, image categorization, and object detection in images. It has also been applied to screen content image segmentation, and proven its application to semantic pixel-wise labeling [25]. Badrinarayanan et al. [26] proposed SegNet, a method for semantic pixel-wise segmentation of road scenes, and tested their algorithm using the CamVid road scenes dataset. They used three popular performance evaluation parameters: global accuracy, class average accuracy, and mean intersection over union (MIoU) over all classes.

Minaee and Wang [25] introduced an algorithm for segmentation of screen content images into two

layers (foreground and background). The foreground layer mainly consists of text and lines and the background layer consists of smoothly varying regions. They compared their results with two algorithms (hierarchical k-means clustering in DjVu, and a shape primitive extraction and coding (SPEC) method) in terms of precision and recall values using five test images. The proposed approach scored 91.47% for precision and 87.73% for recall.

Chen et al. [4] proposed a technique that embeds multiscale features in a fully connected convolutional neural network to perform pixel-based semantic segmentation through pixel-level classification. They introduced an attention model to softly determine the weight of multi-scale features at each pixel location. They trained the FCN with multiscale features obtained from multiple resized images using a shared deep network. The attention model played the role of average pooling and max-pooling. Besides feature reduction, the attention model overcame one of the challenges of deep learning: it enabled the authors to visualize the features at different positions along with their level of importance. They proved the effectiveness of their approach using three datasets: PASCAL-Person-Part, VOC 2012, and MS-COCO 2014.

As reviewed by Minaee and Wang [27], various algorithms are in use to separate text from its background. Approaches include clustering-based algorithms, sparse decomposition based methods, and morphological operations. In their paper, they proposed an alternating direction method of Lagrange multipliers (ADMM) for this problem. They adopted the proposed algorithm to separate moving objects from the background. In a comparison made with the hierarchical k-means approach and sparse decomposition, their proposed method scored higher precision (95%), recall (92.5%), and $F1$ (93.7%) values. The sparse decomposition approach was proposed by themselves in Ref. [28].

3 Materials and methods

3.1 Dataset

Images used in this paper are from different sources. Some are from the MATLAB image database and some from the Berkeley segmentation datasets (BSD) [29]. The same images were used in Refs. [12] and [30] to evaluate their segmentation algorithms. The images used for the first experiment are MRI (mri.tif), Bag (bag.png), Cameraman (cameraman.tif), Coins (coins.png), Moon (moon.tif), Pout (pout.tif), and Glass (glass.png). We ran the second experiment using AT (AT3_1m4_01.tif), Lena (lena.png), Valley (valley.jpg), Airplane (airplane.jpg), Mountain (mountain.jpg), and Breast (bet06.jpg) images [31]. Further experimental analysis was done to measure the effectiveness of our proposed segmentation algorithm using the VOC2012 challenge datasets [32].

3.2 Proposed approach

Our proposed segmentation approach is convolution based, as indicated in Fig. 1. First, the histogram

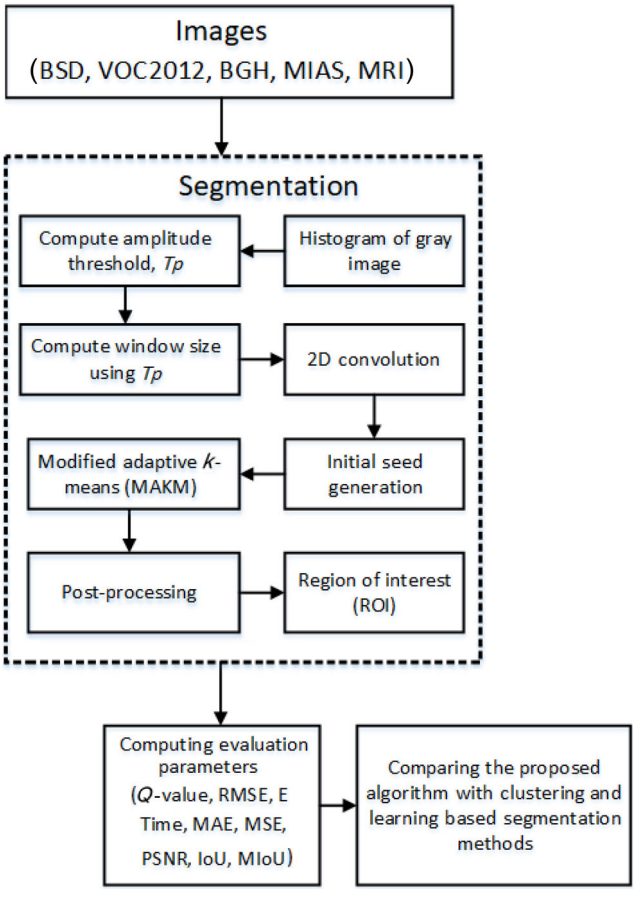


Fig. 1 Flowchart of our convolution-based segmentation algorithm. First, histograms of the grayscale image of the original image are generated. Second, amplitude thresholds, Tp, are computed using the histogram levels. Third, the dynamic window size is computed using an amplitude threshold for each image. This is followed by a 2D convolution operation. Finally, the mean of the convolution is set as the initial seed to generate other new seed values that can be used as the centers of clusters, which are then used to perform clustering.

distribution (a_{li}) of the grayscale image is generated for each image. Second, out of the generated histogram values we select only those with $a_{li} \geqslant 1$. Third, we computed the ratio of the sum of the selected histograms to the number of occurrences (l_0) of such histogram values to obtain the amplitude threshold (Tp). Fourth, we added the histogram values $\geqslant Tp$ and divided by α to get a window size. See Algorithm 1; α is computed as indicated at the end of Algorithm 1. Finally, the convolution operation is performed and the result is converted to an array to compute the mean value used as the initial seed point of the convolution-based modified adaptive k-means (MAKM) segmentation algorithm. The parameters used in the proposed algorithm are constant for each image and the segmentation result is consistent for a number of iterations which is not true for the other clustering algorithms (FCM, HBK, SC, and k-means++) used for comparative analysis. The pseudocode of the proposed algorithm is given in Algorithm 2.

Algorithm 1 Pseudocode for window size generation

```
Get input image and convert to gray image
 Image = readimg()
if (channels(Image)⩾ 3) then
   grayimage = rgb2gray (Image)
else
   grayimage = Image
end if
Calculate the Amplitude Histogram Threshold
 Hist = histogram (grayimage)
for i = 0:255 do
   if (a_li ⩾ 1) then
      l_0 = l_0 + 1
      Sum_{a_li} = Sum_{a_li} + a_li
   end if
end for
 Tp = Sum_{a_li}/l_0
Window size determination, w
for i = 0 : 255 do
   if (a_li ⩾ Tp) then
      Sum_{a_li⩾Tp} = Sum_{a_li⩾Tp} + a_li
   end if
end for
 w = Sum_{a_li⩾Tp}/α
Determine α:
 hist = imhist(img, k),   k = 256
 hist2 = max(hist−Tp + 1, 0)
nonzeros2 = find(hist2)
lt = length(nonzeros2)
α = 1+floor(min(9,lt/2))
```

Algorithm 2 Pseudocode for modified adaptive k-means

```
Input: window size w and gray image
Perform Convolution operation
 gray=conv2(gray,ones(w)/(w^2),'same')
Place the convolution result in an array
 array = gray(:)
Initialize iteration Counters: i = 0,   j = 0
while true do
   Initialize seed point, seed = mean(array)
   Increment counter for each iteration, i = i + 1
   while true do
      Initialize counter for each iteration, j = j + 1
      Compute distance between seed and gray value
          dist = sqrt((array−seed)^2)
      Compute bandwidth for cluster center
          distth= sqrt(sum((array−seed)^2)/numel(array))
      Check values are in selected bandwidth or not
                qualified = dist<distth
      Update mean
                newseed = mean(array(qualified))
          condition for termination
      if (seed = newseed or j>10) then
         j = 0
         Remove values assigned to a cluster
         Store center of cluster
                   center(i) = newseed
         break
      end if
      Update seed: seed = newseed
          check maximum number of clusters
      if (isempty(array) or i > 10) then
         Reset counter: i = 0
         break
      end if
   end while
   Sort centers
   Compute distances between adjacent centers
   Find minimum distance between centers
   Discard cluster centers less than distance
   Make a clustered image using these centers
end while
```

3.3 Evaluation

In this paper we use the Q-value criterion [30], computational cost [30], root mean squared error (RMSE) [33, 34], standard deviation, mean absolute error (MAE), intersection over union (IoU) [35], mean intersection over union (MIoU) [35], entropy (E), and peak signal to noise ratio (PSNR) to assess our proposed, convolution-based, modified adaptive k-means segmentation algorithm.

The Q-value measures image segmentation quality taking into consideration both small and large regions

in the final segmented images. The Q value evaluation function used in this paper is given by

$$Q = \frac{\sqrt{R}}{(NM)} \sum_{i=1}^{R} \left[\frac{e_i^2}{1 + \log A_i} + \left(\frac{R(A_i)}{A_i} \right)^2 \right] \quad (1)$$

where N and M are the numbers of rows and columns in the image respectively, R is the total number of regions in the segmented image, e_i is the color difference between the original and segmented image, A_i is the area of region i, and $R(A_i)$ is the number of regions with the same area as A_i. The area of each region A_i is the number of pixels constituting that region; empty regions and regions with 1 pixel are left unconsidered. Smaller values of Q represent better segmentation results whereas higher values indicate higher color errors due to either under-segmentation or over-segmentation. e_i is given by

$$e_i = \sqrt{((S(x_{si}, y_{si}) - O(x_{oi}, y_{oi}))^2}, \quad i = 1 : R \quad (2)$$

where S and O are points in 2D Euclidean space with coordinates $S(x_{si}, y_{si})$ for the segmented image and $O(x_{oi}, y_{oi})$ for the original (raw) image.

RMSE measures how much the output image deviates from the input image. Mean squared error (MSE) is given by

$$\text{MSE} = \frac{1}{n_x m_y} \sum_{x=0}^{n_x - 1} \sum_{y=0}^{m_y - 1} [r(x, y) - t(x, y)]^2 \quad (3)$$

where $r(x, y)$ and $t(x, y)$ are grayscale values at position x, y in the raw and segmented images. RMSE is the square root of the MSE [33, 34].

$$\text{RMSE} = \sqrt{\text{MSE}} \quad (4)$$

A smaller value means higher image segmentation quality.

For the VOC2012 challenge datasets, popular performance evaluation parameters include global accuracy, class average accuracy, and mean intersection over union (MIoU) for all classes [26]. Global accuracy is the percentage of pixels correctly classified in the dataset whereas the mean of the predictive accuracy over all classes is class average accuracy. In this paper, we use MIoU to compare the performance of our algorithm with learning-based segmentation methods. MIoU is defined as

$$\text{MIoU} = \sum \text{IoU}/N \quad (5)$$

where intersection over union (IoU) is defined in Eq. (6) and N is the number of objects considered from the dataset for a particular experiment.

$$\text{IoU} = \frac{\text{Area(overlap)}}{\text{Area(union)}} \quad (6)$$

Algorithm 3 Pseudocode for mean Q values and standard deviation

```
Initialize Q = 0
Perform 10 iterations
for j = 1:10 do
    Identify the unique label from the indexed image
    Calculate the number of unique labels and initialize to number of regions (R)
    Convert the indexed images to gray-scale using mat2gray
    Compute color error (e_i) and area (A_i) for each region
    for i = 1:R do
        Find number of regions with the same area as A_i and initialize it to R(A_i)
        Compute Q_i for each region
        Q = Q + Q_i
    end for
end for
Compute mean of Q and standard deviation (σ)
```

Algorithm 4 Pseudocode for computation of RMSE

```
Initialize Squared Error, E = 0
Perform 10 iterations
for j = 1:10 do
    Get labeled images with their respective centers
    for i = 1: R do
        Compute Squared Error (SE) for each region
    end for
    for i = 1:R do
        Compute sum of E,
        E = E + E_i
    end for
    Compute RMSE for each iteration
end for
Compute mean RMSE and standard deviation (σ)
```

where area of overlap is the area between the predicted bounding box and the ground-truth bounding box, and area of union is the area covered by both the predicted bounding box and the ground-truth bounding box.

4 Results and discussion

The outcomes of the experiments we conducted using the proposed technique show that gray image segmentation task can be carried out efficiently while initialization of parameters is done automatically. All experiments were performed in MATLAB version 2016b and run on a 3.00 GHz Intel Core i7-4601M CPU, under the Microsoft Windows 10 operating system. The performance of the proposed segmentation algorithm was evaluated using RMSE,

Q-value, computation time, MAE, E, PSNR, IoU, and MIoU. RMSE and MAE were used for standard quality measurement of the segmented output image. It tells us the degree of deviation between the output image and the input image. The same evaluation parameters were used for other selected clustering segmentation algorithms for comparative analysis, but for the deep learning based segmentation algorithm, only IoU and MIoU were used for comparison with the proposed segmentation algorithm. Some performance aspects of the proposed method are discussed in this section.

To evaluate the proposed image segmentation approach, we used images that were also used in Refs. [12] and [30]. In the first experiment, we used MRI (mri.tif), Bag (bag.png), Cameraman (cameraman.tif), Coins (coins.png), Moon (moon.tif), Pout (pout.tif), and Glass (glass.png). The results obtained are indicated in Figs. 2–6 and Tables 2–4.

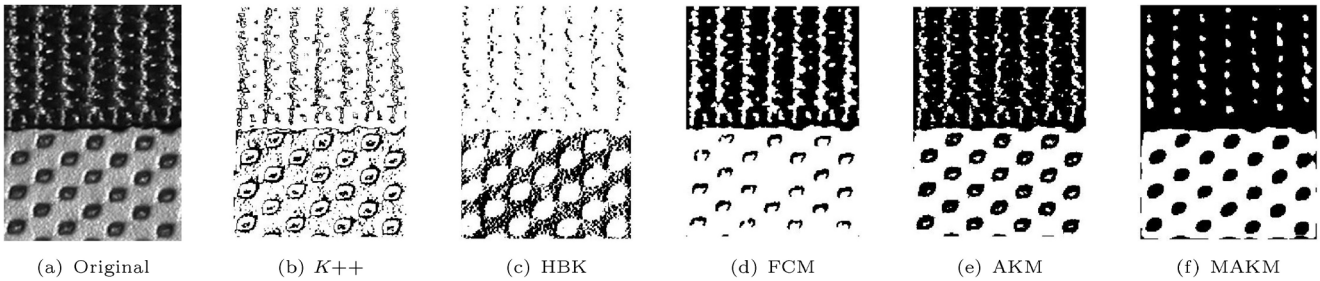
In Table 1, we list images with their respective sizes and number of cluster for every clustering algorithm considered in this paper. In Table 2, we compare other clustering algorithms with our proposed algorithm in terms of segmentation quality. Since the cluster centers varying for K++, FCM, and HBK, the Q-value in Table 2, RMSE in Table 3, and computation cost in Table 4 are computed 10 times and their mean value and standard deviation are determined. However, in the proposed modified adaptive k-means clustering method, the cluster centers are consistent for any number of iterations. Compared to other clustering algorithm, histogram-based k-means (HBK) had the lowest segmentation quality for moon.tif, as indicated in Table 2. Comparing K++ and FCM with HBK shows that K++ and FCM had lower Q score. However, adaptive k-means and modified adaptive k-means methods outperform these three algorithms, even if the adaptive k-means algorithm needs post-processing image to find the region of interest. In some cases, MAKM performs better than AKM, for example for images glass.png, pout.tif, and mri.tif.

In terms of RMSE, FCM had highest score for pout.tif which shows that it is the worst performing clustering algorithm: see Table 3. However, our proposed approach had the minimum RMSE value for all images used in the experiment.

To evaluate the attainment of the proposed technique in image segmentation for other gray images, further experiments were conducted on some commonly used images: AT3_1m4_01.tif, lena.png, valley.jpg, airplane.jpg, mountain.jpg, and bet06.jpg. The subtractive k-means algorithm has the highest computation cost for all images, although in some cases it shows better segmentation quality than K++, HBK, and FCM.

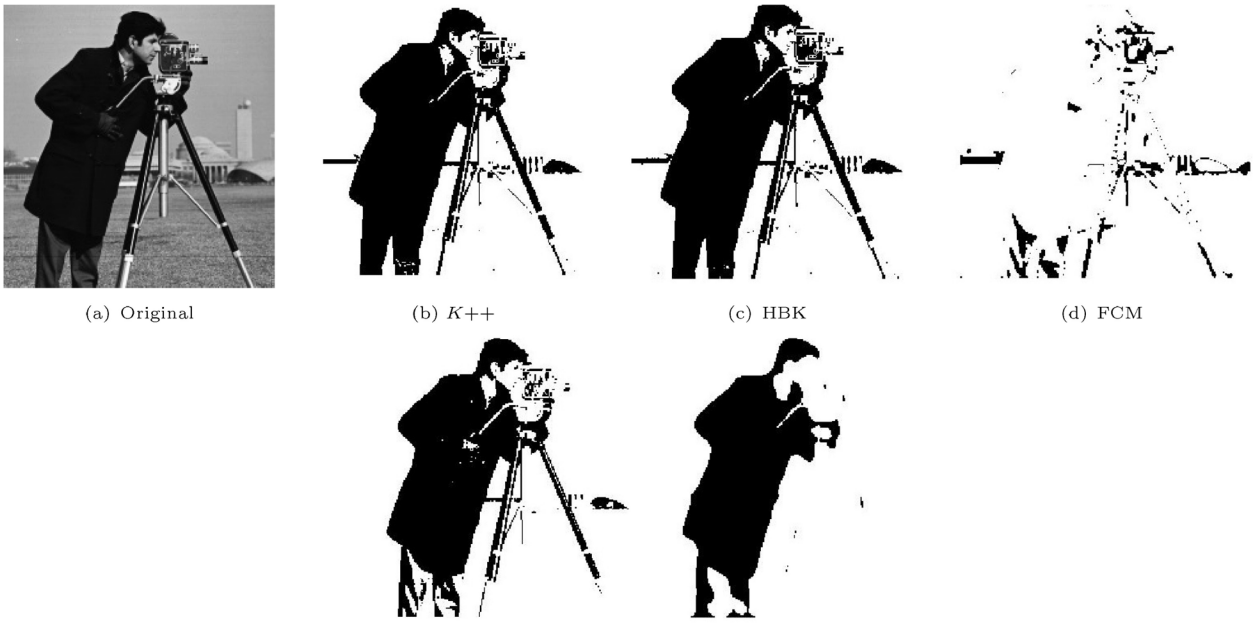
(a) Original (b) K++ (c) HBK (d) FCM (e) AKM (f) MAKM

Fig. 2 MRI-labeled image segmented using various approaches.



(a) Original (b) K++ (c) HBK (d) FCM (e) AKM (f) MAKM

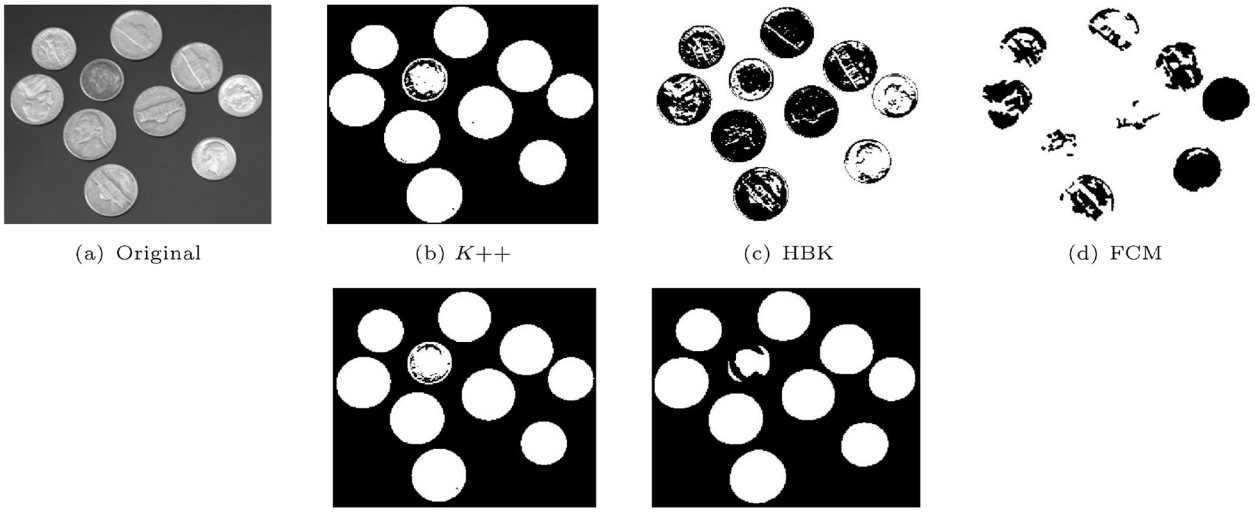
Fig. 3 Bag-labeled image segmented using various approaches.



(a) Original (b) K++ (c) HBK (d) FCM

(e) AKM (f) MAKM

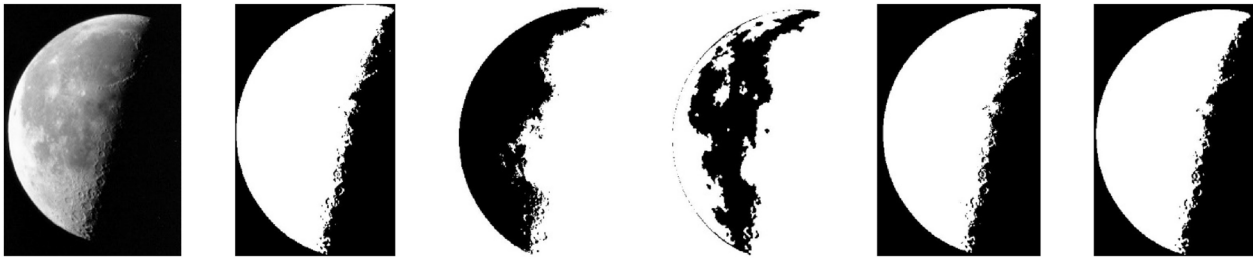
Fig. 4 Cameraman-labeled image segmented using various approaches.



(a) Original (b) K++ (c) HBK (d) FCM

(e) AKM (f) MAKM

Fig. 5 Coins-labeled image segmented using various approaches.



(a) Original (b) K++ (c) HBK (d) FCM (e) AKM (f) MAKM

Fig. 6 Moon-labeled image segmented using various approaches.

Table 1 Number of clusters used in each algorithm for each respective images

	Image	Size	K++	HBK	FCM	SC	AKM	MAKM (proposed)
Experiment-1	MRI	128×128	3	3	4	2	3	3
	Bag	250×189	3	3	4	4	5	4
	Cameraman	256×256	3	3	4	3	4	4
	Coins	246×300	3	3	4	2	4	3
	Moon	537×358	3	3	4	2	4	4
	Pout	291×240	3	3	4	2	3	4
	Glass	181×282	3	3	4	5	3	3
Experiment-2	AT3_1m4_01 (AT)	480×640	3	3	4	1	5	5
	Lena	512×512	3	3	4	4	3	3
	Valley	321×481	3	3	4	5	4	4
	Airplane	321×481	3	3	4	2	4	4
	Mountain	321×481	3	3	4	2	4	3
	Breast	384×512	3	3	4	2	6	5

Table 2 Comparison of algorithms in terms of mean Q-value and standard deviation (Q-value, σ)

Image	Size	K++	HBK	FCM	SC	AKM	MAKM (proposed)
MRI	128×128	0.40, 0.18	1.06, 0.60	0.87, 0.64	0.31, —	0.21, 0.00	0.20, 0.00
Bag	250×189	0.44, 0.35	0.50, 0.43	0.39, 0.32	0.13, —	0.07, 0.00	0.12, 0.00
Cameraman	256×256	0.39, 0.21	0.44, 0.28	0.48, 0.25	0.70, —	0.02, 0.00	0.05, 0.00
Coins	246×300	0.26, 0.23	0.54, 0.37	0.42, 0.18	0.13, —	0.13, 0.00	0.13, 0.00
Moon	537×358	0.25, 0.36	0.87, 0.57	0.60, 0.51	0.06, —	0.01, 0.00	0.01, 0.00
Pout	291×240	0.28, 0.04	0.35, 0.14	0.40, 0.13	0.29, —	0.20, 0.00	0.03, 0.00
Glass	181×282	0.55, 0.24	0.46, 0.41	0.44, 0.29	0.62, —	0.15, 0.00	0.14, 0.00

Table 3 Comparison of algorithms in terms of mean RMSE and standard deviation (RMSE, σ)

Image	Size	K++	HBK	FCM	SC	AKM	MAKM (proposed)
MRI	128×128	1.14, 0.00	1.14, 0.004	0.88, 0.006	0.91, —	1.03, 0.00	1.03, 0.00
Bag	250×189	1.27, 0.004	1.27, 0.01	1.13, 0.001	1.36, —	1.29, 0.00	0.90, 0.00
Cameraman	256×256	1.63, 0.001	1.62, 0.005	1.44, 0.01	1.63, —	1.47, 0.00	1.10, 0.00
Coins	246×300	1.54, 0.002	1.54, 0.01	1.47, 0.004	1.50, —	1.61, 0.00	1.31, 0.00
Moon	537×358	0.86, 0.001	0.86, 0.004	0.91, 0.001	0.76, —	0.88, 0.00	0.86, 0.00
Pout	291×240	1.81, 0.04	1.79, 0.02	1.91, 0.005	1.77, —	1.50, 0.00	1.48, 0.00
Glass	181×282	1.51, 0.00	1.51, 0.001	1.62, 0.005	1.77, —	1.44, 0.00	1.37, 0.00

Table 4 Comparison of algorithms in terms of mean computation time and standard deviation (time (s), σ)

Image	Size	K++	HBK	FCM	SC	AKM	MAKM (proposed)
MRI	128×128	0.01, 0.005	0.02, 0.006	0.02, 0.009	5.00, 0.50	0.003, 0.001	0.004, 0.001
Bag	250×189	0.08, 0.07	0.09, 0.03	0.06, 0.02	33.70, 1.02	0.012, 0.002	0.018, 0.011
Cameraman	256×256	0.04, 0.02	0.09, 0.005	0.08, 0.01	62.77, 1.25	0.0116, 0.002	0.016, 0.004
Coins	246×300	0.09, 0.07	0.11, 0.005	0.08, 0.01	75.90, 1.16	0.015, 0.004	0.022, 0.024
Moon	537×358	0.20, 0.11	0.26, 0.01	0.20, 0.007	1025.01, 5.69	0.041, 0.003	0.042, 0.005
Pout	291×240	0.02, 0.008	0.09, 0.005	0.06, 0.01	74.03, 0.75	0.015, 0.004	0.015, 0.004
Glass	181×282	0.02, 0.006	0.07, 0.004	0.05, 0.02	39.44, 0.05	0.012, 0.008	0.011, 0.004

Table 5 Comparison of proposed algorithm with K++, HBK, FCM, and SC in terms of mean Q-value for AT, LE, VA, AI, MT, and Breast images

Image	Size	K++	HBK	FCM	SC	MAKM (proposed)
AT3_1m4_01 (AT)	480×640	0.53	0.36	0.49	0.50	0.04
Lena (LE)	512×512	0.42	0.35	0.48	0.39	0.06
Valley (VA)	321×481	0.45	0.45	0.55	0.45	0.15
Airplane (AI)	321×481	0.40	0.29	0.36	0.41	0.10
Mountain (MT)	321×481	0.34	0.50	0.53	0.64	0.19
Breast	384×512	0.46	0.39	0.73	0.72	0.01

Table 6 Comparison of proposed algorithm with K++, HBK, FCM, and SC in terms of mean RMSE for AT, LE, VA, AI, MT, and Breast images

Image	Size	K++	HBK	FCM	SC	MAKM (proposed)
AT3_1m4_01 (AT)	480×640	0.83	0.83	0.82	0.74	0.69
Lena (LE)	512×512	1.00	1.00	1.10	1.12	0.84
Valley (VA)	321×481	1.23	1.23	1.20	1.42	1.06
Airplane (AI)	321×481	1.45	1.45	1.51	1.28	1.22
Mountain (MT)	321×481	0.95	0.95	0.88	0.78	0.72
Breast	384×512	0.56	0.58	0.59	0.50	0.56

Table 7 Comparison of proposed algorithm with K++, HBK, FCM, and SC in terms of mean computation cost (s) for AT, LE, VA, AI, MT, and Breast images

Image	Size	K++	HBK	FCM	SC	MAKM (proposed)
AT3_1m4_01 (AT)	480×640	0.34	0.42	0.33	2846.45	0.14
Lena (LE)	512×512	0.21	0.34	0.27	2074.85	0.12
Valley (VA)	321×481	0.20	0.27	0.18	693.66	0.06
Airplane (AI)	321×481	0.12	0.23	0.14	648.50	0.05
Mountain (MT)	321×481	0.09	0.24	0.17	658.37	0.07
Breast	384×512	0.10	0.20	0.17	1092.93	0.05

From the data analysis, we observe that the time taken by subtractive k-means becomes very expensive for some images. The three image samples with highest time cost are AT3_1m4_01.tif (2846 s), lena.png (2075 s), and moon.tif (1565 s).

In the experimental results presented in Table 2, K++ shows better performance than FCM and HBK, except for some image samples: glass.png for both HBK and FCM, and bag.png for FCM. However, in the second experiment using images like AT3_1m4_01.tif, lena.png, valley.jpg, airplane.jpg, mountain.jpg, and bet06.jpg, HBK proved to provide the best image segmentation quality except for mountain.jpg. The modified adaptive k-means algorithm has better image segmentation quality, and minimum RMSE for all cases discussed. It scored well for the breast image compared to other images. The low computation cost of our proposed approach makes it more suitable for image segmentation. The sample indexed images in the second experiment are given in Fig. 7. The final results of the experiment show that the overall achievement of the proposed modified adaptive k-means is superior to other clustering algorithms in terms of image segmentation quality (Q-value), computational cost, and RMSE.

For further analysis, we considered additional images from the VOC2012 challenge dataset and mammography images from Bethezatha General Hospital (BGH) and MIAS.

Four randomly selected images (dog, airplane, plant, and person) from VOC2012 were used to compare the proposed algorithm to three clustering algorithms (AKM, FCM, K++) in terms of Q, time, MAE, E, and PSNR. For all images our proposed algorithm scored better for Q, computation time, and PSNR compared to other clustering algorithms, but not for MAE and entropy: see Table 8. In the case of the "person" image, our proposed algorithm scored minimum MAE compared to other algorithms, indicating good performance for this particular image.

Comparative performance of the proposed algorithm for two randomly selected MRI images is

(a) Original (b) K++ (c) HBK (d) FCM (e) AKM (f) MAKM

Fig. 7 Lena-labeled image segmented using various approaches.

Table 8 Comparison of proposed method with clustering algorithms in terms of Q, computation time, MAE, entropy, PSNR, precision (P), recall (R), and F-score ($F1$) using VOC2012 dataset

Image	Algorithm	Q	Time (s)	MAE	Entropy	PSNR	P	R	$F1$
Dog	AKM	0.042	0.106	6.07	1.53	47.25	89.3	61.7	72.9
	FCM	0.431	0.345	3.22	1.83	46.98	86.9	92.2	89.5
	k-means++	0.209	0.248	2.76	1.94	46.64	84.8	99.5	91.6
	Proposed	0.040	0.052	3.89	1.79	47.28	90.3	80.4	85.1
Plane	AKM	0.031	0.041	0.407	1.54	43.55	95.0	89.6	92.2
	FCM	0.743	0.212	0.302	1.71	43.11	91.2	98.7	94.8
	k-means++	1.00	0.120	0.252	1.63	43.16	94.2	96.6	95.4
	Proposed	0.030	0.036	0.337	1.58	43.59	95.3	92.4	93.8
Plant	AKM	0.012	0.054	0.331	1.63	45.78	71.1	81.7	76.0
	FCM	0.041	0.286	40.60	1.64	45.73	0.0	1.3	0.0
	k-means++	0.621	0.107	43.30	1.64	45.65	0.0	0.0	0.0
	Proposed	0.011	0.061	0.332	1.63	45.84	72.1	82.8	77.2
Person	AKM	0.095	0.11	12.72	1.92	48.67	83.6	88.8	86.1
	FCM	0.139	0.58	16.30	1.62	47.72	92.2	68.1	78.3
	k-means++	0.35	0.65	14.09	1.72	47.34	92.9	72.8	81.6
	Proposed	0.094	0.050	5.11	2.06	48.71	84.9	99.9	91.8

given in Table 9. The proposed algorithm performs better in terms of MSE, Q, and computation time for both MRI images. However, the second MRI recorded a higher IoU value than the first image, as indicated in Table 10. Segmentation results for the second MRI image are given in Fig. 11.

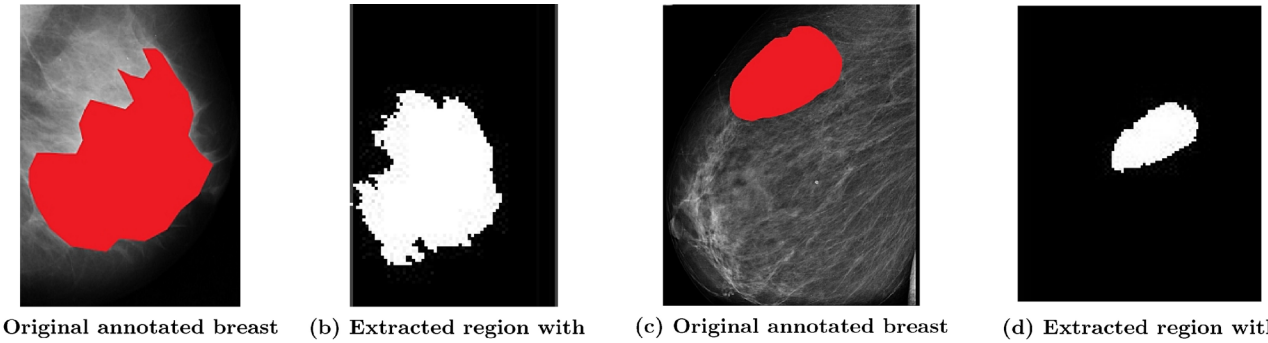
A comparison of the proposed algorithm with learning-based and clustering algorithms is presented in Table 11. The comparison terms of IoU and MIoU indicate that the proposed algorithm scored higher IoU and MIoU for plant and person images, but for the dog and plane images, k-means++ scored

Table 9 Comparison of proposed algorithm with clustering image segmentation algorithm in terms of MSE, Time, and Q for two MRI images

Image	Method	MSE	Time	Q
MRI_1	AKM	0.67	0.03	0.009
	FCM	0.74	0.16	0.238
	k-means++	0.77	0.37	0.053
	Proposed	**0.66**	**0.02**	**0.008**
MRI_2	AKM	0.74	0.03	0.020
	FCM	0.85	0.16	1.585
	k-means++	0.91	0.30	1.492
	Proposed	**0.73**	**0.02**	**0.019**

Table 10 Comparison of proposed algorithm with clustering algorithm in terms of IoU and MIoU for two MRI images

Method	MRI_1	MRI_2	MIoU
AKM	63.19	58.19	60.69
FCM	**79.21**	62.93	71.07
k-means++	77.90	83.96	**80.93**
Proposed	64.22	**86.33**	75.28



(a) Original annotated breast from MIAS dataset **(b) Extracted region with cancer from (a)** **(c) Original annotated breast from BGH dataset** **(d) Extracted region with cancer from (c)**

Fig. 8 Examples of annotated and extracted region with cancer for breast mammographic images from BGH and MIAS datasets using proposed method.

(a) Original annotated dog (b) Segmentation result

Fig. 9 Annotated and respective segmentation result for dog from VOC2012 challenge datasets using proposed method.

higher values of IoU and MIoU. Figures 8–10 present segmentation results for the proposed algorithm for images randomly selected from VOC2012, BGH, and MIAS datasets.

5 Conclusions

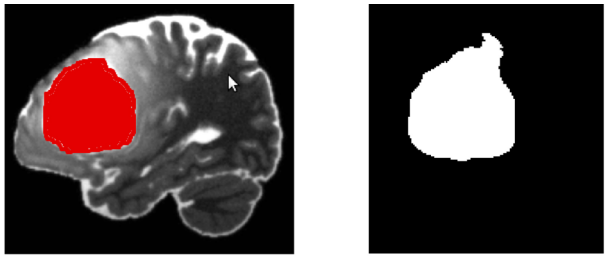
In this study, we presented a convolution-based modified adaptive k-means algorithm, to get the best out of the normal k-means method during image segmentation. Firstly, an automatic window size generation approach was designed to perform the convolution process to get the central value for every convolution step, and the mean of these values is assigned as the initial seed point. Then, using this seed point, the cluster centers and number of clusters are determined as initial parameters and given to the adaptive k-means algorithm. A comparative analysis of the proposed modified adaptive k-means with K++, HBK, and SC methods was made in terms of image segmentation quality (Q), RMSE, and time. The results obtained confirmed the advantages of our proposed modified adaptive k-means algorithm.

(a) Original annotated plane (b) Segmentation result (c) Original annotated person (d) Segmentation result

Fig. 10 Annotated and respective segmentation result for plane and person from VOC2012 challenge datasets using proposed method.

Table 11 Related works from learning-based methods and clustering algorithms for comparison with proposed method in terms of IoU and MIoU for selected images from VOC2012 dataset

Author(s), year, and citation	Method	Dog	Plane	Plant	Person	MIoU
Long et al., 2015 [5]	FCN-8s	71.8	76.8	45.2	73.9	66.92
Chen et al., 2016 [4]	DeepLab	68.7	72	50.8	73.6	66.28
Chen et al., 2016 [4]	DeepLab-Msc	68.4	74.9	51.7	75.0	67.5
Dai et al., 2015 [6]	MSRA-CFM	69.1	75.7	50.4	67.5	65.68
Mostajabi et al., 2015 [7]	TTI-Zoomout-16	74.0	81.9	68.8	44.3	67.25
Chen et al., 2016 [4]	DeepLab-CRF	75.2	78.4	54.7	77.6	71.48
Chen et al., 2016 [4]	DeepLab-MSc-CRF	74.3	80.4	56.9	79.0	72.65
Chen et al., 2016 [4]	DeepLab-CRF-7x7	78.9	83.9	60.3	80.6	75.93
Chen et al., 2016 [4]	DeepLab-MSc-LargeFOV	78.5	83.5	58.2	79.7	74.98
Chen et al., 2016 [4]	DeepLab-MSc-CRF-LargeFOV	79.0	84.4	59.7	80.8	75.98
Liu et al., 2015 [36]	ParseNet Baseline	73.1	82.6	51.9	78.6	71.58
Liu et al., 2015 [36]	ParseNet	77.1	84.1	52.6	78.2	73.0
Reproduced by Ref. [36]	Hypercolumn	72.1	68.7	52.6	72.9	66.58
Yu and Koltun, 2015 [8]	Front-End Module	73.1	82.2	56.6	79.1	72.75
Reproduced	AKM	57.4	85.6	61.30	75.6	69.98
Reproduced	FCM	81.0	90.1	0.0	64.3	58.85
Reproduced	k-means++	**84.5**	**91.2**	0.0	68.9	61.15
Proposed method	MAKM	74.0	88.16	**61.32**	**85.0**	**77.12**



(a) Original annotated MRI (b) Extracted region with tumor from (a)

Fig. 11 Examples of annotated and extracted region with tumor for MRI image using proposed method.

Furthermore, an objective comparison of the proposed modified adaptive k-means algorithm with another soft clustering algorithm, FCM, also proved the advantages of our proposed technique.

To evaluate the robustness of our algorithm we ran additional experiments using the VOC2012 challenge dataset and MRI images, comparing the proposed segmentation algorithm with learning-based methods in terms of IoU and MIoU. They found that our algorithm outperforms learning-based methods for the VOC2012 challenge dataset.

In work, we hope to apply our method to breast cancer image analysis. After segmentation, texture features (quantized compound change histogram, Haralick descriptors, edge histogram MPEG-7, Gabor features, gray-level c-occurrence matrix, and local binary patterns) and shape features (centroid distance function signature, chain code histogram, Fourier descriptors, and pyramid histogram of oriented gradients) can be extracted and used as input to various classifiers to distinguish between normal and abnormal mammograms.

Acknowledgements

The corresponding author would like to thank the Ethiopian Ministry of Education (MoE) and the Deutscher Akademischer Auslandsdienst (DAAD) for funding this research work (funding number 57162925).

References

[1] Zaitoun, N. M.; Aqel, M. J. Survey on image segmentation techniques. *Procedia Computer Science* Vol. 65, 797–806, 2015.

[2] Jaglan, P.; Dass, R.; Duhan, M. A comparative analysis of various image segmentation techniques. In: *Proceedings of the 2nd International Conference on Communication, Computing and Networking. Lecture Notes in Networks and Systems, Vol. 46.* Krishna, C.; Dutta, M.; Kumar, R. Eds. Springer Singapore, 359–374, 2019.

[3] Schwenker, F.; Trentin, E. Pattern classification and clustering: A review of partially supervised learning approaches. *Pattern Recognition Letters* Vol. 37, 4–14, 2014.

[4] Chen, L.-C.; Yang, Y.; Wang, J.; Xu, W.; Yuille, A. L. Attention to scale: Scale-aware semantic image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3640–3649, 2016.

[5] Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3431–3440, 2015.

[6] Dai, J.; He, K.; Sun, J. Convolutional feature masking for joint object and stuff segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3992–4000, 2015.

[7] Mostajabi, M.; Yadollahpour, P.; Shakhnarovich, G. Feedforward semantic segmentation with zoom-out features. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3376–3385, 2015.

[8] Yu, F.; Koltun, V. Multi-scale context aggregation by dilated convolutions. *arXiv preprint* arXiv:1511.07122, 2015.

[9] Chen, L. C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint* arXiv:1706.05587, 2017.

[10] Shi, P. F.; Fan, X. N.; Ni, J. J.; Wang, G. R. A detection and classification approach for underwater Dam cracks. *Structural Health Monitoring: An International Journal* Vol. 15, No. 5, 541–554, 2016.

[11] Khanmohammadi, S.; Adibeig, N.; Shanehbandy, S. An improved overlapping k-means clustering method for medical applications. *Expert Systems with Applications* Vol. 67, 12–18, 2017.

[12] Dhanachandra, N.; Manglem, K.; Chanu, Y. J. Image segmentation using k-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science* Vol. 54, 764–771, 2015.

[13] Faußer, S.; Schwenker, F. Clustering large datasets with kernel methods. In: Proceedings of the 21st International Conference on Pattern Recognition, 501–504, 2012.

[14] Razavi Zadegan, S. M.; Mirzaie, M.; Sadoughi, F. Ranked k-medoids: A fast and accurate rank-based partitioning algorithm for clustering large datasets. *Knowledge-Based Systems* Vol. 39, 133–143, 2013.

[15] Dixit, A. Adaptive kmeans clustering for color and gray image. 2014. Available at https://www.mathworks.com/matlabcentral/fileexchange/45057-adaptive-kmeans-clustering-for-color-and-gray-image.

[16] Bezdek, J. C. Modified objective function algorithms. In: *Pattern Recognition with Fuzzy Objective Function Algorithms. Advanced Applications in Pattern Recognition*. Springer Boston MA, 155–201, 1981.

[17] Faußer, S.; Schwenker, F. Parallelized kernel patch clustering. In: *Artificial Neural Networks in Pattern Recognition. Lecture Notes in Computer Science, Vol. 5998*. Schwenker, F.; El Gayar, N. Eds. Springer Berlin Heidelberg, 131–140, 2010.

[18] Benaichouche, A. N.; Oulhadj, H.; Siarry, P. Improved spatial fuzzy c-means clustering for image segmentation using PSO initialization, Mahalanobis distance and post-segmentation correction. *Digital Signal Processing* Vol. 23, No. 5, 1390–1400, 2013.

[19] Lei, T.; Jia, X. H.; Zhang, Y. N.; He, L. F.; Meng, H. Y.; Nandi, A. K. Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering. *IEEE Transactions on Fuzzy Systems* Vol. 26, No. 5, 3027–3041, 2018.

[20] Arthur, D.; Vassilvitskii, S. k-means++: The advantages of careful seeding. In: Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms, 1027–1035, 2007.

[21] Zhang, Y.; Huang, D.; Ji, M.; Xie, F. D. Image segmentation using PSO and PCM with Mahalanobis distance. *Expert Systems with Applications* Vol. 38, No. 7, 9036–9040, 2011.

[22] Purohit, P.; Joshi, R. A new efficient approach towards k-means clustering algorithm. *International Journal of Computer Applications* Vol. 65, No. 11, 7–10 2013.

[23] Yedla, M.; Pathakota, S. R.; Srinivasa, T. M. Enhanced k-means clustering algorithm with improved initial center. *International Journal of Science and Information Technologies* Vol. 1, No. 2, 121–125, 2010.

[24] Küçükkülahlı, E.; Erdoğmuş P.; Polat, K. Histogram-based automatic segmentation of images. *Neural Computing and Applications* Vol. 27, No. 5, 1445–1450, 2016.

[25] Minaee, S.; Wang, Y. Screen content image segmentation using least absolute deviation fitting. In: Proceedings of the IEEE International Conference on Image Processing, 3295–3299, 2015.

[26] Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A deep convolutional encoder–decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 39, No. 12, 2481–2495, 2017.

[27] Minaee, S.; Wang, Y. An ADMM approach to masked signal decomposition using subspace representation. *IEEE Transactions on Image Processing* Vol. 28, No. 7, 3192–3204, 2019.

[28] Minaee, S.; Wang, Y. Screen content image segmentation using robust regression and sparse decomposition. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* Vol. 6, No. 4, 573–584, 2016.

[29] Martin, D.; Fowlkes, C.; Tal, D.; Malik, J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of the 8th IEEE International Conference on Computer Vision, 416–423, 2001.

[30] Khan, Z.; Ni, J.; Fan, X.; Shi, P. 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, 1509–1525, 2017.

[31] Suckling, J. P. The mammographic image analysis society digital mammogram database exerpta medica. *Digital Mammo* 375–386, 1994.

[32] Everingham, M.; Van Gool, L.; Williams, C. K. I.; Winn, J.; Zisserman, A. The Pascal visual object classes (VOC) challenge. *International Journal of Computer Vision* Vol. 88, No. 2, 303–338, 2010.

[33] Saffor, R.; Ramli, A. R.; Ng, K-H. A comparative study of image compression between JPEG and wavelet. *Malaysian Journal of Computer Science* Vol. 14, No. 1, 39–45, 2001.

[34] Gonzalez, R. C.; Woods, R. E. *Digital Image Processing* New York: Addison-Wesley, 1992.

[35] Kamran, S. A.; Sabbir, A. S. Efficient yet deep convolutional neural networks for semantic segmentation. In: Proceedings of the International Symposium on Advanced Intelligent Informatics, 123–130, 2018.

[36] Liu, W.; Rabinovich, A.; Berg, A. C. Parsenet: Looking wider to see better. *arXiv preprint* arXiv:1506.04579, 2015.

Taye Girma Debelee is a sandwich program Ph.D. student in Ulm University and Addis Ababa Science and Technology University. He holds an M.Sc. degree in computer engineering from Addis Ababa University, Ethiopia. His research interests are in digital imaging processing, pattern recognition, data mining and deep learning.

Friedhelm Schwenker is a senior lecturer and researcher at the Institute of Neural Information Processing, Ulm University, Germany. His research interests are in artificial neural networks, pattern recognition, data mining, and affective computing.

Samuel Rahimeto is an M.Sc. student at Addis Ababa Science and Technology University, Ethiopia. His research interests are in digital image processing and machine learning.

Dereje Yohannes is a director of the Artificial Intelligence Excellence Center. He holds M.Sc. and Ph.D. degrees in computer engineering. His main research interests are in network security and wireless. He is also working in the area of data mining and machine learning.