Egyptian Informatics Journal 22 (2021) 133–143

Contents lists available at ScienceDirect

Egyptian Informatics Journal

journal homepage: www.sciencedirect.com

A new fusion of mutual information and Otsu multilevel thresholding technique for hyperspectral band selection

Nandhini K. *, Porkodi R.

Department of Computer Science, Bharathiar University, Coimbatore 641046, India

ARTICLE INFO

Article history:
Received 5 August 2019
Revised 20 March 2020
Accepted 26 June 2020
Available online 16 July 2020

Keywords:
Hyperspectral band selection
Information theory
Mutual information
Otsu multilevel threshold
SVM

ABSTRACT

Hyperspectral data are a curse with huge dimensionality, high redundancy of spectral information, and are noisy in nature. Hundreds of narrow adjacent bands are present in HS (Hyperspectral) data with high spectral information and it always leads to a computational complexity in space and time. The information theoretic methods are used for hyperspectral band selection to avoid computational complexity. In order to address this issue, the new fusion of Mutual Information (MI) with Otsu (MI_Otsu) threshold method is proposed for hyperspectral band selection by employing three different entropy measures such as joint, conditional, and relative. The proposed approach identifies the probabilities, entropy, and mutual information between two hyperspectral bands. The optimal threshold is obtained using Otsu multi-threshold technique and highly informative bands will be selected. In addition, the SVM (Support Vector Machine) classification technique is adapted for further classification of selected bands to analyze the performance of the proposed algorithm. The experimental analysis is carried out using the real-time dataset from the test site 'Indian Pines' in Northwestern Indiana recorded by AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) sensor that demonstrates the effectiveness of this proposed approach. It is proved that the proposed work shows the competitive performance even with less selected bands and the relative MI_Otsu method shows a higher accuracy of 92.16% with the comparison of joint and conditional MI_Otsu.

1. Introduction

Hyperspectral imaging spectrometers are adept at collecting a huge number of bands with different wavelengths channels for a given spatial area on the topological surface and with huge cursed dimensionality. These bands contain high spectral information and can identify similar materials [1,2]. Every band of pixel value is continuous spectra with the different classes of labels present on it. Over the decades, researchers have stated that hyperspectral dataset creates the Hughes phenomenon effect which decreases the classification accuracy due to the high dimensionality of spectral bands with the redundancy of information [3,4]. The redundancy of information needs to be sorted to overcome the above difficulties by choosing any one of the two broad categorized concepts: 1) Feature extraction or 2) Feature selection methods. Feature extraction techniques such as PCA (Principal Component Analysis) and ICA (Independent Component Analysis) are projected from the high to low dimensionality. Feature selection creates the subset from the original data to reduce the dimensionality with fewer features. In hyperspectral remote sensing, several bands are considered as features. The feature or band selection techniques involve selecting the highly informative bands, and that reduces the dimensionality and increases the classification accuracy. This paper is a study of band selection methods.

Information theoretic approaches, decision rule methods, spectral comparison-based approaches are widely used for hyperspectral band selection [5]. Information theory is one of the front running approaches in the band selection methods. Mutual Information (MI) doesn't require any ground truth reference or background data for identifying the highly informative bands. It gains the information from measuring the image of all bands and selection will occur based on the ranking of the entropy estimated values [6]. The amount of uncertainty in a variable generated through estimation is referred to as entropy. The information from the other variable are considered and measured for reducing the

* Corresponding author.
E-mail addresses: nandhinik2@gmail.com (K. Nandhini), porkodi_r76@buc.edu.in (R. Porkodi).

https://doi.org/10.1016/j.eij.2020.06.002
1110-8665/

uncertainty is MI [7]. There are many research studies employed MI techniques which often proved the high accuracy even with fewer bands. This paper concentrates on MI for band selection. The main contribution of the work is to select higher informative band from hyperspectral data with minimum number of selected bands and the classification method is used to prove the performance of proposed algorithm. The major advantages of the proposed algorithm for band selection include the reduction in false rate, delay in band selection and occurred higher accuracy compared to other previous works.

The rest of this paper is arranged as follows: Section 2 provides a survey of related works in the areas of hyperspectral band selection and classification. Section 3 depicts the architecture of the mutual information and Otsu method (MI-Otsu) and also explains the proposed algorithm. Section 4 illustrates the results obtained from this work and provides relevant discussions on them. Section 5 concludes the paper with future work.

2. Related work

The adaptive measurement of MI concept is proposed by deriving entropy and measuring the dependency of two random variables by statistical methods for hyperspectral image fusion. AVIRIS hyperspectral dataset is experimented by selecting the different cutoff of the bands with kernel classification techniques [8]. MI technique is enhanced by using the same experimental data and concentrating on the neighboring band to avoid redundancy with the application of complementary threshold technique. If the information retains the same mutual informative values in the adjacent band, then rejection of bandwidth will take place [9]. Spatial entropy based mutual information is introduced to collect highly informative bands, which is the extended version of Shannon entropy and mutual information. Spatial entropy is measured by intra-distance (average between pixels) and extra-distance (average distance of pixels) respectively [6].With the continuous effort from [8,9], the authors constructed an estimated reference map by applying mutual information for the experimental dataset AVIRIS Indian pines scene as the improved accuracy of 86.18% [10].

There are many approaches for the hyperspectral band selection on search-based methods. In this case, the hyperspectral bands are considered as a time series. Initially, clustering of spectral is done with training samples which offer the spectral curves for each class. The next step involved the creation of a candidate band subset using key points extracted from the spectral curves, and it follows the search procedure. The last stage is to filter the spectral bands through conditional MI, and bound search algorithms are adapted to optimize the band information [11].The band selection techniques used to detect the blueberry fruit are the combined algorithm of Kullback-Leibler divergence with pair wise class discriminability, hierarchical dimensionality reduction, and non-Gaussianity. K-nearest neighbor, AdaBoost and Support Vector Machine are used for the classification and to test the performance of the above three algorithms [12].

Jeffries-Matusita distance method is used for maximizing the classes separability. Highly correlated neighbor bands are merged to select efficient bands. To find the performance for selected bands the bagger algorithm, SVM, and KNN (k-nearest neighbors) are used to classify the classes. Other than that, this work was extended by implementing a post-classification algorithm which helps to identify the misclassified pixels namely classification error correction [3]. The Interband correlation coefficient method is used to select the bands automatically and is followed by SVD (Singular Value Decomposition) and QR decomposition. The kernel classifiers GURLS (Grand Unified Regularized Least Squares) and are employed to analyze the performance of these algorithms [13]. The neighborhood rough set method is combined with Shannon's entropy mutual information. A forward greedy search method was constructed with neighborhood rough set for significance band selection. Extreme Machine Learning (ELM) and Random Forest (RF) are used for assessing the band selection performance through classification accuracy [5]. The [5] research work got extended by analyzing the stability of the band selection algorithms. The Jaccard Index is used to estimate the sensitivity of the algorithms to the variations in the training set [14]. The representative band selection method is another technique in which all the spectral bands are grouped into clusters. The main aim of this technique is used to minimize its distance inside respective clusters, and the different classes' bands distances are maximized [15].

The clone selection algorithm-based band selection method, initially estimate the affinities of individuals by criterion. Next step is to select the individuals, and it is cloned and mutated till it achieves the best individual affinities which are then combined with Mutual information [4]. A Chaotic Binary Coded Gravitational Search Algorithm (CBGSA), a band selection method, is introduced to reduce the dimensionality. It is the most appropriate algorithm, and it is found by comparing with other algorithms such as the genetic algorithms, binary coded particle swarm optimization, binary coded differential evolution, and binary coded cuckoo search algorithms [16]. From the above background study, it was identified that MI techniques are used widely by hybridizing or involving the other concepts for band selection.

The MIMR-DGSA (Maximum-Information-Minimum-Redundancy - Discrete Gravitational Search Algorithm) adapts the variable bandwidth fast pair wise mutual information algorithm on enabling the neighborhood concepts to select the informative bands by increasing the hyperspectral bands entropy and by minimizing the mutual information between hyperspectral bands in each subset. And it is depending on the MIMR criterion goal to increase the entropy of bands and diminish the mutual information between the bands in a subset. The effect of this is that the achieved classification accuracy is subject to high variance [21]. An efficient clustering method based on Shared Nearest Neighbor (SNNC) to select the most representative bands from the original HS has also been introduced [22]. Based on improved subspace decomposition (ISD) and the artificial bee colony (ABC) algorithm, a band selection technique known as ISD-ABC to address the problem of dimensionality reduction in HIS (Hyperspectral Image) classification is executed. Subspace decomposition is achieved by calculating the correlation coefficients between adjacent bands and using the visualization result of the HSI spectral curve. This provides good classification accuracy compared with six other state-of-the-art band selection techniques [23].

In [24], multi objective based models helps to identify the hyperspectral band subsets with the different number of bands. The hyperspectral bands are selected by using Weakly-pareto-Optimal problem along with the novel boundary intersection adaptive penalty based approach. Basically there are three steps involved in the multi objective hyperspectral band selection method are the initialization of the subset, through the iteration the ideal point need to be identified and the current population and ideal points need to be updated. To select the highly potential hyperspectral bands can be identified by estimating Gaussian and triangular objectives functions to invoke the similarity fuzzy relationship through various parameters. The information measure (IM) is involved to reduce the uncertainty of hyperspectral bands and the highly mutual informative bands helps to classify the class labels well. Thus [25], combines both IM and Fuzzy rough set (FRS) to select the potential bands. In [26], Spectral and spatial pixel information of hyperspectral band correlations are measured through the double graph model, then Maximum information and minimum noise (MIMN) criterion is used to increase the maximum entropy of bands by reducing the noise. Then the determinantal point process (DPP) search algorithm helps to identify the subset of the hyperspectral bands. Thus [26] developed the MIMN-DPP algorithm by combining both MIMN and DPP to select the highly informative bands to improve the classification accuracy.

Recently, the cluster based mutual information approaches are used widely for hyperspectral band selection. In [27], the weighted entropy and the mutual information are calculated for hyperspectral bands and then the K-Means and Fuzzy-K-means pre-clustering algorithms are adapted to select the highly informative bands. In [28], the objective functions such as normalized cut criterion (NC)/top rank cut criterion is measured to develop the Optimal Clustering Framework (OCF) for hyperspectral band selection. The cluster Ranking strategy is applied with NC-OC-MVPCA (Maximum-variance principal component analysis) is used to determine the potential bands. In [29] ONR (Optimal Neighboring Reconstruction) method was developed by introducing the neighboring reconstruction based criterion objective function to identify the highly correlated bands and then based on the search strategy the optimal hyperspectral bands are identified. In [30] S&M (Split and Merge) method developed for hyperspectral band selection by without infringing the spectral data which means of spatially, the adjacent bands are split to find the potential sub-bands. These potential sub-bands are again merged to reduce the dimensional features in order to select the high dimensional data.

3. Proposed methodology

In the hyperspectral dataset, there will be a presence of noisy bands which affects the classification accuracy. The preprocessing stage is enabled to remove the necessary noisy bands from the raw data to overcome the accuracy of the classification. After removing the noisy bands, the estimation of probability is found and the entropy measurement is calculated. Then the optimal value is found by applying Otsu threshold technique and the highly informative band is extracted. Fig. 1 shows the Hyperspectral band selection framework which provides the process flow of MI_Otsu.

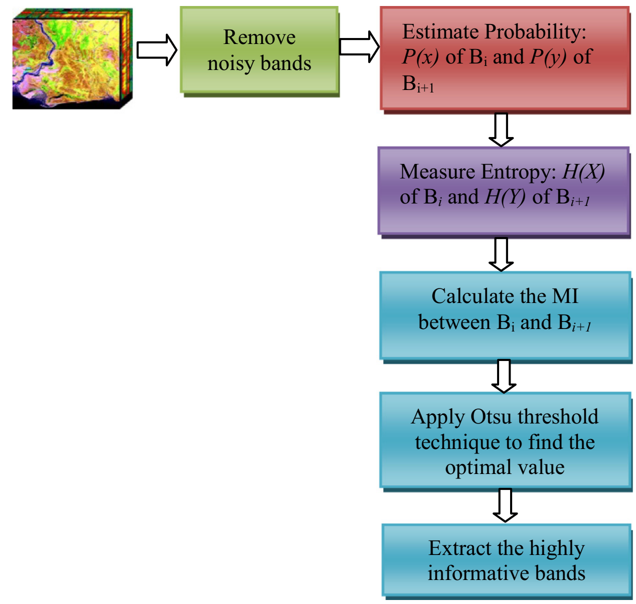

Fig. 1. Hyperspectral Band Selection Framework.

3.1. Band selection using MI_Otsu

The literature study of information theory in band selection shows that MI plays a major role in band selection over the decade which often shows the proven result. This research study took the Shannon entropy based mutual information. The hyperspectral data contains several number of bands B_i {where $i = 1{:}\ n$}. The first step is to estimate the probability using normalized distribution for each band. Consider the $P(x)$ and $P(y)$ is the probability distributions for B_i and $B_i + 1$ respectively. Then the entropy is estimated using three entropy measures such as Joint, Conditional and Relative.

3.1.1. Joint entropy

After identifying the $P(x)$ and $P(y)$ the joint entropy is calculated as follows,

$$\mathrm{H}(X,Y) = -\sum_{x}\sum_{y} P_{BiBi+1}(x,y)\log P_{BiBi+1}(x,y) \tag{1}$$

The joint entropy contains information about the randomness of the two bands. After estimating joint entropy for the two bands the mutual information, I $(B_iB_i + 1)$ of two bands needs to be calculated as follows:

$$I(BiBi+1) = \sum_{xy} P_{BiBi+1}(x,y) log \frac{P_{BiBi+1}(x,y)}{P_{Bi}(x)\cdot P_{Bi+1}(y)} \tag{2}$$

From the above Eq. (2), it derives that MI is linked to entropy. The hyperspectral bands are estimated through the following equations:

$$\begin{aligned} I\ (B_i,\ B_{i+1})\ &=\ H(X) +\ H(Y)\ --H\ (X,\ Y) \\ &= H(X)\ --H\ (X\mid Y) \\ &=\ H(Y)\ --H\ (Y\mid X) \end{aligned} \tag{3}$$

The steps involved in the proposed fusion of MI_Otsu for Joint entropy research work are crafted as an algorithm 1.

Input:HSData D1 (n number of bands)
Output: MI, Threshold, SelectedBands
Joint ()
for i= 1: no_bands (D1)
    CalculateProbability normalized distribution function for B_i, B_{i+1}|P(x), P(y)
Measure the joint entropy using equation (1)
Calculate **MI** using equation(3)
End
Extract the mutual information of the bands MI
Divide into 1: n-1 MI bands into M groups as equation (7)
Apply equation (8), (9) & (10) to find the optimal **Threshold** value
Extract the highly mutual informative **SelectedBands** based on a threshold.

3.1.2. Conditional entropy

The conditional entropy for two bands are estimated as follows,

$$\mathrm{H}(X,Y) = -\sum_{x}\sum_{y} P_{BiBi+1}(x,y)\log P_{BiBi+1}(y|x) \tag{4}$$

The steps involved in the proposed fusion of MI_Otsu for Conditional entropy research work are crafted as an algorithm 2.

Input: HSData D1 (n number of bands)
Output: MI, Threshold, SelectedBands
Conditional ()
for i= 1: no_bands(D1)
Calculate Probability normalized distribution function for B_i, B_{i+1}|P(x), P(y)
Measure the Conditional entropy using equation (4)
Calculate **MI** using equation (5)
End
Extract the mutual information of the bands MI
Divide into 1: n-1 MI bands into M groups as equation (7)
Apply equation (8), (9) & (10) to find the optimal **Threshold** value
Extract the highly mutual informative **SelectedBands** based on a threshold

3.1.3. Relative entropy

The relative entropy or Kullback-Leibler distance *D* is estimated as the distance between the two distributions of bands [20]. The relative entropy of mutual information is derived from equation (2) and calculated as follows:

$$I(B_i, B_{i+1}) = D(P(x,y)||P(x)(P(y)) \tag{5}$$

$$= R_{P(x,y)} log \frac{P_{BiBi+1}(x,y)}{P_{Bi}(x) \cdot P_{Bi+1}(y)} \tag{6}$$

The steps involved in the proposed fusion of MI_Otsu for Relative entropy research work are crafted as an algorithm 3.

Input: HSData D1 (n number of bands)
Output: MI, Threshold, SelectedBands
Relative ()
for i= 1: no_bands(D1)
Calculate Probability normalized distribution function for Bi, Bi+1|P(x),P(y)
Measure the Conditional entropy using equation (6)
Calculate **MI** using equation (2)
End
Extract the mutual information of the bands MI
Divide into1: n-1 MI bands into M groups as equation (7)
Apply equation (8), (9) & (10) to find the optimal **Threshold** value
Extract the highly mutual informative **SelectedBands** based on a threshold

From equation (2) and (3) the hyperspectral bands are formulated using the Joint entropy and MI for identifying informative bands. The equation (6) helps to formulate the relative entropy of mutual information between the bands. The conditional entropy of mutual information is calculated from equation (4) and (2). The 1D (Dimensional) and 2D based thresholding techniques are often used for the image matrix. Otsu method increases the separability between the class variances which yields a high accuracy in selecting bands based on the optimal threshold measure. This multilevel thresholding technique is employed for this hyperspectral band selection and combined with the MI function. In previous [4,8,9,20] research works, after identifying the mutual information for an $n - 1$ number of bands, the histogram-based thresholding measure wasapplied. In this proposed research work, Otsu multilevel thresholding technique is used to identify the threshold value for the *1: n – 1* mutual informative bands. It works on the maximizing between bands variance and minimizing within bands variance using the weighted mean method. The highly informative bands will be selected based on the threshold value. The selected bands (SB) are divided into M groups as $SB_1[1,\ldots t]$, $SB_2[t_1 + 1,\ldots m]$,…$SB_M[t_{M-1} + 1,\ldots K]$are considered as input for Otsu method, and probabilities were estimated and the optimal thresholds $[t_1^*, t_2^* \ldots t_{M-1}^*]$ are maximized σ_{SB}^2 as follows,

$$\{t_1*, t_2*\cdots\cdots..t_{M-1}*\} = argMax\{\sigma_{SB}{}^2(t_1, t_2, \cdots .t_{M-1}), \tag{7}$$

$\{1 \le t_1 < \cdots < t_{M-1} < K\}$ with the probabilities of two selected bands and band means are estimated as follows

$$q_k = \sum_{i \in SBk} p_i \tag{8}$$

$$\mu_k = \sum_{i \in SBk} ip_i/q_k \tag{9}$$

$$\sigma_{SB}{}^2 = \sum_{k=1}^{M} q_k \mu_k{}^2 \tag{10}$$

The above equations (8), (9) and (10) are used to identify the optimal threshold value for the selected bands. Further, the selection of bands through the proposed algorithm of MI_Otsu; the accuracy of band selection needs to be justified by classifying the individual classes using classification techniques. In this framework, SVM kernel classifier is chosen based on the study. In addition, extended morphological profiles are obtained from the study of [2,17–19].

4. Experimental results

The dataset used for this research work is the real-time dataset from the test site, Indian pines in northwestern Indiana recorded by AVIRIS sensor. The dataset contains 145 × 145 pixels per220 bands with 16 different classes in the wavelength of 0.4–2.5 μm. Preprocessing is carried out before band selection; the noisy bands are removed and reduced to 200 bands. The respective classes labels with the number of training samples (10366) are presented in the ground truth image are shown in Table 1 with reference to [2] and Fig. 2.

The proposed algorithm needs to find the mutual information between two bands, optimal threshold value and followed those highly informative bands to be selected. The MI_Otsu algorithm uses the joint entropy measure to measure the mutual information between the two bands. AVIRIS Indian pine dataset contains 145 × 145 × 200 dimensionality after removing the noisy bands. The proposed algorithm first evaluated the probability between the two bands and entropy measure has been done between two

Table 1
Indian pines ground truth dataset with respective number of samples for each class.

Label	Class	No. of Samples
1	Alfalfa	54
2	Corn-no till	1434
3	Corn-min till	834
4	Corn	234
5	Grass Pasture	497
6	Grass-Trees	747
7	Grass-Pasture-Mowed	26
8	Hay-Windrowed	489
9	Oats	20
10	Soybean-no till	968
11	Soybean-min till	2468
12	Soybean-clean	614
13	Wheat	212
14	Woods	1294
15	Buildings-Grass-Trees-Drives	380
16	Stone steel –Towers	95

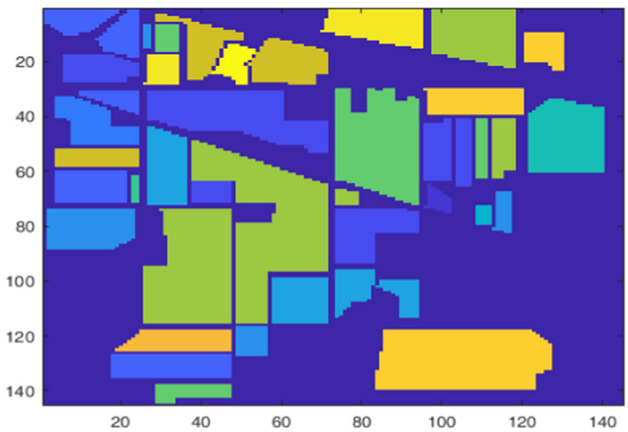

Fig. 2. Ground truth for Indian pines dataset.

adjacent bands using joint, conditional and relative. There are 199 mutual informative bands are generated with respect to the entropy measure. The Otsu multilevel thresholding technique is applied to find the highly informative bands from an estimated 199 MI bands. The training samples are taken from the AVIRIS Indian Pine ground truth (145 × 145) with class labels structure reshaped into the dimension of 2x × 10366. The five-sequence set of training samples are randomly picked with reference to ground truth labels from selected bands on each execution to study the worth of selected bands. The samples are 10, 20, 30, 40 and 50 taken per each class is considered kernel technique and is employed to justify the MI_Otsu algorithm classification accuracy (see Table 2).

The 94 bands are the drawn using joint entropy MI_Otsu method as highly informative bands out of 200 in Indian pines HS data. The selected bands are considered for further SVM classification. The Table 3 shows the training sample set of 50/class achieved 91.64% high average accuracy than another set of samples even with individual classes. The overall accuracy (OA) for 10, 20, 30, 40 and 50 are 68.13%, 77.29%, 81.31%, 86.01% and 86.46% respectively. The kappa statistics evaluated for different training samples are 71.61%, 79.84%, 83.49%, 87.71%, and 88.12%.

The results of a different set of training samples for Joint entropy MI_Otsu are showed in Fig. 3. Green pasture mowed, oats, wheat and stone steel power classes have randomly achieved 100% of accuracy in different sets of training samples. The 78 bands are drawn using conditional entropy MI_Otsu method as highly informative bands out of 200 in Indian pines HS data. Table 3 shows the training sample set of 50/class achieved 92.04% high average accuracy than another set of samples even with individual classes. The overall accuracy (OA) for 10, 20, 30, 40 and 50 are 70.56%, 80.21%, 82.32%, 85.76% and 88.52% respectively. The kappa statistics evaluated for different training samples are 66.72%, 77.67%, 79.90%, 83.83%, and 86.92%. Fig. 4 shows the results of Conditional entropy MI_Otsu.

Fig. 5 shows the results of relative entropy MI-Otsu. The 36 bands are drawn using relative entropy MI_Otsu method as highly informative bands out of 200 in Indian pines HS data. Table 4 shows the training sample set of 50/class achieved 92.16% high average accuracy than another set of samples even with individual classes. The overall accuracy (OA) for 10, 20, 30, 40 and 50 are 71.93%, 73.74%, 85.98%, 83.95%, and 85.21% respectively. The kappa statistics evaluated for different training samples are 75.09%, 76.53%, 87.66%, 85.80%, and 86.97%. The whole comparative study results are shown in Fig. 6. Fig. 7 shows the results average individual accuracy of Joint, Conditional and Relative entropy MI_Otsu. From the above results of joint, conditional, and relative MI_Otsu with SVM classification identified that relative entropy MI_Otsu shows the best results when compared to the other two algorithms. Fig. 7 shows the comparative study of individual classes averages for 50/class training samples of HS Data. The relative entropy achieved high accuracy at 92.16%. All three proposed algorithms yield a high accuracy in the implementation of Indian pines HS data. Even though results are with slight changes with the proposed work, the relative entropy MI_Otsu shows the high accuracy with less selected bands of 36.

4.1. Inference of support vector machine in hyperspectral classification

This classification paradigm requires the potential informative bands for the high classification performance in the accuracy wise. This pixel-wise classification after retrieving the highly informative bands through the mutual information and Otsu method, the SVM classifier is employed to assess the classification performance. The samples are 10, 20, 30, 40 and 50 taken per each class is considered kernel technique is employed to justify the MI_Otsu algorithm classification accuracy. The comparative study conducted with the proposed method by employing relative entropy with the histogram thresholding methods. The SVM classifier is

Table 2
Joint entropy MI_Otsu.

No. of Selected Bands					94
Class	10/class	20/class	30/class	40/class	50/class
1	90.91%	81.82%	90.91%	93.18%	79.55%
2	71.98%	77.09%	85.33%	85.22%	91.04%
3	58.50%	80.59%	82.71%	82.75%	82.27%
4	93.30%	92.52%	92.65%	94.33%	96.74%
5	83.98%	86.37%	82.66%	93.65%	94.63%
6	91.18%	94.09%	97.49%	96.75%	97.56%
7	100.00%	93.75%	100.00%	100.00%	93.75%
8	83.51%	99.36%	99.78%	99.33%	99.09%
9	100.00%	100.00%	100.00%	100.00%	100.00%
10	72.23%	77.00%	78.89%	81.36%	88.24%
11	56.10%	66.83%	71.41%	81.88%	77.71%
12	58.28%	69.19%	81.34%	85.71%	85.82%
13	99.50%	97.92%	99.45%	98.84%	100.00%
14	86.92%	88.38%	90.66%	93.70%	94.94%
15	61.89%	87.22%	84.29%	94.71%	89.39%
16	90.59%	97.33%	100.00%	100.00%	95.56%
Average	**81.18%**	**86.84%**	**89.85%**	**92.59%**	**91.64%**
Overall	**68.13%**	**77.29%**	**81.31%**	**86.01%**	**86.46%**
Kappa	**71.61%**	**79.84%**	**83.49%**	**87.71%**	**88.12%**

Table 3
Conditional entropy MI with OTSU.

No. of Selected Bands					78
Class	10/class	20/class	30/class	40/class	50/class
1	93.18%	93.18%	100.00%	79.55%	68.18%
2	59.62%	74.96%	74.50%	81.71%	83.82%
3	40.90%	75.80%	71.77%	80.48%	89.16%
4	73.66%	86.92%	92.16%	95.88%	97.83%
5	78.44%	93.08%	91.01%	97.37%	94.85%
6	88.33%	94.09%	95.68%	96.75%	96.41%
7	87.50%	93.75%	100.00%	87.50%	100.00%
8	85.39%	98.08%	99.13%	99.33%	99.32%
9	100.00%	100.00%	100.00%	80.00%	100.00%
10	72.44%	81.33%	68.55%	80.39%	82.90%
11	67.58%	67.57%	76.83%	77.39%	81.43%
12	58.44%	66.67%	78.77%	85.54%	84.22%
13	96.53%	97.92%	100.00%	98.84%	99.38%
14	84.35%	92.70%	96.20%	92.19%	97.59%
15	75.14%	85.83%	86.29%	95.88%	97.58%
16	92.94%	88.00%	100.00%	98.18%	100.00%
Average	**78.40%**	**86.87%**	**89.43%**	**89.19%**	**92.04%**
Overall	**70.56%**	**80.21%**	**82.32%**	**85.76%**	**88.52%**
Kappa	**66.72%**	**77.67%**	**79.90%**	**83.83%**	**86.92%**

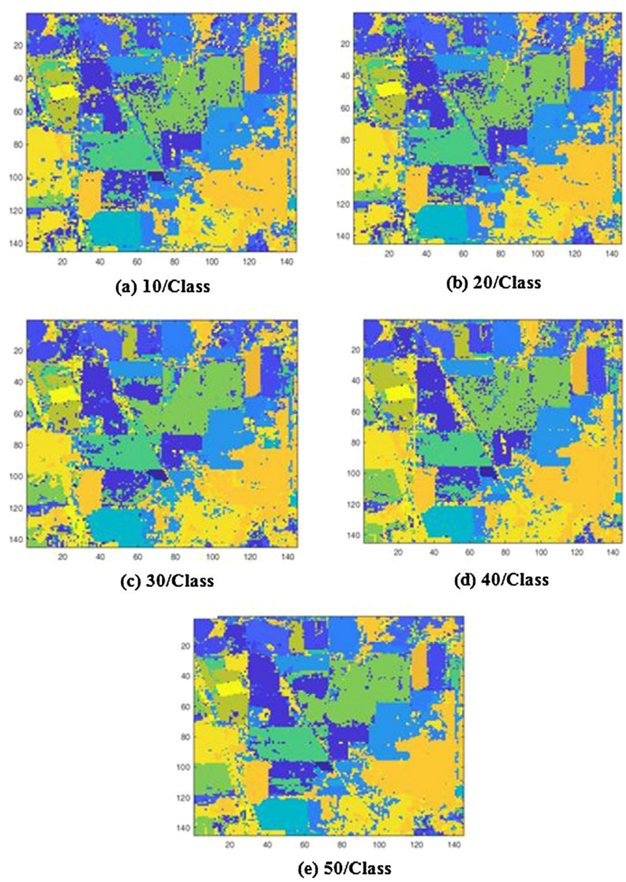

Fig. 3. Joint Entropy with Otsu method band selection Classification results for different training samples (a) 10/ class (b) 20/class (c) 30/class (d) 40/class (e) 50/class.

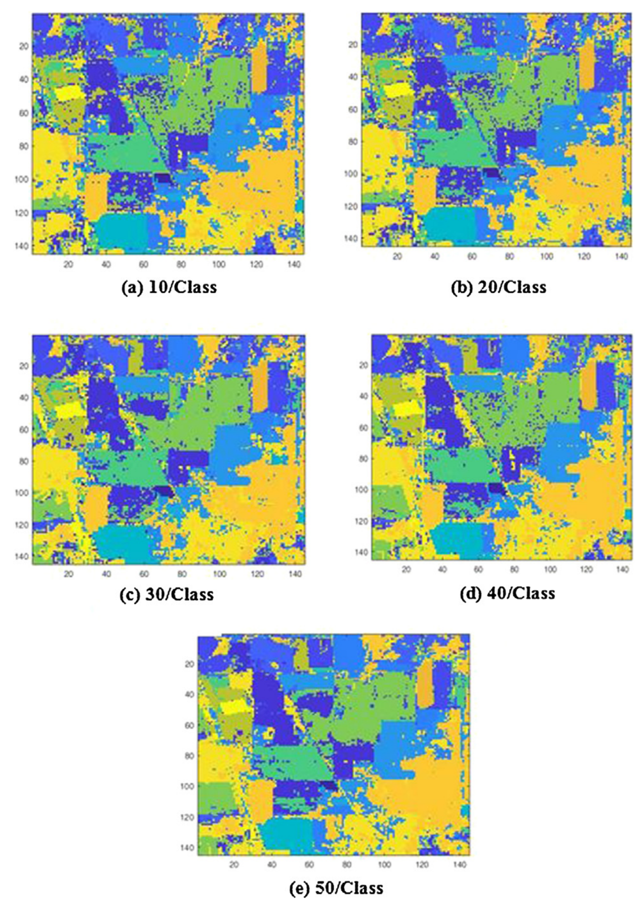

Fig. 4. Conditional Entropy with Otsu method band selection Classification results for different training samples (a) 10/ class (b) 20/class (c)30/class (d) 40/class (e) 50/class.

employed to assess the performance for pixel wise classification. The comparative studies are conducted with the proposed method by employing higher accuracy relative entropy with the histogram thresholding methods.

The Fig. 8 shows the results of relative entropy with mutual information using histogram thresholding. The Table 5 shows the results of predicted individual accuracy, OA (Overall Accuracy), AA (Average Accuracy) and Kappa statistics for different set of arbitrary training samples fed in the SVM classifier. Grass pasture mowed achieves high accuracy of 100% even with all training samples per class. Stone steel towers class achieves 100% of accuracy in 40 and 50 training samples per class. All other classes in the relative entropy achieve more than 60% with the all training samples in each class. The accuracy achieved for relative entropy using

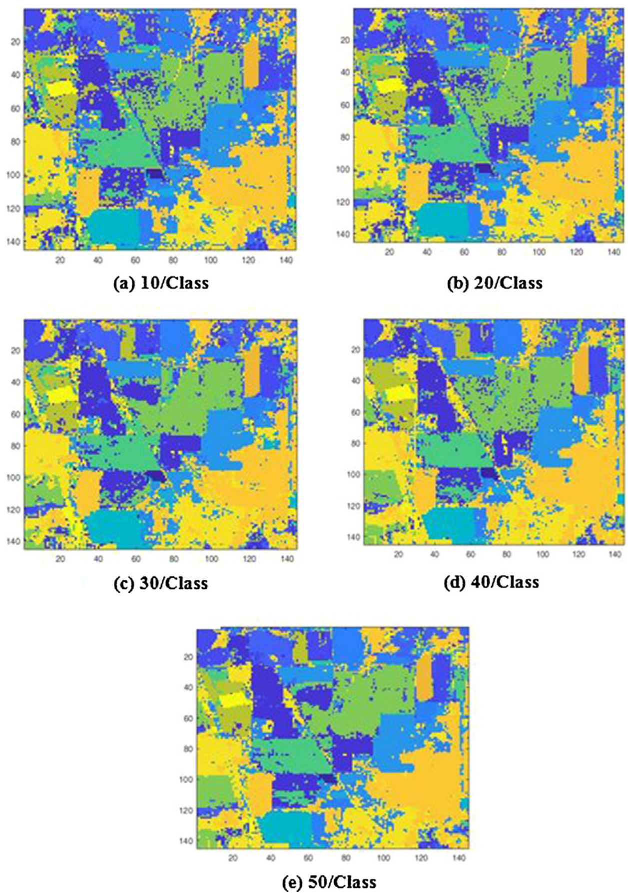

Fig. 5. Relative Entropy with Otsu method band selection Classification results for different training samples (a) 10/ class (b) 20/class (c) 30/class (d) 40/class (e) 50/class.

MI_hist with 50 training samples is OA: 88.35%, AA: 91.49% and Kappa: 86.70.

The 36 bands are drawn using relative entropy MI_Otsu method as highly informative bands out of 200 in Indian pines HS data. Table 6 shows the training sample set of 50/class achieved 92.16% high AA than another set of samples even with individual classes. The overall accuracy (OA) for 10, 20, 30, 40 and 50 are 71.93%, 73.74%, 85.98%, 83.95%, and 85.21% respectively. The Kappa statistics evaluated for different training samples are 75.09%, 76.53%, 87.66%, 85.80%, and 86.97%. The overall results of the selected bands from relative entropy for different set of training samples are shown in the Fig. 9. All classes achieve high accuracy even with the less selected bands in all different set of training samples.

The Figs. 10 and 11 shows the comparative study on both existing and proposed methods. From experimental study, it is identified that the proposed Fusion MI_Otsu algorithms performs well when compared with the existing algorithm even with the less selected bands. The accuracy is high when compared with existing methods. The relative entropy with MI_histogram extracted 47 bands with the AA of 91.49% respectively. The proposed fusion methods of this research work employed relative entropy with MI_Otsu extracted 36 bands with the AA of 92.16% respectively. From the above results of relative MI_Otsu with SVM classification identified that relative entropy MI_Otsu is the best results when compared to the other algorithms. Fig. 10 shows the comparative study of individual class averages for 50/class training samples of hyperspectral Data. The proposed relative entropy with MI_Otsu achieved high accuracy of 92.16%. All other proposed algorithms yield a high accuracy in the implementation of Indian pines HS data. In this research work at each stage the proposed research works are justified for further research process.

Fig. 11 shows the comparison of proposed MI-Otsu method with MI-hist approaches assessed by SVM classifier with 50 training samples per class. For this experimental study, the existing techniques Joint MI_histogram, Conditional MI_histogram, and Relative MI_histogram are employed and compared with the proposed approaches. But the higher accuracy of Relative entropy is demonstrated in the SVM classification section. From the experimental study, it is proved that the proposed methods achieve high accuracy when compared with the existing techniques. the Relative MI_Otsu method achieved the high performance results with less selected bands of 36 of 92.16% average accuracy respectively. The Table 7 shows the comparative study of proposed Relative entropy MI_Otsu method, OCF [28], ONR [29], and S&M [30]. OCF and ONR are the clustering approaches where the selected bands are depend on the cluster k value. Thus, this experimental study given clustered value as k = 36 to retrieve the optimal band in each subset of the cluster framework.

Table 4
Relative entropy MI with Otsu.

No. of Selected Bands					36
Class	10/class	20/class	30/class	40/class	50/class
1	95.45%	93.18%	90.91%	93.18%	95.45%
2	58.01%	74.61%	83.83%	79.12%	82.59%
3	58.86%	79.36%	84.83%	88.66%	83.42%
4	76.34%	90.19%	94.61%	96.39%	94.57%
5	80.29%	83.86%	93.15%	95.84%	92.62%
6	97.56%	92.43%	97.35%	97.60%	97.85%
7	100.00%	100.00%	100.00%	87.50%	93.75%
8	95.20%	97.87%	98.91%	99.11%	99.09%
9	100.00%	100.00%	100.00%	100.00%	100.00%
10	77.87%	80.59%	82.73%	84.81%	92.92%
11	68.92%	55.88%	82.44%	72.08%	77.50%
12	57.62%	71.38%	83.39%	88.85%	82.27%
13	98.02%	100.00%	97.80%	100.00%	98.15%
14	91.67%	84.38%	93.04%	95.14%	91.80%
15	81.62%	91.67%	92.29%	95.59%	96.97%
16	97.65%	88.00%	98.46%	96.36%	95.56%
Average	**83.44%**	**86.46%**	**92.11%**	**91.89%**	**92.16%**
Overall	**71.93%**	**73.74%**	**85.98%**	**83.95%**	**85.21%**
Kappa	**75.09%**	**76.53%**	**87.66%**	**85.80%**	**86.97%**

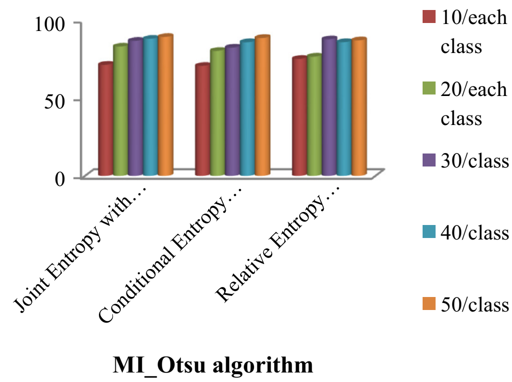

Fig. 6. Comparative analysis of OA for proposed MI_Otsu algorithms with different set of training samples.

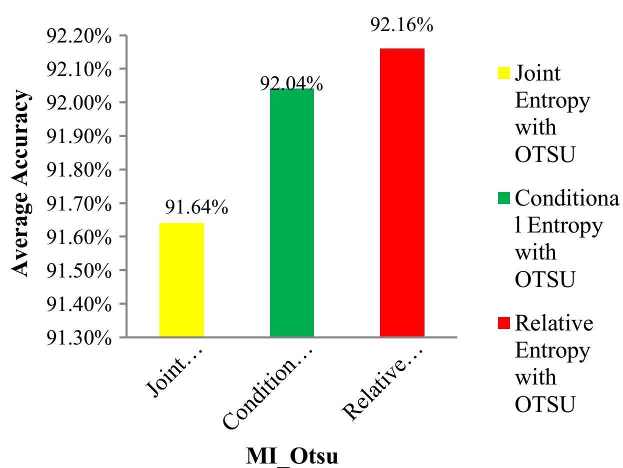

Fig. 7. Comparative analysis of average Individual classes of proposed MI_Otsu with 50 training samples per class.

The experimental study shows that the clustering based approaches ONR and OCF based on the cluster value 36 bands are selected and the classification performance are measured through the evaluation metrics OA, AA and Kappa are 80.51%, 88.24%, 77.99% and 72.90%, 82.48%, 69.54% respectively. The S&M hyperspectral band selection method selected 40 potential bands and the SVM classification accuracy OA, AA and Kappa is 82.85%, 87.78% and 80.59% respectively. The Fig. 12 shows that the proposed Relative MI_Otsu hyperspectral band selection method performs better in classification accuracy when compared with the other algorithms.

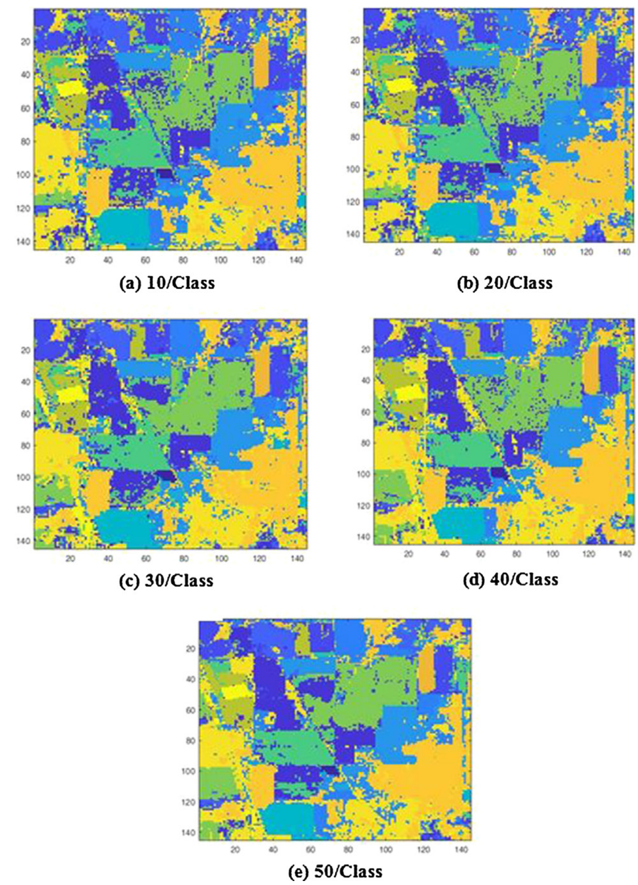

Fig. 8. The OA Results of Relative Entropy MI_hist + SVM with different training samples (a) 10 samples/class (b) 20 samples/class (c) 30 samples/class (d) 40 samples class (e) 50 samples/class.

Table 5
Relative entropy MI_Hist + SVM.

No. of Selected Bands					47
Class	10/class	20/class	30/class	40/class	50/class
1	93.18%	88.64%	97.73%	84.09%	84.09%
2	67.84%	64.43%	75.28%	80.56%	82.15%
3	71.97%	65.60%	82.21%	87.66%	84.03%
4	89.29%	93.46%	89.22%	96.91%	92.59%
5	70.02%	89.73%	93.36%	93.87%	93.14%
6	91.32%	96.15%	98.74%	97.31%	96.72%
7	100.00%	100.00%	87.50%	87.50%	87.50%
8	97.49%	97.65%	98.91%	99.55%	99.32%
9	100.00%	100.00%	100.00%	100.00%	100.00%
10	85.07%	68.14%	78.46%	80.50%	82.56%
11	63.06%	61.89%	75.80%	79.98%	84.77%
12	69.21%	78.62%	90.58%	87.98%	83.66%
13	98.02%	97.92%	97.80%	97.67%	100.00%
14	71.65%	87.76%	93.59%	95.53%	96.00%
15	82.16%	90.28%	94.29%	96.18%	97.31%
16	88.24%	97.33%	96.92%	100.00%	100.00%
AA	**83.66%**	**86.10%**	**90.65%**	**91.58%**	**91.49%**
OA	**74.34%**	**75.69%**	**84.78%**	**87.35%**	**88.35%**
Kappa	**71.22%**	**72.78%**	**82.79%**	**85.58%**	**86.70%**

Table 6
Relative entropy MI_Otsu + SVM.

No. of Selected Bands					36
Class	10/class	20/class	30/class	40/class	50/class
1	95.45%	93.18%	90.91%	93.18%	95.45%
2	58.01%	74.61%	83.83%	79.12%	82.59%
3	58.86%	79.36%	84.83%	88.66%	83.42%
4	76.34%	90.19%	94.61%	96.39%	94.57%
5	80.29%	83.86%	93.15%	95.84%	92.62%
6	97.56%	92.43%	97.35%	97.60%	97.85%
7	100.00%	100.00%	100.00%	87.50%	93.75%
8	95.20%	97.87%	98.91%	99.11%	99.09%
9	100.00%	100.00%	100.00%	100.00%	100.00%
10	77.87%	80.59%	82.73%	84.81%	92.92%
11	68.92%	55.88%	82.44%	72.08%	77.50%
12	57.62%	71.38%	83.39%	88.85%	82.27%
13	98.02%	100.00%	97.80%	100.00%	98.15%
14	91.67%	84.38%	93.04%	95.14%	91.80%
15	81.62%	91.67%	92.29%	95.59%	96.97%
16	97.65%	88.00%	98.46%	96.36%	95.56%
AA	**83.44%**	**86.46%**	**92.11%**	**91.89%**	**92.16%**
OA	**71.93%**	**73.74%**	**85.98%**	**83.95%**	**85.21%**
Kappa	**75.09%**	**76.53%**	**87.66%**	**85.80%**	**86.97%**

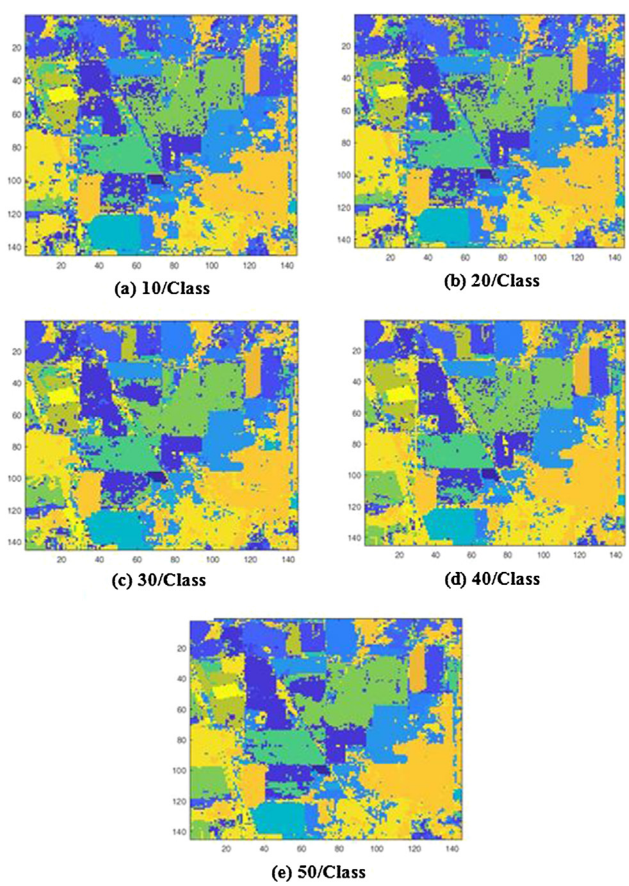

Fig. 9. The OA Results of proposed Relative Entropy MI_Otsu + SVM with different training samples (a) 10 samples/class (b) 20 samples/class (c) 30 samples/class (d) 40 samples/class (e) 50 samples/class.

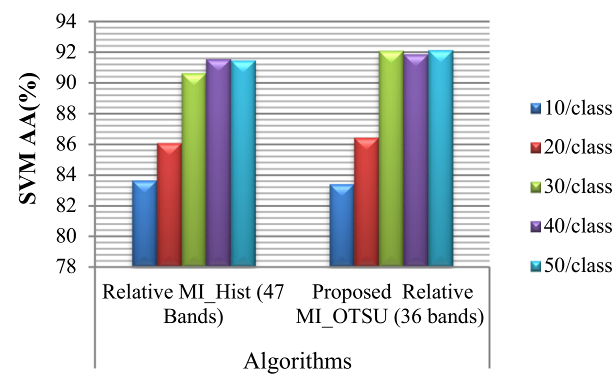

Fig. 10. Comparative study on entropy measures with proposed and histogram based algorithms using SVM classifier (Relative Entropy: MI_Histogram & MI_Otsu).

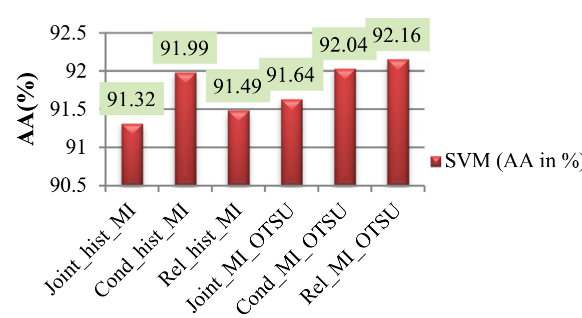

Fig. 11. The proposed MI_Otsu method are compared with MI_hist approaches assessed by SVM classifier with 50 training samples per class.

5. Conclusion and future work

The information theory proved the best in many research studies, especially in hyperspectral band selection methods. A new fusion algorithm MI_Otsu is proposed for band selection algorithm for high accuracy with less selected bands. Three different entropy measures are employed for hyperspectral band selection. The selected bands for joint, conditional, and relative entropy with mutual information are 94, 78, and 36 respectively. The experimental study on the AVIRIS dataset concluded that with less selected bands, the SVM classification technique could be employed to justify the performance. The relative entropy MI_Otsu algorithm offers high accuracy of 92.16% when compared to the other two entropy measures, OCF, ONR and S&M methods. The pro-

Table 7
Comparison of hyperspectral band selection methods with SVM classification.

Class	ONR 36 Bands	OCF 36 Bands	S&M 40 Bands	Relative MI_OTSU 36 Bands
1	86.36%	97.73%	86.36%	95.45%
2	64.29%	49.37%	65.63%	82.59%
3	79.36%	77.67%	85.26%	83.42%
4	95.33%	83.04%	90.65%	94.57%
5	92.03%	92.20%	93.29%	92.62%
6	96.84%	88.60%	96.84%	97.85%
7	100.00%	100.00%	93.75%	93.75%
8	98.08%	93.53%	98.51%	99.09%
9	100.00%	100.00%	100.00%	100.00%
10	73.31%	70.88%	74.16%	92.92%
11	73.86%	64.32%	80.76%	77.50%
12	74.58%	64.57%	78.96%	82.27%
13	97.92%	97.52%	97.92%	98.15%
14	90.74%	88.01%	90.89%	91.80%
15	89.17%	65.14%	82.22%	96.97%
16	100.00%	87.06%	89.33%	95.56%
AA	**88.24%**	**82.48%**	**87.78%**	**92.16%**
Kappa	**77.99%**	**69.54%**	**80.59%**	**86.97%**
OA	**80.51%**	**72.90%**	**82.85%**	**85.21%**

Note: Please check the algorithm style.

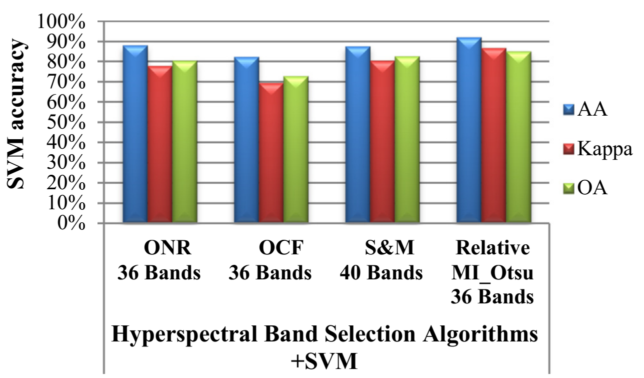

Fig. 12. Comparative study of the different Hyperspectral Band Selection algorithms with the proposed relative MI_Otsu.

posed joint and conditional entropy MI_Otsu classification accuracies are near to the relative entropy MI_Otsu are 91.64% and 92.04% respectively. Also, the results on individual classes of accuracy indicate that there could be a potential issue on spectral unmixing in a pixel which would lead to lower accuracy. In the future, this research work taken to the further novel extraction or classification with the selected bands and nonlinear spectral unmixing will be considered to improve the classification better.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research study was supported under the University Research Fellowship (URF NO: C2/2016/2170) from the Bharathiar University, Coimbatore, India. Authors would like thank the Editor and Reviewer for their valuable comments and suggestions.

References

[1] MahdiKhodadadzadeh, Jun Li, Antonio Plaza, Hassan Ghassemian Jose M. Bioucas-Dias. Spectral-spatial classification of Hyperspectral data using local and global probabilities for missed pixel characterization. IEEE Trans Geosci Remote Sens 52(10):2014. 6298–6314. DOI: 10.1109/TGRS.2013.2296031.
[2] Jun Li, Prashanth Reddy Marpu, Antonio Plaza, Jose M. Bioucas-Dias and Jon AtliBenediktsson, "Generalized Composite Kernel Framework for Hyperspectral Image Classification", IEEE Transactions on Geoscience and Remote Sensing, Vol 51(9),2013, 4816-4829, DOI: 10.1109/TGRS.2012.2230268.
[3] SeyyidAhemedMedjahed, TamazouztAitsaadi, Abdelkader Benyettou, and Mohammed Ouali, "A new post-classification and band selection frameworks for hyperspectral image classification", The Egyptian Journal of Remote Sensing and Space Sciences, Elsevier publication vol (19), 2016, 163-173, https://doi.org/10.1016/j.ejrs.2016.09.003.
[4] Wenbo Yu, Miao Zhang, Yi Shen, "Mutual information and clone selection algorithm based hyperspectral band selection method". Proceedings of the 36th Chinese control conference, Dalian China, IEEE, DOI: 10.23919/ChiCC.2017.8029188
[5] Yao Liu, HongXie, Yuehua Chen, Kezhu Tan, Liguo Wang, WuXie, "Neighborhood mutual information and its application on hyperspectral band selection for classification", Chemometrics and intelligent laboratory systems, Elesevier publications, 2016, Vol (157), 140-151, https://doi.org/10.1016/j.chemolab.2016.07.009.
[6] Wang Baijie, Wang Xin, Chen Zhangxin. Spatial entropy based mutual information in hyperspectral band selection for supervised classification. Int J Numer Anal Model 2012;9(2):181–92.
[7] Information theory https: //www.cc.gatech.Edu/~ksubrama/teaching/cs4641/files/Information Theory.pdf
[8] Guo Baofeng, Gunn Steve, Damper Bob, Nelson James. Adaptive Band Selection for hyperspectral image fusion using mutual information. IEEE 2005. doi: https://doi.org/10.1109/ICIF.2005.1591913.
[9] Guo Baofeng, Gunn Steve, Damper Bob, Nelson James. Band Selection for hyperspectral image classification using mutual information. IEEE Geosci Remote Sens Lett 2006;3(4). doi: https://doi.org/10.1109/LGRS.2006.878240.
[10] Guo Baofeng, Gunn Steve, Damper Bob, Nelson James. Improving hyperspectral Band Selection by constructing an estimated reference map. J Appl Remote Sens 2014;8. doi: https://doi.org/10.1117/1.JRS.8.083692.
[11] Shijin Li, JianbinQiu, Xinxin Yang, Huan Liu, Dingsheng Wan, and Yuelong Zhu, "A novel approach to hyperspectral band selection based on spectral shape similarity analysis and fast branch and bound search", Engineering applications of artificial intelligence, Elsevier Publications, 2014, Vol (27), 241-250, https://doi.org/10.1016/j.engappai.2013.07.010.
[12] Yang Ce, Lee Won Suk, Gader Paul. Hyperspectral band selection for detecting different blueberry fruit maturity stages. Comput Electron Agric 2014;109:23–31. doi: https://doi.org/10.1016/j.compag.2014.08.009.
[13] Reshma R, Sowmya V, Soman KP. Dimensionality reduction using band selection technique for kernel based hyperspectral image classification. Procedia Comput Sci 2016;93:396–402. doi: https://doi.org/10.1016/j.procs.2016.07.226.
[14] Yao Liu, Hong xie, Yuehua Chen, Kezhu Tan, Liguo Wang and WuXie, "Stability analysis of hyperspectral band selections algorithms based onNeighborhood rough set theory for classification", Chemometrics and intelligent laboratory systems, Elsevier publications, 2017, Vol (169), 35-44, https://doi.org/10.1016/j.chemolab.2017.08.005.
[15] Ronglu Yang, Lifansu, Xibin Zhao, Hai Wan, and Jiaguang Sun, "Representative Band selection for hyperspectral image classification", J. Vis. Commun. Image R. 2017, Vol (48), 396-403, https://doi.org/10.1016/j.jvcir.2017.02.002.
[16] . Neurocomputing 2018;273:57–67. doi: https://doi.org/10.1016/j.neucom.2017.07.059.
[17] Camps-Valls G, Gomez-Chova L, Munoz-Mar J, Vila-Francs J, Calpe-Maravilla J. Composite kernels for hyperspectral image classification. IEEE Geosci. Remote

Sens. Lett. Jan. 2006;3(1):93–7. doi: https://doi.org/10.1109/LGRS.2005.857031.

[18] Mura MD, Benediktsson JA, Waske B, Bruzzone L. Morphological attribute profiles for the analysis of very high-resolution images. IEEE Trans. Geosci. Remote Sens. Oct. 2010;48(10):3747–62. doi: https://doi.org/10.1109/TGRS.2010.2048116.

[19] . IEEE Trans. Geosci. Remote Sensing 2011;49(10):3947–60. doi: https://doi.org/10.1109/TGRS.2011.2128330.

[20] Thomas M. Cover, Joy A. Thomas, ''Elements of Information Theory'', Chapter 2, ISBN 0-471-06259-6.

[21] Tschannerl, Julius, et al. ''MIMR-DGSA: Unsupervised hyperspectral band selection based on information theory and a modified discrete gravitational search algorithm.'' Information Fusion 51 (2019): 189-200

[22] Li Qiang, Wang Qi, Li Xuelong. An efficient clustering method for hyperspectral optimal band selection via shared nearest neighbor. Remote Sensing 2019;11(3):350.

[23] Xie Fuding et al. Unsupervised band selection based on artificial bee colony algorithm for hyperspectral image classification. Appl Soft Comput 2019;75:428–40.

[24] Pan B, Shi Z, Xu X. Analysis for the Weakly Pareto Optimum in Multiobjective-Based Hyperspectral Band Selection. IEEE Trans Geosci Remote Sens 2019;1–12. doi: https://doi.org/10.1109/tgrs.2018.2886853.

[25] Liu Y, Wu T, Yang J, Tan K, Wang S. Hyperspectral band selection for soybean classification based on information measure in FRS theory. Biosyst Eng 2019;178:219–32. doi: https://doi.org/10.1016/j.biosystemseng.2018.12.002.

[26] Chen W, Yang Z, Ren J, Cao J, Cai N, Zhao H, et al. MIMN-DPP: Maximum-Information and Minimum-Noise Determinantal Point Processes for Unsupervised Hyperspectral Band Selection. Pattern Recogn 2020;107213. doi: https://doi.org/10.1016/j.patcog.2020.107213.

[27] Varade D, Maurya AK, Dikshit O. Unsupervised band selection of hyperspectral data based on mutual information derived from weighted cluster entropy for snow classification. Geocarto International 2019;1–23. doi: https://doi.org/10.1080/10106049.2019.1665717.

[28] Wang Q, Zhang F, Li X. Optimal Clustering Framework for Hyperspectral Band Selection. IEEE Trans Geosci Remote Sens 2018;1–13. doi: https://doi.org/10.1109/tgrs.2018.2828161.

[29] Zhang F, Wang Q, Li X. Optimal Neighboring Reconstruction for Hyperspectral Band Selection. IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium. 2018. doi: https://doi.org/10.1109/igarss.2018.8517884.

[30] Rashwan S, Dobigeon N. A Split-and-Merge Approach for Hyperspectral Band Selection. IEEE Geosci Remote Sens Lett 2017;14(8):1378–82. doi: https://doi.org/10.1109/lgrs.2017.2713462.