Fused 3-D spectral-spatial deep neural networks and spectral clustering for hyperspectral image classification

Akrem Sellami, Ali Ben Abbes, Vincent Barra, Imed Riadh Farah

▶ **To cite this version:**

Akrem Sellami, Ali Ben Abbes, Vincent Barra, Imed Riadh Farah. Fused 3-D spectral-spatial deep neural networks and spectral clustering for hyperspectral image classification. Pattern Recognition Letters, 2020, 10.1016/j.patrec.2020.08.020 . hal-02927062

HAL Id: hal-02927062
https://hal.science/hal-02927062v1

Submitted on 21 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S016786552030324X
Manuscript_f8cdfa1378411f77e2087fbb1d35d1a0

Pattern Recognition Letters
journal homepage: www.elsevier.com

Fused 3-D spectral-spatial deep neural networks and spectral clustering for hyperspectral image classification

Akrem Sellami[a,b,**], Ali Ben Abbes[b,c], Vincent Barra[d], Imed Riadh Farah[b,e]

[a] *INRIA Nancy Grand Est, LORIA, CNRS, UMR 7503, Vandoeuvre-lès-Nancy, F54506, France*
[b] *University of Manouba, ENSI, RIADI laboratory, Campus Universitaire de la Manouba, 2010 Manouba, Tunisia*
[c] *Centre d'applications et de Recherches en Télédétection (CARTEL), Université de Sherbrooke, Sherbrooke, Canada*
[d] *Clermont Auvergne University, CNRS, UMR 6158, LIMOS, Clermont-Ferrand, F63000,France*
[e] *IMT Atlantique, ITI Department, Brest, F29238, France*

ABSTRACT

Recently, classification and dimensionality reduction (DR) have become important issues of hyperspectral image (HSI) analysis. Especially, HSI classification is a challenging task due to the high-dimensional feature space, with a large number of spectral bands, and a low number of labeled samples. In this paper, we propose a new HSI classification approach, which is called fused 3-D spectral-spatial deep neural networks for hyperspectral image classification. We propose an unsupervised band selection method to avoid the problem of redundancy between spectral bands and automatically find a set of groups C_k each one containing similar spectral bands. Moreover, the model uses the different groups of selected bands to extract spectral-spatial features in order to improve the classification rate. Each group is associated with a 3-D CNN model, which are then fused to improve the precision of classification. The main advantage of the proposed method is to keep the initial spectral-spatial features by automatically selecting relevant spectral bands, which improves the classification of HSI using a low number of labeled samples. Experiments on two real HSIs, Indian Pines and Salinas datasets, are performed to demonstrate the effectiveness of the proposed method. Results show that the proposed method reaches competitive good performances, and achieves better classification rates compared to various state-of-the-art techniques.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Nowadays, hyperspectral imaging sensors can acquire data from wavelength range in the spectrum, which is typically between 0.4 and 0.5μm. Hyperspectral image (HSI) can be seen as a 3-D cube (one spectral dimension, and two spatial dimensions), possibly containing hundreds of narrow and contiguous spectral bands with high spectral resolution. These hyperspectral data provide a rich spectral information, which can then be used for identification, or surface classification (Tu et al., 2019). Therefore, they have been used in spectral unmixing (Liu et al., 2010), anomaly detection (He et al., 2008), etc. All these applications require the classification of each hyperspectral pixel. However, the classification of HSI still faces some majors issues (Chen et al., 2017a), especially: i) the large number of spectral bands makes it difficult to get sufficient labeled training samples, producing the Hughes phenomenon (i.e, the classification rate can be deteriorated in high-dimensional space, when the number of samples is very limited) (Feng et al., 2016); ii) the spectral bands are very correlated; iii) spectral signatures may have high spatial variability. Therefore, dimensionality reduction can be applied as a preprocessing phase before HSI classification step, since it discards the redundant features and allows reducing the large number of contiguous spectral bands.

Usually, dimensionality reduction techniques in HSI analysis can be divided into two main approaches: feature extraction (Zhou et al., 2014) and band selection (Bai et al., 2015). Feature extraction (Ma et al., 2016a; Feng et al., 2018) seeks to transform high-dimensional feature space into a low-dimensional space through linear or non-linear projections, such as Principal component analysis (PCA) (Deepa and Thilagavathi, 2015),

**Corresponding author: Tel.: +33-753-908-573;
e-mail: akrem.sellami@inria.fr (Akrem Sellami)

© 2020 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

locality preserving projection (LPP) (Li et al., 2012) projection pursuit, or local discriminant embedding (Zhou et al., 2014). Alternatively, band selection aims to preserve the physical meaning of hyperspectral data by selecting the most relevant, informative, discriminative and distinctive spectral bands.

According to the availability of training samples, band selection can be supervised, unsupervised or semi-supervised (Li et al., 2016; Sellami et al., 2018). Supervised band selection selects discriminant spectral bands using the class separability of labeled training samples and class label (Yan et al., 2016). Unsupervised band selection evaluates the importance of spectral bands using clustering quality assessment (Martinez-Uso et al., 2007) and various statistical measures (Chang and Wang, 2006). Therefore, the aim is to select the most informative spectral bands based on band ranking (Datta et al., 2012) or band clustering (Xu et al., 2018).

Semi-supervised band selection finds relevant spectral bands by taking advantage of the information of labeled and unlabeled samples (Sellami et al., 2018). Usually, most semi-supervised band selection methods are based on manifold learning (Zhao et al., 2008) or hypergraph models (Bai et al., 2015).

In this paper, we propose a novel unsupervised band selection technique based on hierarchical clustering algorithm in order to preserve the physical meaning of hyperspectral data i.e., keep the initial spectral-spatial information without discarding or distorting the crucial original information. It aims to partition all spectral bands using the similarity criterion, which minimizes the high correlation between spectral bands. The aim is to enhance the classification accuracy.

2. Related work

Usually, HSI classification algorithms use spectral and spatial features to perform the classification. In this context, several issues have been presented in the literature (Ma et al., 2016b; Sellami et al., 2019), such as the high-dimensionality of HSI and the limited number of training samples. Therefore, in order to enhance the HSI classification accuracy, several works have been developed to consider the spectral-spatial features. In (Xia et al., 2015), Xia et al. have proposed an Extended morphological profiles (EMPs) to take into account the spectral and spatial features for classification. Gabor filtering (GF) model has been introduced to extract spatial features, i.e., textures and edges (Chen et al., 2017b). Some approaches also used support vector machine (SVM) (Sellami and Farah, 2016; Li et al., 2017). All these studies cannot detect all spatial properties of the objects (Zhao and Du, 2016).

Several approaches based on deep learning (DL) have recently been developed for spectral-spatial HSI classification, and showed their high effectiveness and performance (Han et al., 2018; Sellami et al., 2019). Many techniques have indeed been developed to fuse the dimensionality reduction paradigm with DL models in order to improve the HSI classification. The logistic regression (LR), PCA, and stacked autoencoders (SAE) have for example been combined for HSI classification, where the SAE were applied for the spectral-spatial features extraction (Chen et al., 2014).

Deep belief networks (DBN) have also been used for HSI classification (Chen et al., 2015). In (Chen et al., 2015), authors have proposed a hybrid framework of PCA and DBN. Zhou et al as for them proposed in (Zhou et al., 2017) a novel approach based on band selection and DBN for HSI classification. However, DBN and SAE, which aim to extract the deep features hierarchically in a layer training, cannot extract the spatial features from HSI because the training samples have to be flattened into a 1-D vector before training. As a result, spatial information is neglected by flattened training samples.

Quite recently, convolutional neural networks (CNN) (Lee and Kwon, 2017) have been used for HSI spectral-spatial classification. CNN can extract the spatial features from HSI without any flattening of the training samples. A CNN was used in (Han et al., 2018) to extract the spatial features from HSI, giving satisfactory results for the classification. An approach based on PCA, CNN and LR was proposed in (Zhao and Du, 2016) for HSI classification. Also, Chen et al. (Chen et al., 2017a) have developed a method for HSI classification, combining Gabor filters (GF) with convolutional filters to mitigate the problem of overfitting. In (Sellami et al., 2019), a novel HSIs classification method was proposed, which combined the adaptive dimensionality reduction and semi-supervised 3-D Convolutional Neural Network (3-D CNN). It can extract the spatial and spectral features of HSI simultaneously. 3-D CNN needs a very large number of training samples in the training phase to obtain appropriate weights between the different nodes. Unfortunately, the number of training samples in the HSI is very limited, which degrades the classification performance of most supervised classification approaches.

3. Proposed methodology

We propose to improve the classification of HSI by performing a band clustering to avoid the high redundancy of spectral bands and fusing a multiple 3-D CNN to keep the spectral-spatial features. Our method can be decomposed into three steps; (i) Band clustering, (ii) Spectral-spatial feature extraction from k-clusters of spectral bands with 3-D CNN and, (iii) Classification based on the fusion of view-pooling. Fig. 1 illustrates the proposed methodology.

3.1. Band clustering based on spectral clustering (BCSC)

The first phase of the proposed approach, i.e., band clustering based on spectral clustering (BCSC), aims to find informative and distinctive spectral bands from the original HSI. Formally, let $X = [x_1, ..., x_N] \in \mathbb{R}^{M \times N}$ be an HSI, where $x_i \in \mathbb{R}^M$ represents a spectral vector of a pixel $i \in \{1, 2, ..., N\}$ and M is the number of spectral bands $B = \{B_1, B_2, .., B_M\}$. Each column X_j of X^t corresponds to a spectral band j. The proposed unsupervised band selection method seeks to select a reduced number $d \ll M$ from all spectral bands based on mutual information (MI).

Mutual information is the measure of independence between spectral bands, which can be considered as a correlation measure. In other words, MI quantifies the statistical dependence between all spectral bands or how much a spectral band can

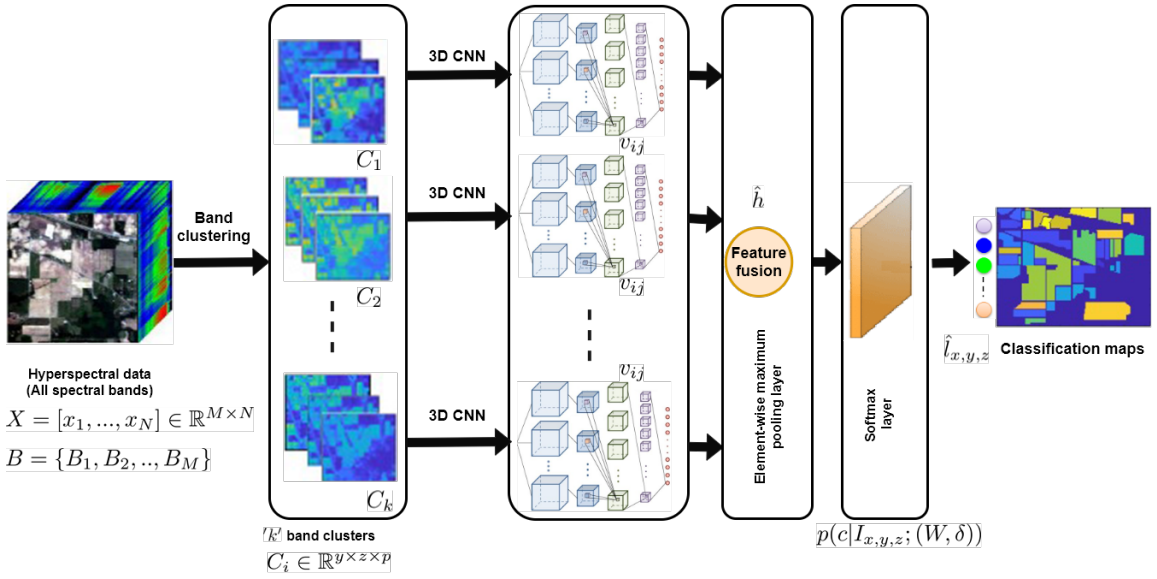

Fig. 1. Flowchart of the proposed methodology

predict another one. Therefore, considering a spectral band $\{B\}$, $MI(B_k, B_l)$ is defined as:

$$MI(B_k, B_l) = H(B_k) + H(B_l) - H(B_k, B_l) \quad (1)$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} p(b_i^k, b_j^l) \log \frac{p(b_i^k, b_j^l)}{p(b_i^k) p(b_j^l)} \quad (2)$$

where $H(B_k, B_l)$ is the joint entropy. Furthermore, the similarity $SIM(B_k, B_l)$ between two spectral bands can be expressed as

$$SIM(B_k, B_l) = \frac{2MI(B_k, B_l)}{H(B_k) + H(B_l)} \quad (3)$$

This similarity function is symmetric and non-negative, and the corresponding similarity matrix is denoted by $S = (s_{ij})_{i,j} = 1, ..., M$.

In order to select the most relevant spectral bands, i.e., informative and distinctive bands, we used the spectral clustering technique (Ng et al., 2002), which seeks to cluster data points using the eigenvalues of the similarity matrix. Therefore, the proposed band clustering model (see Fig. 2) constructs a similarity graph $G = (V, E, W)$, where V is a vertex, i.e. a spectral band, E is an edge (E exists if similarity $S(.) > 0$), and W is the edge weights matrix (similarity $S(.)$). Then, we compute the first k eigenvectors $v_1, ..., v_k$ of the graph Laplacian matrix $L = D - W$, where k is the number of clusters, i.e., the set of spectral band groups, and D is the degree matrix of S. The generalized eigenproblem can be defined as

$$Lv = \lambda Dv \quad (4)$$

where λ is an eigenvalue. The algorithm then clusters the data points with the k-means algorithm into clusters $C_1, ..., C_i, ..., C_k$, where each cluster C_i contains similar spectral bands. Hence, spectral bands in different clusters are dissimilar. Supervised band selection methods need a priori information to select informative spectral bands, where the labeled samples are not always available and very limited. Unsupervised techniques can

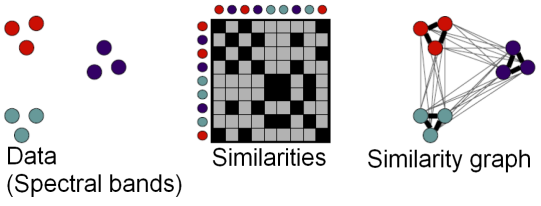

Fig. 2. Construction of similarity graph of spectral bands

select redundant or correlated bands, reducing the classification rate. Our BCSC aims to automatically find relevant spectral bands by grouping similar bands using hierarchical clustering based on mutual information measures without discarding the initial spectral-spatial information of HSI, i.e., keep the spectral signatures of pixels. It allows then to select informative, discriminative, and distinctive spectral bands by minimizing the high correlation between spectral bands without requiring labeled samples, such as a class label. The proposed BCSC model is detailed in Algorithm 1.

3.2. 3-D spectral-spatial Convolutional Neural Network

Let each group of similar spectral bands be denoted by $C_i \in \mathbb{R}^{y \times z \times p}$, where y and z are the spatial dimensions (width and height) and p is the number of selected bands of each cluster C_i. Therefore, each pixel is considered as 3D neighboring pixel of size $P \in \mathbb{R}^{s \times s \times p}$ from C_i, centered at the spatial location (α, β), covering the $s \times s$ spatial window. The number of created 3D neighboring pixels n from C_i can be obtained by $(y - s + 1) \times (z - s + 1)$. Therefore, the 3D-patch $P_{\alpha,\beta}$ covers the height from $\beta - (s-1)/2$ to $\beta + (s-1)/2$, width from $\alpha-(s-1)/2$ to $\alpha+(s-1)/2$, and the number of selected bands for each cluster C_i. Fig. 3 shows the architecture of the 3-D CNN for HSI spectral-spatial classification. Each layer of 3-D CNN is composed of both a 3-D convolution and a pooling operation. The input of the 3-D CNN is a group of selected bands C_i obtained with the BCSC method, where the patch is a neighboring

Algorithm 1 Band Clustering based on Spectral Clustering (BCSC)

Input: $B_1, ..., B_M, B_i \in \mathbb{R}^N, i \in [\![1 \cdots M]\!]$
k: number of clusters
Output: Clusters $A_1, ..., A_k$ with $A_i = \{j | y_j \in C_i\}$
for $i \leftarrow 1, M$ **do** ▷ M is the number of spectral bands
for $j \leftarrow i, M$ **do**
$S_{i,j} \leftarrow SIM(B_i, B_j)$ ▷ Compute the similarity matrix
$S_{j,i} \leftarrow S_{i,j}$
end for
end for
Build similarity graph $G = (V, E, W)$ using the similarity matrix $S(.)$
Compute the Laplacian graph L
Compute the first k-eigenvectors $v_1, ..., v_k$ of L
Let $V \in \mathbb{R}^{M \times k}$ be the matrix containing the vectors $v_1, ..., v_k$ as columns.
for $i \leftarrow 1, D$ **do**
Let $y_i \in \mathbb{R}^k$ be the vector corresponding to the i-th row of V
end for
cluster the data points $(y_i)_{i=1,...,n}$ in $\mathbb{R}^k$ using the k-means algorithm into clusters $C_1, ..., C_i, ..., C_k$.

pixels $P \in \mathbb{R}^{s \times s \times p}$. Generally, the 1-D convolution operation

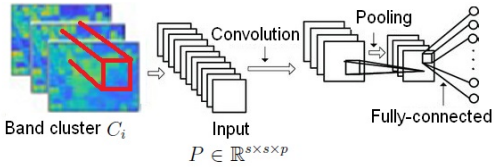

Fig. 3. Architecture of the 3-D CNN-based HSI classification

aims to extract spectral features while 2-D convolution aims to extract only spatial features of each pixel. 3-D CNN convolution extracts both spectral and spatial features from a cluster of spectral bands C_i. Formally, the 3-D convolution operation can be defined as:

$$v_{ij}^{xyz} = \phi\left(\sum_h \sum_{l=0}^{L_{i-1}} \sum_{k=0}^{K_{i-1}} \sum_{m=0}^{R_{i-1}} v_{(i-1)h}^{(x+l)(y+k)(z+m)} * w_{ijh}^{lkm} + \delta_{ij}\right) \quad (5)$$

where v_{ij}^{xyz} is the value of the neuron at position (x, y, z), h is the index of the feature map in the $(i-1)^{th}$ layer connected to the current j^{th} feature map. w_{ijh}^{lkm} is the weight of kernel at position (l, k, m) connected to the h^{th} feature map, with K_i and L_i are the height and width of the convolution kernel. R_i is the size of the 3D kernel along the spectral dimension, and δ_{ij} is the bias.

3.3. Fusion of 3D-CNNs with element-wise maximum pooling layer

In order to fuse the features maps v_{ik} obtained with $\{3-DCNN_1, 3-DCNN_2, ..., 3-DCNN_k\}$, the main goal is to aggregate all pooling layers with an element-wise maximum pooling layer operation. The motivation to use this operator is to improve the classification of HSI without redundant spectral-spatial features. Formally, the element-wise maximum pooling layer can be obtained as follows

$$\hat{h} = MAX\left(ReLU\left\{\sum_h \sum_{l=0}^{L_{i-1}} \sum_{k=0}^{K_{i-1}} \sum_{m=0}^{R_{i-1}} v_{(i-1)h}^{(x+l)(y+k)(z+m)} * w_{ijh}^{lkm} + \delta_{ij}\right\}\right) \quad (6)$$

To perform the spectral-spatial classification, we applied the *softmax(.)* function to the top layer of the proposed fused 3-D CNN. Usually, *softmax(.)* aims to measure the correlation between a reference value (true value) and an output value (predicted value) by a probability score. Formally, let W and δ denote all the parameters of our fused 3-D CNN model. The output of our network f with trainable parameters (W, δ), for a given input patch $I_{x,y,z}$ can be formulated as follows:

$$f(I_{x,y,z}; (W, \delta)) = W\hat{h} \quad (7)$$

The fused 3-D CNN model is trained by transforming the scores $f_c(I_{x,y,z}; (W, \delta))$ of each class of interest $c \in \{1, ..., T\}$ into conditional probability by applying the following *softmax(.)* function:

$$p(c|I_{x,y,z}; (W, \delta)) = \frac{e^{f_c(I_{x,y,z}; (W,\delta))}}{\sum_{t\in\{1,...,T\}} e^{f_t(I_{x,y,z}; (W,\delta))}} \quad (8)$$

Moreover, the model is trained by maximizing the likelihood of training samples, i.e., the parameters (W, δ) are learned by minimizing the negative log-likelihood based on training samples set

$$L(W, \delta) = -\sum_{I_{x,y,z}} ln\ p(l_{x,y,z}|I_{x,y,z}; (W, \delta)) \quad (9)$$

where $l_{x,y,z}$ is the correct pixel label at position (x, y) of the spectral band z. In order to optimize the objective function, the stochastic gradient descent (SGD) with back-propagation algorithm was applied. Finally, at the testing time, the model can predict the label of the pixel $\hat{l}_{x,y,z}$ located at position (x, y) in band B_z using

$$\hat{l}_{x,y,z} = \underset{t\in\{1,...,T\}}{\text{argmax}}\ p(c|I_{x,y,z}; (W, \delta)) \quad (10)$$

4. Experimental Results

4.1. HSI Description

In this paper, two real HSIs are processed to evaluate the performance and effectiveness of the proposed method, including the Indian Pines Dataset and the Salinas Dataset [1].
The Indian Pines HSI collected by the Airborne Visible/ Infrared Imaging Spectrometer (AVIRIS) sensor, which represents the north-western Indiana. It consists of 145×145 pixels with a spatial resolution of 20 m per pixel and 220 spectral bands in the wavelength range from 0.4 to 2.5 μm. The

[1] `http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes`

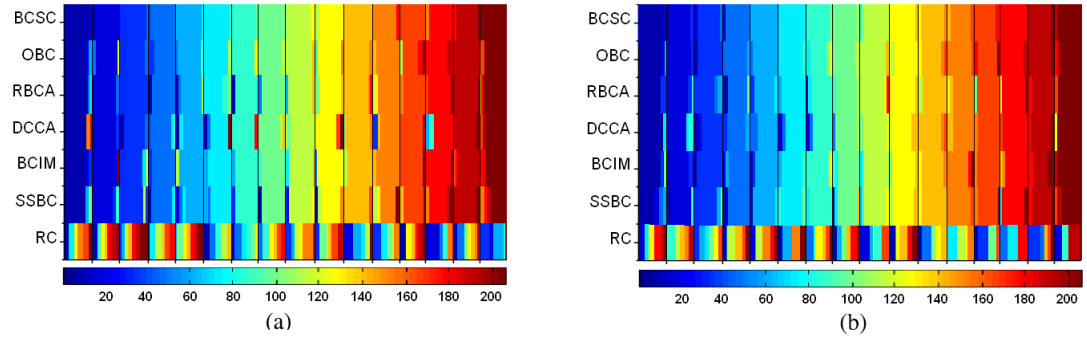

Fig. 4. Groups of spectral bands obtained with different band clustering methods. (a) Indian Pines, (b) Salinas (k = 16 clusters). The blue color denotes the lower spectral band indexes and the red color indicates higher spectral band indexes

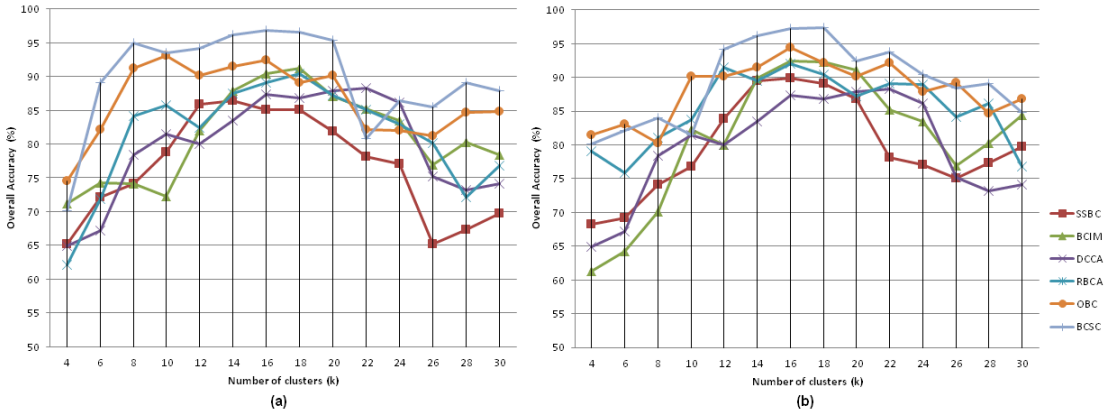

Fig. 5. Comparison of different band clustering techniques. (a) Classification results on Indian Pines. (b) Classification results on Salinas.

ground truth contains 16 classes. The Salinas image collected by the AVIRIS sensor over Salinas, California, which consists of 512 × 217 pixels with a spatial resolution of 3.7 m per pixel and 224 spectral bands. The ground truth contains 16 classes.

4.2. Comparison to other methods

In order to evaluate our algorithm, we compared its results with the following recent band clustering methods: semi-supervised band clustering (SSBC) (Su et al., 2011), band clustering using information measures (BCIM) (Martinez-Uso et al., 2007), dual-clustering band by contextual analysis (DCCA) (Yuan et al., 2015), ranking-based band clustering approach (RBCA) (Jia et al., 2015), and optimal band clustering (Wang et al., 2018) (OBC). Furthermore, we compared our proposed method with various recent and relevant methods based on DL for HSI classification, including DL using pseudo labels (PL) (Wu and Prasad, 2018), dual-strategy sample selection (DS) (Fang et al., 2018), convolutional neural network (CNN) (Liu et al., 2017), graph convolutional neural network (GCN) (Qin et al., 2018), multi-decisions labeling (MDL) (Ma et al., 2016b), 3D deeo learning (3DL) (Hamida et al., 2018), Fast 3D CNN (FST) (Ahmad, 2020), Multi-scale 3D convolutional neural network (MSC) (He et al., 2017), and 3D convolutional neural network (3DCNN) Li et al. (2017) .

Table 1. Band clustering performances using AMI, HI, and ARI

Method	Score value			Score value		
	AMI	HI	ARI	AMI	HI	ARI
SSBC	0.743	0.797	0.595	0.750	0.801	0.598
BCIM	0.781	0.828	0.655	0.805	0.847	0.697
DCCA	0.757	0.808	0.601	0.731	0.762	0.582
RBCA	0.763	0.814	0.612	0.771	0.820	0.633
OBC	0.827	0.864	0.743	0.853	0.884	0.787
BCSC	0.897	0.919	0.867	0.926	0.942	0.905
RC	0.555	0.650	0.379	0.520	0.622	0.345
HSI	*Indian Pines*			*Salinas*		

4.3. Performance Evaluation Metrics and Parameters Setting

It is important to set the evaluation parameters before spectral-spatial classification. Therefore, we fixed $k = 2, \ldots, 30$ clusters of spectral bands from the two HSIs. To analyze each band clustering method, we used three clustering performance metrics, including adjusted mutual information (AMI), homogeneity index (HI), and the adjusted Rand index (ARI). Moreover, we evaluated all band clustering methods on the classification using the 3-D CNN model (Sellami et al., 2019). We also set the learning rate lr to 10^{-3} and the batch size to 200 after several tests. The diagonal degree matrix was fixed with a window size $p = 3$, considering the 8-neighbors spectral pixels, and the number of epochs was set to 400 for Indian Pines

Table 2. Classification performances (PL, DS, GCN, CNN, MDL, 3DL,FST,MSC,3DCNN, and Fused 3-D CNN (Ours)) : Indian Pines HSI (k = 16)

Class	Samples		Method									
	Train	Test	PL	DS	CNN	GCN	MDL	3DL	FST	MSC	3DCNN	Ours
Alfalfa (C1)	5	49	88.05	94.18	**96.40**	94.02	95.01	95.22	95.95	94.15	95.47	96.32
Building-G (C2)	38	342	92.03	91.11	96.28	95.07	92.04	94.61	96.03	92.51	**96.51**	96.26
Corn (C3)	23	211	94.21	92.01	95.03	93.04	92.12	95.81	94.89	94.51	90.12	**96.97**
Corn-M (C4)	83	751	94.92	89.05	93.08	**97.41**	96.95	94.91	95.48	93.96	94.36	96.99
Corn-N (C5)	143	1291	94.95	93.28	95.61	95.48	96.01	96.43	96.32	95.74	92.25	**97.42**
Grass-P-W (C6)	5	21	86.17	88.51	94.02	95.18	95.03	94.96	95.21	89.41	94.91	**95.52**
Grass-P (C7)	50	447	91.25	81.35	95.10	**97.28**	94.53	93.39	96.74	96.12	97.01	97.16
Grass-T (C8)	75	672	92.25	94.02	**97.10**	96.41	93.05	94.21	97.02	94.95	96.31	96.89
Hay-W (C9)	49	440	95.10	92.03	92.01	96.25	95.43	92.98	93.69	91.73	92.06	**96.56**
Oats (C10)	2	18	92.08	89.92	96.98	94.85	97.02	93.52	96.25	90.14	94.85	**97.07**
Soybeans-C (C11)	62	552	87.25	89.09	95.92	96.74	96.94	96.21	94.31	95.22	96.07	**97.31**
Soybeans-M (C12)	247	2221	91.65	91.42	92.05	94.91	95.92	93.22	93.91	91.64	92.82	**96.84**
Soybeans-N (C13)	97	871	89.22	87.42	94.25	94.18	95.35	94.25	95.94	88.49	93.74	**96.21**
Stone-S (C14)	10	85	93.25	92.83	95.90	94.65	95.04	92.98	97.08	95.10	94.97	**97.26**
Wheat (C15)	21	191	92.76	93.54	95.24	95.61	94.05	93.72	96.14	94.42	96.21	**96.61**
Woods (C16)	130	1164	93.67	93.24	95.21	96.22	94.93	94.03	96.87	92.33	95.61	**96.95**
OA (%)	–	–	92.34	91.74	94.29	95.42	95.65	93.84	95.83	94.28	94.81	**96.98**
AA (%)	–	–	92.22	91.38	94.83	95.28	95.39	93.67	95.68	94.08	94.62	**96.85**
$kappa \times 100$	–	–	92.28	91.45	94.15	95.47	92.40	93.72	95.79	94.15	94.77	**96.89**
Time (s)	–	–	381	362	287	322	254	485	321	389	366	274

and 300 for Salinas. Moreover, for the two HSIs datasets, we randomly chose 10% of the samples per class for training and the rest for testing.

In order to get stable classification results, we repeated the experiments 10 times with different sets of training samples. Three performance metrics were computed to evaluate the classification accuracy: kappa coefficient (*kappa*), overall accuracy (OA) and average accuracy (AA).

4.4. Band Clustering Performance Evaluation

The obtained spectral band groups with the proposed BCSC and other existing band clustering techniques, including CBC, RBCA, DCCA, BCIM, SSBC, and RC are reported in Fig. 4. Moreover, we compare also all band clustering techniques with a random clustering method denoted by RC. The aim is to assess the effectiveness of the clustering results. For both HSIs, i.e., Indian Pines and Salinas, 16 clusters of spectral bands are extracted and presented in Fig. 4(a) and (b). The color bars at the bottom of Fig. 4 give the correspondence between the spectral bands Ids and their corresponding wavelengths, which help to analyze the extracted clusters from each HSI produced by the various band clustering methods. The visual interpretation of Fig. 4 shows the potential and the efficiency of the BCSC method for the clustering of relevant spectral bands. For both HSIs, clusters computed with BCSC are more homogeneous than the one obtained with other band clustering methods. Moreover, poorly defined BCSC spectral bands belong to adjacent clusters, i.e., selected spectral bands overlap between two adjacent regions of the spectrum. Groups processed with all but BCSC band clustering methods are heterogeneous and the selected spectral bands are not adjacent, i.e., do not belong to the neighboring regions of the spectrum. We reported also in the table 1 score values using AMI, HI, and ARI. The best score values are obtained using BCSC, where e.g., HI is equal to 0.919 for Indian Pines and 0.942 for Salinas. Furthermore, according to the comparative study, we can notice that the band clustering methods, which uniformly better cluster the entire region of the spectrum, produce better classification results.

4.5. Comparison of band clustering methods using 3-D CNN

The performances of band clustering techniques, i.e., SSBC, BCIM, DCCA, RBCA, OBC and BCSC are studied in this section. We applied the 3-D CNN model on the clusters of selected bands to perform the spectral-spatial classification. Fig. 5 reports the obtained classification using different band clustering techniques.

The OA measures show that our BCSC method outperforms SSBC, BCIM, DCCA, RBCA, OBC: the OAs with BCSC are 96.21% for Indian Pines, and 97.09% for Salinas, with k = 16. Furthermore, the OBC approach gave satisfactory classification results, with the best OA value(92.45% for Indian Pines and 92.86% for Salinas). The SSBC OA's were respectively equal to 85.14% and 89.94%. The proposed approach BCSC succeeded in selecting relevant spectral bands for both datasets, i.e. each cluster contained the most similar bands, which improve the spectral-spatial classification by avoiding the problem of redundancy and the high correlation.

4.6. Classification Performance

In this section, we compare the classification results obtained with our method with other DL-based techniques, including, PL, DS, CNN, GCN, MDL, 3DL, FST, MSC, and 3DCNN.

Table 2 reports the OA measures computed for the Indian Pines HSI. Based on these results, fused 3D-CNN gave better

Table 3. Classification performances (PL, DS, GCN, CNN, MDL, 3DL, FST, MSC, 3DCNN, and Fused 3-D CNN (Ours)) : Salinas HSI (k = 16)

Class	Samples		Method									
	Train	Test	PL	DS	CNN	GCN	MDL	3DL	FST	MSC	3DCNN	Ours
Brocoli-G-W-1 (C1)	200	1809	95.28	96.15	96.92	97.17	95.23	95.99	96.89	96.22	96.71	**97.01**
Brocoli-G-W-2 (C2)	372	3354	95.03	95.91	96.32	**96.82**	96.10	96.05	96.74	95.03	95.99	96.68
Fallow (C3)	197	1779	94.25	95.21	**97.09**	96.23	96.81	95.82	96.89	93.48	94.49	96.88
Fallow-R-P (C4)	139	1255	94.92	94.61	95.92	96.08	96.21	96.11	95.21	95.72	**96.27**	95.98
Fallow-S (C5)	267	2411	94.81	96.28	95.99	96.42	96.12	96.23	96.12	96.41	96.33	**96.75**
Stubble (C6)	395	3564	96.03	95.24	96.14	96.23	95.61	95.89	97.44	96.98	97.24	**97.88**
Celery (C7)	357	3222	96.24	96.12	95.96	96.45	95.08	96.17	97.12	96.94	97.12	**97.51**
Grapes-U (C8)	1127	10144	96.96	96.89	95.96	**97.30**	96.85	97.02	96.92	96.26	96.74	96.94
Soil-V-D (C9)	620	5583	96.01	95.53	95.98	96.49	97.19	96.98	97.03	97.18	97.02	**97.26**
Corn-S-G-W (C10)	327	2951	95.99	94.12	96.23	96.18	96.01	96.52	96.41	96.42	96.16	**96.73**
Lettuce-R-4wk (C11)	106	962	95.41	95.32	96.92	96.26	97.18	97.15	95.40	95.84	**97.25**	97.10
Lettuce-R-5wk (C12)	192	1735	94.18	94.72	95.99	94.83	94.87	96.11	94.95	94.31	95.89	**96.28**
Lettuce-R-6wk (C13)	91	825	95.32	96.13	95.98	96.24	96.28	95.77	96.40	96.22	97.14	**97.21**
Lettuce-R-7wk (C14)	107	963	96.48	94.73	95.98	**97.54**	96.19	96.99	96.85	95.65	97.28	97.16
Vineyard-U (C15)	726	6542	95.28	95.96	97.25	97.31	96.09	97.02	96.99	96.69	97.10	**97.28**
Vineyard-V-T (C16)	180	1627	96.74	96.07	96.62	97.00	97.41	97.32	96.84	97.32	96.94	**97.62**
OA (%)	–	–	95.94	96.89	97.55	96.84	96.93	96.72	96.75	96.42	96.97	**97.65**
AA (%)	–	–	94.72	96.87	96.20	96.63	96.82	96.54	96.69	96.24	96.81	**97.52**
$kappa \times 100$	–	–	95.81	96.60	96.48	96.71	96.88	96.65	96.93	96.35	96.86	**97.49**
Time (s)	–	–	321	301	195	210	245	402	364	394	371	289

classification rates, compared to other DL-based approaches. with OA (resp. AA and $kappa$) equal to 96.98% (resp. 96.85%, and 96.89%). However, the GCN model gave the best classification performance for the following two classes: 'Corn-M' and 'Grass-P', with an OA of 97.41% and 97.28%, respectively, and the CNN gave better classification rates for these three classes: 'Alfafa-C', and 'Grass-T', with an OA of 96.40%, and 97.10%, respectively. For the remaining 11 out of 16 Indian Pines classes, fused 3-D CNN had the best performance indexes. We find that the CNN and GCN models can give better classification rates for some vegetation classes than the fused 3-D CNN model. Furthermore, the 3-D CNN has better classified some urban areas, e.g, the 'Building-G' class, than the rest of classifiers. However, our algorithm gave better results for most of land use, e,g, woods, vegetation, etc. This proves that the band clustering method selects representative spectral bands throughout the spectrum and thus improves the spectral-spatial classification. The selected bands for each cluster can be considered as informative and discriminative for each specific class.

For the Salinas HSI, the OA measure obtained with the fused 3-D CNN model is 97.65%, AA is 97.52%, and $kappa$ is 97.39% as shown in Table 3. Also, we can observe that the fused 3-D CNN approach gave better classification rates for 13 out of 16 classes, with a number of clusters of spectral bands k = 16. Furthermore, most classification rates were greater than 97%. Here again, the GCN method gave the best classification performances for the following three classes: 'Brocoli-G-W-2', 'Grapes-U', and 'Lettuce-R-7wk', with an OA of 96.82%, 97.30% and 97.54%, respectively. 3-D CNN provided better classifications for the two classes: 'Fallow-R-P', and 'Lettuce-R-7wk', with OA of $96,27\%$, and 97.25%, respectively. And finally, the CNN model gave a higher classification rates for a one class: 'Fallow', with an OA of 97.09%.

According to the classification results, we can state that the proposed approach fused 3-D CNN is more effective than many other state-of-the-art DL-based methods. Also, it can select the relevant spectral bands of HSI with limited labeled samples by preserving the physical meaning of data, i.e., the spectral and spatial information simultaneously, and it also provides a good spectral-spatial classification by exploiting the semi-supervised graph model. The CNN and GCN models can give better OAs for some vegetation classes. The MDL method has better classified some vegetation classes. In the case where other DL-based methods performed best OA for some classes, our method needed only a few labeled samples for classification with low computation time. In general, we can affirm that band clustering and the fused 3-D CNN models can find most relevant spectral bands and get a high classification rate by preserving the spectral-spatial features.

5. Conclusion

In this paper, we proposed a novel method for HSI spectral-spatial classification based on spectral band clustering and 3-D convolutional neural network (3-D CNN). The proposed method seeks to find the most discriminative, informative, and distinctive, spectral bands using spectral clustering. Moreover, a fused 3-D CNN model is proposed to extract the deep spectral and spatial hyperspectral features from the different clusters, which contain similar spectral bands. The main advantage of the proposed model is to jointly preserve the spatial and spectral information using 3-D CNN, which improves the classification performances of HSI. Experimental results have shown that the

method is more effective compared to state-of-the-art DL-based classification methods, including CNN-based methods. As future work, we will develop a multimodal graph convolutional network (MGCN) to incorporate the spatial feature in order to improve the spectral-spatial classification task.

References

Ahmad, M., 2020. A fast 3d cnn for hyperspectral image classification. arXiv preprint arXiv:2004.14152 .

Bai, X., Guo, Z., Wang, Y., Zhang, Z., Zhou, J., 2015. Semisupervised hyperspectral band selection via spectral–spatial hypergraph model. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8, 2774–2783.

Chang, C.I., Wang, S., 2006. Constrained band selection for hyperspectral imagery. Geoscience and Remote Sensing, IEEE Transactions on 44, 1575–1585.

Chen, P., Jiao, L., Liu, F., Gou, S., Zhao, J., Zhao, Z., 2017a. Dimensionality reduction of hyperspectral imagery using sparse graph learning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10, 1165–1181.

Chen, Y., Lin, Z., Zhao, X., Wang, G., Gu, Y., 2014. Deep learning-based classification of hyperspectral data. IEEE Journal of Selected topics in applied earth observations and remote sensing 7, 2094–2107.

Chen, Y., Zhao, X., Jia, X., 2015. Spectral–spatial classification of hyperspectral data based on deep belief network. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8, 2381–2392.

Chen, Y., Zhu, L., Ghamisi, P., Jia, X., Li, G., Tang, L., 2017b. Hyperspectral images classification with gabor filtering and convolutional neural network. IEEE Geoscience and Remote Sensing Letters 14, 2355–2359.

Datta, A., Ghosh, S., Ghosh, A., 2012. Clustering based band selection for hyperspectral images, in: Communications, Devices and Intelligent Systems (CODIS), 2012 International Conference on, IEEE. pp. 101–104.

Deepa, P., Thilagavathi, K., 2015. Feature extraction of hyperspectral image using principal component analysis and folded-principal component analysis, in: Electronics and Communication Systems (ICECS), 2nd International Conference on, IEEE. pp. 656–660.

Fang, B., Li, Y., Zhang, H., Chan, J.C.W., 2018. Semi-supervised deep learning classification for hyperspectral image based on dual-strategy sample selection. Remote Sensing 10, 574.

Feng, J., Jiao, L., Liu, F., Sun, T., Zhang, X., 2016. Unsupervised feature selection based on maximum information and minimum redundancy for hyperspectral images. Pattern Recognition 51, 295–309.

Feng, J., Liu, L., Cao, X., Jiao, L., Sun, T., Zhang, X., 2018. Marginal stacked autoencoder with adaptively-spatial regularization for hyperspectral image classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11, 3297–3311.

Hamida, A.B., Benoit, A., Lambert, P., Amar, C.B., 2018. 3-d deep learning approach for remote sensing image classification. IEEE Transactions on geoscience and remote sensing 56, 4420–4434.

Han, M., Cong, R., Li, X., Fu, H., Lei, J., 2018. Joint spatial-spectral hyperspectral image classification based on convolutional neural network. Pattern Recognition Letters .

He, L., Pan, Q., Di, W., Li, Y., 2008. Anomaly detection in hyperspectral imagery based on maximum entropy and nonparametric estimation. Pattern Recognition Letters 29, 1392–1403.

He, M., Li, B., Chen, H., 2017. Multi-scale 3d deep convolutional neural network for hyperspectral image classification, in: 2017 IEEE International Conference on Image Processing (ICIP), IEEE. pp. 3904–3908.

Jia, S., Tang, G., Zhu, J., Li, Q., 2015. A novel ranking-based clustering approach for hyperspectral band selection. IEEE Transactions on Geoscience and Remote Sensing 54, 88–102.

Lee, H., Kwon, H., 2017. Going deeper with contextual cnn for hyperspectral image classification. IEEE Transactions on Image Processing 26, 4843–4855.

Li, L., Ge, H., Gao, J., 2017. A spectral-spatial kernel-based method for hyperspectral imagery classification. Advances in Space Research 59, 954–967.

Li, S., Zheng, Z., Wang, Y., Chang, C., Yu, Y., 2016. A new hyperspectral band selection and classification framework based on combining multiple classifiers. Pattern Recognition Letters 83, 152–159.

Li, W., Prasad, S., Fowler, J.E., Bruce, L.M., 2012. Locality-preserving dimensionality reduction and classification for hyperspectral image analysis. IEEE Transactions on Geoscience and Remote Sensing 50, 1185–1198.

Liu, B., Yu, X., Zhang, P., Tan, X., Yu, A., Xue, Z., 2017. A semi-supervised convolutional neural network for hyperspectral image classification. Remote Sensing Letters 8, 839–848.

Liu, L., Wang, B., Zhang, L., 2010. An approach based on self-organizing map and fuzzy membership for decomposition of mixed pixels in hyperspectral imagery. Pattern Recognition Letters 31, 1388–1395.

Ma, X., Wang, H., Geng, J., 2016a. Spectral–spatial classification of hyperspectral image based on deep auto-encoder. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9, 4073–4085.

Ma, X., Wang, H., Wang, J., 2016b. Semisupervised classification for hyperspectral image based on multi-decision labeling and deep feature learning. ISPRS Journal of Photogrammetry and Remote Sensing 120, 99–107.

Martinez-Uso, A., Pla, F., Sotoca, J.M., García-Sevilla, P., 2007. Clustering-based hyperspectral band selection using information measures. IEEE Transactions on Geoscience and Remote Sensing 45, 4158–4171.

Ng, A.Y., Jordan, M.I., Weiss, Y., 2002. On spectral clustering: Analysis and an algorithm, in: Advances in neural information processing systems, pp. 849–856.

Qin, A., Shang, Z., Tian, J., Wang, Y., Zhang, T., Tang, Y.Y., 2018. Spectral-spatial graph convolutional networks for semisupervised hyperspectral image classification. IEEE Geoscience and Remote Sensing Letters .

Sellami, A., Farah, I.R., 2016. High-level hyperspectral image classification based on spectro-spatial dimensionality reduction. Spatial Statistics 16, 103–117.

Sellami, A., Farah, M., Farah, I.R., Solaiman, B., 2018. Hyperspectral imagery semantic interpretation based on adaptive constrained band selection and knowledge extraction techniques. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11, 1337–1347.

Sellami, A., Farah, M., Farah, I.R., Solaiman, B., 2019. Hyperspectral imagery classification based on semi-supervised 3-d deep neural network and adaptive band selection. Expert Systems with Applications 129, 246–259.

Su, H., Yang, H., Du, Q., Sheng, Y., 2011. Semisupervised band clustering for dimensionality reduction of hyperspectral imagery. IEEE Geoscience and Remote Sensing Letters 8, 1135–1139.

Tu, B., Li, N., Fang, L., Yang, X., Wu, J., 2019. Hyperspectral image classification with a class-dependent spatial–spectral mixed metric. Pattern Recognition Letters 123, 16–22.

Wang, Q., Zhang, F., Li, X., 2018. Optimal clustering framework for hyperspectral band selection. IEEE Transactions on Geoscience and Remote Sensing 56, 5910–5922.

Wu, H., Prasad, S., 2018. Semi-supervised deep learning using pseudo labels for hyperspectral image classification. IEEE Transactions on Image Processing 27, 1259–1270.

Xia, J., Dalla Mura, M., Chanussot, J., Du, P., He, X., 2015. Random subspace ensembles for hyperspectral image classification with extended morphological attribute profiles. IEEE Transactions on Geoscience and Remote Sensing 53, 4768–4786.

Xu, X., Li, J., Wu, C., Plaza, A., 2018. Regional clustering-based spatial preprocessing for hyperspectral unmixing. Remote Sensing of Environment 204, 333–346.

Yan, C., Bai, X., Ren, P., Bai, L., Tang, W., Zhou, J., 2016. Band weighting via maximizing interclass distance for hyperspectral image classification. IEEE Geoscience and Remote Sensing Letters 13, 922–925.

Yuan, Y., Lin, J., Wang, Q., 2015. Dual-clustering-based hyperspectral band selection by contextual analysis. IEEE Transactions on Geoscience and Remote Sensing 54, 1431–1445.

Zhao, J., Lu, K., He, X., 2008. Locality sensitive semi-supervised feature selection. Neurocomputing 71, 1842–1849.

Zhao, W., Du, S., 2016. Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach. IEEE Transactions on Geoscience and Remote Sensing 54, 4544–4554.

Zhou, X., Li, S., Tang, F., Qin, K., Hu, S., Liu, S., 2017. Deep learning with grouped features for spatial spectral classification of hyperspectral images. IEEE Geoscience and Remote Sensing Letters 14, 97–101.

Zhou, Y., Peng, J., Chen, C.P., 2014. Dimension reduction using spatial and spectral regularized local discriminant embedding for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing 53, 1082–1095.