IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

CLASSIFICATION OF HYPERSPECTRAL IMAGE USING CNN

[1]R.V.V Lohit, [2]V.S.l Sai Krupa, [3]P.Prudhvi Krishna, [4]S. Sai Mounish, [5]K.V.Satya Kumar
[1]UG Student, [2]UG Student, [3]UG Student, [4]UG Student,[5]Assistant Professor
[1]Electronics and Communications Engineering,
[1]GITAM University, Visakhapatnam, India.

Abstract: There are many methods that are so far developed to classify hyperspectral images. Different from traditional spectral-spatial classification frameworks where the spectral feature extraction (FE), spatial FE, and classifier training are separated, these processes are integrated into a unified network in our model. In this project we are proposing a model using stationary wavelet transform (SWT), principal component analysis (PCA) and convolutional neural networks (CNN) for HSI classification. SWT is used to extract the meaningful features from the hyperspectral Data. PCA selects a subset of transformed coefficients for classification by reducing redundancy which can be also called as dimension reduction (DR) and reshapes the image data to into 2D from 3D. A CNN will be designed to classify the selected transformed coefficients and predict the accurate output. In this project, we are planning to use popular online hyperspectral image dataset to show better accuracy when compared with other frame works which are already in implemented.

***Index Terms* - Convolutional neural network (CNN), dimension reduction (DR), feature extraction (FE), hyperspectral image (HSI), Principal component analysis (PCA), Stationary wavelet transform (SWT).**

I. INTRODUCTION

Hyperspectral image (HSI) classification is more challenging because it contains hundreds of bands in it. In order to alleviate this problem, dimensionality reduction methods are proposed, which can be divided into feature selection [16] and feature extraction (FE)[15] methods. The main aim of feature selection is to preserve the most representative and crucial bands from the original data set and discard those making no contribution to the classification [2]. By designing suitable criteria, feature selection methods can eliminate redundancies among adjacent bands and improve the discriminability of different targets [2]. The FE, on the other hand, is used to find an appropriate feature mapping to transform the original high-dimensional feature space into a low-dimensional one. After reducing the dimensions to the required values, we apply for the spatial classification and spectral classification to an Image. The goal of hyperspectral imaging is to obtain the spectrum for each pixel in the image of a scene, with the purpose of finding objects, identifying materials, or detecting processes [8]. Engineers build hyperspectral sensors and processing systems for applications in astronomy, agriculture, molecular biology, biomedical imaging, geosciences, physics, and surveillance.

Apart from the Hughes phenomenon, the complex spatial distribution and the spectral heterogeneity of the objects also make it difficult to achieve a high classification accuracy in HSI. Recent works try to incorporate the spatial feature into consideration. Those approaches include gray- level co-occurrence matrix, wavelet transform, Gaborfilter, and so on. Among these spatial features, mathematical morphological features have received a great attention.Combining opening and closing operations with structural elements of different sizes, morphological profiles (MPs) are built and applied to the high-resolution satellite imageryin. Considering the high dimensionality of hyperspectral data, the extended MPs (EMPs) are proposed, and MPs are computed on the first several principal components of hyperspectral data rather than the original bands, which are then fed into a support vector machine (SVM) classifier [8]. The disadvantage of hand-crafted features like EMPs lies on the dependence on the prior information (empirical spatial filter parameters) of the designers, resulting in a poor adaptation to different data sets. Besides, handcrafted features are regarded as shallow features, lacking of robustness when confronted with complex circumstances, where the imaging environment varies so sharply that images may change a lot even within a short interval. From the perspective of deep learning, the components of an image are hierarchical. Specifically speaking, pixels are first assembled to form edges; edges are then assembled to form parts; parts are finally assembled to form different objects. Therefore, the essential theory of deep learning is to learn deep features in a hierarchicalway [8]. Compared with the shallow features, high-level features have the capability to represent more abstract and complex structure information, thus possessing a greater robustness andinvariance toward local changes of the image. In the remote sensing area, the neural network-based methods have been utilized to deal with many challenging tasks. Typical deep neural network models include stacked auto encoders (SAEs), deep belief networks (DBNs), and convolutional neural net- works (CNNs) based on which a series of methods have been proposed, aiming at the HSI classification task In order to alleviate this problem, the CNN-based methods are introduced into the HSI classification task. In a 3-D CNN model is proposed, where the convolution operation is applied on the hyperspectral cube to extract deep features.

However, 3-D CNN fails to take full advantage of spectral information, bringing about the over smoothing phenomenon, which will cause the misclassification for small objects and class boundaries [9]. Zhao and Du proposed a spectral–spatial feature-based classification (SSFC) method, using the balanced local discriminant embedding and CNN to extract spectral and spatial features, respectively, which are then combined together and fed into the SVM or logistic regression classifier. The SSFC method can be summarized as. However, in this framework, FE and classifier training are separated. Both processes have their own objective functions and the classifier training may not help to optimize the features under such a condition. The performances of FE and classifier training will also influence the classification results of the whole framework, making it hard to obtain a stable accuracy.

To overcome the aforementioned drawbacks, we propose the spectral–spatial unified networks (SSUNs) for HSI classification. As shown in, there are two characteristics of the SSUN [8]. FE and classifier training can share a uniform objective function, since all the processes are integrated into a whole neural network. In this way, the training error of the classifier will be passed to the features through back propagation, which will make the learned features become more discriminative. Second, different from the traditional framework where the spectral FE and the spatial FE are separated, we train the spectral features and the spatial features simultaneously in the SSUN. the long short-term memory (LSTM) model is adoptedas the spectral feature extractor, which is an updated versionof recurrent neural networks [8]. As to the spatial FE,a multiscale CNN (MSCNN) is proposed in this paper. The major contributions of this paper are summarized as follows.

1) An end-to-end framework named the SSUN is proposed for HSI classification which integrates the spectral FE, spatial FE, and classifier training into a unified neural network [8].

2) A band grouping-based LSTM algorithm is proposed forspectral FE. Two novel grouping strategies are proposedto better learn the contextual features among adjacent bands for the LSTM. Compared with the traditional band-by-band strategy, the proposed methods can prevent a too deep network for the HSI [8].

3) The MSCNN is proposed for spatial FE. Different from the traditional CNN which only utilizes the deepest convolutional features to complete the classification, the proposed MSCNN combines both shallow and deep convolutional layers in the classification. This may be a promising way to deal with the information loss during the convolution and pooling operations in the CNN.Besides, the MSCNN also has the property of multi-scale,since features in different layers correspond to different scales [8].

The rest of this paper is organized as follows. Section II describes the proposed band grouping-based LSTM algorithm in detail. Section III presents the SSUN framework. The informationdata sets used in this study and the experimental results are given in Section IV. Conclusions and other discussions are summarized in Section V.

II. BAND GROUPING-BASED LSTM ALGORITHM

In this section, we will first make a brief introduction to the LSTM. Then, the proposed grouping strategies for processing the spectral information with the LSTM are presented.

1. LSTM

The main challenge when training the RNN is the long-termdependence that gradients tend to either vanish or explode during the back propagation stage. To mitigate this problem,a gated RNN called the LSTM is proposed in, whichhas been successfully applied to many tasks, such as natural language processing, handwriting recognition, and land cover change detection. The core component of the LSTM is the memory cell which replaces the hidden unit in traditional RNNs. As shown in Fig. 1, there are four main elements in the memory cell, including an input gate, a forget gate, an output gate, and a self-recurrent connection [1]. The inputgate can either allow the income signal to update the state of the memory cell or block it. The output gate will control. whether the cell state will have an effect on other neurons atthe next time step. The forget gate, on the other hand, will modulate the self-recurrent connection of the memory cell, making the cell remember or forget its previous state [8]. The forward propagation of the LSTM for time step t is defined as follows.

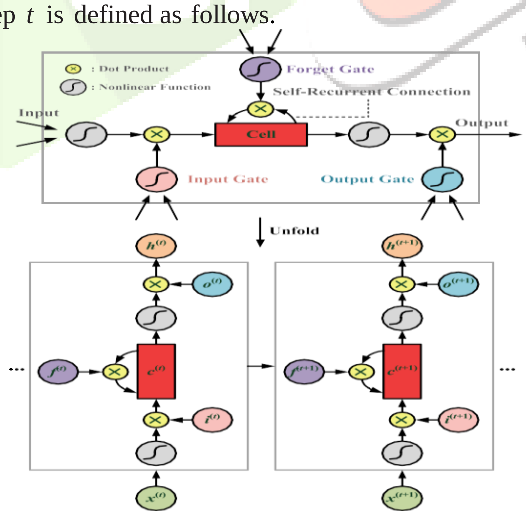

Fig. 1. Illustration of an LSTM model.

Input Gate:

$I(t)=\sigma(W_i x(t) + U_i h(t-1) + b_i)$. (1)

Forget Gate:

$f(t) = \sigma(W_f x(t) + U_f h(t-1) + b_f)$. (2)

Output Gate:

$o(t) = \sigma(W_o x(t) + U_o h(t-1) + b_o)$. (3)

Cell State:

$C(t)=i(t)\times g(W_C x(t)+U_C h(t-1)+b_C)+f(t)\times c(t-1)$ (4)

LSTM Output:

$H(t) = o(t) \times g(c(t))$.

Here, W_i, W_f, W_o, W_c, U_i, U_f, U_o, and U_c are weightmatrices. b_i, b_f, b_o, and b_c are bias vectors. $\sigma(x)$ 1/(1+exp(-x)) is the sigmoid function and denotes thedot product. Similar to traditional RNNs, the LSTM network can be trained by the mini-batch stochastic gradient descent with the BPTT algorithm[8]. Refer to for more detailed descriptions.

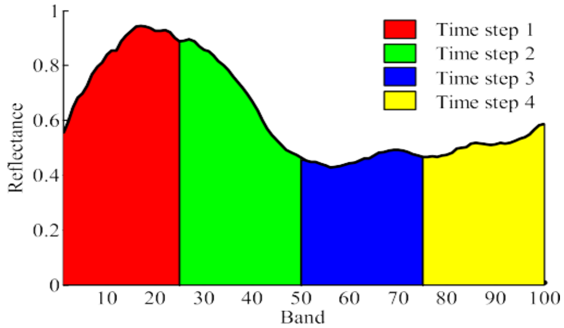

Fig. 2. Grouping strategy 1. Adjacent bands are divided into the same sequence according to the spectral orders. The bands marked with the same color will be fed into the LSTM network at each time step.

2. Band Grouping-Based LSTM Algorithm

Previous studies have shown that the deep architecture possesses better generalization ability when dealing with the complicated spectral structure. By feeding the entirespectral vector into the SAE or DBN, deep spectral features can be learned automatically in an unsupervised way. While existing methods focus on the integrality of spectra, the LSTMnetwork pays more attention to the contextual information among adjacent sequential data [8]. Therefore, how to divide the hyperspectral vector into different sequences in a proper wayis crucial to the performance of the network [2]-[6]. A natural ideais to consider each band as a time step and input one band ata time. However, hyperspectral data usually have hundreds of bands, making the LSTM network too deep to train in such a circumstance. Thus, a suitable grouping strategy is needed [9].

Let n be the number of bands and τ be the number of time steps in the LSTM. Then, the sequence length of each time step is defined as m floor(n/τ), where floor(x) denotes roundingdown x [4]. For each pixel in the HSI, let z,$z_1, z_2, \ldots z_i, \ldots z_n$ be the spectral vector, where z_i is the reflectance of thei^{th} band. The transformed sequences are then denoted by$x$,$x_1, x_2, \ldots x_i, \ldots x_\tau$, where x_i is the sequence at thei^{th} time step. In what follows, we use two grouping strategies in this paper.

Grouping Strategy 1:

$x(1) = [z_1, z_2, \ldots z_m]$

$x(2) = [z_{m+1}, z_{m+2}, \ldots z_{2m}]\ldots$

$x(i) = [z(i-1)_{m+1}, z(i-1)_{m+2}, \ldots z_{im}]$

$x(\tau) = [z(\tau-1)_{m+1}, z(\tau-1)_{m+2}, \ldots z_{\tau m}]$----(6)

where x(i) is the sequence at time i[8]. As shown in Fig. 2, strategy 1 focuses on the local features and makes the signals.

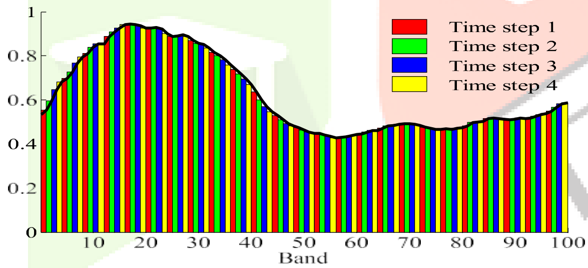

Fig. 3. Grouping strategy 2. Every group in this case will cover a large spectral range. The bands marked with the same color will be fed into the LSTM network at each time step.

inside a group continuous without any intervals. Each group concentrates on a narrow spectral range. The spectral distance between different time steps will be relatively longer under such circumstances [8].

Grouping Strategy 2:

$x(1) = [z1, z1+\tau, \ldots z1+\tau(m-1)]$

$x(2) = [z2, z2+\tau, \ldots z2+\tau(m-1)]$

$x(i) = [zi, zi+\tau, \ldots zi+\tau(m-1)]$

$x(\tau) = [z_\tau, z_{2\tau}, \ldots z_{\tau m}]$------------ (7)

3. MATERIALS AND METHODS

The materials and methods related to our discussion has three functions and is concerned to Stationary Wavelet Transform, Principal Component Analysis and Convolutional Neural Networks [2].

A. Stationary Wavelet Transform (SWT):

Stationary wavelet transform helps to identify the image edge features coefficients. By this image edge features we can differentiate the parts in the image and this will make Feature Extraction much easier [2],[8],[19].

B. Principal Component Analysis (PCA):

Principal Component Analysis (PCA) is a Widely using dimensionality-reduction method. This is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information of the large set [20].

C. Convolutional Neural Network (CNN):

A Convolutional Neural Network (CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.

III. SSUN (SPECTRAL AND SPATIAL UNIFIED NETWORK)

Spectral FE and spatial FE are very important processes for the HSI classification. Previous works usually extract the spectral features and spatial features independently. Besides, the FE and classifier training are also separated from each other. Then both spectral FE, spatial FE, and classifier training will have their own objective functions, making the whole framework more complicated. In this section, we will introduce the proposed SSUN framework, which integrates thespectral FE, spatial FE, and it will classifier training into a unified neural network. the true label, and m is the size of training set. The whole network is trained in an end-to-end manner, where all the parameters are optimized by the mini-batch stochastic gradient descent algorithm at the same time. The implementation details of the proposed SSUN are shown in Algorithm 1.

Algorithm 1:- Proposed SSUN Framework

Input:
1) A HSI with ground-truth.
2) The number of time steps τ , the number of reserved principal components (PCs) p, the size of patches w.
Step 1: For each pixel in the HSI, divide the hyperspectral vector into τ sequences by (6) or (7) as the spectral features.
Step 2: Apply PCA to the HSI and reserve the first p PCs.
Step 3: Around each pixel in the dimension reduced HSI, extract a patch with the size of $w \times w \times p$ as the spatial features.
Step 4: Initialize the weights in the network with random values which are subject to Gaussian distribution with a mean of 0 and a standard deviation of 0.1. The bias terms are initialized with 0.
Step 5: Input the training samples to the network and optimize it with the mini-batch stochastic gradient descent algorithm.
Step 6: For each pixel in the HSI, input the corresponding spectral and spatial features to the network to complete the classification of the whole image.
Output: A two dimensional matrix which records the labels of the HSI[8].

IV. RESULTS AND DISCUSSION

4.1 Results of Descriptive Statics of Study Variables

Data Description and Experimental Design In our experiments, classical hyperspectral dataset, Indian Pines, is utilized to evaluate the performance of the proposed method. The data set is gathered by the Airborne Visible/ Infrared Imaging Spectrometer (AVIRIS) sensor over the Indian Pines test site in Northwestern Indiana [11]-[12]. After the removal of the water absorption bands, the image consists of 200 spectral bands with 145 × 145 pixels. It has a spectral coverage from 0.4 to 2.5 µm and a spatial resolution of 20 m. The false-color composite of the Indian Pines image and the corresponding ground truth map are shown in Fig.4. The training and test set are listed in Table I. The experiments are conducted in three parts. The first two parts analyze the LSTM and the MSCNN, respectively [2]. The third part reports the performance of the proposed SSUN and comparing methods [8]. All the experiments in this paper are randomly repeated 25 times with random training and test data. In each repetition, we first randomly generate the training set from the reference data. Then, the remaining reference samples make up the test set. The OA and Kappa coefficient are utilized to quantitatively estimate different methods. Both the average value and the standard deviation are reported.

4.1.1 Figures and Tables

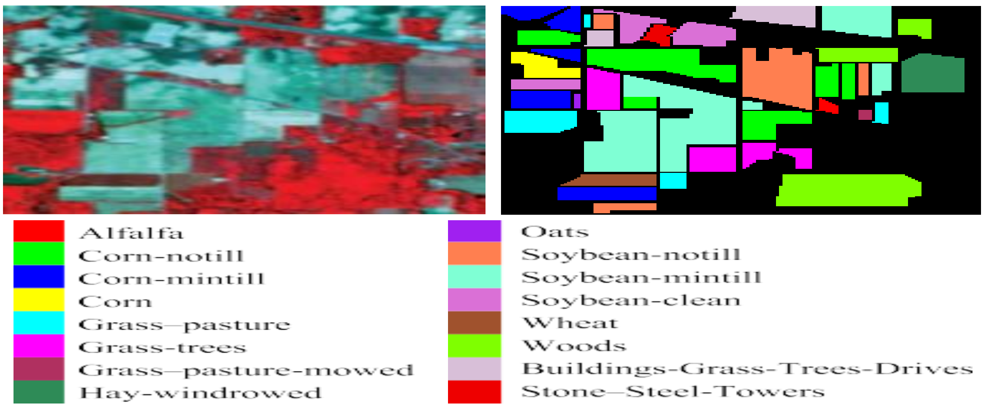

FIG. 4. FALSE-COLOR IMAGE AND GROUND-TRUTH MAP FOR THE INDIAN PINES DATA SET.

TABLE I

Number of training and test samples used in the Indian pines data set

#	Class	Training	Testing
1	Alfalfa	30	16
2	Corn-notill	1278	150
3	Corn-mintill	680	150
4	Corn	137	100
5	Grass-pasture	333	150
6	Grass-trees	580	150
7	Grass-pasture-mowed	20	8
8	Hay-windrowed	328	150
9	Oats	15	5
10	Soybean-notill	822	150
11	Soybean-mintill	2305	150
12	Soybean-clean	443	150
13	Wheat	150	55
14	Woods	1115	150
15	Buildings-Grass-Trees-Drives	336	50
16	Stone-Steel-Towers	43	50
	Total	8615	1634

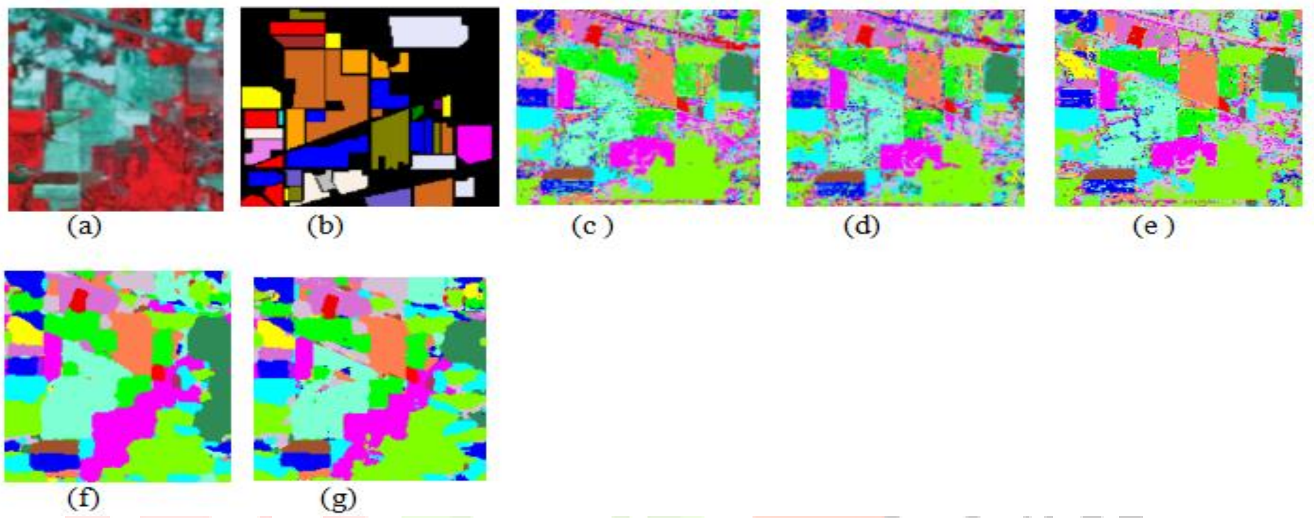

FIG. 5. CLASSIFICATION MAPS FOR THE INDIAN PINES DATA SET. (A) FALSE-COLOR IMAGE. (B) GROUND-TRUTH MAP. (C) RAW. (D) PCA. (E) LSTM (F) MSCNN. (G) SSUN.

Table II
Classification results

Class	RAW	PCA	LSTM	MSCNN	SSUN
1	83.12	83.54	81.17	99.79	100
2	73.86	67.79	76.22	96.68	94.96
3	76.79	72.74	72.95	98.61	96.36
4	83.11	80.87	82.97	99.73	96.85
5	94.73	93.88	87.26	99.39	98.01
6	97.75	96.02	91.03	99.27	99.55
7	87.91	87.91	82.08	100	99.58
8	98.1	97.9	97.3	99.96	100
9	96	80.66	93.66	100	100
10	83.47	73.23	84.86	97.85	98.15
11	68.18	62.04	71.25	95.67	97.65
12	82.11	74.45	82.82	99.17	99.62
13	99.33	98.3	96.45	100	100
14	93.21	93.4	91.69	99.19	99.59
15	57.93	46.54	59.66	98.78	99.22
16	95.34	91.39	93.02	99.37	100
OA(%)	79.68±0.87	74.72±0.99	81.71±1.54	96.82±0.99	98.13±1.05
Kappa *100	76.68±0.97	71.05±1.18	78.46±1.12	95.15±0.55	96.27±0.99
Runtime	4.93±0.06	0.95±0.03	39.9±0.32	55.03±0.38	104.36±0.60

V. CONCLUSION

In this project, we are planning for a model for classifying hyperspectral images using SWT, PCA and CNN. Instead of directly applying the raw image to the CNN, priority is made to extract the SWT coefficients which provide better spectral information. PCA is applied to reduce the dimension of resulting coefficients. CNN works on the reduced coefficients to provide a classification map. Experiments are conducting on a standard hyperspectral image dataset (Indian pines) to validate the model and it is expected that SWT-PCA-CNN model will achieve a classification accuracy of 98.13% Which is better than its counterparts that are already in existing with other models.

VI. ACKNOWLEDGEMENT

We would like to thank our project guide Mr. K.V. Satya Kumar, Assistant Professor, Dept. of EECE, GITAM University, Visakhapatnam, India, for his stimulating guidance and profuse assistance. We shall always cherish our association with him/her for his/her guidance, encouragement, and valuable suggestions throughout the progress of this work. We consider it a great privilege to work under his guidance and constant support.

REFERENCES

[1] IEEE paper Published in: 2017 Chinese Automation Congress (CAC)
https://ieeexplore.ieee.org/document/8243039

[2] IEEE paper Published in: 2019 International Conference on Information Technology (ICIT)
https://ieeexplore.ieee.org/document/9031937

[3] https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks

[4] https://www.intechopen.com/books/processing-and-analysis-of-hyperspectraldata/hyperspectral-image classification

[5] Yonghao Xu , Liangpei Zhang ,Bo Du "Spatial spectral classification of hyperspectral image". in:(2018)

[6] Xuchu Yu,Anzhu Yu,Pengqiang Zhang&Gang Wan "Spectral-spatial classification of hyperspectral imagery based on recurrent neural networks". in:(2018)

[7] Andrew G. Howard, "Some Improvements on Deep Convolutional Neural Network Based Image Classification".

[8] IEEE paper Published in: 2018 At IEEE Transactions on Geoscience and Remote Sensing
https://ieeexplore.ieee.org/document/8356713

[9] C. Rodarmel and J. Shan, "Principal component analysis for hyperspectral image classification," Surv. Land Inf. Sci., vol. 62, no. 2, p. 115, 2002.

[10] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality- preserving dimensionality reduction and classification for hyperspectral image analysis," IEEE Trans. Geosci. Remote Sens., vol. 50, no. 4,pp. 1185–1198, Apr. 2012.

[11] Hyperspectral Remote Sensing Scenes. Accessed: Mar. 10, 2017.[Online]. Available:
http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes

[12] https://www.kaggle.com/abhijeetgo/indian-pines-hyperspectral-dataset[Online]

[13] https://github.com/gokriznastic/HybridSN[Online]

[14] https://github.com/YonghaoXu/SSUN[Online]

[15] M. Dalla Mura, J. Benediktsson, B. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," Geoscience and Remote Sensing, IEEE Transactions on, vol. 48, pp.3747 – 3762, 11 2010

[16] Y. Yuan, X. Zheng, and X. Lu, "Discovering diverse subset for unsupervised hyperspectral band selection," IEEE Transactions on Image Processing, vol. 26, no. 1, pp. 51–64, Jan 2017.

[17] A. M. Rufai, G. Anbarjafari, and H. Demirel, "Lossy image compression using singular value decomposition and wavelet difference reduction," Digital Signal Processing, vol. 24, pp. 117 – 123.

[18] H. Demirel and G. Anbarjafari, "Image resolution enhancement by using discrete and stationary wavelet decomposition," IEEE Transactions on Image Processing, vol. 20, no. 5, pp. 1458–1460, May 2011.
http://www.sciencedirect.com/science/article/pii/S0167865518307748.

[19] M. Han, R. Cong, X. Li, H. Fu, and J. Lei, "Joint spatial spectral hyperspectral image classification based on convolutional neural network," Pattern Recognition Letters, 2018.[Online].
Available :http://www.sciencedirect.com/science/article/pii/S0167865518307748

[20] Principal component analysis. [Online]. Available: www.dezyre.com/ data-science-in-python-tutorial/principal-component-analysis-tutorial.