UDC 62-50
DOI: 10.20535/SRIT.2308-8893.2023.4.07

NOVEL MODIFIED KERNEL FUZZY C-MEANS ALGORITHM USED FOR COTTON LEAF SPOT DETECTION

PRADIP M. PAITHANE, SARITA JIBHAU WAGH

Abstract. Image segmentation is a significant and difficult subject that is a prerequisite for both basic image analysis and sophisticated picture interpretation. In image analysis, picture segmentation is crucial. Several different applications, including those related to medicine, facial identification, Cotton disease diagnosis, and map object detection, benefit from image segmentation. In order to segment images, the clustering approach is used. The two types of clustering algorithms are Crisp and Fuzzy. Crisp clustering is superior to fuzzy clustering. Fuzzy clustering uses the well-known FCM approach to enhance the results of picture segmentation. KFCM technique for image segmentation can be utilized to overcome FCM's shortcomings in noisy and nonlinear separable images. In the KFCM approach, the Gaussian kernel function transforms high-dimensional, nonlinearly separable data into linearly separable data before applying FCM to the data. KFCM is enhancing noisy picture segmentation results. KFCM increases the accuracy rate but ignores neighboring pixels. The Modified Kernel Fuzzy C-Means approach is employed to get over this problem. The NMKFCM approach enhances picture segmentation results by including neighboring pixel information into the objective function. This suggested technique is used to find "blackarm" spots on cotton leaves. A fungal leaf disease called "blackarm" leaf spot results in brown leaves with purple borders. The bacterium can harm cotton plants, causing angular leaf blotches that range in color from red to brown.

Keywords: Cluster Accuracy Rate (CAR), Clustering, Cotton Leaf Disease, Fuzzy Clustering Method (FCM), Kernel Fuzzy C-means Algorithm (KFCM), Novel Modified Kernel Fuzzy C-Means Clustering Algorithm (NMKFCM).

INTRODUCTION

Cotton is the most significant cash crop farmed in Maharashtra, India. The primary issue reducing cotton output is disease on the plant. Because a minute difference in color pattern might be caused by a different disease that is present on a cotton leaf, we know that the human eye's perception is not powerful enough to enable it to recognize minute variations in the diseased region of an image. The cotton plant's leaf is the disease's primary source. The leaves of the cotton plant are where 80–90% of the illness is located. One crucial technique for separating a picture into its backdrop and its objects is image segmentation. Clustering is one of the crucial phases in picture segmentation. In the early portion of the season, affected crops may develop slowly or be stunted. Blackening of the roots is a symptom of the illness, which results in the destruction of the root cortex (outer layer). Thielaviopsis basicola does not kill seedlings on its own, however some roots may perish. Significant black root rot exposes the root to Pythium or Rhizoctonia infection. When growth begins in warmer temperatures, the dead cells of the root cortex are shed, and plants that were severely harmed earlier in the season may not continue to exhibit symptoms later in the season.

- Diseases on Leaves of Cotton.
- The diseases on the cotton leaves are classified as:
  – Bacterial disease: e.g. Bacterial Blight, Crown Gall, Lint Degradation;
  – Fungal diseases: e.g. Anthracnose, Leaf Spot;
  – Viral disease: e.g. Leaf Curl, Leaf Crumple, Leaf Roll;
- Diseases like alternaria leaf spot, Bacterial blight, Bacterial stunt, Black root, Boll rot/tight lock.

The collection of observations is divided into smaller groups so that observations within each group are somewhat comparable to one another. Multivariate data analysis typically uses clustering as a routine practice. It is intended to investigate the data objects' innate natural structure, where items in the same cluster are as similar as possible to one another and objects in separate clusters are as distinct as possible from one another. The method used to arrange items or patterns so that samples from the same group resemble one another more than samples from other groups. There have been many different clustering techniques employed, including the hard clustering scheme and the fuzzy clustering scheme, each of which has unique particular traits. Each data point can only belong to one cluster when using the traditional hard clustering approach. As a result, when using this method, the segmentation results are frequently quite precise, meaning that every pixel in the image belongs to exactly one class. Yet, in many actual scenarios, problems with pictures like inadequate contrast, noise, overlapping intensities, and insufficient spatial resolution make this hard (crisp) segmentation a challenging process.

Types of Clustering:

- Hard: same object can only belong to single cluster.
- Soft: same object can belong to different clusters.

The current days, deep learning approach is used for cotton leaf segmentation. The CNN, VGG-16, VGG-19, ResNet-50 and some hybrid model has been used for this problem. The deep learning approach has been improved the accuracy of cotton leaf image segmentation as compared to state-of-art. The NMKFCM model is also gives stable result as compared to deep learning approaches. In deep learning approaches, training time period is major constraint for this problem. In the experimental analysis, the training time is near about 1 hour to 2 hour and in the proposed method the training process is not required.

MATERIAL AND METHODS

Fuzzy Clustering

In image segmentation, a soft segmentation technique has received extensive study and effective application. Since it has resilient qualities for ambiguity and can preserve significantly more information than hard segmentation methods, the Fuzzy C-Means (FCM) algorithm is the most widely used fuzzy clustering approach in picture segmentation. The typical FCM method has a severe flaw in that it lacks spatial context information, making it vulnerable to noise and imaging artefacts even while it performs well on the majority of noise-free pictures.

Fuzzy C-Means Algorithm. A well-liked and practical image division algorithm is FCM. The FCM algorithm was created by Dunn and enhanced by Bezdek [1]. This method is intended to scale back an objective goal [2]. Because each quality vector may only belong to one cluster and the quality vectors of the data set can be separated into solid clusters, this method outperforms the k-mean technique. Instead, the FCM loosens the restriction and enables the quality vector to assign a range of association scores to diverse clusters. Suppose a set of data with related clusters. A data value is equidistant from both clusters while also being near to them.

Activity in the clustering loop is FCM. By reducing the intragroup biased sum of the squared error task J_m function, it produces the best c partitions [3]:

$$J_m = \sum_{j=1}^{C} \sum_{i=2}^{N} U_{i,j}^{m} d_{i,j}^{2} ,$$

where N — the number of patterns in X; C — the number of clusters; U_{ij} — the degree of membership; W_j — the center of cluster j; d_{ij} — distance between object X_i and cluster center W_j; m — the biased value.

The FCM algorithm focuses on minimizing J_m, subject to the following constraints on U:

$$U_{ij} \in [0,1], \; i = 1,2,3,\dots,N, \; \text{and} \; j = 1,2,3,\dots,C;$$

$$\sum_{j=1}^{C} U_{ij} = 1, \quad i = 1,2,3,\dots,N, \quad 0 < \sum_{i=1}^{N} U_{ij} < 1, \quad j = 1,2,3,\dots,C.$$

Objective function J_m describe a constrained optimization problem, which can be converted to an unconstrained optimization problem by using Lagrange multiplier technique. By using this calculates membership function and update cluster center separately:

$$U_{ij} = \frac{1}{\sum_{i=1}^{c} \left(\frac{d_{ij}}{d_{il}} \right)^{\frac{2}{(m-1)}}}, \quad i = 1,2,\dots,N, \text{ and } j = 1,2,\dots,C;$$

$$\text{If } d_{ij} = 0 \text{ then } U_{ij} = 1 \text{ and } U_{ij} = 0 \text{ for } 1 \neq j .$$

And calculate cluster center using following step

$$w_j = \frac{\sum_{i=1}^{N} (U_{ij})^m x_i}{\sum_{i=1}^{N} (U_{ij})^m} , \quad j = 1,2,\dots,C .$$

The FCM algorithm focuses on minimizing objective function Jm. It fails in noisy image to detect accurate and sharp image segmentation process.

Kernel Fuzzy C-Means Algorithm. The FCM algorithm calculates the distance between the cluster center and the data item using Euclidian distance. FCM fails in noisy and nonlinear data sets because Euclidian distance does not perform as intended in noisy data. Kernel Fuzzy C-means technique is used to address this flaw. Kernel information is used with FCM in the KFCM approach [4]. KFCM works by mapping input data into a higher-dimensional feature space and utilizing the Kernel technique to transform nonlinearly separable data into linearly separable data. While using the kernel approach, the data set was complicated and nonlinear before becoming simple and separable when using the FCM method [5].

KFCM classifies noisy objects into clusters with greater clarity than FCM and with greater accuracy in noisy images. The value of the KFCM membership matrix U may range from 0 to 1

KFCM is iterative clustering methods that generate optimal c partition by using minimize objective function J_{kfcm}:

$$J_{km}(U,W) = 2\sum_{j=1}^{C}\sum_{i=1}^{N} U_{ij}^{m}(1-K(X_i,W_j)) .$$

In this objective function Gaussian kernel function is used:

$$K(X,Y) = \exp\left(-\frac{x-y^2}{\sigma^2}\right).$$

In KFCM clustering algorithm choose initial cluster randomly and perform following step.

1. Provide Gaussian kernel function for input image.
2. Evaluate membership function between object and cluster center.
3. Evaluate new updated cluster center.
4. Repeat step iteratively until no new cluster found.

KFCM it work properly in noisy image but KFCM not focus on neighborhood term.

Modified Kernel Fuzzy C-Means (MKFCM). This method is intended to scale back an objective goal [6]. Because each quality vector may only belong to one cluster and the quality vectors of the data set can be separated into solid clusters, this method outperforms the k-mean technique. Instead, the Fuzzy C-Mean loosens the restriction and enables the quality vector to assign a range of association scores to diverse clusters. Suppose a set of data with related clusters. A data value is equidistant from both clusters while also being near to them.

FCM is looping clustering activity. It generates optimum c partitions by abating the intragroup biased sum of the squared error task J_m function [7].

Kernel Method. The kernel methodology is a method that, by replacing the internal product with an appropriate Mercer Kernel, generates an implicit non-line map of the feedback information to a high-level quality space [8]. The kernel may be used in any method that solely depends on the dot product between two vectors. Every time a kernel is applied, a dot product is replaced. When two data are planned into a high-level-dimensional space, the space metrical that calculates the space between them is simplified. It is easier to tell apart and more distinctly differentiated [9].

Feature Space Mapping. Consider a non-line map task $\Phi : I = \mathbb{R}^2 \to F = \mathbb{R}^3$ from the 2-dimensional input space *I* into the 3-dimensional feature space *F*:

$$\varnothing(\vec{x}) = (x_1^2, \sqrt{2x_1x_2}, x_2^2)^T . \quad (1)$$

Hyperplane is represented by Eq. for separable dataset:

$$\vec{w}^T \vec{x} + b = 0 . \quad (2)$$

Consider the splitting hyperplane Eq. (1) into a linear task in $\mathbb{R}^3$:

$$\vec{w}^T \varnothing(\vec{x}) = w_1x_1^2 + w_2\sqrt{2x_1x_2} + w_3x_2^2 = 0 .$$

Eq. (2) is an elliptic job when as usual value of a constant *c* and assessed in $\mathbb{R}^2$.

In Fig. 1 any nonlinear separable data is converted into linear separable, so every pixel is classified on the basis of a feature.

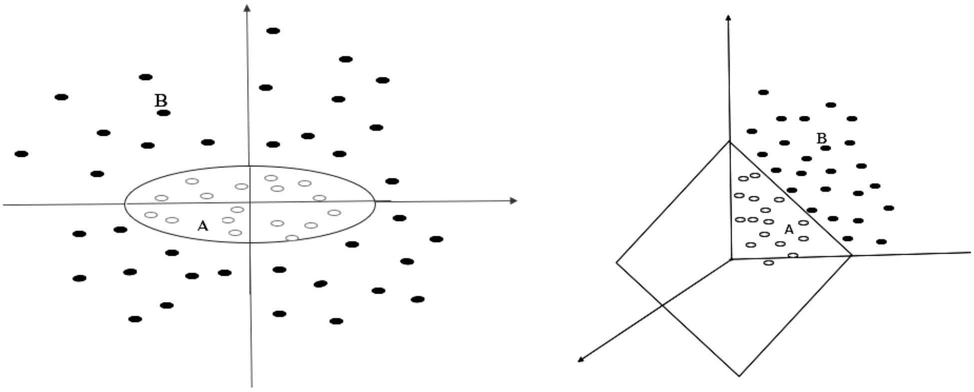


Fig. 1. Conversion of Non-line Distinguishable Data into Line Distinguishable Data

Use the appropriate mapping function to use *F*'s linear classifier with the converted form of the data to find a non-straight classifier without hassle. After mapping the non-line distinguishable data to a high-level space, *I*, locate a hyperplane that distinguishes linearly. For sensitive learning consider Fig. 1.

It depends only on the data mapped by the inner product of the feature space *F*. Defining a function $K(\vec{x}_i, \vec{x}) = \varnothing(\vec{x}_i)^T \varnothing(\vec{x})$, called kernel, that directly calculates the dot product of the mapping data places in the quality space eliminates the need for even the explicit coordinates of *F* or the mapping task [10]. The subsequent standard sample of a kernel "*K"* shows the computation of the dot product in the quality space applying $K(\vec{X}, \vec{Z}) = (\vec{X}^T, \vec{Z})^2$. It is encouraging the map task $\Phi(\vec{x}) = (x_1^2, \sqrt{2}x_1, x_2, x_2^2)^T$:

$$\vec{x} = (x_1, x_2), \quad \vec{z} = (z_1, z_2);$$

$$K(\vec{X}, \vec{Z})(\vec{X}^T, \vec{Z})^2 = (x_1z_1 + x_2z_2)^2 =$$

$$= (x_1^2z_1^2 + 2x_1z_1x_2z_2 + x_2^2z_2^2) = (x_1^2, \sqrt{2x_1x_2}, x_2^2)^T (z_1^2, \sqrt{2z_1z_2}, z_2^2) .$$

The advantage of such a kernel operation is that the complexness of the improvement of drawback continues solely reliant on the spatial property of the "*input space*" and not of the "*quality space*".

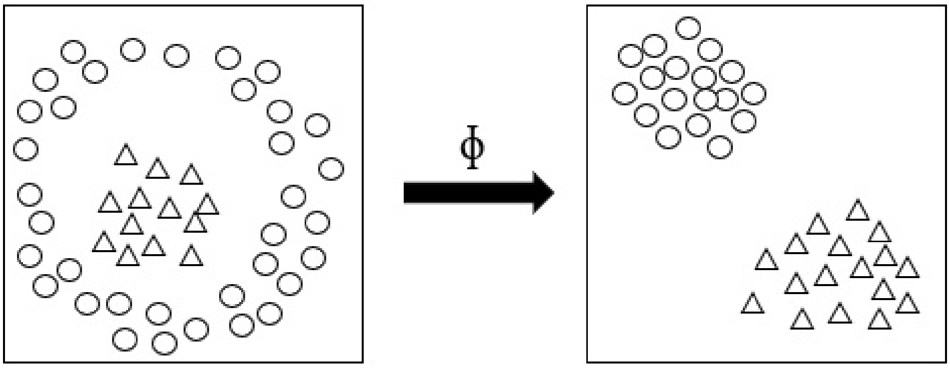
Different types of Kernels are mentioned below [11]:

Linear Kernel function: $K(x,z) = x^T z$;

Polynomial Kernel function: $K(x,z) = (x^T z + \theta)^d$;

Gaussian Kernel function: $K(x,z) = \exp\left(-\frac{x - z^2}{\sigma^2}\right)$;

Sigmoid Kernel function: $K(x,z) = \tanh(\alpha((x^T z) + \theta))$.



Input Space Kernel Feature Space

Fig. 2. Kernel Feature Space

Figure 2 is depicted the kernel working process for liner separable process to detect correct segmented regions. NMKFCM method is integrating closer pixel quantity in objective function [12]. NMKFCM method is a revised form of KFCM. KFCM is unsuitable for images damaged by instinct disturbance. KFCM has operated accurately in indistinct and nonlinear separable data, but it doesn't consist of information of closer pixel, to overcome this drawback, introduced NMKFCM is integrating closer pixel cost by applying "3×3" or "5×5" window window. A closer pixel quantity is included in objective task [13; 14]. Thec "α" constraint is applied to manage the impact of closer's term. It is having upper cost with growth of image disturbance. Scale of α cost rests within "0 to 1", if ratio of disturbance is minimal then take cost of α between "0 and 0.5". Ratio of disturbance is above average then take cost of α is "0.5 and 1.0". It is a beneficial and useful algorithm as compared to other algorithms. It has achieved sharp outcomes in disturbance images.

It is a looping procedure. It reduces the cost of objective tasks through closer pixel. In this objective task, present window across pixel and "α"parameter [15]:

$$J_{NMKFCM_{obj}}(\mathrm{U},\mathrm{W}) = \sum_{y=1}^{Q}\sum_{x=1}^{P} U_{xy}^{m}(1 - \mathrm{K}_T(\mathrm{Z}_x,\mathrm{W}_y))\left(\frac{N_R - \alpha \sum_{k\in N_i} U_{yk}}{N_R}\right), \quad (3)$$

where N_R — the cardinality; N_i — set of closer pixel value include into a window across pixel Z_i. Objective task j_{nmkn} illustrate a constrained optimization dilemma. Eq. 3 is applied for conversion into an unconstrained optimization dilemma. In Eq. 3, Lagrange multiplier technique is used.

By applying this computes membership function and update cluster center separately:

$$U_{xy} = \left(\frac{\left((1 - \mathrm{K}_T(\mathrm{Z}_x, \mathrm{W}_y)) \left(\frac{N_R - \alpha \sum_{l \varepsilon N_i} U_{yl}}{N_R} \right) \right)^{-\frac{1}{m-1}}}{\sum \left((1 - \mathrm{K}_T(\mathrm{Z}_x, \mathrm{W}_y)) \left(\frac{N_R - \alpha \sum_{l \varepsilon N_l} U_{kl}}{N_R} \right) \right)^{-\frac{1}{m-1}}} \right).$$

And calculate cluster center using following step:

$$W_y = \frac{\sum_{y=1}^{P} U_{xy}^m K_T(\mathrm{Z}_x, \mathrm{W}_y)\, \mathrm{Z}_x}{\sum_{x=1}^{Q} U_{xy}^m (\mathrm{Z}_x, \mathrm{W}_y)}.$$

Algorithm: Objective Function of NMKFCM

INPUT

1. $Z = \{Z_1, Z_2, \ldots, Z_N\}$, Data set
2. P, $2 \le P \le y$, y is number of cluster
3. Define cost of ε, used to terminate loop
4. Set membership function U_{xy}^0 using input data and cluster.
5. Determine cluster center $W_0 = (w_{01}, w_{02}, \ldots, w_{0p})$

OUTPUT

$W_j = \{W_0, W_2, \ldots, W_p\}$, targeted center of clusters.

begin
  for
    t=0
  if $\{U^t - U^{t+1}\} < \varepsilon$
      Update center W_p^t *with* U^t by using Eq.
      Update membership function U^{t+1} by using Eq.
      t+1
  else
    segmented output
  end

This method is advantageous to integrate closer pixel information. Standard FCM and IFCM methods are responsive to disturbance and preliminary cluster centers. It is ignoring the 3-D correlation of pixels, leading to inaccurate clustering outcomes [16]. NMKFCM work very fit in neighborhood pixel material.

Goal and Objectives:

- Choosing value of alpha to improve accuracy of image segmentation.

- Add Gaussian kernel method and RBF function to give more accuracy and also work in noisy and noiseless.
- Determine required number of clusters for image segmentation.
- Improving CAR value in all image formats.

Hills Climbing Algorithm

Image segmentation is a crucial step in the processing of images. Applications like image segmentation, adaptive compression, and region-based image retrieval benefit from the detection of conspicuous picture areas. Saliency is measured by comparing an image region's local contrast to its surrounding area at different scales [17]. It is using a contrast determination filter that runs at various scales to produce saliency maps with saliency values per pixel for the purpose of identifying salient locations. These separate maps come together to form the final saliency map [18]. We employ a rather straightforward segmentation approach to show how the final saliency map may be used to segment whole objects [19].

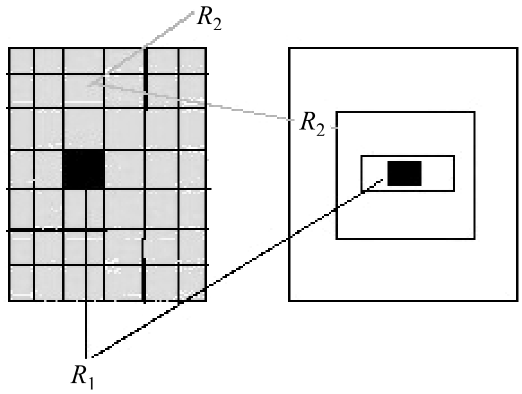


Fig. 3. Saliency map with R_1 inner and R_2 outer region

In this, compare the distance between the average feature vectors of the pixels in a subregion of the picture with the pixels in the area around it. Instead of merging separate saliency maps for scalar values of each feature, this enables the creation of a combined feature map at a particular scale utilizing feature vectors for each pixel [20]. The distance D between the average vectors of pixel characteristics of the inner area R_1 and that of the outer region R_2 is what determines the contrast-based saliency value $c(i, j)$ for a pixel at location $I(j)$ in the picture:

$$c_{i,j} = D\left[\left(\frac{1}{N_1}\sum_{p=1}^{N_1} v_p\right), \left(\frac{1}{N_2}\sum_{q=1}^{N_2} v_q\right)\right],$$

where v is the vector of feature elements corresponding to a pixel and N_1 and N_2 are the number of pixels in R_1 and R_2, respectively. If v is a vector containing uncorrelated feature items, then the distance D is a Euclidean distance; if the vector's elements are correlated, then the distance D is a Mahalanobis distance. In this study, feature vectors for color and brightness are generated using the CIELab color space and RGB photographs. Although the CIELab colour space's perceptual differences are roughly Euclidian, D in equation [13]:

$$c_{i,j} = v_1 - v_2\,,$$

where $v_1 = [L_1; a_1; b_1]T$ and $v_2 = [L_2; a_2; b_2]T$ are the average vectors for regions R_1 and R_2, respectively.

The final saliency map is determined as a sum of saliency values across the scales S:

$$m_{i,j} = \sum_{s} c_{i,j} .$$

The hill-climbing technique may be thought of as a search window that is ran through the d-dimensional histogram's space to locate the biggest bin inside of it. Each bin in the colour histogram has $3d - 1 = 26$ neighbors because the CIELab feature space is three-dimensional, where d is the number of dimensions in the feature space. The values of these bins serve as the starting seeds, and the number of peaks obtained reveals the value of K [21]. By adding up values in the final saliency map M that correspond to pixels in the segmented picture, the average saliency value V per segmented region is determined:

$$V_k = \frac{1}{|r_k|} \sum_{i,j \in r_k} m_{i,j} ,$$

$|r_k|$ is the segmented region's size in pixels. The segments with an average saliency value greater than a predetermined threshold T are maintained, while the other segments are removed, according to a straightforward threshold-based procedure. As a consequence, the output only includes the segments that make up the salient item.

The $L*a*b*$ color space enables us to quantify these differences. The $L*a*b*$ color space is derived from the CIE *XYZ* tristimulus values. The $L*a*b*$ space comprises of a luminosity layer '$L*$', chromaticity-layer '$a*$' indicating where color falls along the red-green axis, and chromaticity-layer '$b*$' indicating where the color falls along the blue-yellow axis [22].

Algorithm: Hill-climbing Based Segmentation.

Input: An Image.

Output: a group of aesthetically connected segments.

1. Create the image's color histogram.
2. Ascend the color histogram's slope from a non-zero bin to the apex as shown below:

2.1. The amount of pixels in the current histogram bin should be compared to the numbers in the adjacent (left and right) bins.

2.2. The algorithm moves upwards towards the neighboring bin with the greater number of pixels if the surrounding bins have differing amounts of pixels.

2.3. The algorithm checks the next nearby bins if the immediate neighbors have the same amount of pixels, and so on, until two neighboring bins with different numbers of pixels are discovered. Next, a shift upward is performed to the bin with the most pixels.

2.4. Repeat steps 2.1–2.3 to continue going upwards until you reach a point from which you can travel no further uphill. When the adjacent bins contain less pixels than the current bin, that is the situation. As a result, the present bin is considered a high.

Choose a different unclimbed bin to use as your starting bin, then follow step 2 to locate another summit. This process is repeated until the color histogram's non-zero bins are all climbed (associated with a peak). The discovered peaks are preserved since they indicate the input image's original number of clusters.

2.5. The halting bin is designated as the peak of a hill if no upward progress is made, and all bins going to this peak are connected to it.

3. Choose a different unclimbed bin to use as your starting bin, then follow step 2 to locate another summit. This process is repeated until the histogram's non-zero bins are all climbed (associated with a peak).

4. The recognized peaks are preserved because they show how many clusters there were in the input picture at the beginning.

5. The same peak's neighboring pixels are clustered together.

Lastly, pixels that are close to one another and lead to the same peak are grouped together, assigning each pixel to a different peak. Hence, create the input image's clusters.

EXPERIMENTAL RESULT

Evaluation Parameter

1. Cluster Accuracy Rate

$$CAR = \left|\frac{A \cap S}{A \cup S}\right|.$$

2. Dice

$$dice(A, S) = 2 * \frac{|A \cap S|}{\||A| + |S|\|}.$$

3. IOU

$$IOU = \frac{|A \cap S|}{|A \cup S|}.$$

4. Bfscore

$$bfscore = \frac{2 * precision * recall}{(recall + precision)},$$

where $A =$ output image; $S =$ input image.

Detail comparison of proposed method with traditional method (see Table 1–3)

Table 1. Detail Comparison of Proposed Method with Traditional Method

Image Name	Approach	Evaluation Parameter		
		IOU	bfscore	dice
Image 1	FCM	55.78	33.51	71.62
	KFCM	73.81	26.36	84.93
	NMKFCM	81.55	41.25	89.83
Image 2	FCM	68.61	21.28	81.38
	KFCM	84.47	36.81	91.58
	NMKFCM	89.88	42.29	98.81

Continued Table 1

Image Name	Approach	Evaluation Parameter		
		IOU	bfscore	dice
Image 3	FCM	90.22	37.14	94.86
	KFCM	80.66	34.71	89.29
	NMKFCM	90.75	36.98	95.15
Image 4	FCM	72.94	18.11	84.35
	KFCM	90.06	40.33	94.77
	NMKFCM	94.51	54.46	97.18
Image 5	FCM	74.95	26.51	85.68
	KFCM	73.17	25.37	84.51
	NMKFCM	80.36	27.42	87.99

Table 2. Detail Comparison of Proposed Method with Traditional Method

Image Name	Approach	Cluster Accuracy Rate(CAR)
Image 1	FCM	63.71
	KFCM	71.78
	NMKFCM	74.98
Image 2	FCM	57.9021
	KFCM	64.8723
	NMKFCM	69.9572
Image 3	FCM	67.2396
	KFCM	64.3482
	NMKFCM	70.2246
Image 4	FCM	58.6222
	KFCM	66.3623
	NMKFCM	69.184
Image 5	FCM	86.8277
	KFCM	86.0188
	NMKFCM	95.337

Table 3. Detail comparison of proposed method with traditional method

Image Name	Approach	Time Period
Image 1	FCM	12.24
	KFCM	10.44
	NMKFCM	8.47
Image 2	FCM	14.24
	KFCM	08.37
	NMKFCM	06.54
Image 3	FCM	11.61
	KFCM	11.29
	NMKFCM	09.24
Image 4	FCM	13.19
	KFCM	12.66
	NMKFCM	07.29
Image 5	FCM	12.24
	KFCM	08.59
	NMKFCM	07.76

The Table 4 depicts the detail comparison of NMKFCM approach with deep learning approaches. The NMKFCM is providing strong and stable result as compare to CNN model. The CAR value of NMKFCM approach is 98.80 which higher than other approaches. The IOU and Precision value of NMKFCM achieved higher result value as compared to deep learning models. The NMKFCM is having less value The Dice value of NMKFCM improvised the result as compare to state-of-art. in bfscore as compare to other approaches.

T a b l e 4 . Detail comparison of proposed method with traditional method

Approach	CAR	IOU	Precision	Dice	bfscore	Time	Training Time
CNN [23]	95.37	92.00	0.8750	46.0	87.50	8~9 second	94 Minute
VGG16 [23]	98.10	92.18	0.9583	46.0	95.16	5~6 second	54 Minute
ResNet-50 [23]	98.32	91.49	0.9482	50.0	95.65	2~3 second	53 Minute
Menon Model[23]	98.53	94.23	0.9579	50.0	96.42	5~6 second	77 Minute
NMKFCM	98.80	94.51	1.0000	98.81	54.46	4~6 second	Not Required

The execution time of NMKFCM is less than CNN, VGG-16 and Menon Model, but higher than ResNet-50. The training time is not required for NMKFCM (Fig. 4).

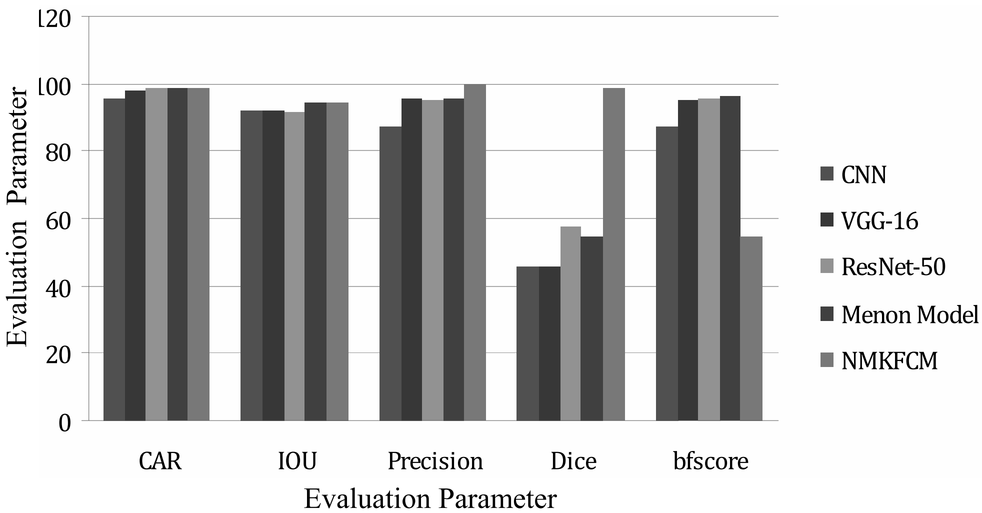


Fig. 4. Comparison of NMKFCM with Deep Learning Approaches

In above image, sub image (A), (E), (I), (M) and (Q) are original image of cotton leaf. Sub image (B), (F), (J), (N) and (R) are segmented by FCM approach, Sub image (C), (G), (K), (O) and (S) are segmented by KFCM approach, Sub image (D), (H), (L), (P) and (T) are segmented by NMKFCM approach. The sub image (A) is affected by Bacterial Blight disease, The sub image (E) is affected by Leaf Curl, The sub image (I) is affected by alternaria leaf spot , The sub image (M) is affected by fungal disease.

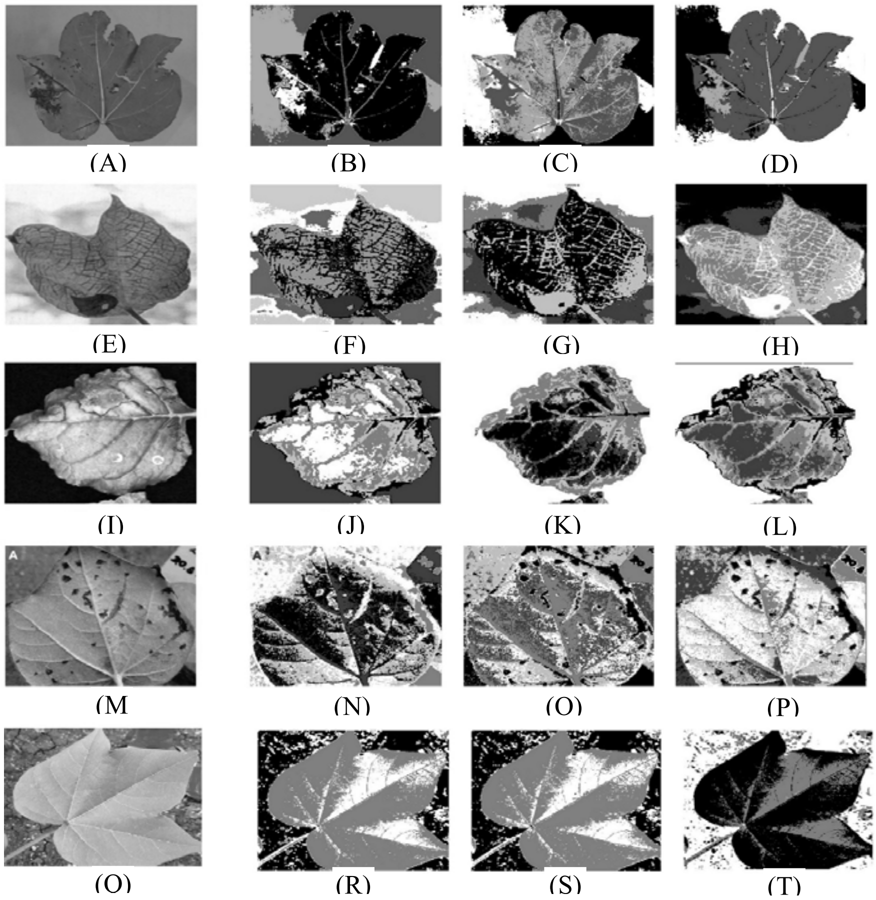


Fig. 5. Cotton Leaf Image Segmentation using FCM, KFCM and NMKFCM

CONCLUSION

The crucial and indispensable element of an image analysis system, image segmentation is a key area of study for many image processing researchers. Four methods — Clustering, Thresholding, Region Extraction, and Edge Detection — are used to segment images. Clustering is the downgrouping of related data elements. Here, we've used techniques for clustering like crisp and fuzzy. In this system, Fuzzy C-Means, Kernel Fuzzy C-Means, and Modified Kernel Fuzzy C-Mean Clustering are all used as clustering techniques. In comparison to FCM and Crips clustering methods, MKFCM is a suggested system that provides accurate picture segmentation while also enhancing segmentation performance by adding the influence of neighbor pixel information. The MKFCM method can automatically identify the necessary cluster number for picture segmentation with the use of the Hill climbing algorithm. The suggested technique can automatically estimate the cluster number for a noisy picture, but this number is not helpful for image segmentation since the proposed algorithm has formed a cluster for noisy pixels, making image segmentation less effective than for noiseless pixels. In the future, we will be able to select an alpha value to increase the precision of picture segmentation and CAR (Cluster Accuracy Rate) values across all image formats.

The proposed method is not required training time but in deep learning approaches training is mandatory. The proposed method is improvising the IOU, precision, Dice and CAR value as compared to deep learning approaches.

No conflict of interest.

REFERENCES

1. P.M. Paithane and S.N. Kakarwal, "Automatic determination number of cluster for multi kernel NMKFCM algorithm on image segmentation," in *International Conference on Intelligent Systems Design and Applications*, pp. 870–879. Springer, Cham, 2018.
2. Chun-Yan Yu, Ying Li, Ai-lian Liu, and Jing-hong Liu, "A novel modified kernel fuzzy c- means Clustering algorithms on Image segmentation," *2011 14th IEEE International Conference*. doi: 10.1109/CSE.2011.109.
3. Saiful Islam and Dr. Majidul Ahmed, "Implementation of Image Segmentation for Natural Images using Clustering Methods," *IJETAE*, vol. 3, issue 3, March 2013.
4. L.A. Zadeh, "Fuzzy Sets", *Information and Control*, 8, pp. 338–353, 1965.
5. Songcan Chan and Daoqiang Zhang, "Robust Image Segmentation Using FCM With Spatial Constraints Based on New Kernel Induced Distance Measure," *IEEE transactions on Systems, MAN and Cybernetics-Part B: Cybernetics*, vol. 34, no. 4, August 2004.
6. P.M. Paithane, S.N. Kakarwal and D.V. Kurmude, "Automatic Seeded Region Growing with Level Set Technique Used for Segmentation of Pancreas," in *International Conference on Soft Computing and Pattern Recognition*, pp. 374–382. Springer, Cham, 2020.
7. P.M. Paithane and S.N. Kakarwal, "Automatic Pancreas Segmentation using A Novel Modified Semantic Deep Learning Bottom-Up Approach," *International Journal of Intelligent Systems and Applications in Engineering*, 10(1), pp. 98–104, 2022.
8. Pradip Mukundrao Paithane, "Yoga Posture Detection Using Machine Learning," *Artificial Intelligence in Information and Communication Technologies, Healthcare and Education: A Roadmap Ahead*, 2022.
9. S. Kakarwal and Pradip Paithane, "Automatic pancreas segmentation using ResNet-18 deep learning approach," *System Research and Information Technologies*, no. 2, pp.104–116, 2022.
10. Elnomery Zanaty and Sultan Aljahdali,"Improving Fuzzy Algorithms for Automatic Magnetic Resonance Image Segmentation," *The International Arab Journal of Information Technology*, vol. 7, no. 3, pp. 271–279, July 2009.
11. Kaur Prabhjot, Gupta Pallavi, and Sharma Poonam, "Review and Comparison of kernel Based Fuzzy Image Segmentation Techniques," *I.J. Intelligent Systems and Applications*, 7, pp. 50–60, 2012.
12. Robert L. Cannon, Jintendra V. Dave, and James C. Bezdek, "Efficient Implementation of the Fuzzy c-Means Clustering Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, issue 2, March 1986.
13. Martin Hofman, *Support vector Machines-Kernel and the Kernel Trick,* pp. 1–16, 2006.
14. Daoqiang Zang and Songcang Chen, "Fuzzy Clustering Using Kernel Method," *Proceedings of the 2002 International Conference on Control and Automation,* Xiemen, China, June 2002. doi: 10.1109/ICCA.2002.1229535.
15. Daoqiang Zang and Songcang Chen, "A novel kernalized fuzzy C-means algorithm with Application in medical image segmentation," *Artificial Intelligence in Medicine*, 32, pp. 37–50, 2004.
16. E.A. Zanaty, Sultan Aljahdli, and Narayan Debnath, "A Kernalized Fuzzy C-Means Algorithm for Automatic Magnetic Resonance Image Segmentation," *Journal of Computational Methods in Sciences and Engineering Archive*, vol. 9(1,2S2), pp. 123–136, April 2009. doi: 10.3233/JCM-2009-0241.
17. Shailash Kochra and Sanjay Joshi, "Study on Hill-Climbing Algorithm for Image Segmentation Technology," *International Journal of Engineering Research and Applications (IJERA*), vol. 2, issue 3, pp. 2171–2174, May-Jun 2012.
18. Garima Goyal, "TEM Color Image Segmentation using Hill Climbing Algorithm," *International Journal of Computer Science and Information Technologies*, vol. 5, pp. 3457–3459, 2014.
19. A. Abirami Shri, E. Aruna, and Ajanthaa Lakkshmanan,"Image segmentation and recognition," *International Journal of Computer Applications; 3rd National Conference on Future Computing, February 2014.*

20. Pradip M. Paithane, S.N. Kakarwal, and D.V. Kurmude, “Top-down method used for pancreas segmentation,” *Int. J. Innov. Exploring Eng. (IJITEE)*, vol. 9, issue 3, pp. 1790–1793, 2020.
21. Sarita Jibhau Wagh, Pradip M. Paithane, and S.N. Patil, “Applications of Fuzzy Logic in Assessment of Groundwater Quality Index from Jafrabad Taluka of Marathawada Region of Maharashtra State: A GIS Based Approach,” *Hybrid Intelligent Systems: 21st International Conference on Hybrid Intelligent Systems (HIS 2021), December 14–16, 2021*. Cham: Springer International Publishing, 2022.
22. Pradip M. Paithane and Sarita Jibhau Wagh, “Automatic Quality Control Scrutiny of Sugar Crystal using K-Means Clustering Algorithm Image Processing,” *American Scientific Research Journal for Engineering, Technology*, and Sciences, 9(12):2395-0056, 2022.
23. M.S. Memon, P. Kumar, and R. Iqbal, “Meta Deep Learn Leaf Disease Identification Model for Cotton Crop,” *Computers*, 11(7), 102, 2022. Available: https://doi.org/10.3390/computers11070102
24. Pradip Paithane, Sarita Jibhau Wagh, and Sangeeta Kakarwal, “Optimization of route distance using k-NN algorithm for on-demand food delivery,” *System Research and Information Technologies*, no. 1, pp. 85–101, 2023. doi: 10.20535/SRIT.2308-8893.2023.1.07.

Received 01.04.2023

INFORMATION ON THE ARTICLE

Pradip M. Paithane, ORCID: 0000-0002-4473-7544, Vidya Pratishthan’s Kamalnayan Bajaj Institute of Engineering and Technology, India, e-mail: paithanepradip@gmail.com

Sarita Jibhau Wagh, ORCID: 0000-0003-4798-2147, T.C. College Baramati, India

НОВИЙ МОДИФІКОВАНИЙ АЛГОРИТМ ЯДРА FUZZY C-MEANS, ЩО ВИКОРИСТОВУЄТЬСЯ ДЛЯ ВИЯВЛЕННЯ ПЛЯМ НА ЛИСТУ БАВОВНИКА / Прадіп М. Пейтане, Саріта Джібхау Ваг

Анотація. Сегментація зображення є важливою та складною темою, яка є необхідною умовою як для базового аналізу зображення, так і для складної інтерпретації зображення. В аналізі зображень сегментація зображення має вирішальне значення. Кілька різних програм, зокрема ті, що стосуються медицини, ідентифікації обличчя, діагностики хвороби Коттона та виявлення об’єктів на карті, отримують переваги від сегментації зображення. Для сегментації зображень використовується підхід кластеризації. Існує два типи алгоритмів кластеризації: чіткий і нечіткий. Техніка чіткості перевершує нечітку кластеризацію. Нечітка кластеризація використовує добре відомий підхід FCM для поліпшення результатів сегментації зображення. Техніка KFCM для сегментації зображення може бути використана для усунення недоліків FCM у зашумлених і нелінійних роздільних зображеннях. У підході KFCM ядрова функція Гауса використовується для перетворення високовимірних нелінійно розділених даних у лінійно розділені дані перед застосуванням FCM до даних. KFCM поліпшує результати сегментації зображення із шумом, підвищує рівень точності, але ігнорує сусідні пікселі. Щоб подолати цю проблему, використовується модифікований підхід нечіткого C-середнього ядра. Підхід NMKFCM поліпшує результати сегментації зображення шляхом включення інформації про сусідні пікселі до цільової функції. Цей запропонований метод використовується для виявлення плям «чорної шкірки» на листу бавовника. Грибкове захворювання листя під назвою «чорна плямистість» призводить до коричневого листя з фіолетовими краями. Бактерія може завдати шкоди рослинам бавовника, спричиняючи кутасті плями на листу, які мають колір від червоного до коричневого.

Ключові слова: коефіцієнт точності кластера (CAR), кластеризація, хвороба листя бавовника, метод нечіткої кластеризації (FCM), алгоритм нечіткого C-середнього ядра (KFCM), новий модифікований алгоритм кластеризації нечіткого C-середнього ядра (NMKFCM).