Arabian Journal of Geosciences (2021) 14:777
https://doi.org/10.1007/s12517-021-07069-4

ORIGINAL PAPER

Color-based SAR image segmentation using HSV+FKM clustering for estimating the deforestation rate of LBA-ECO LC-14 modeled deforestation scenarios, Amazon basin: 2002–2050

Kalaiyarasi Murugesan[1] · Perumal Balasubramani[2] · Pallikonda Rajasekaran Murugan[2] · Saravanan Sankaranarayanan[3]

Received: 8 May 2018 / Accepted: 1 April 2021
© Saudi Society for Geosciences 2021

Abstract

Image segmentation is an essential process for image evaluation tactics. Synthetic aperture radar (SAR) image segmentation can be implemented for a vast magnificence of various problems. Due to the existence of speckle noise in SAR images, the edge attributes of the images are unknown. This leads over or under segmentation and produces poor quality of segmented results. In this paper, the deforestation rate of Amazon, South America, has been determined by using the color space-based SAR image segmentation. Since the fuzzy K-means (FKM) clustering technique is robust to speckle noise, it is combined with different color spaces for better segmentation. The proposed method has been compared with different existing methods, and it gives better segmentation results with a segmentation accuracy of 99.5%, Jaccard index of 99%, recall of 99.5%, and specificity of 99.7% with the minimum error rate of 0.004. The segmentation accuracy of 14.9% and the Jaccard index of 10.3% have been improved when compared with FKM clustering technique. This paper suggests that HSV+FKM is a suitable technique for the segmentation of LBA-ECO LC-14 modeled deforestation scenarios. Brazil's National Institute of Spatial Research (INPE) discovered that the span of Amazon could be diminished to 50% by 2050. This paper also predicts that the Amazon rainforest will be reduced to 43.26% at the end of 2050.

Keywords SAR image · Segmentation · Fuzzy K-means clustering · HSV color space · Deforestation rate

Introduction

The Amazon basin is the chunk of South America depleted by the Amazon River and its estuaries. The greater part of basin is protected with the aid of Amazon rainforest, called as Amazonia. The Amazonia is one of the world's biggest and most prominent normal assets. The total area of Amazon forest is 5.5 million km^2. The forest is being destroyed nowadays, for wood both for timbers and for making fires, for agriculture both small- and large-scale farms, land for poor farmers who do not have land anywhere to live, and grazing land for cattle and for road construction. PRODES (Amazon Deforestation Monitoring Project) and the Brazilian National Institute for Space Research (INPE) oversee the deforestation over the satellite images of the categorical deforestation in Amazon. The annual deforestation rate is predicted depending on the desertification expansion identified in every satellite image integument the Brazilian region known as legal Amazon. About 7.3% of the Amazon was deforested between 1976 and 2003 and further 2.6% between 2000 and 2010

Responsible Editor: Nilanchal Patel

✉ Kalaiyarasi Murugesan
mukalaiyarasi@gmail.com

Perumal Balasubramani
perumal@klu.ac.in

Pallikonda Rajasekaran Murugan
m.p.raja@klu.ac.in

Saravanan Sankaranarayanan
saravananklu@gmail.com

1 Department of Electronics and Communication Engineering, V.S.B Engineering College, Karur, Tamil Nadu, India

2 Department of Electronics and Communication Engineering, Kalasalingam Academy of Research and Education, Virudhunagar, Tamil Nadu, India

3 Department of Electrical and Electronics Engineering, B V Raju Institute of Technology, Telangana, India

Published online: 24 April 2021

(Guimberteau et al. 2017). According to INPE and Food and Agriculture Organization (FAO), till now totally 768,935 km^2 are deforested in Amazon basin (Kalaiyarasi et al. 2017).

Recently, the National Aeronautics and Space Administration (NASA) provides the results of the two modeled scenarios for future patterns of deforestation across the Amazon basin from 2002 to 2050 (LBA-ECO LC-14). Nowadays, SAR images are used to find the deforestation occurred in tropical rainforest. It has been estimated by comparing the two or more SAR images which are taken at different time periods. In this paper, the segmentation-based approach, namely, color space-based FKM clustering technique has been applied on LBA-ECO LC-14 dataset to calculate the deforestation rate of Amazon. In division 2, the existing deforestation monitoring approach and segmentation techniques as consistent with the prevailing papers in image segmentation is reviewed. Division 3 gives the detailed description of the methodology used for segmentation, whereas Division 4 shows the simulation results of SAR images and the estimation of deforestation rate of Amazon. Division 5 summarizes the findings and conclusions.

Related works

Object-based approach and RF algorithm are used on the Terra SAR-X (X-band) and Sentinel-1 (C-Band) SAR data to detect the degraded forest in the Sumaco Biosphere Reserve, Ecuador (Delgado-Aguilar et al. 2017). DETERring-Satellite–based system is utilized for monitoring the Amazon forest (Assuncao et al. 2017). NODEF scenario, LCC scenarios, and SSP of the AR5 RCP 8.5 scenarios over 2009–2099 are used to focus the future changes in the river of Amazon basin (Guimberteau et al. 2017). Quasi-experimental impact evaluation method is used to estimate the changes in forest cover in Ecuadorian Amazon (Kelly et al. 2017). Till now, different approaches and various scenarios have been used for monitoring the tropical rainforest. In this paper, segmentation-based approach is used to find the deforestation rate, for this is essential to find the better segmentation method. Therefore, the further related works is the review of exiting segmentation technique.

Fuzzy thresholding is used for SAR image segmentation. This method has been verified on Indian Remote Sensing (IRS) and Satellite Pour l'Observation de la Terre (SPOT) satellite imageries (Pal et al. 2000) (Kalaiyarasi et al. 2016). Color-based segmentation with genetic algorithm is proposed (Rathore et al. 2012), where the original color image is first converted to L*a*b* color space. Then, three individual components of L*a*b* color spaces are detached individually. Among this, a single component has been chosen depending on the color concealed by deliberation. Genetic algorithm is performed on the chosen sin component image to obtain the segmented image. This algorithm is very impressive for segmenting intricate background images. Clustering-based image segmentation is proposed where K-Means clustering technique with cosine distance measure is applied to the L*a*b color space image (Dibya and Kumar Gupta 2014). The initial clustered image is somewhat blurry. Sobel filter has been used to remove the blur present in the initial clustered image, and watershed algorithm is used for the final segmentation. This method is gratifyingly quite based on mean square error (MSE) and peak signal-to-noise ratio (PSNR) values. Improved color-based K-means clustering algorithm has been proposed for satellite images (Yadav and Biswas 2017). It consists of two stages. In the first stage, the interactive selection process is used to find the initial cluster which is then given to fuzzy K-means (FKM) clustering algorithm. The experimental results are evaluated through Dunn's index and Davis bound index. A fuzzy weighted active contour model is used for synthetic aperture radar image segmentation. This method improves the segmentation accuracy (Javed et al. 2016). In Hang et al. (2016), the authors used multi-feature ensemble–based segmentation for synthetic aperture radar images where the heterogeneous features are extracted and then integrated into the feature and similarity level. The fuzzy clustering algorithm is used for the segmentation. This method is validated on synthetic and real SAR images. In Modava and Akbarizadeh (2016), fuzzy clustering and Otsu's thresholding are carried out initially on SAR image. To eliminate spurious segments after binarization, morphological filters are utilized, and the coastline is extracted using active contour level set method. This method is based on contour model, and it does not require any preprocessing steps for speckle noise reduction. Experimental results on low- resolution and high-resolution SAR images show good performance for coastline extraction. In Bansal and Aggarwal (2011), ant colony optimization technique is applied to the CIELAB color space. Combination of multiple classifiers (CMC) distance measure is utilized to determine the distance between the image pixels. This technique has been simulated on test images, and the performance of this technique is evaluated with MSE values. In Lei et al. (2018), fast and robust fuzzy C-means (FRFCM) clustering technique has been proposed to improve the segmentation accuracy and also to suppress the noise present in the image. In this method, morphological reconstruction operation has been performed, it suppress the noise as well as preserves the edges. Additionally, local spatial information has been used to improve the segmentation outcome. In Lei et al. (2019), super pixel-based fast fuzzy C-means (SFFCM) clustering technique has been developed for color image segmentation. In this method, fastest segmentation has been achieved through color histogram. The main drawback of this method is the number clusters must be initialized prior. In Raval et al. (2017), color-based segmentation using FCM has been developed where they concluded that RGB color space will not provide good segmentation results also FCM (fuzzy C-means) technique produce cluster overlapping when less number of cluster center is selected.

The major drawback of K-means clustering is that it suffers from local minima, multiplicative noise, and over-segmentation. The computational complexity of K-means clustering is high when calculating the cluster centers. Color space model has been used to overcome the drawbacks of K-means clustering technique. It is very much beneficial for various fields such as multimedia applications, medical applications, color maps, remote sensing applications. In order to reduce the computational complexity of K-means, histogram quantization is used in HSV color space. HSV color space reduces the over-segmentation of K-means technique. In addition with, fuzzy K-means is used to overcome the cluster overlapping occurred in FCM. Also kernel-based FKM is robust to multiplicative noise. Therefore, various color space models and fuzzy K-means clustering are used together for SAR image segmentation, and the results are evaluated using various evaluation parameters.

Contribution of the paper

The contribution of the paper has been summarized as follows:

- Proposed a novel color space-based kernel fuzzy K-means clustering technique for SAR image segmentation. The brief explanation about the proposed segmentation technique is described in Section 3.
- The comparative analysis of proposed method with state of the art methods have been presented based on different performance evaluation metrics.
- From the segmentation results of HSV+FKM technique, the deforestation rate has been estimated using compound interest formula.

Color-based segmentation using fuzzy K-means clustering

To determine the deforestation rate of Amazon, it is necessary to segment the forestry region separately. Here, color space model and kernel-based FKM clustering techniques are used together to segment the forestry region. After segmenting the forestry region, compound interest formula has been used to calculate the annual deforestation rate of Amazon.

Color space gives a standard way to define a specific color by characterizing 3D coordinate scheme and a subspace that includes all the constructible hues within a specific model. Generally, the digital images are found in RGB format; other than this different color models are available which are most suitable for certain applications. To quantify the color difference accurately, the RGB color representation has been transformed into various color spaces. Then the transformed color spaces are segmented using kernel-based fuzzy K-means clustering process. Segmentation using color space provides better identification of objects in the segmented image compared to those generated using grayscale segmentation methods. The segmentation results using color space are very close to human perception. Different color space models are available in image processing. In this paper, HSV, YCbCr color spaces, and grayscale image segmentation have been used along with the kernel fuzzy K-means clustering technique to segment the forested region separately. Initially, the original SAR images are converted into different color spaces such as HSV and YCbCr. Then the new color space images are segmented using FKM technique. The FKM clustering technique segregates the original image into k number of clusters, and each cluster is correlated with the cluster center. Here, *3* clusters have been selected for LBA ECO scenarios. The block diagram of color-based segmentation using kernel-based fuzzy K-means clustering is described in Fig. 1. FKM technique provides good segmentation accuracy if the clusters in the image are well separated (Cebeci and Yildiz 2015). Also, HSV color space provides better color constancy and image similarity. When FKM clustering technique is used with HSV, YCbCr color spaces, HSV+FKM technique provides better segmentation accuracy when compared to other color spaces.

Color spaces

Color space is the specific organization of colors. It describes the range of colors as tuples of numbers, typically as 3 or 4 values or color components. Each color in an image is represented by a single value. The frequently used color space model is RGB model. Even though there exist various color space models, all the other color space models other than RGB are basically derived from RGB model. Here, the original images of the LBA-ECO LC-14 dataset appear in RGB color space. Since different color space models are suitable for certain applications, for forestry change detection analysis, it has been found that HSV color space is well suitable when compared with other color space models.

HSV (hue saturation value) color space

The HSV model is commonly used by researchers as it can be distinctly characterized by human perception, which is not consistent in the case of RGB and cyan, magenta, yellow, and key (CMYK) color space. This color space separates the intensity from chromaticity and represents the value of the color independently. Hue describes the location of the color in a 360° spectrum. Saturation describes the purity of the color

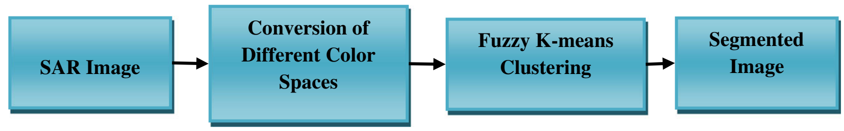


Fig. 1 Color-based segmentation using fuzzy K-means clustering

which measures the aberration amid the color and a grayscale value of equal intensity. Value is the luminance of the image.

The HSV components of LBA ECO LC-14 (2002 and 2050) images are given in Fig. 2, and the color conversion of RGB to HSV color space model images before segmentation are given in Fig. 3. The HSV color conversion is achieved by using the below formula (Dina et al. 2014):

$$H=\begin{cases}0 & \text{if}\quad Max=Min\\ \left(60^\circ\times\frac{G-B}{Max-Min}+360^\circ\right)\times \mathrm{mod}360^\circ & \text{if}\quad Max=R\\ 60^\circ\times\frac{B-R}{Max-Min}+120^\circ & \text{if}\quad Max=G\\ 60^\circ\times\frac{R-G}{Max-Min}+240^\circ & \text{if}\quad Max=B\end{cases}\tag{1}$$

$$S=\begin{cases}0 & \text{if}\quad Max=0\\ \frac{Max-Min}{Max} & \text{otherwise}\end{cases}\tag{2}$$

$$V=Max\tag{3}$$

YCbCr color space

The YCbCr model is acclimated for digital video. Luminance is represented as an individual component (Y), and chrominance is represented as two components (Cb and Cr). Cb is the aberration amid the blue component and reference value. Cr is the distinction amid the red component and reference value. Cb is strong in case of parts of the image containing the sky (blue), whereas Cb and Cr are anemic in case of blush like green and Cr is powerful in the places of the prevalence of reddish colors. The YCbCr components of LBA ECO LC-14 (2002 and 2050) images are given in Fig. 4, and the color conversion of RGB to YCbCr color space model images are given in Fig. 5. The formulae used for converting from RGB color space to YCbCr color space are stated below.

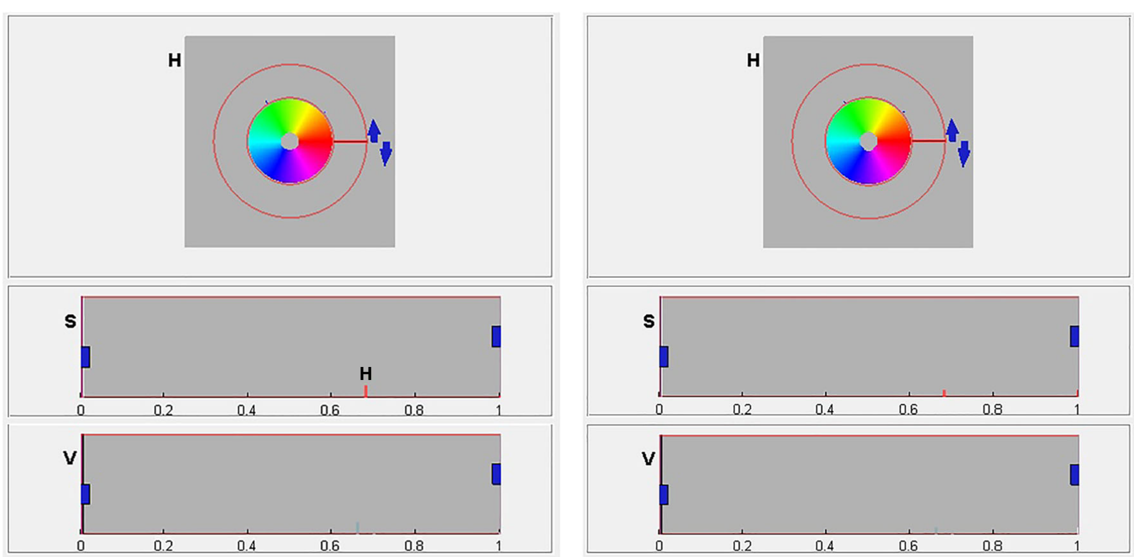
$$Y=\left(\frac{77}{256}\right)R+\left(\frac{150}{256}\right)G+\left(\frac{29}{256}\right)B\tag{4}$$

$$Cb=-\left(\frac{44}{256}\right)R-\left(\frac{87}{256}\right)G+\left(\frac{131}{256}\right)B+128\tag{5}$$

$$Cr=\left(\frac{131}{256}\right)R-\left(\frac{110}{256}\right)G-\left(\frac{21}{256}\right)B+128\tag{6}$$

Fuzzy K-means (FKM) clustering

Fuzzy-based image segmentation is derived from fuzzy thresholding and clustering techniques (Baldevbhai and Anand 2012). This technique is used to associate the pixels depending upon the specific characteristics of the image into *K* number of different groups. Each group has a separate cluster center (Pham 2015). The relationship amid the image



(i) 2002 **(ii) 2050**

Fig. 2 HSV (Hue, Saturation and Value) components of LBA ECO LC-14 2002 and 2050 images

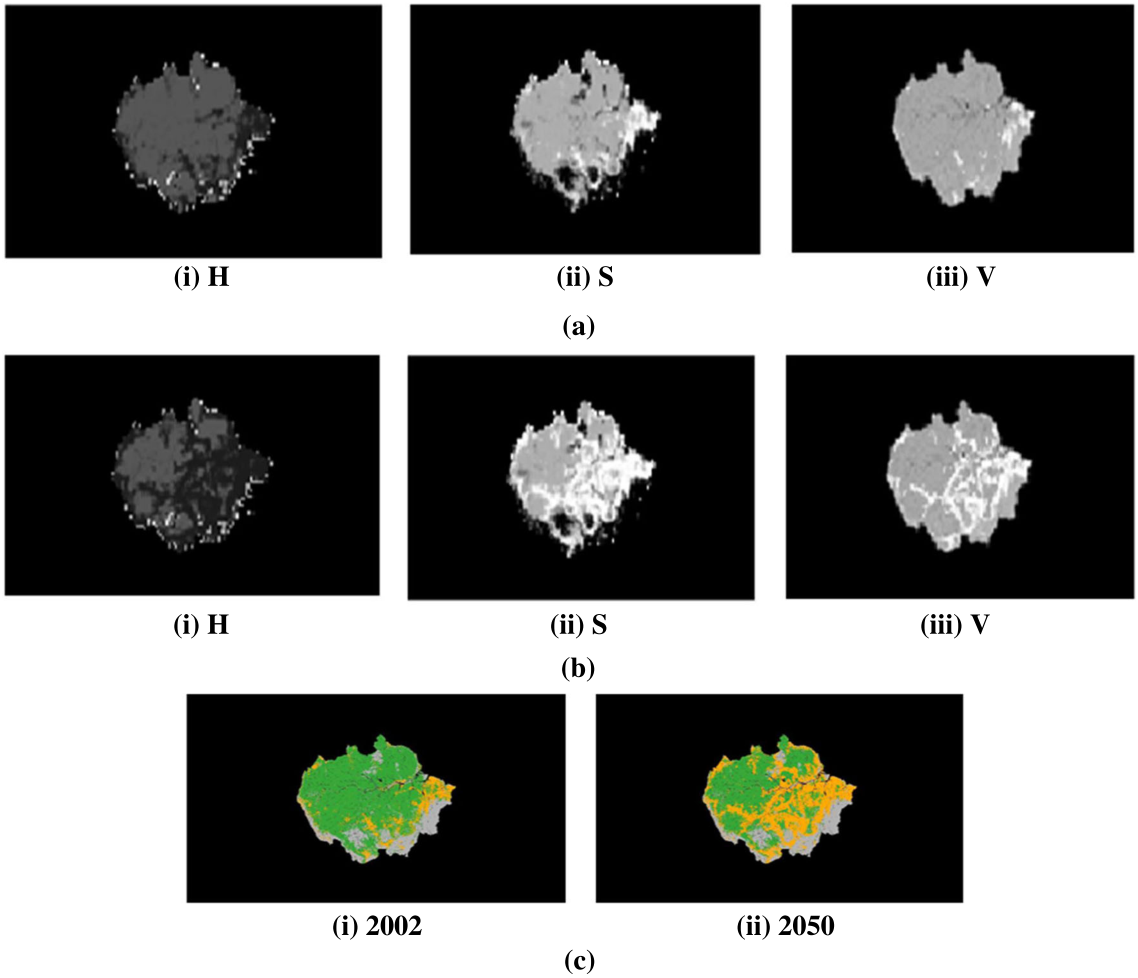


Fig. 3 Color conversion of RGB to HSV (Hue, Saturation and Value) color space model before segmentation

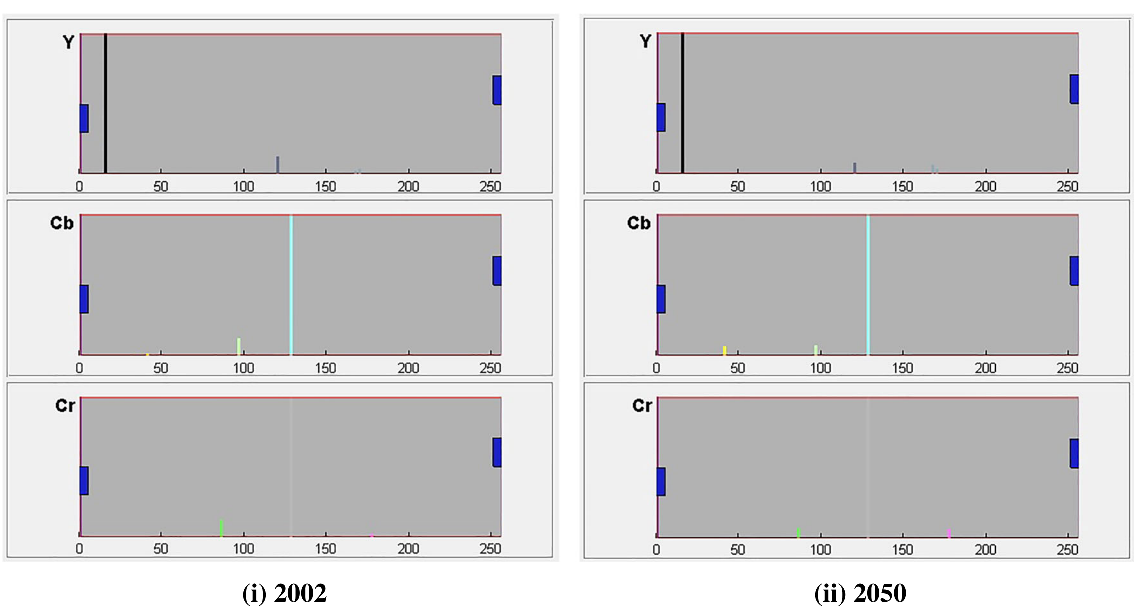


Fig. 4 YCbCr (Luminance and Chrominance (blue and red)) components of LBA ECO LC-14 2002 and 2050 images

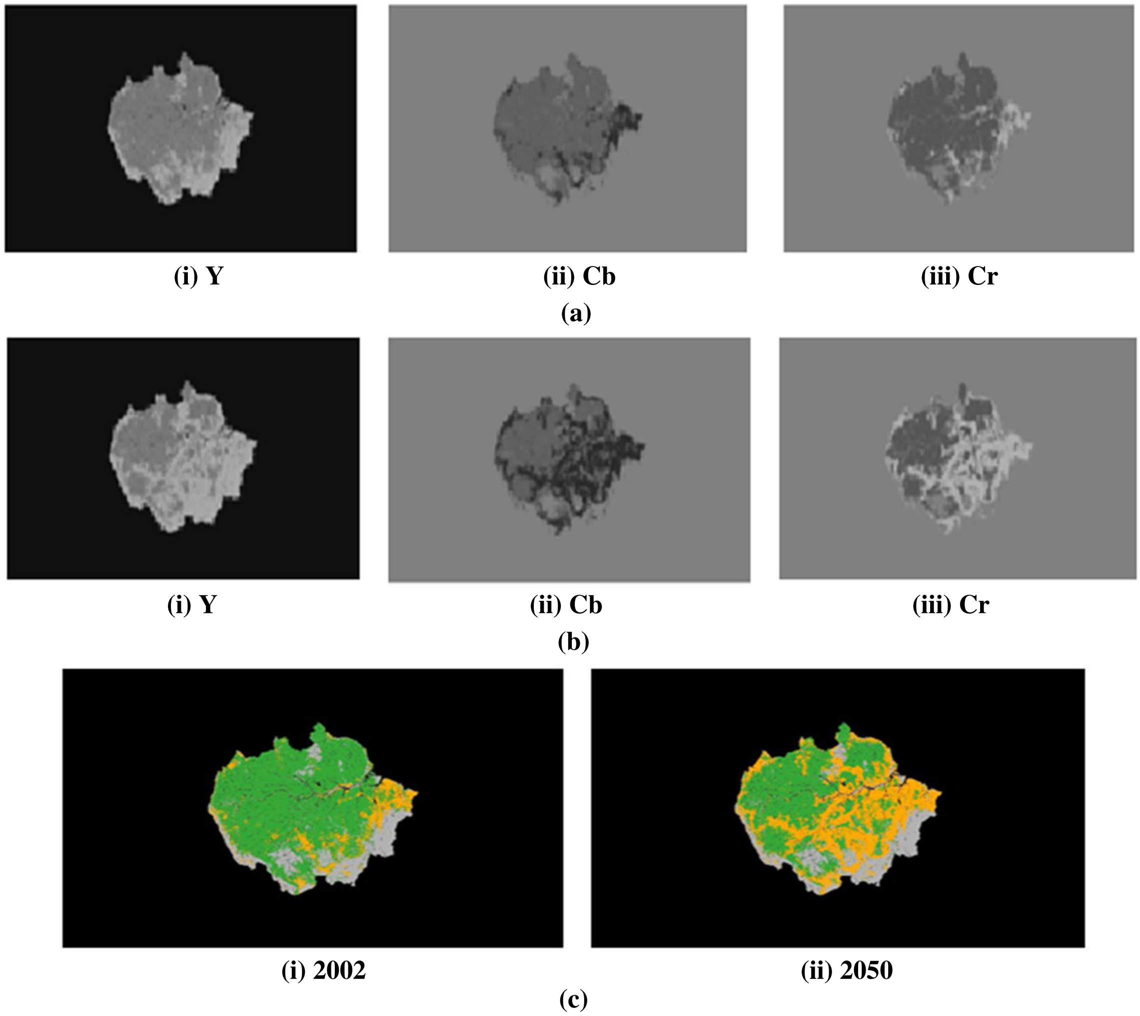


Fig. 5 Color conversion of RGB to YCbCr (Luminance and Chrominance (blue and red)) color space model before segmentation

pixels and cluster center is fuzzy. The membership function $u_{i,j} \in [0, 1]$ is used to represent the degree of accouterments of data point X_i and cluster center C_j. The fuzzy K-means algorithm is used to minimizing the following distortion (Emami and Derakhshan 2015):

$$J = \sum_{j=1}^{k} \sum_{i=1}^{N} u_{i,j}^{m} d_{ij} \tag{7}$$

where N is the number of data points, m is the fuzzifier parameter, k is the number of clusters, and d_{ij} is the squared

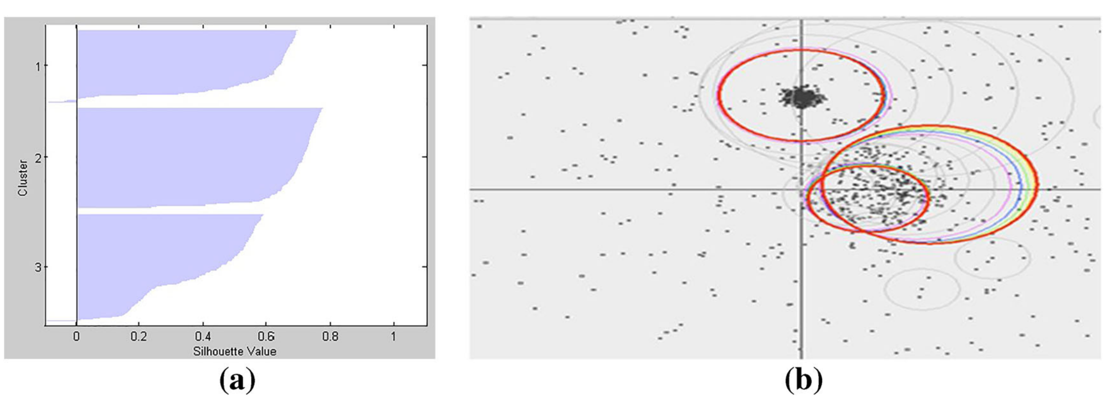


Fig. 6 **a** Number of clusters. **b** Fuzzy K-means clustering

Euclidean distance (ED) amid the image pixels and the cluster center (Kalaiyarasi et al. 2020). The selection of number of clusters and fuzzy K-means clustering is given in Fig. 6. Totally 3 clusters have been selected for LBA-ECO LC-14 deforestation dataset.

The membership $u_{i,j}$ must satisfy the following constraints (Chih-Tang et al. 2011):

$$\sum_{j=1}^{N} u_{ij} = 1, for\ i = 1\ to\ N \tag{8}$$

Since the kernel distance measure is robust to speckle noise and effective for SAR image segmentation (Deliang et al. 2014), in this paper kernel-based distance measure has been performed. The objective function depends on the kernel distance; it is most essential to reduce the distance measure. If the distance is minimum, then the distortions in the objective function is also minimum. FKM begins with a set of initial cluster centers chosen randomly, and no two or more clusters have the same cluster centroid. Then update the membership function using kernel distance for calculating new centroids. After having K-new centroids, a new grouping must be done between the same image pixels and the nearest new cluster center. Thus, an iterative process has been performed. The K-new cluster centers change their location for every iteration until the cluster center becomes stable. In such a way, FKM clustering minimizes the kernel distance between images pixels and cluster center. When the distance is minimum, the distortion is also minimum. Hence the distortions in the objective function have been reduced. The fuzzy K-means algorithm is as follows:

Step 1. Choose a set of initial clusters randomly and set $p = 1$.
Step 2. Calculate the kernel based distance d_{ij}.

The squared distance is computed in the kernel space using a kernel function such that (Shen et al. 2006; Zhang and Chen 2003)

$$d_{ij} = \Phi(x_k) - \Phi(v_i)^2 \tag{9}$$

It can be represented as

$$\begin{aligned} \Phi(x_k) - \Phi(v_i)^2 &= \Phi(x_k)^T \Phi(x_k) + \Phi(v_i)^T \Phi(v_i) - 2\Phi(x_k)^T \Phi(v_i) \\ &= K(x_k, x_k) + K(v_i, v_i) - 2K(x_k, v_i) \end{aligned} \tag{10}$$

Step 3. Update the membership function $u_{i,j}$ using the below equation (Chih-Tang et al. 2011):

$$u_{i,j} = \left((d_{ij})^{1/_{m-1}} \sum_{l=1}^{k} \left(\frac{1}{d_{il}} \right)^{1/_{m-1}} \right)^{-1} \tag{11}$$

Here, $l \neq j$. If $d_{ij} < \eta$, then set $u_{i,j} = 1$, where η is the small positive number.

Step 4. Compute the new set of cluster centers for each initial cluster center by using the equation given below (Chih-Tang et al. 2011):

$$C_j = \frac{\sum_{i=1}^{N} u_{ij}^m X_i}{\sum_{i=1}^{N} u_{ij}^m} \tag{12}$$

Step 5. If $\|C_j - C_{j-1}\| < \varepsilon\ for\ j=1\ to\ k$, then stop the iteration. Otherwise set $p + 1 \rightarrow p$ and repeat step 2.

The computational complexity of step 3 is higher than step 4. Here, 12 iterations have been performed. During 12th iteration, it satisfies the above condition. Therefore the iteration process has been stopped.

Experimental results on SAR images

It is necessary to separate the color components from intensity for differentiating the forestry land cover from the deforested and the non-forest land cover. Among different color spacing models, HSV color space is often used in image processing for different applications when compared with other color spaces. In order to ensure the performance of HSV+FKM technique, it is compared with other color space methods such as YCbCr+FKM and Gray+FKM. Here, the simulation procedures for these three methods are same. They differ only with the color space models. In this method, initially the RGB model is converted into different color space models (HSV, YCbCr, and gray), and then the FKM clustering technique has been performed on the image. It is explained with two steps. The steps are as follows:

Step 1. Convert the RGB model into HSV, YCbCr, and gray color space model using Eqs. 1 to 6.
Step 2. Apply FKM clustering technique

Dataset description

LBA-ECO LC-14 modeled deforestation scenarios, Amazon basin: 2002–2050 have been used for this experiment, and this scenario is downloaded from the website: https://daac.ornl.gov. This scenario gives the results of two modeled situations for future patterns of deforestation throughout the

Amazon basin from 2002 to 2050. It includes two types of eventualities named as business-as-usual and governance scenario. In this paper, the business-as-usual scenario has been considered for the experiment. In this scenario, the deforestation trends across the basin are projected through historical figures.

In LBA-ECO dataset, it is required to segment the forestry land cover and deforestry land cover separately. For this purpose, 3 clusters are more sufficient instead of choosing a higher number of clusters. Therefore only 3 clusters have been utilized for the fuzzy K-means clustering technique. Three clusters represent forest, de-forest, and non-forest land cover. Initially, three cluster centers have the values of 78,128 and 153, respectively. Histogram analysis has been performed to find the initial value of cluster center. The minimum and maximum values of clusters are obtained from the histogram. The average of minimum and maximum value of a single cluster gives the initial cluster center. The values are shown in Table 1.

In FKM, 12 iterations have been performed. The cluster centers (centroids) have new value for every new iteration, and at the end of 12th iteration, they have the values of 174,128 and 208, respectively. The computational complexity of step 3 is high when compared with step 4. The computational complexity of FKM is represented in terms of Euclidean distance calculation, that is $O(Nkt)$, where t is the number of iterations. When the number of iterations is less, the computational complexity is also reduced. Since FKM performs an iterative process to generate a new set of cluster centers, it has extensive convergence rate. Anyway, the convergence rate of FKM is much lower than that of the hard K-means clustering technique (Fan et al. 2003). Therefore, the convergence time of this proposed method is also less. The segmented results of LBA-ECO LC-14 modeled deforestation scenarios, Amazon basin 2002–2050 using HSV+FKM clustering, YCbCr+FKM clustering and grayscale image segmentation with FKM clustering are shown in Table 2.

Performance evaluation

The color-based segmentation along with FKM is applied on the SAR images. Here, the segmented SAR images and ground truth images are used to compare the performance metrics of the abovementioned three methods. The performance analysis is commonly evaluated using confusion matrix. From the analysis of performance metrics, it can be seen that the color-based segmentation using HSV and fuzzy K-means clustering method provides better segmentation results.

Confusion matrix is a table, where the number of correct and incorrect predictions are summarized with count values and broken down by each class. It is also referred to a contingency table. The confusion matrix is given in Table 3. The entries in the confusion matrix are described as follows: true positive (TP) is the number of correct prediction that a pixel belongs to a particular cluster; true negative (TN) is the number of correct prediction that a pixel does not belong to a particular cluster; false positive (FP)is the number of incorrect predictions that a pixel belongs to cluster; and false negative (FN) is the number of incorrect predictions that a pixel does not belong to a cluster. Various performance measures can be derived from the confusion matrix.

Evaluation parameters

Before one gets to know the performance of an algorithm, knowing comprehensively the definitions of these matrices is inevitable. Various performance parameters used for evaluation of SAR image segmentation are as follows.

Recall (sensitivity or true positive rate)

It is calculated as the ratio of correct positive predictions divided by the total number of positives. High recall indicates the class is correctly recognized (small number of FN). The best sensitivity is 1, whereas the worst is 0. It is calculated by using the below formula (Aqil Burne and Tariq 2014):

$$\mathrm{Re}call = \frac{TP}{TP + FN} \tag{13}$$

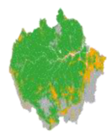
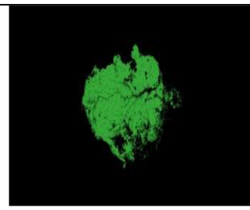
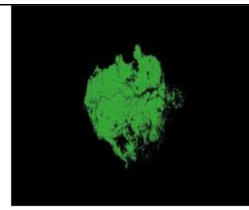
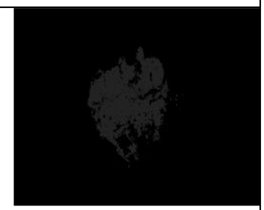
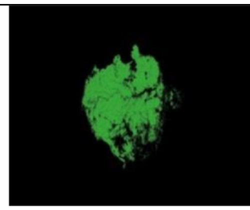
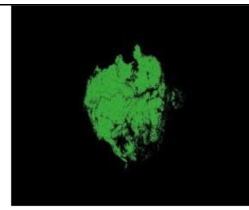
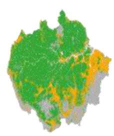
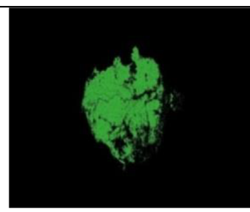
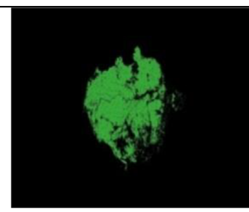
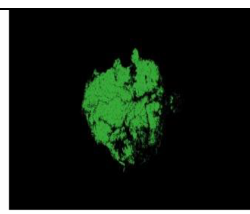
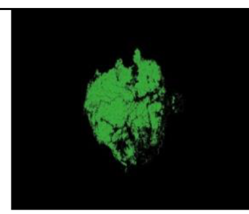
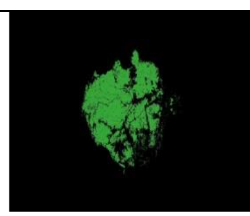
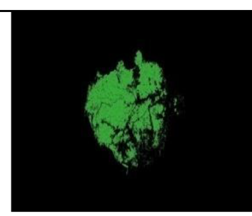
Specificity (true negative rate)

Specificityreflectsthetruenegativerateofclass.Itistheratioof correctnegativepredictionsdividedbythetotalnumberofnegatives. The best specificity is 1, whereas the worst is 0. It is obtained by using the below formula (Aqil Burne and Tariq 2014):

Table 1 Minimum, maximum, and initial value of cluster center

Cluster	Area (pixels)	Name	Minimum	Cluster center	Maximum
1	50736	51–105.82	51	78.409091	105.818182
2	13494	105.82–150.36	105.818182	128.090909	150.363636
3	9607	150.36–156.18	150.363636	153.272727	156.181818

Table 2 Segmented results of LBA-ECO LC-14 modeled deforestation scenarios, Amazon basin: 2002–2050 (a) Original image (b) HSV+FKM Clustering (c) YCbCr+FKM clustering and (d) grayscale +FKM clustering

Images (Year)	Input Images	HSV+FKM	YCbCr+FKM	Gray+FKM
2002				
2010				
2015				
2020				
2025				

$$Specificity = \frac{TN}{TN + FP} \quad (14)$$

Accuracy

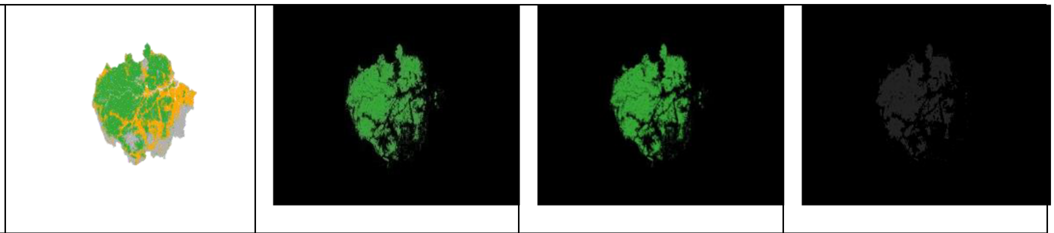
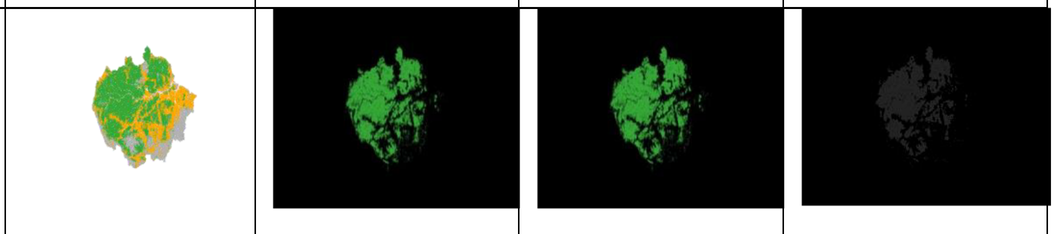
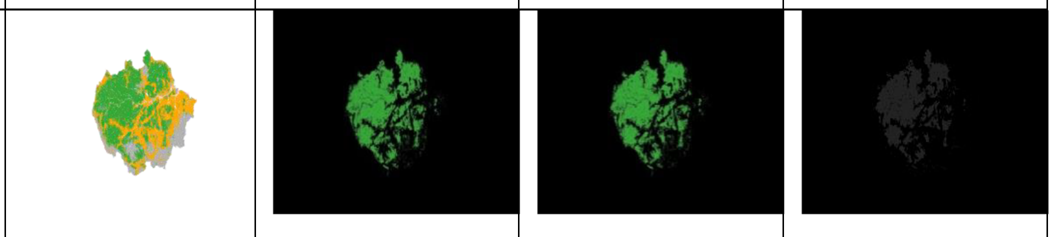
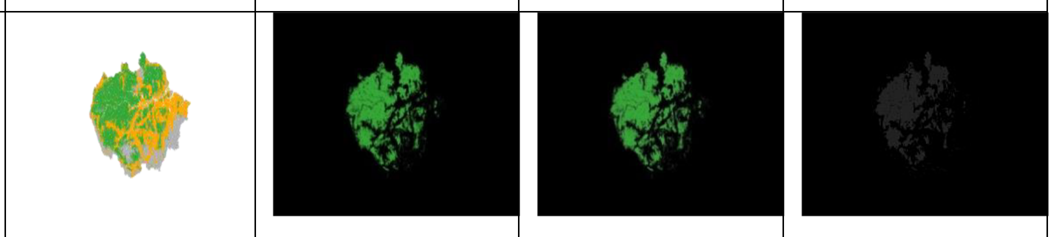
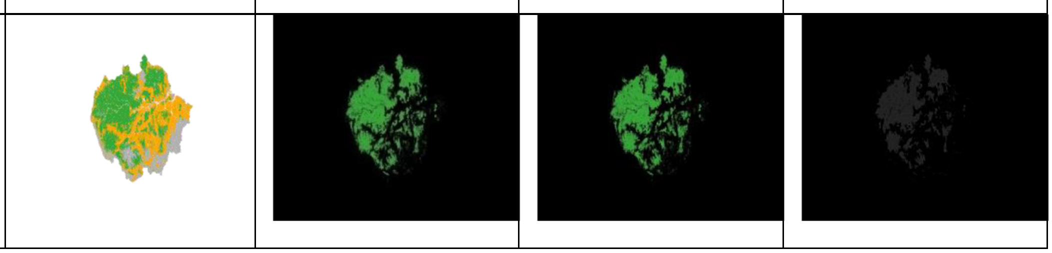
Accuracy is the proportion of the total number of predictions that are correct. Ninety-nine percent accuracy can be excellent. It is determined by using the subsequent formula (Aqil Burne and Tariq 2014):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (15)$$

Error rate

Error rate is the ratio of a total number of two incorrect predictions (FN + FP) to the total number of a dataset (P + N). The best error rate is 0.0, whereas the worst is 1.0 which is calculated by the formula mentioned below:

Table 2 (continued)

2030	
2035	
2040	
2045	
2050	

(a) **(b)** **(c)** **(d)**

$$Error\ \ Rate = \frac{FP+FN}{TP+TN+FN+FP} = \frac{FP+FN}{P+N} \tag{16}$$

$$Jaccard\quad Index\quad = \frac{TP}{TP+FP+FN} \tag{17}$$

Jaccard index

Jaccard index is a similarity coefficient used to measure the similarity between the two images, returned as a numeric scalar or vector with values in the range [0, 1]. `Similarity` of 1 proves that the segmentation in the two images is a perfect match.

Table 3 Confusion matrix

	Predicted No	Predicted Yes
Actual No	TN	FP
Actual Yes	FN	TP

Table 4 Performance evaluation of accuracy and Jaccard index for HSV+FKM, YCbCr+FKM, and Gray+FKM methods

Images (year)	Accuracy			Jaccard index		
	HSV+FKM	YCbCr+FKM	Gray+FKM	HSV+FKM	YCbCr+FKM	Gray+FKM
2002	0.992565	0.99532	0.986727	0.948088	0.963372	0.939765
2010	0.995464	0.993781	0.986871	0.964378	0.962059	0.940348
2015	0.99519	0.993713	0.986371	0.969741	0.962208	0.941787
2020	0.995181	0.995415	0.986268	0.966618	0.963019	0.939904
2025	0.995227	0.993690	0.986445	0.969741	0.963330	0.940030
2030	0.995161	0.993676	0.986189	0.976351	0.965316	0.943618
2035	0.995423	0.993038	0.986182	0.972658	0.964993	0.944463
2040	0.995342	0.993103	0.986411	0.978575	0.963871	0.946482
2045	0.995211	0.992750	0.986269	0.976867	0.967833	0.947506
2050	0.995263	0.992913	0.986486	0.978599	0.96878	0.945479
Average	0.995002	0.993739	0.986421	0.970161	0.964478	0.942938

The obtained values of various parameter metrices are listed in Tables 4, 5 and 6. These values are used to compare the performance of segmentation algorithm. Table 4 gives the performance evaluation of accuracy and Jaccard index of three methods. From the table, it is inferred that HSV+FKM provides higher value of Jaccard index and accuracy.

Table 5 defines the recall and specificity values. For better segmentation, recall and specificity must be high. HSV+fuzzy K-means clustering technique produces the highest value for recall and specificity. The performance evaluation of error rate is given in Table 6. The lowest error rate gives better segmentation result.

The comparison of proposed method with different existing segmentation methodologies are given in Table 7. The HSV+FKM method has been compared with different existing methodologies such as SFFCM (Lei et al. 2019), FRFCM (Lei et al. 2018), spatial constraint-based fuzzy clustering (Singh and Garg 2014), fuzzy K-means (Chih-Tang et al. 2011), and K-means clustering (Pham 2015) segmentation techniques. In this table, the average value of Jaccard index is 0.97 for HSV+FKM method, 0.964 for YCbCr+FKM method, and 0.94 for Gray+FKM method. For a good segmentation technique, the value of Jaccard index, accuracy, recall, and specificity must be a higher value and lower for error rate. Both the HSV and YCbCr+fuzzy K-means clustering techniques produce the accuracy of 99%, but HSV method produces little bit higher than the YCbCr method, whereas the grayscale segmentation produces the accuracy of 98%. The proposed HSV+FKM method produces better segmentation results with the accuracy of 99.5%, Jaccard index of 97%, recall of 99.5%, and specificity of 99.7% with the minimum error rate of 0.4%. The SFFCM clustering technique produces

Table 5 Performance evaluation of recall/sensitivity and specificity for HSV+FKM, YCbCr+FKM, and Gray+FKM methods

Images	Recall			Specificity		
(year)	HSV+FKM	YCbCr+FKM	Gray+FKM	HSV+FKM	YCbCr+FKM	Gray+FKM
2002	0.991621	0.987087	0.983388	0.994656	0.997647	0.988358
2010	0.994640	0.989377	0.983184	0.997906	0.996141	0.988623
2015	0.996527	0.988680	0.982961	0.997878	0.996320	0.988314
2020	0.997544	0.985761	0.982872	0.998064	0.998289	0.988411
2025	0.995490	0.987917	0.982850	0.998106	0.996473	0.988579
2030	0.994372	0.987229	0.982625	0.998181	0.996581	0.988471
2035	0.997416	0.988179	0.982631	0.998377	0.995863	0.988422
2040	0.996320	0.988116	0.982412	0.998421	0.996017	0.988792
2045	0.996336	0.987147	0.982256	0.998464	0.995866	0.988842
2050	0.997301	0.986648	0.982268	0.998590	0.996107	0.989104
Average	0.995756	0.987614	0.982744	0.997864	0.996530	0.9885916

Table 6 Performance evaluation of error rate for HSV+FKM, YCbCr+FKM, and Gray+FKM methods

Images	Error rate		
	HSV+FKM	YCbCr+FKM	Gray+FKM
2002	0.007435	0.004681	0.013273
2010	0.004536	0.006219	0.013129
2015	0.004811	0.006287	0.013629
2020	0.004819	0.004585	0.013732
2025	0.004773	0.006314	0.013555
2030	0.004839	0.006324	0.013811
2035	0.004577	0.006962	0.013818
2040	0.004658	0.006897	0.013589
2045	0.004789	0.007252	0.013731
2050	0.004737	0.007087	0.013514
Average	0.004997	0.006260	0.013578

Table 7 Comparison of proposed method with different existing methods

Methods	Accuracy	Jaccard Index	Recall	Specificity
HSV+FKM	99.5	97.0	99.5	99.7
YCbCr+FKM	99.3	98.7	98.7	99.6
Gray+FKM	98.6	94.2	98.2	98.8
SFFCM (Lei et al. 2019)	96.3	92.6	98.1	98.6
FRFCM (Lei et al. 2018)	92.4	91.1	97.3	98.2
FKM (Chih-Tang et al. 2011)	84.6	86.7	96.8	97.1
K-means (Pham 2015)	75.5	77.8	88.5	95.6
Spatial constrain-based fuzzy clustering (Singh and Garg 2014)	79.6	78.6	97.2	97.3

the accuracy of 96.3%, Jaccard index of 92.6%, recall of 98.1%, and specificity of 98.6%. Similarly, the FRFCM clustering technique produces the accuracy of 92.4%, Jaccard index of 91.1%, recall of 97.3%, and specificity of 98.2%. The spatial constraint-based fuzzy clustering method produces the segmentation accuracy of 79.67%. The FKM clustering technique produces the accuracy of 84.6%, Jaccard index of 86.7%, recall of 96.8%, and specificity of 97.1%. The segmentation accuracy of K-means clustering technique is 75.5%, Jaccard index is 77.8%, recall of 88.5%, and specificity of 95.6%, respectively.

The graphical illustrations of comparison of accuracy, Jaccard index, recall, and specificity for the proposed HSV+FKM method with other existing methodologies are represented in Fig. 7. From the figure, it is inferred that the HSV+FKM has the highest value of Jaccard index, accuracy, specificity, and recall. It can be clearly observed that the HSV+FKM clustering technique yields better segmentation results when compared to YCbCr and grayscale-based fuzzy K-means clustering technique. When comparing the efficiency of proposed method, 19.9% of accuracy has been improved with spatial constraint-based fuzzy clustering, 24% with K-means clustering, and 14.9% with fuzzy K-means clustering techniques. In addition to that, 3.2% of accuracy has been improved with SFFCM and 7.1% has been increased when compared with FRFCM clustering techniques. From Table 8, it is observed that 18.4% of Jaccard index has been improved when compared with spatial constraint-based fuzzy clustering, 10.3% when compared with FKM clustering, and 19.2% when compared with K-means clustering technique. When comparing the Jaccard index of proposed method with the fastest clustering techniques, 4.4% has been improved with SFFCM, and 5.9 have been expanded with FRFCM techniques. A 2.3% of recall has been improved when compared with spatial constraint-based fuzzy clustering, 1.4% has been increased

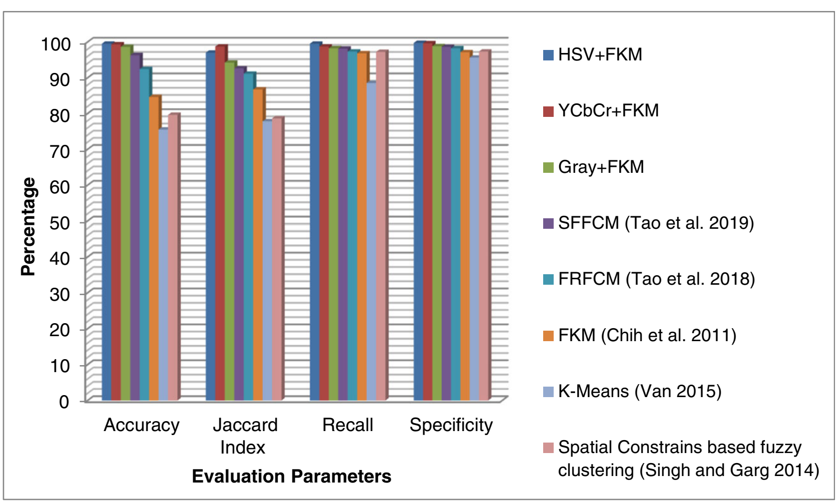


Fig. 7 Graphical illustration of comparison of proposed method with different existing segmentation methodologies

Table 8 Computational time of segmentation

Methods	Time (sec)
HSV+FKM	5.26
YCbCr+FKM	5.34
Gray+FKM	5.52
SFFCM (Lei et al. 2019)	5.58
FRFCM (Lei et al. 2018)	6.38
FKM (Chih-Tang et al. 2011)	6.59
K-Means (Pham 2015)	7.91
Spatial constraint-based fuzzy clustering (Singh and Garg 2014)	8.98

with SFFCM, 2.2% with FRFCM, 2.7% when compared with FKM clustering technique, and 11% when compared with K-means clustering technique. Similarly, 2.4% of specificity has been improved when compared with spatial constraint-based fuzzy clustering, 2.6% when compared with FKM clustering technique, and 4.1% with K-means clustering technique. Also, 1.1 % of specificity has been expanded with SFFCM and 1.5% increased with FRFCM when compared with the proposed HSV+FKM method. From the analysis, it can be clearly observed that the HSV+FKM clustering technique produces better segmentation result when compared with other existing segmentation methods. Therefore, this technique is suitable for the segmentation of LBA-ECO deforestation scenarios.

The comparative analysis of computational time for segmentation has been given in Table 8. The average computational time has been computed in seconds. From the table, it can be said that the HSV+FKM is computationally faster than the other existing segmentation methods. The experiment has been simulated using MATLAB 2014 with the intel-core i3-2330M processor and 2 GB memory. Table 8 suggests that the color-based HSV+FKM segmentation method has a shorter segmentation time than the existing segmentation methods.

Application of deforested area calculation

Estimation of deforestation rate of Amazon

From the above experimental results and parameter evaluation, it can be noticed that the HSV+FKM clustering technique is the best segmentation technique for segmenting LBA-ECO LC-14 SAR images. The deforestation rate of Amazon can also be computed from the segmented results.

The deforestation rate has been computed from the segmented forestry land cover images by using the below equation which is described in Jean-Philippe (2003) and Kalaiyarasi et al. (2017):

$$Deforestation \quad Rate = \frac{1}{t_2 - t} \ln\left(\frac{A_2}{A_1}\right) \tag{18}$$

where t_2, t_1 are the two distinct time periods and A_2, A_1 are the areas of forestry land cover. The estimation of deforestation rate, percentage of annual deforestation, and total deforestation of Amazon scenarios are given in Table 9. The annual deforestation rate of 2050 is not calculated, because it needs 2051 scenario. This paper utilizes 2002–2050 deforestation scenarios. Figure 8 gives the graphical representation of the

Table 9 Estimation of deforestation rate, percentage of annual deforestation, and total deforestation of Amazon scenarios

Images (year)	Estimated remaining forest cover		Annual forest loss		Total forest loss		*Annual deforestation rate*
	Pixels	%	Pixels	%	Pixels	%	
2005	50736	68.7135	306	0.4144	9607	13.0110	-0.00657
2010	49062	66.4463	357	0.4834	11281	15.2782	-0.00697
2015	47281	64.0342	369	0.4997	13062	17.6903	-0.00918
2020	45187	61.1983	402	0.5444	15156	20.5262	-0.00965
2025	42896	58.0955	454	0.6148	17447	23.6290	-0.01128
2030	40625	55.0198	460	0.6229	19718	26.7047	-0.01209
2035	38329	51.9102	483	0.6541	22014	29.8143	-0.01136
2040	36097	48.8874	445	0.6026	24246	32.8372	-0.01218
2045	33918	45.9363	388	0.5254	26425	36.8569	-0.01142
2046	33533	45.4149	385	0.5214	26810	39.7882	-0.0677
2047	31338	42.4421	2195	2.9727	29005	36.2824	0.043919
2048	32745	44.3476	-1407	-1.9055	27598	37.3769	-0.01173
2049	32363	43.8303	382	0.5173	27980	37.8942	-0.01297
2050	31946	43.2655	417	0.5647	28397	38.4590	-

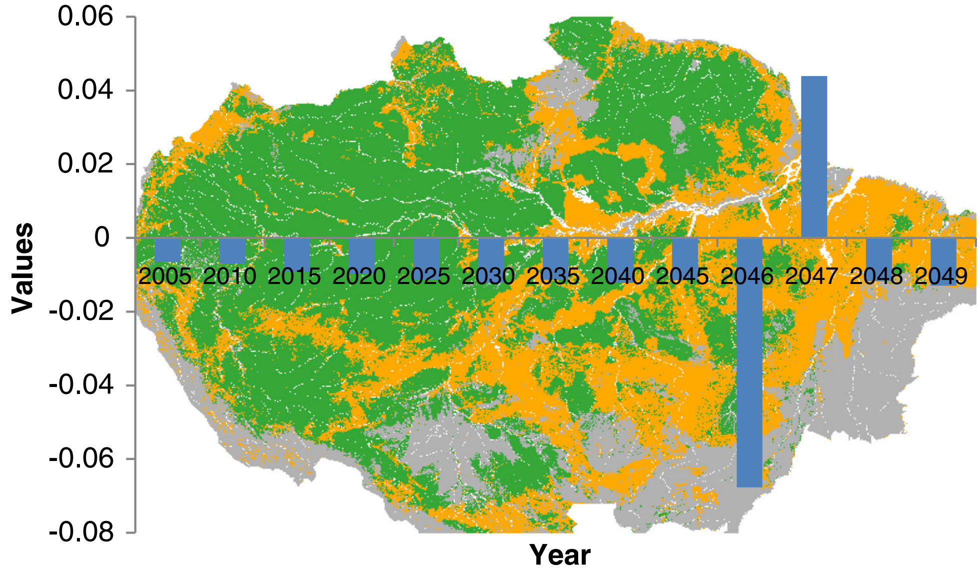


Fig. 8 Annual deforestation rate of LBA-ECO deforested scenarios of Amazon

annual deforestation rate of LBA-ECO LC-14 modeled deforestation scenario in Amazon.

The annual deforestation rate gradually increases from 2002 to 2050. It shows that the Amazon rainforest has been progressively shrinking since 2002. The negative values represent the deforestation, and positive values represent the reforestation. The higher amount of deforestation is about to occur by 2046. Due to the highest annual deforestation rate in 2046, the Amazon forest will be reforested with the rate of + 0.0439 in 2047. Again the forest will be continuously deforested up to 2050. At the end of 2050, only 43.26% of the Amazon rainforest will be left out based on the LBA ECO deforestation scenarios. Also, it should be noted that Brazil's National Institute of Spatial Research discovered that the span of Amazon could be diminished to 50% by 2050.

Conclusion

In this paper, color space-based SAR image segmentation using kernel-based FKM clustering technique has been performed. This proposed technique has been compared with different existing segmentation techniques. The comparative analysis predicts that HSV+FKM clustering technique gives the better segmentation accuracy of 99.5% with minimum error rate of 0.004. Furthermore, the segmentation results of HSV+FKM technique is used to evaluate the deforestation rate of LBA-ECO deforestation scenarios. This paper also gives the accurate percentage of Amazon rainforest by the end of 2050, and it will be 43.26%. Deforestation increases the amount of carbon dioxide (CO_2) and other trace gases in the atmosphere. It also affects the biodiversity and local climate of an area by reducing the evaporative cooling that takes place from both soil and plant life. The annual deforestation rate if 2020 is 0.009 (0.9%). If the current rate of deforestation remains, the world's forest will disappear within 100 years, causing unknown effects on global climate.

Acknowledgements We thank the Department of ECE of Kalasalingam Academy of Research and Education (Deemed to be University), Tamil Nadu, India, for the computational facilities made available in the Centre for Signal Processing Laboratory (supported by Department of Science and Technology (DST), New Delhi under FIST Program) (Reference No: SR/FST/ETI-336/2013 dated November 2013).

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

Aqil Burne SM, Tariq H (2014) K-means cluster analysis for image segmentation. Int J Comput Appl 96(4):1–8

Assuncao J, Gandour C, Rocha R (2017) DETERring deforestation in the Amazon: environmental monitoring and law enforcement. Clim Policy Initiat 9:1–47

Baldevbhai PJ, Anand RS (2012) Color image segmentation for medical images using L*a*b* color space. IOSR J Electron Commun Eng 1(2):24–45

Bansal S, Aggarwal D (2011) Color image segmentation using CIELab color space using ant colony optimization. Int J Comput Appl 29(9): 28–34

Cebeci Z, Yildiz F (2015) Comparison of k-means and fuzzy c-means algorithms on different cluster structures. J Agric Inform 6(3):13–23

Chih-Tang C, Jim ZCL, Jeng M-D (2011) A fuzzy k-means clustering algorithm using cluster center displacement. J Inf Sci Eng 27:995–1009

Delgado-Aguilar MJ, Fassnacht FE, Peralvo M, Gross CP, Schmitt CB (2017) Potential of TerraSAR-X and Sentinel 1 imagery to map deforested areas and derive degradation status in complex rain forests of Ecuador. Int For Rev 19(1):102–115

Deliang X, Tang T, Hu C, Yu L, Yi S (2014) A kernel clustering algorithm with fuzzy factor: application to SAR image segmentation. IEEE Geosci Remote Sens Lett 11(7):1290–1294

Dibya JB, Kumar Gupta A (2014) A novel approach towards clustering based image segmentation. Int J Emerg Sci Eng 2(11):6–10

Dina K, Ebied HM, Hussein AS, Tolba MF (2014) Color image segmentation based on different color space models using automatic GrabCut. Sci World J 2014:126025–126010. https://doi.org/10.1155/2014/126025

Emami H, Derakhshan F (2015) Integrating fuzzy K-means, particle swarm optimization, and imperialist competitive algorithm for data clustering. Arab J Sci Eng 40(12):3545–3554

Fan J, Zhen WZ, Xie WX (2003) Suppressed fuzzy c-means clustering algorithm. Pattern Recogn Lett 24(9-10):1607–1612

Guimberteau M, Ciais P, Ducharne A, Boisier JP, Aguiar APD, Biemans H, Ugent HDD, Galbraith D, Kruijt B, Angerwisch F (2017) Impacts of future deforestation and climate change on the hydrology of the Amazon basin: a multi-model analysis with a new set of land-cover change scenarios. Hyreol Earth Syst Sci 21(3):1455–1475

Hang Y, Jiao L, Liu F (2016) CRIM – FCHO : SAR image two-stage segmentation with multi feature ensemble. IEEE Trans Geosci Remote Sens 54(4):2400–2423

Javed U, Riaz MM, Ghafoor A, Cheema TA (2016) SARimage segmentation based on active contours with fuzzy logic. IEEE Trans Aerosp Electron Syst 52(1):181–188

Jean-Philippe P (2003) Standardizing the calculation of the annual rate of deforestation. For Ecol Manag 177:593–596

Kalaiyarasi M, Perumal B, Pallikonda Rajasekaran M (2016) A review on: de-speckling methods of SAR image. International conference on control, instrumentation, communication and computational technologies, Noorul Islam University, December 16-17

Kalaiyarasi M, Perumal B, Pallikonda Rajasekaran M (2017) Estimation of deforestation rate for Amazonia using fuzzy k-means clustering. International Conference on Data Security (INCODS)

Kalaiyarasi M, Perumal B, Pallikonda Rajasekaran M (2020) Estimation of deforestation rate for LBA-ECO LC-14 modeled deforestation scenarios, Amazon Basin: 2002-2050 using fuzzy k-means clustering. Int J Recent Technol Eng 8(3):5746–5750

Kelly WJ, Margaret BH, Naughton-treves L, Morales M, Suarez L, Keenan K (2017) Forest conservation incentives and deforestation in the Ecuadorian Amazon. Thematic Sect: Forest Ecosyst Serv 44(1):56–65

Lei T, Jia X, Zhang Y, He L, Meng H, Nandi AK (2018) Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering. IEEE Trans Fuzzy Syst 26(5):3027–3041

Lei T, Jia X, Zhang Y, Liu S, Meng H, Nandi AK (2019) Superpixel-based fast fuzzy c-means clustering for color image segmentation. IEEE Trans Fuzzy Syst 27(9):1753–1766

Modava M, Akbarizadeh G (2016) Coastline extraction from SAR images using spatial fuzzy clustering and the active contour method. Int J Remote Sens 38(2):355–370

Pal SK, Ghosh A, Shankar BU (2000) Segmentation of remotely sensed images with fuzzy thresholding and quantitative evaluation. Int J Remote Sens 21(11):2269–2300

Pham VH (2015) An image segmentation approach for fruit detection using k-means clustering and graph-based algorithm. Vietnam J Computer Sci 2(1):25–33

Rathore VS, Kumar MS, Verma A (2012) Colour based image segmentation using L*A*B* colour space based on genetic algorithm. Int J Emerg Technol Adv Eng 2(6):156–162

Raval K, Shukla R, Ankit KS (2017) Color image segmentation using FCM clustering technique in RGB, L*a*b, HSV,YIQ Color spaces. Eur J Adv Eng Technol 4(3):194–200

Shen H, Yang J, Wang S, Liu X (2006) Attribute weighted Mercer kernel based fuzzy clustering algorithm for general non-spherical dataset. Soft Comput 10(11):1061–1073

Singh PP, Garg RD (2014) Classification of high resolution satellite images using spatial constraints-based fuzzy clustering. J Appl Remote Sens 8(1):083526. https://doi.org/10.1117/1.jrs.8.083526

Sangeeta Yadav, Mantosh Biswas (2017) Improved color-based K-means algorithm for clustering of the satellite image. IEEE Conference on Signal Processing and Integrated Networks (SPIN)

Zhang DQ, Chen SC (2003) Clustering incomplete data using kernel-based fuzzy c-means algorithm. Neural Process Lett 18(3):155–162