doi:10.3772/j.issn.1006-6748.2018.04.006

D-SS Frame: deep spectral-spatial feature extraction and fusion for classification of panchromatic and multispectral images①

TEFFAHI HANANE* **, Yao Hongxun②*
(* School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, P. R. China)
(** Algerian Space Agency, Algiers 16342, Algeria)

Abstract

Facing the very high-resolution (VHR) image classification problem, a feature extraction and fusion framework is presented for VHR panchromatic and multispectral image classification based on deep learning techniques. The proposed approach combines spectral and spatial information based on the fusion of features extracted from panchromatic (PAN) and multispectral (MS) images using sparse autoencoder and its deep version. There are three steps in the proposed method, the first one is to extract spatial information of PAN image, and the second one is to describe spectral information of MS image. Finally, in the third step, the features obtained from PAN and MS images are concatenated directly as a simple fusion feature. The classification is performed using the support vector machine (SVM) and the experiments carried out on two datasets with very high spatial resolution. MS and PAN images from WorldView-2 satellite indicate that the classifier provides an efficient solution and demonstrate that the fusion of the features extracted by deep learning techniques from PAN and MS images performs better than that when these techniques are used separately. In addition, this framework shows that deep learning models can extract and fuse spatial and spectral information greatly, and have huge potential to achieve higher accuracy for classification of multispectral and panchromatic images.

Key words: image classification, feature extraction(FE), feature fusion, sparse autoencoder, stacked sparse autoencoder, support vector machine(SVM), multispectral(MS) image, panchromatic(PAN) image

0 Introduction

The development of earth observation (EO) sensors technology allows the development of new and fastidious methods for feature extraction and fusion of remote sensing images with the aim of obtaining more relevant classification maps. Remote sensing image classification has become a challenging task because of the importance of great realization of land cover and land use maps for many multi-temporal studies in different fields as agriculture, urban, geology, security, etc. Satellites have been frequently employed to get land-cover information on the earth surface. Most of the recent works in remote sensing classification topic are realized based on images from optic satellites. These types of satellites can be divided into two categories: hyperspectral (HYP) sensors and multispectral sensors, where the second category can obtain panchromatic (PAN) and multispectral (MS) images.

The advanced classification methods are based on spatial-spectral feature extraction and fusion techniques which improve their efficiency particularly those based on deep learning models. In last decade, feature extraction (FE) and fusion methods are considered as robust techniques in remote sensing image processing topics.

Traditional feature extraction methods usually have limited performance in feature learning. However, in the last twenty years[1], motivated by the hierarchical organization of the human brain, deep learning has been proposed to extract the features in a hierarchical manner, which also provides a promising direction for deep feature-based fusion[1]. The feature extraction and fusion techniques based deep learning is used recently in many frameworks like a fusion of hyperspectral (HYP) and LIDAR data, pan-sharpening, hyperspectral and multispectral data fusion, etc.

The purpose of the work is to develop a new framework based on deep feature extraction and fusion

① Supported by the National Natural Science Foundation of China(No. 61472103, 61772158, U.1711265).
② To whom correspondence should be addressed. E-mail: h.yao@hit.edu.cn
Received on May 23, 2018

of spectral and spatial information for accurate classification of multispectral (MS) and panchromatic (PAN) images. To obtain an accurate land cover and land use map, details of both fine spatial information from PAN image and the rich spectral information of MS image should be simultaneously taken into account to solve the classification problem to get a good land-cover map. The target of pansharpening technique is to integrate the information with different spatial and spectral resolution from a variety of earth observation platforms[2]. In order to get satisfactory results for the image classification task, many image fusion techniques have been proposed to merge PAN and MS images to obtain high spatial resolution and wide spectral information. Usually, the multiresolution image classification process is divided into two parts. The first one is to generate a common resolution image by means of downsampling or upsampling[3,4], unmixing[5] or pansharpening[6,7] algorithms. The second one is to classify the fused data by supervised methods. However, the artifacts or distortions from the geometric corrections and pixel-level fusion procedures[6,7] affect the classification results. There are also some other techniques proposed for multiresolution classification without the procedure of image fusion[8,9].

In this paper, the potential of a new framework based on deep learning models for feature extraction and fusion of panchromatic (PAN) and multispectral (MS) images is investigated. The purpose is to present new successful classification results and prove that the fused features extracted from PAN and MS images are more robust and deep for classification than the features extracted from PAN and MS images separately. In the proposed framework, stacked sparse autoencoder (SSAE) is used for extracting spectral features of MS images, then the sparse autoencoder (SAE) is employed for extracting the spatial features of PAN images. Finally, deep features obtained from PAN and MS images are concatenated directly as a simple fusion feature. The final step is the integration of the fused features into SVM for classification.

The rest of the paper is organized as follows. Section 1 presents the description of the proposed method. Section 2 exposes the experimental results and compares the results with some feature extraction and fusion methods. Section 3 is the conclusion.

1 Methodology

1.1 Sparse autoencoder (SAE) and stacked sparse autoencoders (SSAE)

Hinton and his collaborators proposed the first deep learning (DL) network. One of the major branches of deep learning models is sparse autoencoder which is a bioinspired hierarchical neural network and has an intrinsic ability to extract more abstract features[10].

Chen et al.[11] introduced the first work of deep learning in remote sensing field in 2013. They used a deep learning model named stacked autoencoders based feature extraction for spatial-spectral classification of hyperspectral data[10].

A standard SAE contains three layers: one input layer, one hidden layer, and one reconstruction layer. Commonly, the previous layer neurons are connected to the next layer neurons, but no connections among the same layer neurons[10]. However, according to Refs[11-15], sparse autoencoder is exposed as follows.

The input data (original image) is defined as $\{x(1), x(2), \cdots, x(n), \cdots, x(N)\}$, where $x^i \in \mathbb{R}^N$. To be convenient in the following, x is used for the input and h for the hidden when explaining SAE. In the training of SAE, there are two steps: encoding and decoding.

During the encoding step, the input vector $x \in \mathbb{R}^N$ is processed by applying a linear mapping and a nonlinear activation function to the network:

$$h = f(W_h x + b_h) \tag{1}$$

where, $W_h \in \mathbb{R}^{N\times K}$ is a weight matrix with K features, $b_h \in \mathbb{R}^K$ is the encoding bias, and f is the logistic sigmoid function as

$$f(x) = 1 + \exp(-x)^{-1} \tag{2}$$

A vector is decoded using a separate linear decoding matrix:

$$z = f(\mathbf{W}_z h + b_z) \tag{3}$$

where, $\mathbf{W}_z \in \mathbb{R}^{K\times N}$ is a weight matrix and $b_z \in \mathbb{R}^N$ is the decoding bias. W_h and $\mathbf{W}_z$ denote the input-to-hidden and the hidden-to-output weights, respectively. For rendering the parameterizations identical, $W = W_h = W_z^T$ is restrained and b_h, b_z denote the bias of hidden and output units respectively. By employing the backpropagation algorithm, features in the data are extracted by minimizing the difference between input and its reconstruction, and the objective function of autoencoder is

$$\text{Arg}\ \min_{W,b}[L(x, z)] \tag{4}$$

The reconstruction error can be measured in many ways depending on the appropriate distributional assumptions on the input given the code[13]. The traditional squared error can be used:

$$L(x, z) = \| x - z \|^2 + \lambda \| W \|^2 \tag{5}$$

However, the objective function of SAE architecture with a weight decay term and a sparsity constraint term is defined as

$$L(x, z) = \| x - z \|^2 + \lambda \| W \|^2 + \beta KL(x \| z) \quad (6)$$

The first term on the right side of Eq. (6) is an average sum-of-squares error term which represents the gap between x and z. The second term denotes the weight decay term, which is employed to reduce the autoencoder from overfitting by controlling the amplitude of the weights, λ is a weight decay parameter. The third term denotes a sparsity penalty term, β controls the weight of the term, and KL is a Kullback-Leibler divergence:

$$KL(x \| z) = \sum_{i=1}^{K} [x_i \log z_i + (1 - x_i) \log(1 - z_i)] \quad (7)$$

where, K is the number of neurons in the hidden layer, and index j is summing over the hidden units in the proposed network. For the minimization process of objective function $L(x, z)$, the stochastic gradient descent (SGD) and backpropagation algorithm are used to update parameters W and b in each iteration.

Afterward, $L(x, z)$ is expected to produce a considerably small value which prompts SA to learn abstract features from original data. Stacked sparse autoencoder (SSAE) is a layer-wise encoding neural network in which multiple layers of sparse autoencoders are stacked and pre-trained via greedy methods layer by layer[10]. SSAE yields a deep representation of input data at the output of the last layer. After finishing training a former layer of parameters, the subsequent layer is trained according to the output of its previous layer. Stacking these input-to-hidden layers sequentially constructs a stacked autoencoder[11].

The learning process of SSAE is equivalent to the SAE architecture, with the objective function is to minimize the reconstruction error. In this work, stacked sparse autoencoder is used to compute deep representations of spectral data for extracting high-level features from the original MS data. In another part, sparse autoencoder is employed for extracting spatial information from PAN data.

1.2 General framework

In this section, the structure of the proposed approach is exposed. Before the beginning of the image processing step, pre-processing should be made for the datasets. The general framework of the approach is shown in Fig. 1.

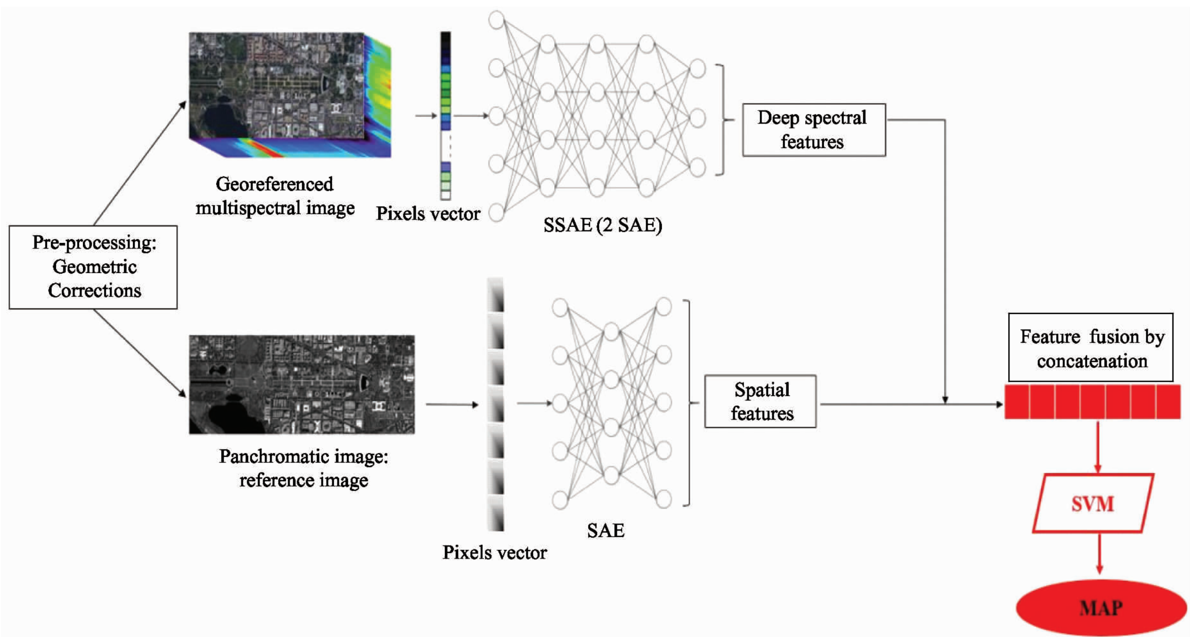


Fig. 1 The general framework

1.2.1 Image pre-processing

Before starting the image processing step, it is necessary to make pre-treatment. The satellite images downloaded from the Digital Globe website are images that have been radiometrically corrected. The multispectral MS image was georeferenced using the panchromatic (PAN) image as a reference. This process of geometric corrections has been realized using the ERDAS software. In addition, the PAN image is resized to MS resolution and dimension.

1.2.2 Image processing

The framework of the proposed method is shown in

Fig. 1. The first step is to extract deep spectral features from the MS images after the pre-processing level, using stacked sparse autoencoder and only two SAE are used. Then the high spatial information is extracted from the original PAN image using one sparse autoencoder. The third step is the feature fusion and classification. The spatial and spectral features obtained from PAN image, MS image respectively, are concatenated directly as a simple fusion in one fused data, and finally, classification is performed using the SVM classifier.

2 Experiments and discussion

In order to study the potential of the proposed framework for PAN and MS image classification, two datasets of different areas are adopted. The experiments are carried out using datasets of WorldView-2 satellite and ground truth files there. The first one is Washington DC data and the second one is Stockholm data used for the first time in Ref. [14] and Ref. [15].

2.1 Datasets

The first data set Washington DC, USA, was acquired with the WorldView-2 satellite on February 9th, 2016. As shown in Fig. 2, the dataset analyzed is available with 8 spectral bands and spatial resolution 0.4 m for panchromatic and 1.6 m for multispectral. The dimension of the MS and PAN images used for the processing is 2 438 × 896 pixels. The ground reference data constructed by visual inspection (photointerpretation), GIS tools and helped by Open Street Map (OSM), consist of 8 classes of interest as described in Table 1.

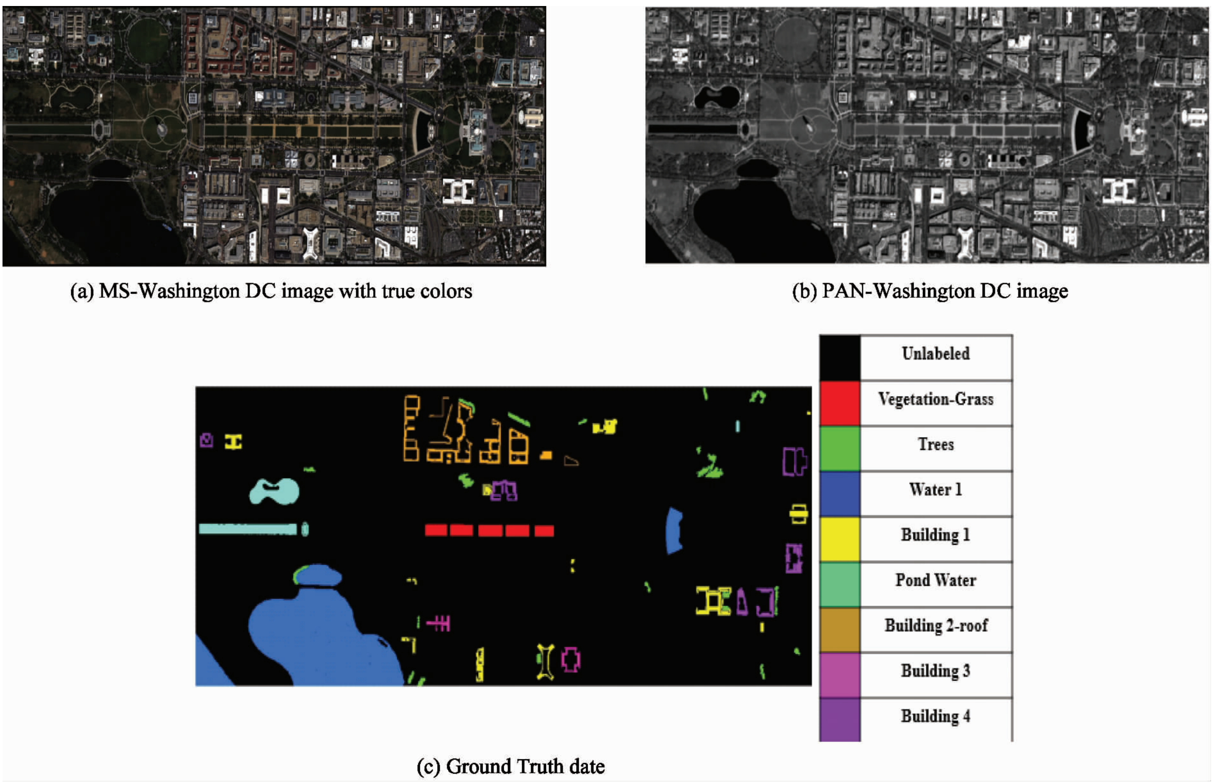
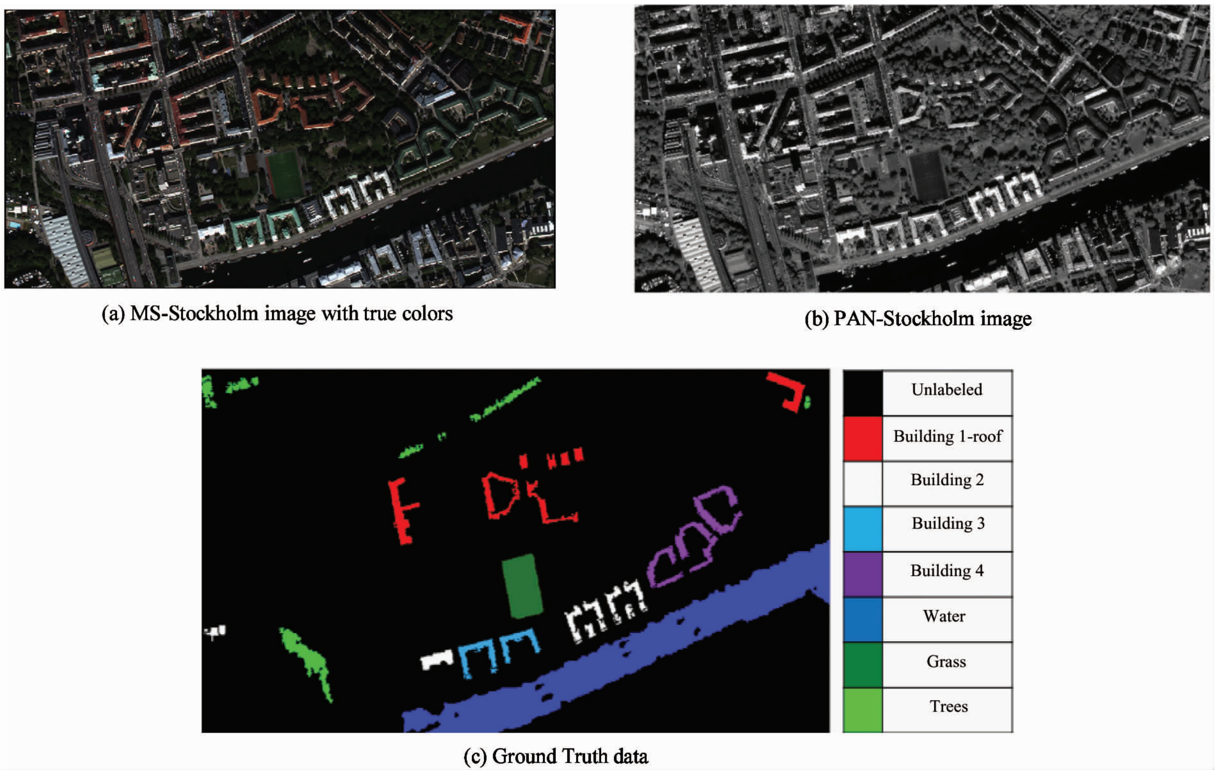


Fig. 2 (a) MS-Washington DC image with true colors, (b) PAN-Washington DC image and (c) Ground Truth date

Table 1 Eight ground reference classes of Washington DC data[15]

Class number	1	2	3	4	5	6	7	8
Class name	Grass	Trees	Water 1	Building 1	Pond water	Building 2 – roof	Building 3	Building 4
Total Samples	14 550	12 846	148 872	18 153	22 859	19 229	3 545	12 468

The second data set Stockholm, Sweden, was acquired with the WorldView-2 satellite on August 27th, 2016. As shown in Fig. 3, The dataset analyzed is available with 8 spectral bands and spatial resolution 0.4 m for panchromatic and 1.6 m for multispectral. The dimension of the MS and PAN images used for the processing is 832 × 416 pixels. The ground reference data constructed by visual inspection (photointerpretation), GIS tools and helped by Open Street Map (OSM), consist of 7 classes of interest (Table 2).



(a) MS-Stockholm image with true colors

(b) PAN-Stockholm image

(c) Ground Truth data

Fig. 3 (a) MS-Stockholm image with true colors, (b) PAN-Stockholm image and (c) Ground Truth data

Table 2 Ground reference classes of Stockholm data[15]

Class number	1	2	3	4	5	6	7
Class name	Building 1-roof	Building 2	Building 3	Building 4	Water	Grass	Trees
Total Samples	3 720	2 295	1 387	2 582	24 162	2 929	3 073

2.2 Experiments and analyses

In the experiments, sparse autoencoder (SAE) is used for extracting spatial features from the original PAN images to make full use of the spatial information around each pixel's neighborhood and keep the spatial information. For each pixel of PAN image, there are $w \times w$ neighbor pixels with a region size of $w = 28$. Since PAN image has only one channel, a pixel can be represented as a box with $w \times w \times 1$ members.

The SSAE architecture has shown high potential to learn a representation of remote sensing data with multiple levels of abstraction. In the proposed framework, SAE is employed to extract deep spectral features from the original MS image (georeferenced).

In the two datasets, the labeled parts of the images are divided into two sets: training samples and testing samples. $R = 0.01$ is chosen randomly as the training-rate (1%) from the labeled samples from each class and the rest is defined as the testing and validating set.

The MS image contains 8 bands, a single pixel can be represented as an 8-dimensional vector. Window structure can increase the discriminant information. Therefore, a pixel can be represented by a box with $w \times w \times 8$ pixels ($w = 28$).

The ground truth map of 'Washington DC' and 'Stockholm' data[14] were constructed. In each labeled scene, the black area represents the unlabeled area. Different values of the SSAE and SAE parameters were experimented, and Tables 3, 4, 5, 6 list the classification accuracy of (SSAE + SVM) and (SAE + SVM) for MS and PAN images of Washington DC and Stockholm data respectively, versus the number of neurons in a hidden layer.

Table 3 Classification accuracies from different of hidden sizes for MS-Washington DC data

Washington DC	Measurement	$h = 50$	$h = 100$	$h = 200$	$h = 300$	$h = 400$	$h = 500$
	OA (%)	98.3949	98.8054	99.0122	99.0822	99.1099	99.0690
MS – SSAE	Kappa	0.9743	0.9808	0.9842	0.9853	0.9857	0.9851
	AA	0.9685	0.9753	0.9774	0.9796	0.9811	0.9805

Table 4 Classification accuracies from different of hidden sizes for PAN-Washington DC data

Washington DC	Measurement	50	100	200	300	400	500
	OA (%)	87.4328	87.48	87.6053	87.8099	87.5511	87.5071
PAN – SAE	Kappa	0.7989	0.7997	0.8000	0.8049	0.80	0.8000
	AA	0.6492	0.6495	0.6616	0.7265	0.6568	0.6496

Table 5 Classification accuracies from different of hidden sizes for MS-Stockholm data

Stockholm	Measurement	$h = 50$	$h = 100$	$h = 200$	$h = 300$	$h = 400$	$h = 500$
	OA (%)	98.6659	98.7877	98.8348	99.0534	99.0562	98.9842
MS – SSAE	Kappa	0.9780	0.9801	0.9808	0.9844	0.9845	0.9833
	AA	0.9692	0.9738	0.9746	0.9803	0.9805	0.9782

Table 6 Classification accuracies from different of hidden sizes for PAN-Stockholm data

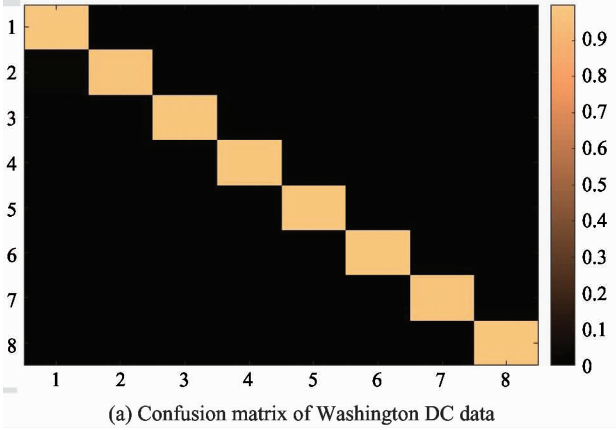
Stockholm	Measurement	50	100	200	300	400	500
	OA (%)	81.4392	81.4725	81.7575	81.8572	81.9374	81.2981
PAN – SAE	Kappa	0.6919	0.6922	0.6976	0.6993	0.7009	0.6892
	AA	0.5572	0.5614	0.5638	0.5658	0.5711	0.5533

The fusion of each instance of MS-SSAE with each instance of PAN-SAE is used, meaning that each instance of MS-SSAE (according to Table 3) is fused with six instances of PAN-SAE (Table 4).

The best classification performance for Washington DC data can be obtained ($h_1 = h_2 = 400$, for MS-SSAE) fused with features of PAN-SAE with hidden size ($h = 300$). On the other hand, the best classification performance for Stockholm data can be obtained with $h_1 = h_2 = 400$, for MS-SSAE and fused with features of PAN-SAE with hidden size ($h = 400$). The highest overall accuracy is estimated at 99.84% for Washington DC data and 99.33% for Stockholm data (Table 7, Fig. 4).

Table 7 Classification accuracies of the proposed method for the two datasets

Data	OA(%)	Kappa	AA
Washington DC	99.84	99.63	0.9938
Stockholm	99.33	98.91	0.9860



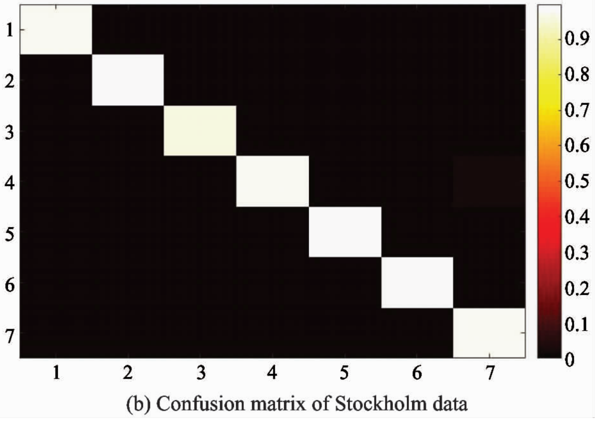


Fig. 4 Confusion matrix

For these two datasets, the proposed method was trained by stochastic gradient descent, with a batch size of 64, the maximum number of iteration is 1 000 for both SAE and SSAE, and sparsity penalty value is 0.001 for both MS-SSAE and PAN-SAE. In addition, there is no important change in the classification accuracy using a big number of neurons for the same training rate.

The SVM classification accuracy after feature fusion (MS-SSAE + PAN-SAE = feature fusion fed into SVM) is exposed in Table 7 such as, in SVM, the kernel function is radial basis function (RBF), the semi-radius of the kernel function $g = 1$ and penalized parameters $c = 100$. The accuracy assessment of the proposed approach is demonstrated by the overall accuracy (OA), average accuracy (AA) and the kappa coefficient. In addition, the individual class accuracies were evaluated by the producer's accuracy (PA) and the user's accuracy (UA) measures (Table 8, 9). The results are compared with different frameworks applied to the same datasets (Table 10). These comparative methods are exposed with their best parameters such as SVM is employed for all the comparative methods for classification with the same parameters like in the approach where the overall classification accuracies are shown in Table 10.

Table 8 Individual class accuracies by PA and UA for Washington DC data

Washington DC	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8
PA	0.9861	0.9878	0.9959	0.9989	0.9958	0.9984	0.9961	0.9973
UA	0.9973	0.9880	0.9976	0.9996	0.9917	0.9973	0.9981	0.9968

Table 9 Individual class accuracies by PA and UA for Stockholm data

Stockholm	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
PA	0.9992	0.9903	0.9905	0.9994	0.9850	0.9775	0.9801
UA	0.9970	0.9927	0.9850	0.9999	0.9800	0.9788	0.9790

Table 10 Comparison of the overall classification accuracies (OA)

	Proposed method	M1	M2	M3	M4	M5	M6	M7	M8
Washington DC (%)	99.84	99.10	87.80	99.56	99.64	98.80	98.58	98.79	99.01
Stockholm (%)	99.33	99.05	81.93	98.70	99.13	98.27	98.43	98.12	98.77

The comparative methods are presented as follows.

Method 1 (M1): Stacked sparse autoencoder for the multispectral images (MS-SSAE) using two SAE where the hidden size is $h_1 = h_2 = 400$ for both Washington DC and Stockholm data (Table 3, 5).

Method 2 (M2): Sparse autoencoder for the panchromatic images (PAN-SAE) with hidden size $h = 300$ for Washington DC data and $h = 400$ for Stockholm data (Table 4, 6).

Method 3 (M3)[14]: (MS-EMAP) + SAE such as EMAP + SAE applied to MS image consists of two steps. The first step is to construct EMAP = [EAP_a EAP_s], where λ_a = [100 500], λ_s = [25 125], exploiting the spatial information. The second step is to use sparse autoencoder on the obtained features to reduce the high dimensionality of EMAP and extract the robust spatial features (with $h = 10$ for Washington DC data and $h = 25$ for Stockholm data). Classification is performed on the reconstruction layer of SAE by using a multiclass SVM.

Method 4 (M4)[15]: [(MS-Spectral + EMAP) + SAE] such as (Spectral + EMAP) + SAE applied to MS image consists of: the first step is to build EMAP = [EAP_a EAP_s] with λ_a = [100 500], λ_s = [25 125]. After that, the spectral information of the original image is combined with spatial information. Then, SAE is used (with hidden size $h = 10$ for Washington DC data and $h = 25$ for Stockholm data) for feature extraction and dimensionality reduction. Finally, the classification is performed on the reconstruction layer of SAE by using SVM.

Method 5 (M5): Fusion of MS-SSAE and (PAN-EAP) + SAE using two SAE where the hidden size for both Washington DC and Stockholm data is $h_1 = h_2 = 400$. The used EAP is EAP_a with $\lambda_a = 500$. The hidden size of SAE is $h = 5$ for Washington DC data and $h = 15$ for Stockholm data. The spectral (MS) and spatial (PAN) features are simply concatenated and the classification is performed by SVM on the fused vector.

Method 6 (M6): Fusion of MS-SSAE and (PAN-Spectral + EAP) where for MS-SSAE two SAE with the hidden size $h_1 = h_2 = 400$ are used for both Washington DC and Stockholm data. For (PAN-Spectral + EAP) + SAE, The used EAP is EAP_a with $\lambda_a = 500$ and the hidden size of SAE is $h = 5$ for Washington DC data and $h = 15$ for Stockholm data. The classification is performed using the concatenated spatial and spectral features as the input of the SVM classifier.

Method 7 (M7): SVM for MS image where the kernel function is radial basis function (RBF), the semi-radius of the kernel function $g = 1$ and penalized parameters $c = 100$.

Method 8 (M8): MS $\oplus$ PAN fed into SSAE using two SAE where the hidden size for both Washington DC and Stockholm data is $h_1 = h_2 = 600$. In addition, SVM is used for performing the classification where the kernel function is radial basis function (RBF), the semi-radius of the kernel function $g = 1$ and penalized parameters $c = 100$.

As it is shown in Table 7, classification results using feature extraction and techniques present high accuracies and the proposed method has the best accuracy.

In addition to the high accuracy obtained by the proposed approach, the UA and PA individual accuracies demonstrate (Table 8, 9) that the proposed framework can extract greatly features from the built classes (buildings). Thus, this approach operated with small training rate in urban areas gave satisfactory results for classification problem.

3 Conclusion

In this paper, a new feature extraction and fusion strategy are proposed for remote sensing image classification based on the spatial-spectral information. The purpose of the work (D-SS Frame) is to develop a new framework based on deep feature extraction and fusion of spectral and spatial information for accurate classification of multispectral (MS) and panchromatic (PAN) imagery.

Firstly, spectral features are extracted from MS image by stacked sparse autoencoder (SSAE) network and get spatial features from PAN image by sparse autoencoder (SAE), respectively. Secondly, two kinds of features are concatenated, as a simple feature fusion manner. Finally, the SVM classifier is used to identify the high features.

The proposed method obtains satisfactory results where the classification results are very close to the ground truth map. D-SS Frame proves its robustness compared to others frameworks based on feature extraction and fusion or only feature extraction for MS and PAN image classification. In addition, the efficiency and the robustness of the algorithm on other optical remote sensing images with different spatial resolutions is a very important factor, which will be researched in the future work.

Acknowledgments

The MS and PAN were downloaded from the website (www. digitalglobe. com) and the ground truth data of the two datasets were constructed by first author and will be published.

References

[1] Chen Y, Li C, Ghamisi P, et al. Deep fusion of remote sensing data for accurate classification[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(8): 1253-1257

[2] Mao T, Tang H, Wu J, et al. A generalized metaphor of Chinese restaurant franchise to fusing both panchromatic and multispectral images for unsupervised classification [J]. *IEEE Transaction on Geoscience and Remote Sensing*, 2016, 54(8): 4594-4604

[3] Forzieri G, Moser G, Vivoni E R, et al. Riparian vegetation mapping for hydraulic roughness estimation using very high-resolution remote sensing data fusion[J]. *Journal of Hydraulic Engineering*, 2010, 136(11): 855-867

[4] Chen D M, Stow D. Strategies for integrating information from multiple spatial resolutions into Land-Use/ Land-Cover classification routines[J]. *Photogrammetric Engineering & Remote Sensing*, 2003, 69(11): 1279-1287

[5] Price J C. Combining multispectral data of differing spatial resolution[J]. *IEEE Transaction on Geoscience and Remote Sensing*, 1999, 37(3): 1199-1203

[6] Konstantinos G. Nikolakopoulos. Comparison of nine fusion techniques for very high-resolution data[J]. *Photogrammetric Engineering & Remote Sensing*, 2008, 74(5): 647-659

[7] Colditz R, Wehrmann T, Bachmann M, et al. Influence of image fusion approaches on classification accuracy-a case study [J]. *International Journal Remote Sensing*, 2006, 27(15): 3311-3335

[8] Wemmert C, Puissant A, Forestier G, et al. Multiresolution remote sensing image clustering[J]. *IEEE Geoscience and Remote Sensing Letters*, 2009, 6(3): 533-537

[9] Redner R A, Walker H F. Mixture densities, maximum likelihood and the EM algorithm [J]. *SIAM Review*, 2012, 26(2): 195-239

[10] Wan X, Zhao C, Wang Y, et al. Stacked sparse autoencoder in hyperspectral data classification using spectral-spatial, higher order statistics and multifractal spectrum

features[J]. *Infrared Physics and Technology*, 2017, 86: 77-89

[11] Chen Y, Lin Z, Zhao X, et al. Deep learning-based classification of hyperspectral data[J]. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, 2014, 7(6):2094-2107

[12] Andrew Ng. Sparse Autoencoder[M]. CS294A Lecture Notes, Stanford University, 2010

[13] Wang L, Zhang J, Liu P, et al. Spectral-spatial multi-feature-based deep learning for hyperspectral remote sensing image classification[J]. *Soft Computing*, 2017, 21(1): 213-221

[14] Teffahi H, Yao H, Belabid N, et al. Feature extraction based on extend multi-attribute profiles and sparse autoencoder for remote sensing image classification[C]. In: Proceedings of the Multispectral Image Acquisition, Processing, and Analysis, 2018,10607:106070W-1

[15] Teffahi H, Yao H, Chaib S, et al. A novel spectral-spatial classification technique of multispectral images using extend multi-attribute profiles and sparse autoencoder [J]. *Remote Sensing Letters*, 2018,10(1):30-38

TEFFAHI HANANE, born in Algiers, Algeria, in 1989. She received the Bachelor degree in 2009 and Master degree in 2011. She majors in the Operational Research (Mathematics). She was Ph. D student in Computer Science and Technology of Harbin Institute of Technology (Harbin, China) and she graduated on October 2018. She is currently a researcher in Algerian Space Agency. Her research interests include the remote sensing image classification, feature extraction and fusion, deep learning and machine learning.