Hyperspectral Image Processing

Nikos Sourlos

October 2020

1 Introduction

There are many methods for image acquisition. One of them is hyperspectral imaging. In hyperspectral imaging, at every pixel location (x,y) in an image and along a range of the electromagnetic spectrum (λ) information is collected and processed. These images are then combined to form a 3D image (x,y,λ). This 3D image has the number of spectral bands as its third dimension (like RGB images has their third dimensions equals to 3). The information from that spectrum is of great importance since certain molecules and materials could be identified by their spectral signature [2].

2 Problem Definition

The goal of this study is to process hyperspectral images taken from a satellite along with their ground truth to classify each pixel in these images in 1 out of 14 different classes, representing the land cover type. In total 2 hyperspectral images consist the training set (Loukia and Dioni) and 3 the test set (Nefeli, Erato and Kirki) [1]. These images have different spatial dimensions but the spectral dimension is the same for all of them and equal to 176. Moreover, not all classes are present in training image 'Dioni'.

3 State-of-the-art Methods

There are many methods used for hyperspectral image classification. These methods are divided in those that have a hand-designed feature extraction and those with a learning-based feature extraction (CNNs) [7]. The most promising are those that use Deep CNNs. Below some of these Deep Learning methods are presented:

1. **1D-CNN with pixelwise spectral data**: If a hyperspectral image (HSI) is denoted as $\mathrm{N_x}$*$\mathrm{N_y}$*N where $\mathrm{N_x}$ and $\mathrm{N_y}$ are the spatial dimensions and N is the spectral one, one-dimensional input vectors are extracted from that HSI, each one of them containing the spectral information at a specific pixel location x,y. The dimension of each on these vectors is N.

Some of these vectors are used for the training of a 1D-CNN and some others for testing its performance [3]. An example of how a vector is extracted is shown in Fig.1

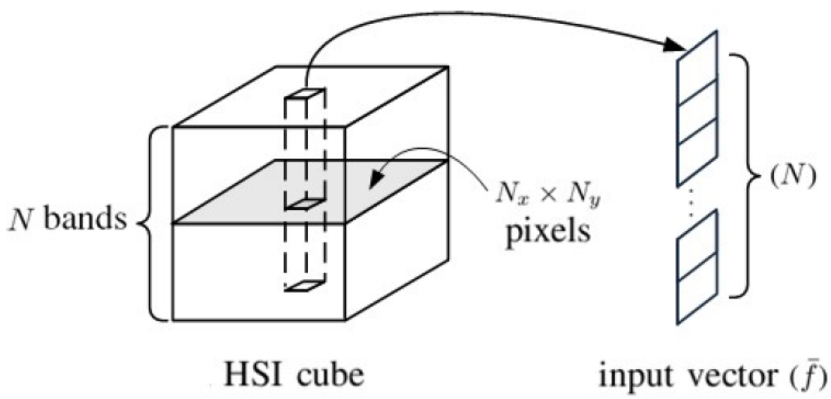

Figure 1: 1D feature vector extraction [3]

2. **1D-CNN with Spectral-Spatial Data**: Here PCA is firstly applied along the spectral dimension of each pixel to extract Q principal components (dimensionality reduction). Then, all these Q components of an area of R*R pixels surrounding our target pixel are collected, vectorized to form a R*R*Q 1D vector and concatenated with the original spectral dimension (N) of that pixel to form a N+R*R*Q 1D vector which is given as input to a CNN. Similarly as before, some pixels are used for training and some others for testing the performance of the CNN [3]. An example of how a vector is extracted in this case is presented in Fig.2

3. **2D-CNN with Principal Components**: PCA is again applied along the spectral dimensions to extract the first Q principal components of each pixel. Then, an area of R*R pixels around a given pixel location from the N_x*N_y*Q representation that emerged forms an input layer for a 2D CNN associated with that pixel [3]. This procedure is depicted in Fig.3

4. **Features Fusion Model**: It combines the features from the 1D-CNN with pixelwise spectral data with those of the 2D-CNN with Principal Components. Basically the 1D and 2D CNN models are trained by using different input data forms and then features are extracted from them, stacked, fused and are used for training a new model which has one additional FC and one additional output layer [4]. The architecture of this model is demonstrated in Fig.4

5. **3D CNN**: In this method the input data, which consist of the spatial and the spectrum dimensions, are directly feed to the network and convolved with 3D kernels [5].

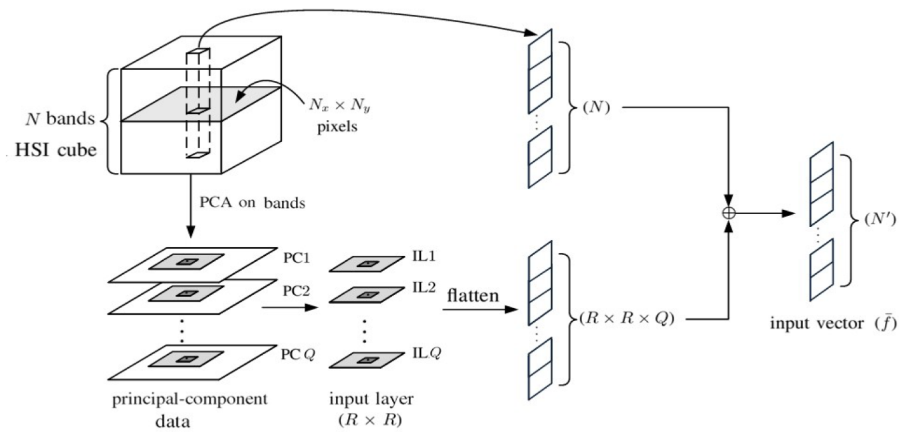

Figure 2: Feature vector with spectral-spatial information [3]

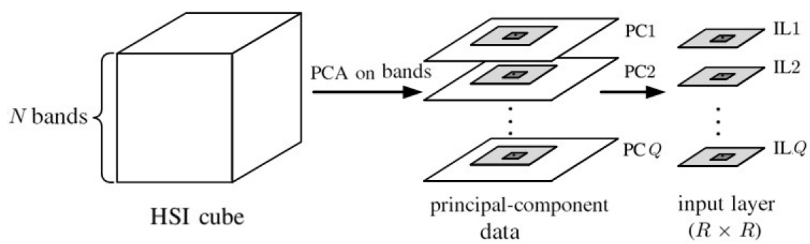

Figure 3: Preparation of layers for a 2D network [3]

6. **HSI-CNN**: For a given pixel and for its neighbors spatial and spectral features are extracted. Convolutions are applied on these features which result in a number of 1D feature maps. These maps are finally stacked and reshaped into a 2D matrix which can make better use of the information contained in spectral and spatial dimensions of the input data. Operations, like convolution, are again applied in this matrix and finally, 2 FC layers along with a softmax one are responsible for the final classification [5]. The architecture of this model is illustrated in Fig.5

7. **Contextual CNN**: There is a number of convolutional layers in this network used for the classification. At the beginning, a 'multi-scale filter bank' is followed by two blocks of convolutional layers associated with residual modules. This filter bank is similar to the inception module as a concept and it is used to exploit diverse local structures of the input image. The last three layers are used for the classification by using local features. Dropout and ReLU activations are also used in this network [6]. The architecture of this model is shown in Fig.6

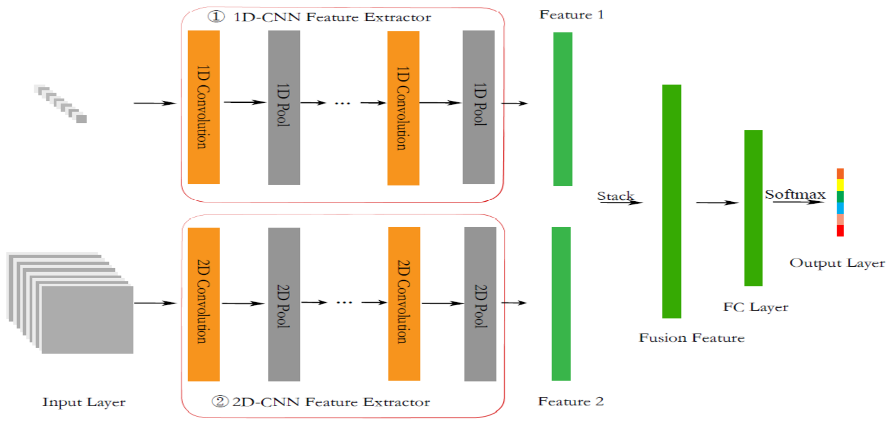

Figure 4: Fusion model architecture [4]

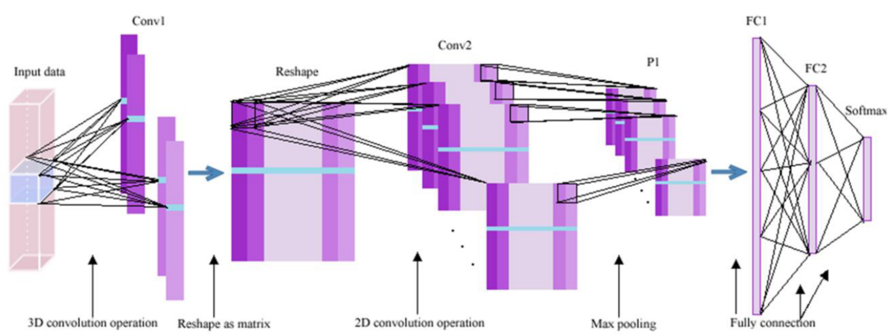

Figure 5: HSI-CNN architecture [5]

8. **HybridSN**: This is a method which has a 3D-CNN, which exploits both the spatial and the spectral information, followed by a 2D CNN which basically learns an abstract representation of the spatial information [7]. In that method, Principal Component Analysis is firstly used to reduce the spectral dimension by keeping only the features that are most relevant. The architecture of this model is shown in Fig.7

9. **Spectral-Spatial Residual Network (SSRN)**: This method takes as input parts of the original hypercube, similar as the patches in HybridSN. In its architecture the residual blocks are used which allow it to learn identity mappings [8]. The architecture of this model is shown in Fig.8

In general, performance of 1D CNN is not as good as the one of 2D and 3D equivalents. The problem with a 2D CNN is that it can only take into account the spatial dimension and not the spectral one. On the other hand, 3D CNNs

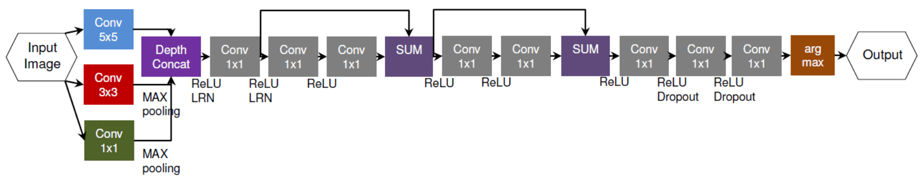

Figure 6: Contextual CNN architecture [6]

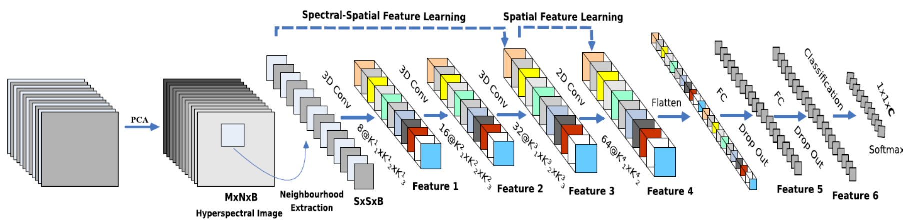

Figure 7: HybridSN architecture [7]

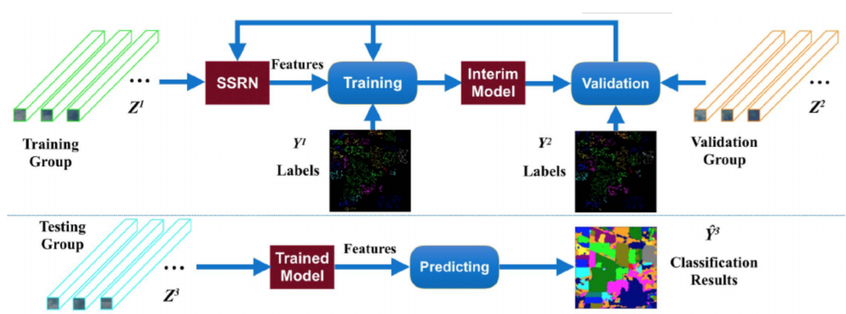

Figure 8: SSRN architecture [7]

are able to overcome that issue but they are very computationally expensive, which make their use prohibited in many applications, and they perform worse for classes that have similar textures over many spectral bands [7].

From the above methods the one that seems to obtain the best results is the HybridSN [7]. Moreover this method reducers the complexity that would result from the use of a 3D CNN alone. Therefore this is the method that will be used in our study.

References

[1] "The ISPRS HyRANK Benchmark", https://www2.isprs.org/commissions/comm3/wg4/hyrank/, accessed: 30-10-2020.

[2] "Hyperspectral Imaging: Spectral Analysis in 3 Dimensions", https://www.findlight.net/blog/2017/05/21/hyperspectral-imaging/, accessed: 30-10-2020.

[3] Hsieh, T. H., Kiang, J. F. (2020). Comparison of CNN Algorithms on Hyperspectral Image Classification in Agricultural Lands. Sensors (Basel, Switzerland), 20(6), 1734. https://doi.org/10.3390/s20061734

[4] Jiang, T. Wang, X.. (2020). HYPERSPECTRAL IMAGES CLASSIFICATION BASED ON FUSION FEATURES DERIVED FROM 1D AND 2D CONVOLUTIONAL NEURAL NETWORK. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XLII-3/W10. 335-341. 10.5194/isprs-archives-XLII-3-W10-335-2020.

[5] Y. Luo, J. Zou, C. Yao, X. Zhao, T. Li and G. Bai, "HSI-CNN: A Novel Convolution Neural Network for Hyperspectral Image," 2018 International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, 2018, pp. 464-469, doi: 10.1109/ICALIP.2018.8455251.

[6] H. Lee and H. Kwon, "Going Deeper With Contextual CNN for Hyperspectral Image Classification," in IEEE Transactions on Image Processing, vol. 26, no. 10, pp. 4843-4855, Oct. 2017, doi: 10.1109/TIP.2017.2725580.

[7] S. K. Roy, G. Krishna, S. R. Dubey and B. B. Chaudhuri, "HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification," in IEEE Geoscience and Remote Sensing Letters, vol. 17, no. 2, pp. 277-281, Feb. 2020, doi: 10.1109/LGRS.2019.2918719.

[8] Zhong, Zilong Li, Jonathan Luo, Zhiming Chapman, Michael. (2017). Spectral-Spatial Residual Network for Hyperspectral Image Classification: A 3-D Deep Learning Framework. IEEE Transactions on Geoscience and Remote Sensing. 56. 847-858. 10.1109/TGRS.2017.2755542.