HYPERSPECTRAL IMAGE CLASSIFICATION VIA SHAPE-ADAPTIVE DEEP LEARNING

Atif Mughees\textsuperscript{1*}, Ahmad Ali\textsuperscript{2} and Linmi Tao\textsuperscript{1}

\textsuperscript{1}Department of Computer Science, Tsinghua University, Beijing, China
\textsuperscript{2}Pakistan Institute of Engineering and Applied Sciences

ABSTRACT

Hyperspectral image (HSI) classification is one of the most prevalent issues in remote sensing area. Recently, application of deep learning in HSI classification has emerged. However, merging spatial features with spectral properties in deep learning is a pervasive problem. This paper presents a discriminative spatial updated deep belief network (SDBN) which effectively utilizes spatial information within spectrally identical contiguous pixels for HSI classification. In the proposed approach, HSI is first segmented into adaptive boundary adjustment based spatially similar regions with similar spectral features, following which an object-level feature extraction and classification is undertaken using deep belief network (DBN) based decision fusion approach that incorporates spatial-segmented contextual and spectral information into a DBN framework for effective spectral-spatial HSI classification. Moreover, for improved accuracy, band preference/correlation based feature selection approach is used to select the most informative bands without compromising the original content in HSI. Usage of local contextual features and spectral similarity from adaptive boundary adjustment based approach, and integration of spatial and spectral features into DBN results into improved accuracy of the final HSI classification. Experimental results on well known hyperspectral data indicate the classification accuracy of the proposed method over several existing techniques.

Index Terms— image classification, hyperspectral image, deep belief network, segmentation

1. INTRODUCTION

Effective and accurate hyperspectral images (HSI) analysis has become more crucial with advancements in remote sensing technology, acquiring data with tens of hundreds of spectral channels along with the detailed spatial information of the scene. The rich spectral and spatial information if efficiently utilized, can produce higher classification accuracies [1]. Moreover, recent advances in remote sensing sensors has enabled to acquire rich spatial resolution [2], hence made it possible to evaluate the small spatial land covers in the HSI. Due to the inherent rich spectral-spatial information, HSI classification has been an active research area. However it also poses immense challenges such as higher dimensionality, limited labeled sample, hughes phenomena. HSI classification has primarily focused on either feature selection (FS) or feature extraction (FE) as a way to feed classifier [3] for classification. FS consists of selection of appropriate subset of spectral bands, while FE involves the transformation of the data into a space of reduced dimensionality [2]. Support vector machines (SVMs) classifier performed well on high-dimensional data with few training samples [3, 4]. In general, feature extraction/ dimensionality reduction techniques such as principle component analysis (PCA) and independent component analysis (ICA) are applied, which results in the loss of detailed information inevitably and hence effect the classification accuracy. Moreover, these traditional FE techniques may also results into the loss of local structural information.

In recent years, researchers have realized the importance of incorporating spatial information along with spectral data in HSI classification process [5, 6] as recent hyperspectral sensors can deliver excellent spatial resolution along with the spectral channels. Recently deep learning architectures (DLA) are employed for HSI classification which have attained historically high classification accuracies [7, 8, 9, 10]. Deep learning architectures extracts more abstract and invariant features. In recent years, two DLAs, deep belief network (DBN) and the stacked auto-encoder (SAE), have outperform SVMs in HSI [11, 12, 13]. However incorporating spatial information in deep network for improved accuracy is still an open research area.

In this paper, we are primarily interested in hyper-segmentation based deep belief network (DBN) architecture that effectively exploits the spatial information of the HSI where pixel by pixel based scanning window is changed by local spatial information. Adaptive boundary adjustment based hyper-segmentation [14] has dual effect, first, it keeps the spectral association in spectral channels and adopts the object boundaries in spatial domain, second it subtly opt for best informative bands and eliminate noisy data without losing the original content. Proposed technique takes full advantage of the available information with restricted training data. Each resulted segmented region [14] can be treated as a local spatial region with similar spatial characteristics in HSI. The

\*Corresponding author. E-mail: maoz14@mails.tsinghua.edu.cn
weighted-segmentation based spatial updated deep belief network (SDBN) first exploits an efficient hyper-segmentation approach [14] to segment the HSI into expressive spatially similar segments. Pixels in each resulted segmented region expected to have pixels with similar spectral properties, and their correlations are exploited via a deep belief network. Proposed technique exploits multi-layer DBN to learn shallow and deep features of HSI. The remaining paper is defined as follows. Section 2 describes proposed methodology. Experimental analysis and data sets are presented in section 3 and conclusion is drawn in section 4.

2. HYPER-SEGMENTATION BASED DBN FOR HSI CLASSIFICATION

It is strongly acknowledged that combining spatial contextual information along with the spectral features can greatly improve the classification accuracy [5, 6]. Feature extraction process for HSI classification should consider two major facts [15] 1) There is a high probability that pixels close in feature space share the same class 2) There is also a high probability that spatially neighboring pixels should also share the same class. In other words, data with higher spectral/feature and spatial signatures have larger probabilities to share the same class. To incorporate these factors, we have enhanced the DBN classification process to extract discriminative spectral-spatial features. DBN considers every input data sample independently without considering any correspondence with neighboring data. In order to device the feature learning phase to incorporate the above mentioned factors, we enhance DBN process by keeping the correlation between sampled data by providing DBN contextual features along with the selected spectral features. That provides an enhanced HSI classification results.

A weighted hyper-segmentation is an adaptive boundary movement based technique [14] that provides spatial contextual structures with flexible shape and size that matches different actual structures present in the HSI. In this paper, the SDBN algorithm adopts the DBN to effectually exploit spectral–spatial information within and among each resulted segmented region from hyper-segmentation. In general, the proposed SDBN approach mainly consists of three phases as shown in Fig. 1: 1) creation of spatially adaptive segments in HSI 2) feature selection 3) exploration of spectral–spatial features of these segmented regions via DBN.

2.1. Spatial Feature Extraction by Hyper-segmentation

The hyper-segments are created by employing an efficient adaptive boundary adjustment based segmentation technique [14] that takes both spectral and spatial information into account through the energy function [14]:

\[
E(a, A_K) = \sqrt{|x_a^2 - g_k|^2 + \lambda h_k(p)|Grad(a)|}
\]  

An energy based attraction force technique is utilized to extract the contextual information which is calculated based on the tri-factor criteria further explained in [14]. Efficient band selection technique [16] is employed to reduce the computational load before the segmentation process to obtain best informative bands that hold the most discriminative information for the whole HSI, it is utilized as the base image for the hyper-segmentation as illustrated in Fig. 1.

2.2. Extraction of Spectral-Spatial Information of Spatial Segments via DBN

A deep belief net can be viewed as a structure of generative neural network based learning modules each of which is a Re-
stricted Boltzmann Machine (RBM) that comprises of a visible layer of input data and a hidden layer that learn to characterize features that capture higher-order correlations in the input data as shown in Fig. 2. The two RBM layers are connected by a matrix of symmetrically weighted connections, \( w \), with no visible-visible or hidden-hidden connections. This restriction makes the hidden units conditionally independent. A combined formation of the energy for hidden unit \( h \) and input units \( v \) is given by [12]:

\[
E(v, h, \theta) = -\sum_{i=1}^{n} \frac{(v_i - b_i)^2}{2\sigma^2} - \sum_{j=1}^{m} a_j h_j - \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} v_i h_j
\]

The conditional distributions are given by:

\[
P(h_j|v; \theta) = g \left( \sum_{i=1}^{n} w_{ij} v_i + a_j \right)
\]

\[
P(v_j|h; \theta) = N \left( \sum_{j=1}^{m} w_{ij} h_j \sigma_i^2 + b_i \right)
\]

where \( \sigma \) is the standard deviation of a Gaussian visible unit, and \( N(.) \) is the Gaussian distribution. RBMs are stacked together and trained in greedy way to form DBN. The whole process can be summarized as follows:

1. Dataset is used to pre-train the single layer of DBN. Constrastive divergence (CD) [17] is used to train RBM.
2. Use the output of the first layer as input for a second layer. Train the second layer as RBM.
3. Repeat 1 and 2 for desired number of layers.
4. Fine tune all parameters with labeled training samples.

In summary, first we utilize hyper-segmented spatial region and the spectral information after FS as an input. Then, a DBN is applied to learn the deep and abstract features from the inputs through multilayer DBN, finally LR is utilized to classify and label the pixels based on learn features.

### Algorithm 1: spectral-spatial Segmentation based DBN classification

**Input:** Hyperspectral image \( I \), having pixels \( p \) with intensity vectors \( x_p \)

**Output:** HSI classification

1. To improve Computation, noisy bands are eliminated based through [16];
2. To obtain the spatial features, weighted hyper-segmentation [14] is applied on to the selected bands and \( R \) spatial segmented regions are obtained.
3. for each spatial dominated region \( R \) do
4.   Calculate the number of pixels \( M \) in each region \( R \);
5.   Flatten the array into a feature vector of size \( M \).
6. end
7. Scale \( M \) into unit interval.
8. Normalize the whole initial image onto unit interval.;
9. for each pixel do
10.   Add spectrum of each pixel on tail of each pixel’s feature vector i.e, rows in \( M \),
11. end
12. Train Deep belief network with \( M \).

### Table 1: Dataset Specifications

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Image Size</th>
<th>Classes</th>
<th>Bands</th>
<th>Labeled Pixels</th>
<th>Wavelength range (nm)</th>
<th>Spatial resolution (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indian Pine</td>
<td>145 × 145</td>
<td>16</td>
<td>200</td>
<td>10240</td>
<td>400 2500</td>
<td>20</td>
</tr>
<tr>
<td>Pavia University</td>
<td>340 × 610</td>
<td>9</td>
<td>103</td>
<td>42776</td>
<td>430 860</td>
<td>1.3</td>
</tr>
<tr>
<td>Houston University</td>
<td>349 × 1905</td>
<td>15</td>
<td>144</td>
<td>15029</td>
<td>380 1050</td>
<td>2.5</td>
</tr>
</tbody>
</table>

### 3. EXPERIMENTAL RESULTS AND PERFORMANCE COMPARISONS

In order to validate the performance of the proposed method, we conducted experiments on three well known and challenging real hyperspectral datasets acquired through different sensors.

#### 3.1. Data Set Description

The three datasets include Indian Pine, Pavia University and Houston University dataset. The detailed specification is described in Table 1. In order to evaluate the classification accuracy, we randomly selected 50% samples each for training and testing purposes in all the datasets as shown in Table 3.

#### 3.2. Parameter Setting

We conducted experiments on windows 7 system, with 4.0 GHz processor having NVIDIA GeForce GTX 970. Theano was used as implementation tool. We choose depth of 2 for Indian Pine and Houston University dataset and 4 for Pavia University dataset as suggested by the experiments in [12].
### Table 2: Classification accuracies for Indian Pine and Pavia University datasets.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Measurement</th>
<th>LORSAL-ML</th>
<th>DBN-LR</th>
<th>SDBN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indian Pine</td>
<td>Overall Accuracy</td>
<td>87.18</td>
<td>95.95</td>
<td>96.87</td>
</tr>
<tr>
<td></td>
<td>Average Accuracy</td>
<td>90.83</td>
<td>95.45</td>
<td>96.75</td>
</tr>
<tr>
<td></td>
<td>Kappa coefficient</td>
<td>0.85</td>
<td>0.9539</td>
<td>0.9567</td>
</tr>
<tr>
<td>Pavia University</td>
<td>Overall Accuracy</td>
<td>98.95</td>
<td>99.05</td>
<td>99.18</td>
</tr>
<tr>
<td></td>
<td>Average Accuracy</td>
<td>97.28</td>
<td>98.48</td>
<td>98.75</td>
</tr>
<tr>
<td></td>
<td>Kappa coefficient</td>
<td>0.9819</td>
<td>0.9875</td>
<td>0.9905</td>
</tr>
</tbody>
</table>

Number of units for each hidden layer for Indian Pine, Pavia University and Houston University is set to 30, 50 and 100 respectively. According to the researcher’s observation, number of hidden units is more imperative than the number of hidden layers. The proposed method is compared with well known existing techniques such as deep belief network with logistic regression (DBN-LR) [12] and augmented Lagrangian - multilevel logistic (LORSAL-MLL) [18].

### 3.3. Spectral-Spatial Classification Results

we used first six components of PCA and a window size of $7 \times 7$ for DBN-LR. Classification results of Indian Pine and Pavia University datasets and their comparison is shown in Table 2 while class level accuracy for Houston University is described in Table 3. Mixed pixel is the major challenge in Indian Pine due to its low spatial resolution and small size. Results confirm that spectral-spatial classification using contextual feature extraction has significant effect on the classification accuracy because spatial features help prevent the salt and paper noise. Fig. 3 shows the classification of three datasets.

Overall, experimental results demonstrates the significant improvement in HSI classification by incorporating spatial information and spectral feature selection. The algorithm has performed significantly well on the low spatial resolution dataset.

### 4. CONCLUSION

In this paper, we have introduced a new SDBN approach to exploit spatial contextual information via hyper-segmentation for effective HSI classification. The size and shape of each hyper-segment (HS) can be flexibly adapted according to the actual HSI spatial structures, which results into an effective exploitation of spatial contexts. SDBN then uses the multi-layer DBN to effectively exploit the HS based spatial features and spectral features within and among segmented hyper-segments. Experimental results demonstrates that proposed approach produces better performance then other existing approaches in HSI classification accuracies, especially in images with small spatial structures. Furthermore, SDBN as one of the robust feature extractor, works well in heterogeneous images, particularly for complex urban scenes.

One of our future research direction is to develop a more systematic method of adaptively selecting the number of hyper-segments based on different spatial/structural information. Moreover, there is a strong motivation to apply the hyper-segment based DBN approach to other HSI applications such as change detection, noise detection, and object recognition.
5. REFERENCES


