RESEARCH ARTICLE

Study the Effect of MRF Model on Fuzzy c Means Classifiers with Different Parameters and Distance Measures

Shilpa Suman[1] · Dheeraj Kumar[1] · Anil Kumar[2]

Abstract

Image classification used to organize pixels of an image into specific land-cover classes. There are various algorithms for classifying images. Fuzzy c-Means (FCM) is one of the essential fuzzy classifiers, but it does not consider the neighborhood pixel information. In this paper, the MRF model is applied in the FCM classifier to enhance the output of image classification. FCM classifier with MRF Model (DA(H1), DA(H2), DA(H3), DA(H4), and SP) has been studied with different distance measures and parameters to identify which parameter combination provides the best result. This best algorithm was identified by classifying the Dense Forest, Eucalyptus, Riverine sand, Water, Wheat, and Grassland classes applying the multispectral image of Landsat-8 and Formosat-2 of the Haridwar area. These classified image's Overall Accuracy (OA) are calculated using the FERM (Fuzzy Error Matrix) technique. It was observed that Mean Absolute Difference distance measures at $m = 1.1$, $\lambda = 0.9$, and $\gamma = 0.5$ for DA (H1) using the base FCM classifier yield the highest overall accuracy (82.06%) irrespective of the classifier.

Keywords MRF · Discontinuity adaptive · Smoothing prior · FCM · Distance measure · Fuzziness factor

Introduction

Image classification discovers the characteristic of an object concerning the pixels in an image. Image classification accuracy mainly depends on the training data set, analysis problem, and classification algorithms (Singha et al., 2015). The classification algorithm utilizes spectral reflectance information for locating the pixel to the corresponding class (Lillesand et al., 2014). Conventional classification assumes that pixels are found in pure form means one pixel represents only one class. But, in a real scenario, due to the restricted spatial resolution of the sensor and dissimilarity between pixels and class boundary, one pixel may represent more than one class known as mixed pixels. Presence of mixed pixels in the data yields less accuracy in classification (Singh & Garg, 2017). A fuzzy based method is utilized in order to reduce crispness

in classification (Singh & Garg, 2014). J. Zhang and Foody (2001) proposed adapting fuzziness in classification techniques to classify the mixed pixels, which comes under the advanced classification technique as fuzzy classification. In this technique, membership of individual pixel may level with partial and multiple classes. Bezdek et al. (1984) introduced a fuzzy-based clustering algorithm. The purpose of this algorithm was to classify the mixed pixel by allocating membership values to individual classes. But, Fuzzy c means (FCM) (Bezdek et al., 1984) fails to provide spatial contextual information. The spatial contextual information refers to the relationship between neighboring pixels. The spatial context of a pixel is the occurrence of respective class labels at neighboring pixels (Schistad Solberg et al., 1996).

In order to overcome the problem associated with FCM, Markov random field (MRF) models were introduced, which is the advanced way to consider spatial contextual information (Geman & Geman, 1984). MRF approach uses the spatial interaction between pixels in the satellite images. MRF, Smoothness Priors (SP) Model contains the contextual information property and considers smoothness is everywhere in the image (Li, 1995), but discontinuity

✉ Anil Kumar
  anil@iirs.gov.in

1 Indian Institute of Technology (Indian School of Mines), Dhanbad, Jharkhand, India

2 Indian Institute of Remote Sensing, Dehradun 248001, India

occurs at the edges or boundaries. Discontinuity Adaptive (DA) (Li, 1995)models were introduced to overcome the over smoothing. The developed MRF models were used for specific purposes; the DA is mainly used for edge enhancement, whereas the SP model does the smoothing applied to remove the noise. There are several applications where these models were used distinctly to extract the required information. MRF models concentrated on edge preservation of class boundaries and smoothing the classified output. Earlier various studies have been conducted related to MRF for improving image classification, such as Baysian Image classification (Berthod et al., 1996), adaptive Bayesian contextual classification (Jackson et al., 2002), image Fusion (Xu et al., 2011). Still, different parameters and distance measures have not been considered in previous studies. This paper introduces Discontinuity adaptive (DA) MRF models and Smoothing prior (SP) MRF models in FCM as a base classifier to improve image classification concerning different parameters and distance measures. The FCM MRF model is used in classification application in different fields such as medical, urban planning, and geohazards(Ahmadvand & Daliri, 2015; Gong et al., 2014; Hao et al., 2013).

This study aimed to show the effect of the MRF model within an FCM classifier concerning different parameters (lambda (λ), beta (β), gamma (γ), and fuzziness factor(m)) and distance measures to obtain the best algorithm. The fuzzy Error Matrix (FERM) proposed by (Binaghi et al., 1999) has been used to calculate the Overall Accuracy of the classified output. Here, Parameter λ varies 0.2–0.9 with an interval of 0.1, β varies 1–9 with a period of 1, γ varies 0.1–0.9 with an interval of 0.1, and m is equal to 1.1–3 with an interval of 0.2. Bray Curtis, Canberra, Chessboard, Correlation, Cosine, Euclidean, Manhattan, Mean Absolute Difference, Median Absolute Difference, and Normalized Square Euclidean are used as distance measures which are described in Sect. 3.

Formulation of FCM and MRF model

Fuzzy c – means (FCM)

Fuzzy c-means (FCM) (Bezdek et al., 1984)is an essential clustering method that works on the principle of fuzzy set theory (Ichoku & Karnieli, 1996). The method allows the pixel to contain partial membership of one or more than one class (Kaymak & Setnes, 2000). Its objective function is described in Eq. (1)

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m d_{ij}^2, 1 < m \leq \infty \tag{1}$$

In this algorithm, the following condition should satisfy.

$$\sum_{j=1}^{C} u_{ij} = 1, u_{ij} \in [0,1], 0 \leq u_{ij} \leq 1$$

$$\sum_{i=1}^{N} \mu_{ij} > 0$$

Here, m = Fuzziness Factor (it contain any real value greater than 1),u_{ij}= Degree of membership represents of ith pixel for cluster j,d_{ij}= distance measures, N = total no of data, and C = Number of classes.

Smoothing Prior (SP)

Smoothing Prior of MRF gives contextual information while using information about the pixel and its neighboring pixel. The smoothness hypothesis was introduced mathematically by prior further probability, described as energy (Li, 1995). The general form of the regularizers is defined in Eq. (2), and smoothing prior standard regularizers are employed in Eq. (3).

$$U(F) = \sum_{n=1}^{N} U_n(f) = \sum_{n=1}^{N} \lambda_n \int_{n}^{b} g(f^n(x))dx \tag{2}$$

U(f) = prior energy, which represent the *nth* order regularizer,λ_n=weighting factor, where($\lambda_n > = 0$)and $g(f^n(x))$ = potential function.

$$g(f^n(x) = (n) = \eta^2 \tag{3}$$

Discontinuity Adaptive (DA)

Discontinuity Adaptive (DA) is a robust algorithm of MRF, which preserves boundaries and edges. The necessary condition for a regularizer to be discontinuity adaptive is given as Eq. (4)(Smits & Dellepiane, 1997).

$$\lim_{n \to \infty} |g'(\eta)| = \lim_{n \to \infty} |2\eta h(\eta)| = c \tag{4}$$

c = constant (cε[0,∞]).
Four DA model described as in equation (5-8).

$$g_{1\gamma}(\eta) = -\gamma e^{\frac{\eta^2}{\gamma}} \tag{5}$$

$$g_{2\gamma}(\eta) = \frac{-\gamma}{1 + \frac{\eta^2}{\gamma}} \tag{6}$$

$$g_{3\gamma}(\eta) = \gamma \ln \left(1 + \frac{\eta^2}{\gamma}\right) \qquad (7)$$

$$g_{4\gamma}(\eta) = \gamma|\eta| - \gamma^2 \ln \left(1 + \frac{|\eta|}{\gamma}\right) \qquad (8)$$

Mathematical Formula of Similarity and Dissimilarity Measures

This study was focused to show the effect of different distance measures on FCM based MRF model; considering this, eight dissimilarity measures such as Braycurtis, Canberra, Chessboard, Euclidean, Manhattan, Mean Absolute Difference, Median Absolute Difference, Normalized Square Euclidean, and two similarity measures Cosine and Correlation have been used. Various similarities and dissimilarities measures studied in FCM classifier as distance criteria to be generated to identify unknown vectors belong to which class. The frequently used distance measures in different applications were selected for analysis in this study. The different distance measures were used to validate and assess the models and examine how different distance measures affect the FCM based MRF model parameters. Mathematical expressions of all dissimilarity and similarity measures are given below. Here x and v represent the vector pixel, c is the mean value, and b represents the no of bands.

Dissimilarity Meaures

Braycurtis

Braycurtis (Bray & Curtis, 1957) dissimilarity measures are frequently applied to calculate the relationship between environmental sciences, ecology, and related field. The equation of Braycurtis is given in Eq. (9);

$$D(x,y) = \frac{\sum_{i=1}^{n} |x_i - y_i|}{\sum_{i=1}^{n} |x_i + y_i|} \qquad (9)$$

Here y is the sample, and n represents the no of data points.

Canberra

Canberra (Agarwal et al., 2009) was introduced in 1966. This measure is mainly applied for positive values. This measure has been used for comparing ranked lists. The equation for Canberra distance measures is given in Eq. (10);

$$D(x_j, v_i) = \frac{|x_{j1} - v_{i1}|}{|x_{j1}| + |v_{i1}|} + \frac{|x_{j2} - v_{i2}|}{|x_{j2}| + |v_{i2}|} + \cdots + \frac{|x_{jb} - v_{ib}|}{|x_{jb}| + |v_{ib}|} \qquad (10)$$

Chessboard

Chessboard (Baccour and John 2015) distance measures represent a vector space to calculate the distance between any co-ordinate dimensions, which is the greatest of the distances along two vectors. It is also known as Chebyshev distance. The equation for chessboard distance is given in Eq. (11);

$$D(x_j, v_i) = \text{Max}\left[|x_{j1} - v_{i1}|, |x_{j2} - v_{i2}|, \ldots, |x_{jb} - v_{ib}|\right] \qquad (11)$$

Euclidean

Euclidean (Hasnat et al., 2013) distance is the distance between two points in the Euclidean space. It calculates the square root of the sum of the squares of the difference among parallel data points values. The equation of Euclidean distance measures is given in Eq. (12);

$$D(x,y) = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} \qquad (12)$$

Manhattan

Manhattan (Hasnat et al., 2013) distance measure is used to compare images. Manhattan distance among two data points is the sum of the difference of their parallel element. Manhattan distance measures equation is given in Eq. (13);

$$D(x_j, v_i) = |x_{j1} - v_{i1}| + |x_{j2} - v_{i2}| + \cdots + |x_{jb} - v_{ib}| \qquad (13)$$

Mean Absolute Difference

Mean Absolute Difference (Vassiliadis et al., 1998) is a statistical measurement of depression. It is defined as the addition of the absolute difference and the variable of two items with a similar location of the same place and divided by the entire number of bands. Mean Absolute Difference distance measures equation is given in Eq. (14);

$$D(x_j, v_i) = \frac{1}{b}\left[|x_{j1} - v_{i1}| + |x_{j2} - v_{i2}| + \cdots + |x_{jb} - v_{ib}|\right] \qquad (14)$$

Median Absolute Difference

Median Absolute Difference (MAD) (Scollar et al., 1984) may use in place of Mean Absolute Difference to reduce the effect of impulse noise on the calculated measures. MAD is mathematically defined as calculating the difference between the absolute intensities of the similar pixels of two images, and after that, we take the median of data. MAD distance measures equation is given in Eq. (15);

$$D(x_j, v_i) = \text{Median}\big[|x_{j1} - v_{i1}|, |x_{j2} - v_{i2}|, \ldots, |x_{jb} - v_{ib}|\big] \tag{15}$$

Normalized Square Euclidean

Normalized Square Euclidean (NSE) (Hasnat et al., 2013) obtains the NSE distance between two vectors. It requires normalization of the pixel's intensities before computing the addition of squared difference between the pixels of two images. NSE distance measures equation is given in Eq. (16);

Similarity Measures

Cosine

Cosine (Senoussaoui et al. 2014) similarity measures determine the angle's cosine along two vectors contained in an inner product space. It gives the measurement of two vectors with respect to each other. The mathematical equation of cosine is given in Eq. (17);

$$D(x_j, v_i) = 1 - \frac{x_{j1}v_{i1} + x_{j2}v_{i2} + \ldots + x_{jb}v_{ib}}{\sqrt{|x_{j1}|^2 + \cdots + |x_{jb}|^2}\sqrt{|v_{i1}|^2 + \cdots + |v_{ib}|^2}} \tag{17}$$

Correlation

Correlation (M. Zhang et al., 2008) similarity is a calculation of obtaining the correlation along two vectors. The similarity along the two vectors is obtained by applying the Pearson-r correlation. The mathematical formula of cosine is given in Eq. (18);

$$D(x_j, v_i) = \frac{\left|\left\{x_{j1} + \frac{1}{b}\left(-x_{j1} - x_{j2} - \cdots - x_{jb}\right) - v_{i1} + \frac{1}{b}(v_{i1} + v_{i2} + \cdots + v_{ib})\right\}^2\right| + \cdots + \left|\left\{x_{jb} + \frac{1}{b}\left(-x_{j1} - x_{j2} - \cdots - x_{jb}\right) - v_{ib} + \frac{1}{b}(v_{i1} + v_{i2} + \cdots + v_{ib})\right\}^2\right|}{2\left[\left[\left|\left\{x_{j1} + \frac{1}{b}\left(-x_{j1} - x_{j2} - \cdots - x_{jb}\right)\right\}^2\right| + \cdots + \left|\left\{x_{jb} + \frac{1}{b}\left(-x_{j1} - x_{j2} - \cdots - x_{jb}\right)\right\}^2\right|\right] + \left|\left\{v_{i1} + \frac{1}{b}\left(-v_{i1} - v_{i2} - \cdots - v_{ib}\right)\right\}^2\right| + \cdots + \left\{v_{ib} + \frac{1}{b}\left(-v_{i1} - v_{i2} - \cdots - v_{ib}\right)\right\}^2\right]} \tag{16}$$

$$D(x_j, v_i) = 1 - \frac{\left[\left\{x_{j1} + \frac{1}{b}\left(-x_{j1} - x_{j2} - \cdots - x_{jb}\right)\right\}\left\{v_{i1} + \frac{1}{b}\left(-v_{i1} - v_{i2} \cdots - v_{ib}\right\} + \cdots + \left\{x_{jb} + \frac{1}{b}\left(-x_{j1} - x_{j2} - \cdots - x_{jb}\right)\right\}\left\{v_{ib} + \frac{1}{b}\left(-v_{i1} - v_{i2} \cdots - v_{ib}\right\}\right]\right.}{\sqrt{\left|\left[x_{j1} + \frac{1}{b}\left(-x_{j1} \cdots - x_{jb}\right)\right]^2\right| + \cdots + \left|\left[x_{jb} + \frac{1}{b}\left(-x_{j1} \cdots - x_{jb}\right)\right]^2\right|}\sqrt{\left|\left[v_{i1} + \frac{1}{b}\left(-v_{i1} \cdots - v_{ib}\right)\right]^2\right| + \cdots + \left|\left[v_{ib} + \frac{1}{b}\left(-v_{i1} \cdots - v_{ib}\right)\right]^2\right|}} \tag{18}$$

Table 1 Specification of Landsat-8 and Formosat2 sensors

Specification	Landsat-8	Formosat-2
Spatial Resolution (m)	30 m	8 m
Spectral Resolution (μm)	B1: 0.450–0.515 μm (Blue)	B1: 0.45–0.52 μm (Blue)
	B2: 0.525–0.600 μm (Green)	B2: 0.52–0.60 μm (Green)
	B3: 0.630–0.680 μm (Red)	B3: 0.63–0.69 μm (Red)
	B4: 0.845–0.885 μm (Near Infrared)	B4: 0.76–0.90 μm (Near Infrared)
Revisit Period	Repeat every 16 days	Daily
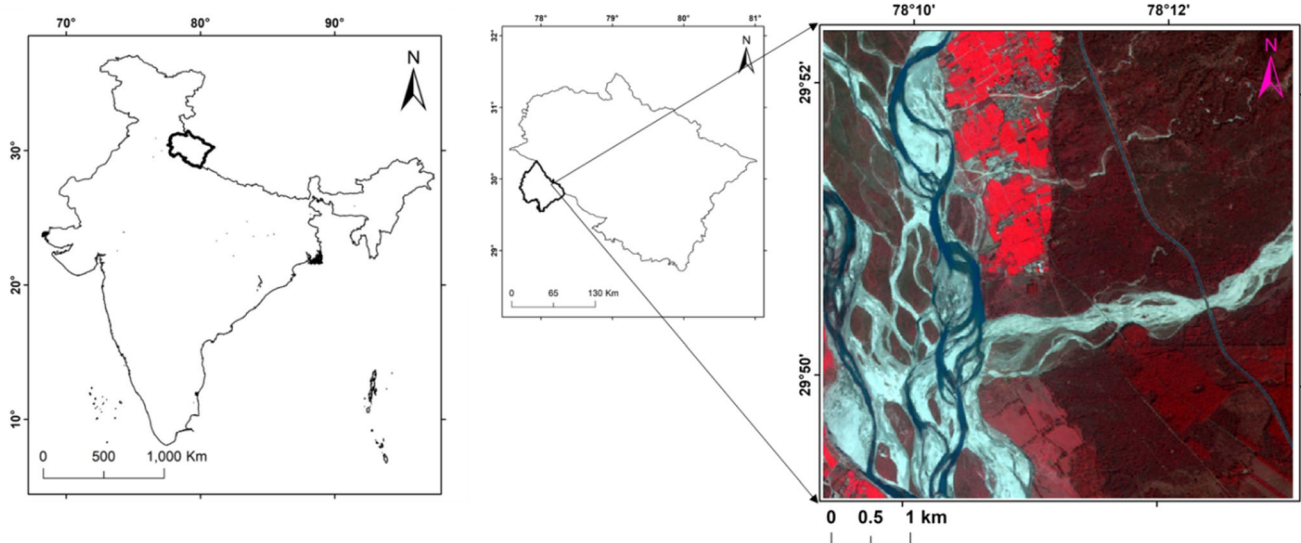
Fig. 1 Study area
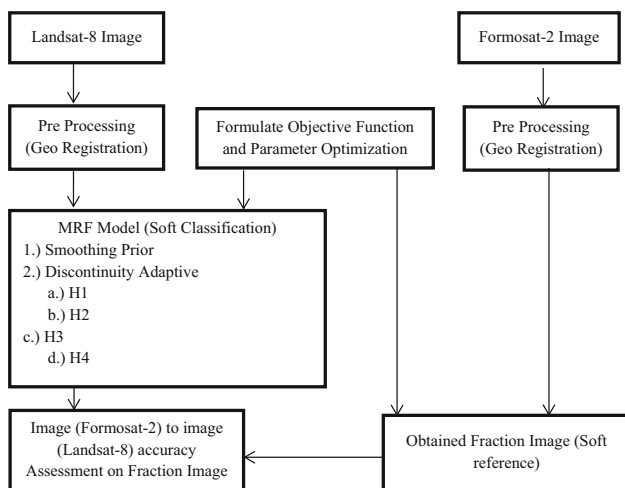
Fig. 2 Methodology adopted

Study Area and Data Used

The religious city Haridwar is situated in the foothills of Himalaya and on the bank of River Ganga in Northern Indian state Uttrakhand. Shivalik ranges of Himalaya have covered the Haridwar city in the Northern side; part of Haridwar was taken as a study area in this work. The temperature in Haridwar varies from 35 to 45 °C in the months of summer and 10–30 °C in winters, and the average rainfall is 1174.3 mm. The study area lies in the flood plains of the Ganga River; this area mainly consists of Water, Wheat, Dense Forest, Eucalyptus, Grassland, and Riverine Sand. Diversity is the main reason to select the area; due to diversity, mixed pixels are possible, which helps in examining the various method base fuzzy classifier methods. In this study, a remotely sensed image of Landsat-8 was selected for soft classification purposes, and Formosat-2 was selected as soft classified reference data. Table 1 shows the specification of Formosat-2 and Landsat-8 sensors. The study area maps were prepared using ArcGIS 10.5 tool shown in Fig. 1.

Methodology

The main objective of this paper is to study the effect of different distance measures and as well as spatial contextual information through two types of MRF models like DA(H1, H2, H3, and H4) models and SP using the base classifier FCM. DA (H1, H2, H3, H4) and SP are MRF
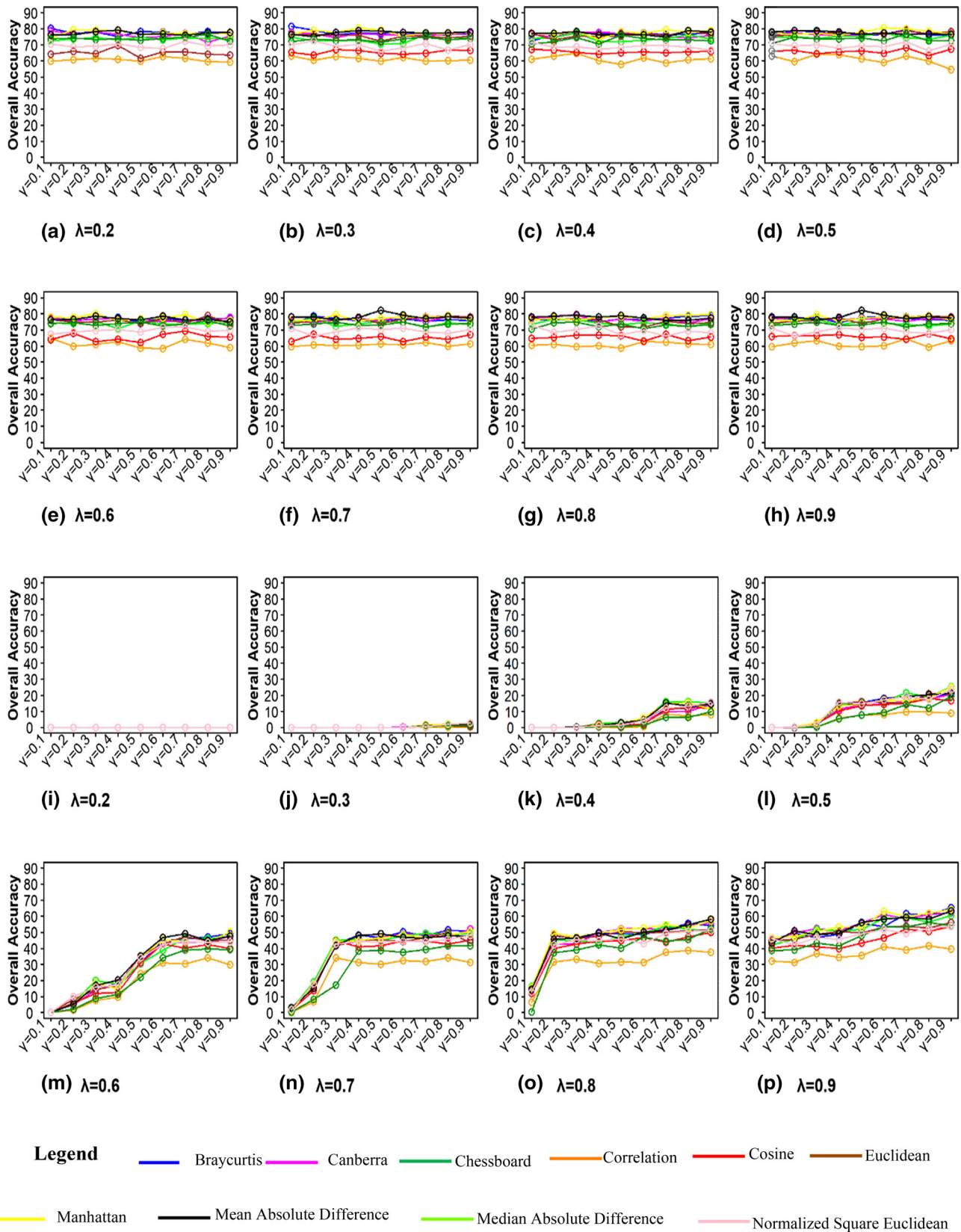
Fig.3 Comparison of overall accuracy in discontinuity adaptive prior (DA) (H1) Fig. 3 (**a–h**) where DA (H2) Fig. 3 (**i–p**) using base classifier FCM for applying different distance measures and m = 1.1
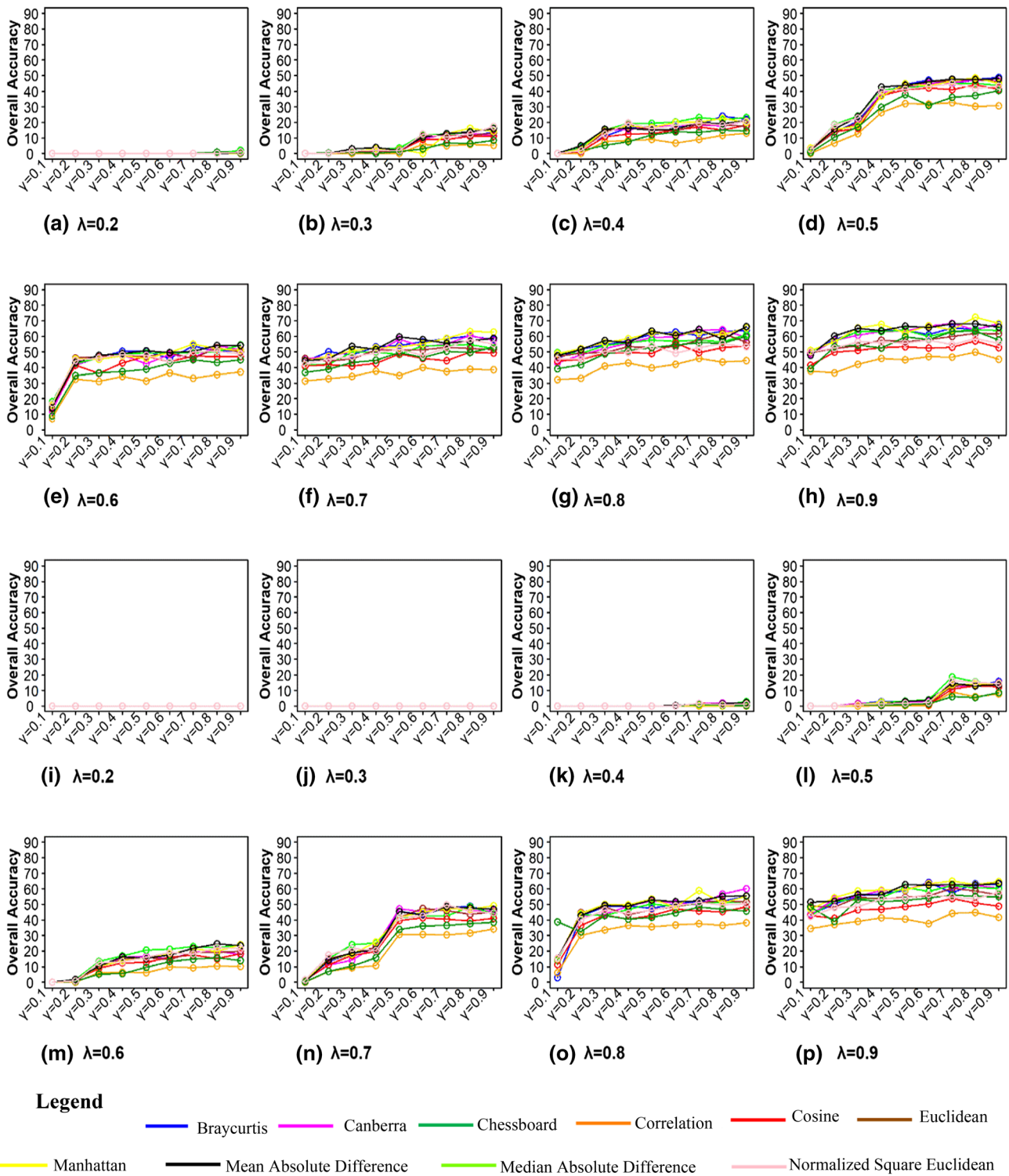
Fig. 4 Comparison of overall accuracy in discontinuity adaptive prior (DA) (H3) Fig. 4 (**a–h**) where DA (H4) Fig. 4 (**i–p**) using base classifier FCM for applying different distance measures and m = 1.1

models which consist of contextual spatial information; in this paper, FCM is used as a base classifier to combine the contextual spatial information with spectral information in MRF based FCM model. Here, Landsat-8 fraction output

images are used as classified data, and Formosat-2 fraction output images are used as a reference dataset.

The methodology adopted for this paper is shown in Fig. 2. Different parameters such as λ (0.2–0.9) with an
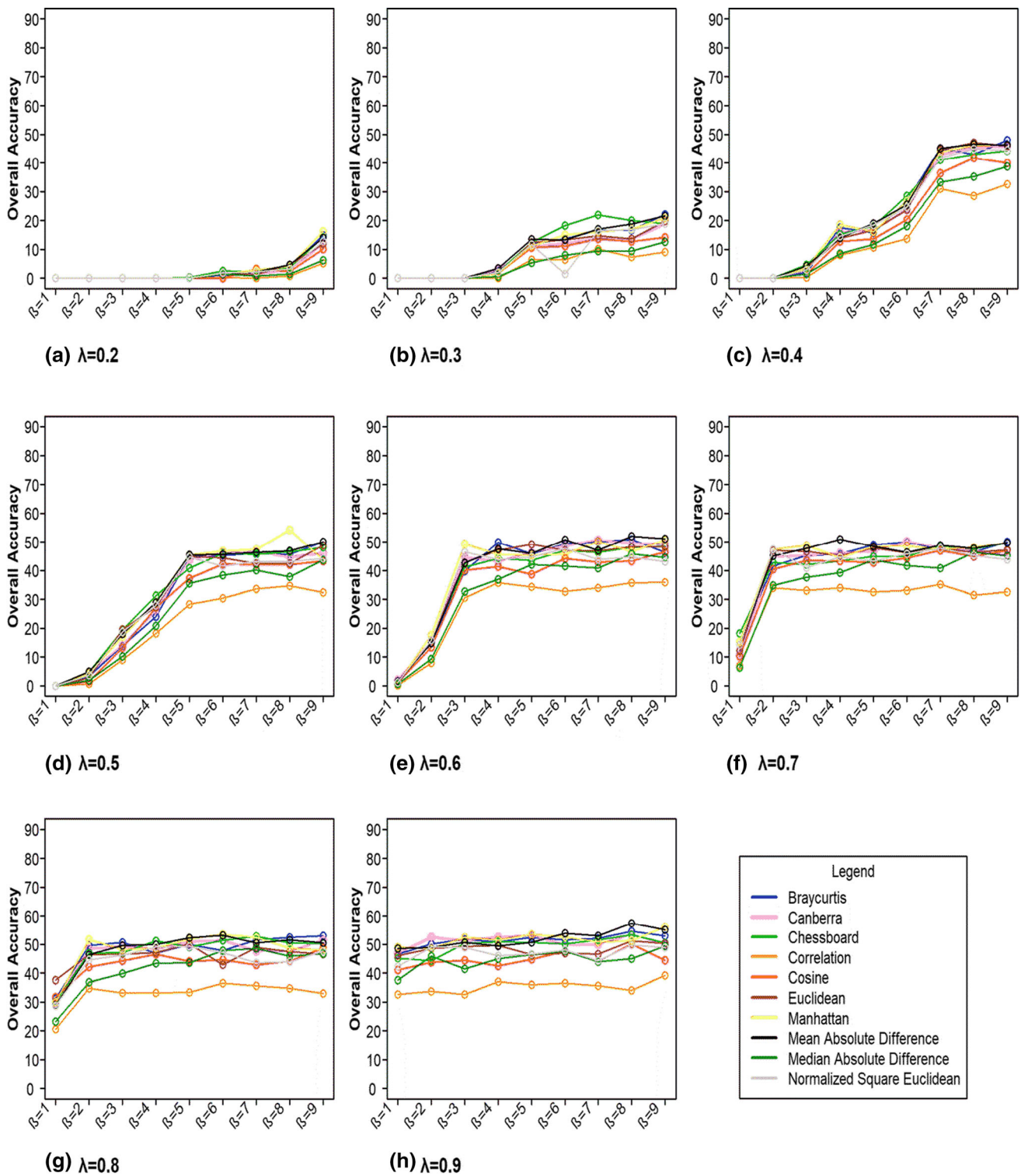
Fig. 5 Comparison of overall accuracy in smoothing prior (SP) using base classifier FCM for applying different distance measures and m = 1.1

interval of 0.1, γ (0.1–9) with a period of 0.1, m (1.1–3) with an interval of 0.2, and β (1–9) with an interval of 1 are used. For DA λ, γ and m parameters are used, while for SP β, γ and m parameter are used. Various distance measures are Bray Curtis, Canberra, Chessboard, Correlation, Cosine, Euclidean, Manhattan, Mean Absolute Difference, Median Absolute Difference, and Normalized Square Euclidean.

Table 2 Maximum overall accuracy of H1, H2, H3, H4 and Smoothing Prior (SP) using base classifier FCM for various distance measures and parameters (λ, γ, and β) where 'm' = 1.1

Maximum Overall Accuracy	Bray Curtis	Canberra	Chessboard	Correlation	Cosine	Euclidean	Manhattan	Mean Absolute Difference	Median Absolute Difference	Normalized Square Euclidean
H1 (λ/ γ)	0.3/0.1	0.8/0.3	0.5/0.6	0.6/0.1	0.6/0.7	0.7/0.8	0.5/0.6	0.9/0.5	0.5/0.7	0.8/0.1
	− 81.31	− 79.39	− 77.42	− 64.74	− 69.5	− 78.99	− 80.45	− 82.06	− 76.9	− 72.85
H2 (λ/ γ)	0.9/0.9	0.9/0.9	0.9/0.8	0.9/0.8	0.9/0.9	0.9/0.9	0.9/0.9	0.9/0.9	0.9/0.8	0.9/0.9
	− 65.14	− 61.97	− 60.39	− 41.49	− 53.84	− 55.92	− 63.99	− 63.19	− 56.11	− 53.63
H3 (λ/ γ)	0.9/0.9	0.9/0.9	0.9/0.8	0.9/0.8	0.9/0.8	0.9/0.8	0.9/0.8	0.9/0.8	0.9/0.8	0.9/0.8
	− 67.71	− 68.23	− 64.28	− 49.72	− 57.11	− 61.87	− 72.28	− 68.08	− 63.36	− 58.33
H4 (λ/ γ)	0.9/0.6	0.9/0.6	0.9/0.7	0.9/0.8	0.9/0.7	0.9/0.7	0.9/0.7	0.9/0.9	0.9/0.7	0.9/0.9
	− 64.08	− 63.39	− 61.98	− 44.76	− 53.7	− 60.43	− 64.96	− 63.47	(56.17)	− 57.82
SP (λ/ β)	0.9/8	0.9/8	0.9/8	0.9/9	0.9/8	0.9/8	0.9/9	0.9/8	0.9/9	0.9/8
	− 54.46	− 53.55	− 53.24	− 39.13	− 50.09	− 51.23	− 56.21	− 57.41	− 49.37	− 49.79

Step 1: Classified the image applying MRF Model as DA (H1, H2, H3, and H4) and SP where m = 1.1, λ, γ, β, and different distance measures.

Step 2: Overall Accuracy of the classified image has been calculated to get the optimized parameter.

Step 3: The optimized parameter has been used in the MRF model to classify the image where m varies 1.1 to 3 with an interval of 0.2.

Step 4: Finally, Accuracy Assessment is calculated, which gives the optimized algorithm in terms of parameters and distance measures.

Results and Discussion

To determine the optimized parameters of DA (H1, H2, H3, H4) and SP using base classifier FCM, graphs were plotted between parameters like (λ, β, γ) and Overall Accuracy (OA) for different distance measures. DA(H1, H2, H3, H4) depended on the parameter λ, γ, and m, whereas the SP on parameters λ, β, and m. The value of parameter λ varied in the range of (0.2–0.9), γ(0.1–0.9) with an interval of 0.1, β (1–9) with a period of 1, and a value of m = 1.1. Overall Accuracy is determined for DA(H1, H2, H3, H4) and SP using base classifier FCM for different parameter values and distance measures. Figure 3–5 shows the various technique as DA(H1, H2, H3, H4) and SP using FCM as a base classifier to find the optimized parameter in different distance measures. DA(H1, H2, H3, H4) and SP using FCM as a base classifier in result section defined as simply DA(H1), DA(H2), DA(H3), DA(H4), and SP, respectively.

From Fig. 3a–h: For the DA (H1) technique, Bray Curtis distance measures provided the highest OA for λ = 0.3 and γ = 0.1. Chessboard and Manhattan distance measure gave the maximum OA for λ = 0.5 and γ = 0.6. The Canberra distance measure offered the highest OA for λ = 0.8 and γ = 0.3. The Correlation and Cosine distance measures provided the highest OA for λ = 0.6 but γ = 0.1 and 0.7, respectively. Euclidean distance measure gave the highest OA for λ = 0.7 and γ = 0.8. Mean Absolute Difference provided the most increased OA for λ = 0.9 and γ = 0.5. Median Absolute Difference provided the most increased OA for λ = 0.5 and γ = 0.7. Normalized Square Euclidean gave the highest OA for λ = 0.8 and γ = 0.1.

From Fig. 3i–p: For the DA (H2) algorithm, Bray Curtis, Canberra, Cosine, Euclidean, Manhattan, Mean Absolute Difference, Normalized Square Euclidean gave the maximum OA for λ = 0.9 and γ = 0.9. Chessboard, Correlation, and Normalized Square Euclidean provided the highest OA at λ = 0.9 and γ = 0.8.

From Fig. 4a–h: For DA (H3) technique, Bray Curtis, Canberra, Chessboard, Correlation, Cosine, Euclidean, Manhattan, Mean Absolute Difference, Median Absolute Difference, and Normalized Square Euclidean gave the maximum OA for λ = 0.9. At the same time, γ has been different which are as follows 0.9, 0.9, 0.8, 0.8, 0.8, 0.8, 0.8, 0.8, 0.8 and 0.8, respectively.

From Fig. 4i–p: For the DA (H4) algorithm, Bray Curtis and Canberra gave the maximum OA for λ = 0.9 and γ = 0.6. Chessboard, Cosine, Manhattan, and Median Absolute Difference provided the OA for λ = 0.9 and γ = 0.7. Correlation showed the highest OA for λ = 0.9 and γ = 0.8. Median Absolute Difference and Normalized Square Euclidean gave the maximum OA at λ = 0.9 and γ = 0.7.

From Fig. 5: For SP technique, Bray Curtis, Canberra, Chessboard, Correlation, Cosine, Euclidean, Manhattan,
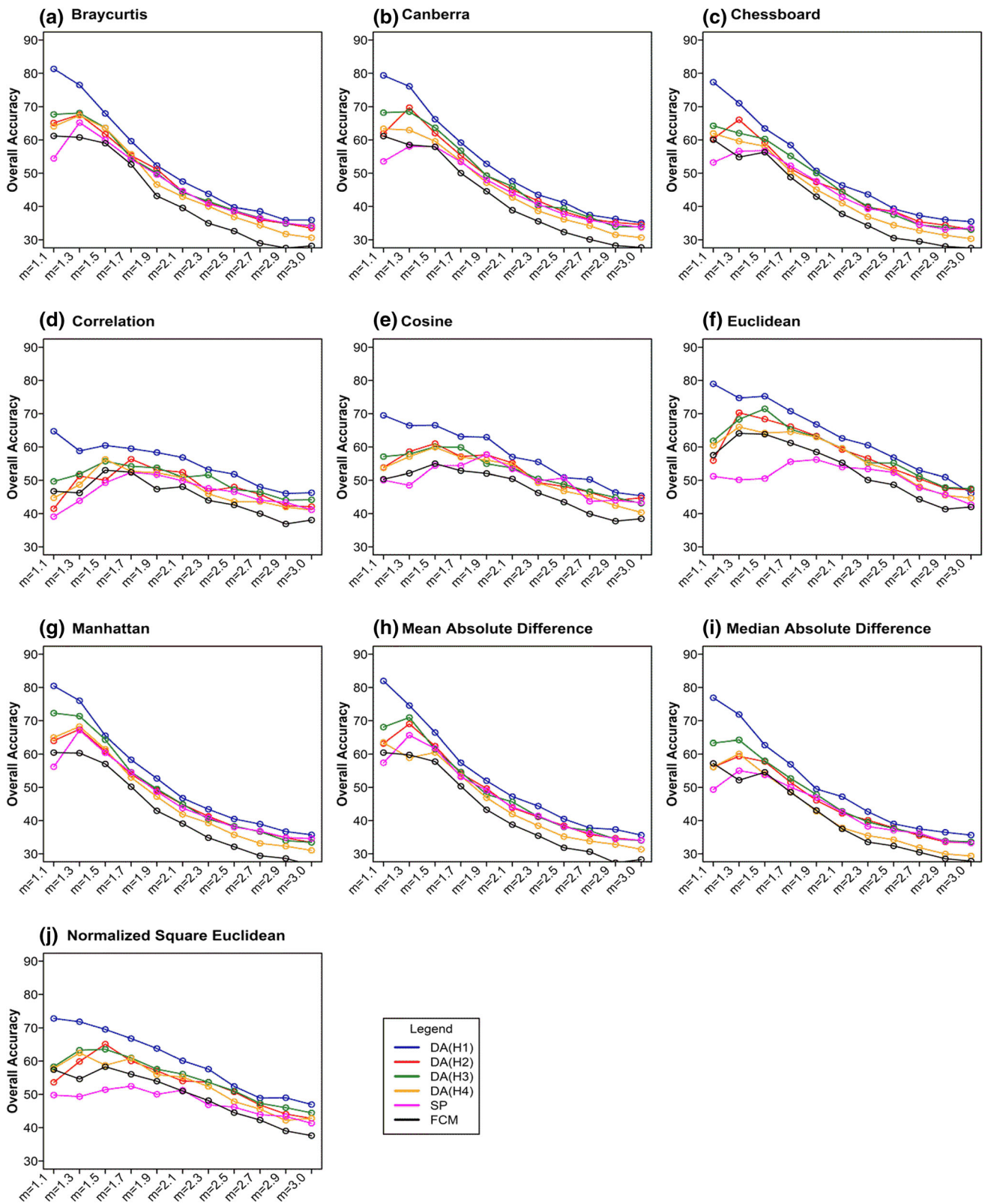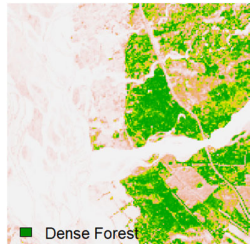
Fig. 6 Comparison of overall accuracy in discontinuity adaptive prior (DA) (H1, H2, H3, H4), Smoothing Prior (SP) and FCM, respectively, for different m(1.1–3.0) and distance measures **a** Bray Curtis **b** Canberra **c** Chessboard **d** Correlation **e** Cosine **f** Euclidean **g** Manhattan **h** Mean Absolute Difference **i** Median Absolute Difference **j** Normalized Square Euclidean

Table 3 Maximum overall accuracy of different algorithms using FCM and FCM based DA and SP Algorithm

Algorithm	m	Distance Measures	Overall Accuracy (%)
DA(H1)	1.1	Mean Absolute Difference	82.06
DA(H2)	1.3	Euclidean	70.29
DA(H3)	1.1	Manhattan	72.28
DA(H4)	1.3	Manhattan	68.26
SP	1.3	Manhattan	67.16
FCM	1.3	Euclidean	64.09

Mean Absolute Difference, Median Absolute Difference, and Normalized Square Euclidean gave the maximum OA for $\lambda = 0.9$. At the same time, β had been different, as follows 8, 8, 8, 9, 8, 8, 9, 8, 9, and 8, respectively.

Table 2 gives the maximum overall accuracy of DA(H1, H2, H3, H4) and Smoothing Prior (SP) using base classifier FCM for various distance measures and parameters (λ, γ, and β) where 'm' = 1.1.
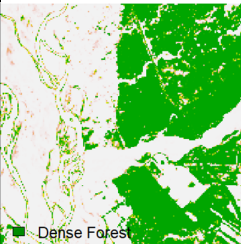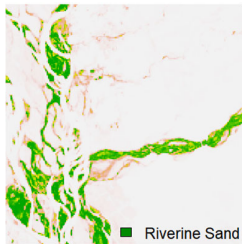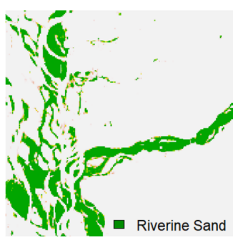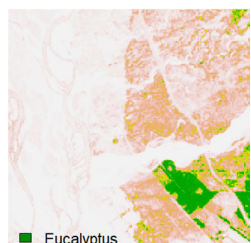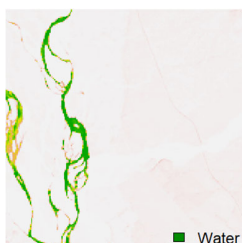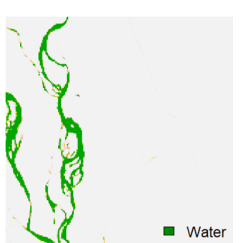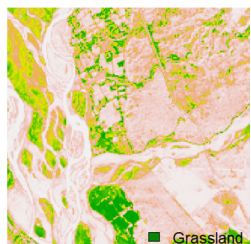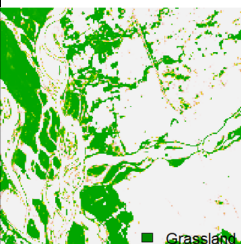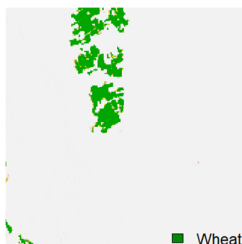
Figure 6 was plotted between m and overall accuracy (OA) for different distance measures. The m value in the range of (1.1–3) with an interval of 0.2 was considered. Figure 6 shows the various techniques as FCM, DA(H1, H2, H3, H4), and SP using FCM as a base classifier to obtain the best algorithm for m value and distance measures, where DA (H1, H2, H3, H4) and SP algorithms used an optimized parameter value of λ, γ, and β concerning distance measures, respectively.

In Fig. 6a, Bray Curtis distance measures, DA(H1), FCM gave the best OA for m = 1.1, whereas for another rest method (DA(H2), DA(H3), DA(H4), and SP), m = 1.3 gave the best result.

In Fig. 6b, Canberra distance measures, DA(H1), FCM, and DA(H4) gave the highest OA for m = 1.1, while for other methods, at m = 1.3, gave the highest OA.

In Fig. 6c, Chessboard distance measures with FCM, DA(H1), DA(H3), and D(H4) gave the highest OA for

Table 4 Classified classes concerning the algorithms

Classified classes	Classified image by FCM classifier	Classified image by DA (H1)using FCM as base classifier	Classified classes	Classified image by FCM classifier	Classified image by DA (H1)using FCM as base classifier
Dense forest	Dense Forest	Dense Forest	Riverine sand	Riverine Sand	Riverine Sand
Eucalyptus	Eucalyptus	Eucalyptus	Water	Water	Water
Grassland	Grassland	Grassland	Wheat	Wheat	Wheat

m = 1.1. DA(H2) provided the best OA for m = 1.3, while SP gives the most increased OA for m = 1.5.

In Fig. 6d, when Correlation distance measures used with DA(H1) gave the highest OA for m = 1.1. DA(H2) and SP provide the best for m = 1.7, while other methods at m = 1.5 provided the highest.

In Fig. 6e, Cosine distance measures with DA(H1) gave the best OA for m = 1.1. SP provided the highest for m = 1.9, while m = 1.5 gave the highest OA for other distance measures.

In Fig. 6f, for Euclidean distance measures with FCM, DA(H1), DA(H2), DA(H3), DA(H4), and SP gave the best OA for m = 1.3, m = 1.1, m = 1.7, m = 1.5, m = 1.3, and m = 1.9, respectively.

In Fig. 6g, at Manhattan distance measures with FCM, DA(H1) and DA(H3) provided the highest OA for m = 1.1. DA(H2), DA(H4), and SP gave the best OA for m = 1.3.

In Fig. 6h, for Mean Absolute Difference distance measures, FCM, DA(H1), and DA(H3) gave the highest OA for m = 1.1. At the same time, m = 1.3 provided the highest OA for the rest method.

In Fig. 6i, When Median Absolute Difference distance measures are used, FCM and DA(H1) provided the highest OA for m = 1.1, while m = 1.3 gave the most increased OA for another method.

In Fig. 6j), Normalized Square Euclidean distance measures, FCM, DA(H1), DA(H2), DA(H3), DA(H4), and SP gave the best OA for m = 1.5, m = 1.1, m = 1.5, m = 1.5, m = 1.3, and m = 1.7, respectively.

Table 3 gives the overall accuracy of the different algorithms concerning various distance measures and 'm' values studied.

Table 4 shows the compared outputs after classifying the Landsat-8 images for six different classes (Dense Forest, Eucalyptus, Grassland, Riverine Sand, Water, and Wheat) applying two methods. The first method is the FCM classifier involving Euclidean as distance measures; in contrast, the second method is the DA (H1) FCM based classifier using the Mean Absolute Difference measures. It was observed from the classified outputs that DA (H1) FCM based classifier given precise mapping of classes, better patches, and favorable outputs.

Conclusion

The conventional classification technique did not provide any information about mixed and neighborhood pixels. In contrast, the MRF model technique takes care of the information of neighborhood pixels to reduce the noisy pixel effect. This study used different MRF models as DA (H1, H2, H3, and H4) and SP to find the best algorithm considering the different parameters (λ, β, γ, and m) and

distance measures. DA (H1), the algorithm has yielded a maximum overall accuracy of 82.06% for Mean Absolute Difference distance measures with m = 1.3, λ = 0.9, and γ = 0.5. This study will help to select the Distance Measures and parameters (λ, γ, β, and m) for Discontinuity Adaptive Prior (H1, H2, H3, H4) and SP using base classifier FCM to classify the images accurately and achieve the highest accuracy.

Declarations

References

Agarwal, S., Burgess, C., & Crammer, K. (2009). Advances in Ranking.In *Workshop, Twenty-third annual conference on neural information processing systems, Whistler, BC.*

Ahmadvand, A., & Daliri, M. R. (2015). Improving the runtime of MRF based method for MRI brain segmentation. *Applied Mathematics and Computation, 256*, 808–818.

Baccour, L., & John, R. I. (2014, August).Experimental analysis of crisp similarity and distance measures.In *2014 6th International conference of soft computing and pattern recognition (SoCPaR)* (pp. 96–100). IEEE.

Berthod, M., Kato, Z., Yu, S., & Zerubia, J. (1996). Bayesian image classification using Markov random fields. *Image and Vision Computing, 14*(4), 285–295.

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences, 10*(2–3), 191–203.

Binaghi, E., Brivio, P. A., Ghezzi, P., & Rampini, A. (1999). A fuzzy set-based accuracy assessment of soft classification. *Pattern Recognition Letters, 20*(9), 935–948. https://doi.org/10.1016/S0167-8655(99)00061-6

Bray, J. R., & Curtis, J. T. (1957). Uma ordenação das comunidadesflorestais de terrasaltas do sul do Wisconsin. *Monografiasecológica, 27*, 325–349.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*, 721–741.

Gong, M., Su, L., Jia, M., & Chen, W. (2013). Fuzzy clustering with a modified MRF energy function for change detection in synthetic aperture radar images. *IEEE Transactions on Fuzzy Systems, 22*(1), 98–109.

Hao, M., Zhang, H., Shi, W., & Deng, K. (2013). Unsupervised change detection using fuzzy c-means and MRF from remotely sensed images. *Remote Sensing Letters, 4*(12), 1185–1194.

Hasnat, A., Halder, S., Bhattacharjee, D., Nasipuri, M., &Basu, D. K. (2013).Comparative study of distance metrics for finding skin color similarity of two color facial images. *ACER: New Taipei City, Taiwan*, 99–108.

Ichoku, C., & Karnieli, A. (1996). A review of mixture modeling techniques for sub-pixel land cover estimation. *Remote sensing reviews, 13*, 161–186.

Jackson, Q., & Landgrebe, D. A. (2002). Adaptive Bayesian contextual classification based on Markov random fields. *IEEE Transactions on Geoscience and Remote Sensing, 40*(11), 2454–2463.

Kaymak, U., &Setnes, M. (2000).Extended fuzzy clustering algorithms. *Available at SSRN 370850*.

Li, S. Z. (1995). On discontinuity-adaptive smoothness priors in computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 17*(6), 576–586.

Lillesand, T., Kiefer, R. W., & Chipman, J. (2015). *Remote sensing and image interpretation*. Wiley.

Scollar, I., Weidner, B., & Huang, T. S. (1984). Image enhancement using the median and the interquartile distance. *Computer Vision, Graphics, and Image Processing, 25*(2), 236–251.

Senoussaoui, M., Kenny, P., Stafylakis, T., & Dumouchel, P. (2013). A study of the cosine distance-based mean shift for telephone speech diarization. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 22*(1), 217–227.

Singh, P. P., & Garg, R. D. (2017). On sphering the high resolution satellite image using fixed point based ICA approach. In *Proceedings of international conference on computer vision and image processing* (pp. 411–419). Springer, Singapore.

Singh, P., & Garg, R. D. (2014). Classification of high resolution satellite images using spatial constraints-based fuzzy clustering. *Journal of Applied Remote Sensing, 8*(1), 083526.

Singha, M., Kumar, A., Stein, A., Raju, P. N. L., & Murthy, Y. K. (2015). Importance of DA-MRF models in fuzzy based classifier. *Journal of the Indian Society of Remote Sensing, 43*(1), 27–35.

Smits, P. C., &Dellepiane, S. G. (1997). Discontinuity adaptive MRF model for remote sensing image analysis.In *IGARSS'97. 1997 IEEE international geoscience and remote sensing symposium proceedings. Remote sensing-a scientific vision for sustainable development* (Vol. 2, pp. 907–909). IEEE.

Solberg, A. H. S., Taxt, T., & Jain, A. K. (1996). A Markov random field model for classification of multisource satellite imagery. *IEEE transactions on geoscience and remote sensing, 34*(1), 100–113.

Vassiliadis, S., Hakkennes, E. A., Wong, J. S. S. M., &Pechanek, G. G. (1998, August). The sum-absolute-difference motion estimation accelerator.In *Proceedings.24th EUROMICRO Conference (Cat.No. 98EX204)* (Vol. 2, pp. 559–566).IEEE.

Xu, M., Chen, H., & Varshney, P. K. (2011). An image fusion approach based on Markov random fields. *IEEE transactions on geoscience and remote sensing, 49*(12), 5116–5127.

Zhang, M., Therneau, T., McKenzie, M. A., Li, P., & Yang, P. (2008). A fuzzy c-means algorithm using a correlation metrics and gene ontology. In *2008 19th International Conference on Pattern Recognition* (pp. 1–4). IEEE.

Zhang, J., & Foody, G. M. (2001). Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: Statistical and artificial neural network approaches. *International Journal of Remote Sensing, 22*(4), 615–628.