Multisource Domain Transfer Learning Based on Spectral Projections for Hyperspectral Image Classification

Bin Yang, Shunshi Hu, Qiandong Guo, and Danfeng Hong, *Senior Member, IEEE*

***Abstract*—Hyperspectral image classification is an important topic for hyperspectral remote sensing with various applications. Hyperspectral image classification accuracy. It has been greatly improved with the introduction of deep neural networks, while the idea of transfer learning provides an opportunity to solve the problem even with the lack of training samples. In this article, we propose an effective transfer learning approach for hyperspectral images, projecting hyperspectral images with different sensors and different number of bands into a general spectral space, preserving the relative positions of each band for spectral alignment, and designing a hierarchical depth neural network for shallow feature transfer and deep feature classification. The experiments show that the proposed method can effectively preserve the source domain features, especially for the scenarios with very few samples in the target domain, which can significantly improve the classification accuracy and reduce the risk of model overfitting. Meanwhile, this strategy greatly reduces the requirement of source domain data, using multisensor data to jointly train a more robust general feature model. The proposed method can achieve high accuracies even with few training samples compared to currently many state-of-the-art classification methods.**

***Index Terms*—Classification, convolutional neural network (CNN), generalized feature extraction network (GFEN), hyperspectral imaging, spectral projection, transfer learning.**

I. Introduction

HYPERSPECTRAL images can capture tens or even hundreds of bands for the object to be measured, with nanometers spectral resolution ranging from the visible light bands to the near-infrared bands, which makes the analysis of remote sensing data can be jointly carried out in both spatial and spectral dimensions, greatly enriching the applications of remote sensing and expanding its use cases [1]. Hyperspectral image classification is a classical problem in hyperspectral applications. The continuous bandwidth obtained by hyperspectral remote sensing is generally within 10 nm, which can distinguish the surface materials with diagnostic spectral characteristics in sufficient spectral resolution. Thus, the materials that are hard to be identified in multispectral images can be distinguished in hyperspectral images. Because of these features, hyperspectral imageries are well suited for automatic feature identification at the pixel level in large scenes [2]–[8].

With the great success of applying deep learning in the field of computer vision, a large number of scholars have introduced deep neural networks into the hyperspectral imagery analysis [2], [4]. Deep neural networks can extract knowledge from large-scale data. Their multilayer network structure implies that higher level semantic information of the samples can be obtained. There are many deep learning models for hyperspectral image classification, among which convolutional neural networks (CNNs) as the main classification models, have achieved impressive performance. Specifically, one-dimensional (1-D) CNN [9] classifier based on the spectrum of a single image element, while 3-D CNN [10], [11] combined both spatial and spectral information and incorporates the influence of neighboring pixels on the category. SSRN [12] added a residual module to the CNN to guarantee that the network can reach deeper depths. Moreover, various optimizations to improve the network's adaptability and increase its operational efficiency [13]–[19]. All these methods can achieve a satisfactory classification accuracy with sufficient training samples. However, for deep neural networks, the dense spectral and spatial features require a large number of labeled samples to ensure that the network is adequately trained. Obtaining accurate labeling of features in hyperspectral images is usually time-consuming and expensive. Thus, the needs for a high volume of training samples have been an important obstacle to the performance of deep learning models.

Transfer learning provides an important means to solve the abovementioned hyperspectral image classification problem. Transfer learning proposed to learn knowledge from existing classification tasks (source domain) and applied it to new scenarios (target domain). The deep networks designed by transfer learning methods aimed to make full use of source domain data to solve the problem of sparse training samples in the target domain, while cross-domain classification models can better

Manuscript received March 16, 2022; revised April 19, 2022; accepted May 6, 2022. Date of publication May 16, 2022; date of current version May 19, 2022. This work was supported in part by the Scientific Research Project by Hunan Provincial Department of Education under Grant 21B0046 and in part by 2019 Hunan Postgraduate High Quality Course Project "Microwave and Hyperspectral Remote Sensing" under Grant 130. *(Corresponding author: Danfeng Hong.)*

Bin Yang and Shunshi Hu are with the Hunan Key Laboratory of Geospatial Big Data Mining and Application, Hunan Normal University, Changsha 410081, China (e-mail: yangbin@hunnu.edu.cn; shunshi.hu@hunnu.edu.cn).

Qiandong Guo is with the School of Geosciences, University of South Florida, Tampa, FL 33620 USA (e-mail: guo1@mail.usf.edu).

Danfeng Hong is with the Key Laboratory of Computational Optical Imaging Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China (e-mail: hongdf@aircas.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2022.3173676

extract invariant features and improve network generalization ability. In recent years, some hyperspectral image transfer learning methods have been proposed. Zhang *et al.* [20] used 3-D CNN structure to learn parameters using hyperspectral images of different sensors and RGB images. He *et al.* [21] combined deep networks with an attention mechanism to migrate VGG deep networks to hyperspectral classification domain. However, transfer learning works only if there are similarities between the source and target domains, and the network needs to effectively use these similarities to apply the source domain knowledge to the target domain. It is well known that hyperspectral images acquired by different scenes and sensors vary significantly, and how to reduce interdomain differences is a key issue to improve the effectiveness of transfer learning. Lin *et al.* [22] extracted significant samples from source and target domains and fine-tuned the network layer to layer by analyzing interdomain correlations. Deng *et al.* [23] proposed a cascaded sparse self-encoder to learn joint spectral space features and migrate them to the target domain. Qu *et al.* [24] proposed a domain adaptation method based on a shared abundance space to reduce the differences between source and target domains. Hong *et al.* [25] believed that in remote sensing image classification, the combination of multimodal deep learning brings better effect. Hang *et al.* [26] proposed a multitask generative adversarial network to undertake the hyperspectral reconstruction and classification task. In [27] and [28], LiDAR and hyperspectral data are fused for multisource classification. All of these methods have achieved solid results.

When studying the invariant characteristics among hyperspectral images, it is worth noting that the central wavelengths (bands) of the spectral values of the features acquired by whatever sensors are known, and there are specific reflection patterns for various features in different bands. It is the mechanism of hyperspectral remote sensing to investigate the reflection characteristics of various features in specific spectral bands. However, in the existing transfer learning models for hyperspectral image classification, few models that take advantage of this known information. In order to use this information, this article presents the following work accomplished.

1) A spectral preprocessing approach was used, i.e., a spectral mapping (resampling) of hyperspectral images acquired by different sensors to a manually selected high-dimensional space. By analyzing the sensor's parameters and available bands of typical hyperspectral datasets, a common spectral space containing different sensors' bands is constructed. Then, the hyperspectral data is projected into this high-dimensional space, with no data or removed bands filled with 0 values to achieve the effect of spectral position preservation and alignment.
2) Corresponding to this preprocessing approach, a hierarchical deep neural network was designed for transfer learning and classification. The shallow layer of the network was used for the extraction and migration of general diagnostic spectral features of hyperspectral images, and the deep one was used for further mining of exclusive features of objects to be classified.
3) Experimentally validate the classification efficiency of the network, compare and analyze the transfer learning performance on different datasets.
4) Use multiple datasets to train the generic feature extraction network together with multitask learning. Then, build a multisource domain integrated network with generalization capability for future classification applications using fewer samples.

The rest of this article is organized as follows. Section II describes proposed spectral projection strategy and introduces the detailed architecture of deep neural network. The experiment design and results are reported in Section III. Finally, Section IV concludes this article.

II. Methodology

A. Spectral Projection Strategy

The ability of hyperspectral remote sensing to effectively identify objects stems from the fact that it has sufficient spectral resolution to acquire diagnostic spectra of targets. However, when using hyperspectral images as input, the exact location distribution of the bands in the spectrum is rarely considered. This produces no performance loss when the classifier is trained and used independently, because a well-designed classification network can find and utilize these features automatically. Thus, there is no significant performance difference in the trained classification model between a dataset with ordered spectra and with randomly disordered spectra (as long as the training and test sets have the same bands). However, using a trained network for new data with different band alignments will make it difficult to conduct the feature extraction.

On the other hand, neural networks usually require the same dimensional input, and existing transfer learning for hyperspectral image classification either uses dimensionality reduction, such as principal component analysis or independent component analysis, or adds additional adaptive layers to have the same dimensionality. Either way is difficult to satisfy the plain idea of maintaining the original spectral segment position.

For this reason, this article designs a spectral reprojection method, which constructs a common spectral space containing different sensor bands by analyzing the band wavelength of popular sensors, arranges the bands from different sensors in the ascending order of their central wavelengths, and merges the bands with wavelength differences less than 2 nm. Taking AVIRIS and Hyperion sensors as an example, there are 224 original bands of AVIRIS and 242 original bands of Hyperion, and the total number of bands after merging is 421, including 45 shared (merged) bands. As shown in Fig. 1(a) and (c) is the original spectral data, and Fig. 1(b) and (d) is the waveforms after spectral projection. The spectral bands obtained by different sensors after projection are aligned. The data-free places are filled with zero values to obtain equal-length sparse spectra.

B. Generalized Feature Extraction Network

From the abovementioned image after spectral projection, this article designs a hierarchical depth neural network named

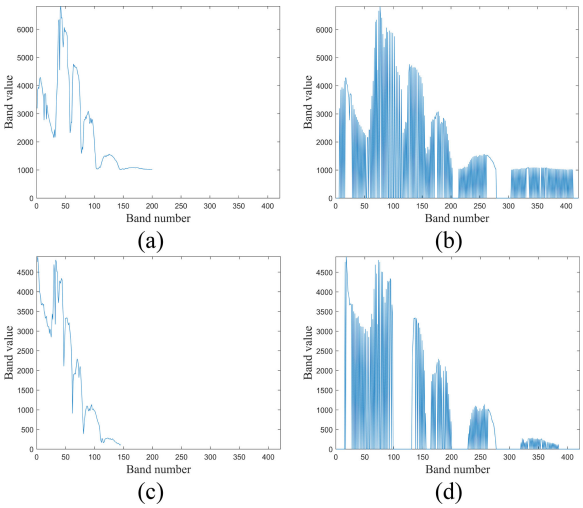

Fig. 1. Spectral projection example diagram. (a) AVIRIS raw data spectrum, 200 bands. (b) AVIRIS Post-projection spectrum. (c) Hyperion raw data spectrum, 145 bands. (d) Hyperion postprojection spectrum.

TABLE I
PARAMETERS OF GFEN NETWORK

Layer Name	Filter Size	Output Shape	BN+Relu
General Feature Extraction Layers			
Input	N/A	7×7×421	N
Conv3D_1	(7×7×419)×128	128×(7×7×3)	Y
Conv3D_2	(1×1×1)×128	128×(7×7×3)	Y
Conv3D_3	(1×1×1)×24	24×(7×7×3)	Y
Deep Feature Extraction Layers (Repeat *N* times, *N*=5)			
Conv3D_n1	(1×1×1)×24	24×(7×7×3)	Y
Conv3D_n2	(1×1×1)×24	24×(7×7×3)	Y
Residual-Add	N/A	24×(7×7×3)	N
Neighborhood Feature Fusion and Classification			
Conv3DOut1	(1×1×3)×128	128×(7×7×1)	Y
Reshape	N/A	128×(7×7)	N
Conv_Gauss	(7×7)×128	128×(1×1)	N
Flatten	N/A	128	N
Output	N/A	Class Number	N

generalized feature extraction network (GFEN). The network is based on the architecture of a residual 3-D convolutional network, and is divided into universal feature extraction layer, depth feature extraction layer, neighborhood feature fusion, and output layer. Its workflow is shown in Fig. 2.

The network input is a 7 × 7 neighborhood size spectral data centered on the image element to be divided, i.e., a 7 × 7× B size tensor is input to the input layer, and B is the number of bands. The first layer uses 128 (1 × 1× B) 3-D convolution kernels to perform global convolution on the input tensor. Considering the influence of spectral projection deviation and spectral shift of the data itself, the global convolution kernel is allowed to make a little sliding operation, so the size of the convolution kernel is adjusted to 1 × 1 × (B-2), i.e., three-step sliding convolution. The second layer uses 128 (1 × 1 × 1) 3-D convolution kernels to combine the feature spectral segments and generate 128 (7 × 7 × 3) feature tensor. The third layer is the transition layer, which uses 24 (1 × 1 × 1) 3-D convolution kernels to bring the number of combined features down to 24 for subsequent deep learning. Unlike a large number of image classification networks that extract image features with smaller convolution kernels sliding (especially in the field of machine vision), the authors believe that hyperspectral images after spectral projection, which have stable positions of shallow spectral dimensional features and do not require sliding to extract repeated features, can extract global features in the spectrum more effectively by directly convolving in the depth direction with 3-D convolution kernels of equal or approximate length to the number of bands (with little sliding to cope with errors). The global features and local features (when the convolution kernels are sparse) are extracted more effectively using the proposed method.

The deep feature extraction layer uses N feature extraction modules with the same structure, each module contains two layers of 24 kernels of (1 × 1 × 1) 3-D convolution operations, and the residual calculation is performed once after each module execution. The introduction of the residual operation can effectively deepen the network structure to avoid gradient disappearance and explosion problems, while a deeper network structure can accommodate more parameters and improve the depth feature extraction capability. Considering the performance and arithmetic power, the number of modules N used in this article is set to 5, i.e., the depth feature extraction layer consists of 5 feature extraction modules in series, containing 10 layers of convolution, and the output is still 24 (7 × 7 × 3) feature tensor.

The neighborhood feature fusion layer first uses a (1 × 1 × 3) convolution kernel to merge spectral dimensional features for final feature classification. Before classification, a 2-D Gaussian convolutional kernel with fixed parameters is designed for the fusion of spectral features in the 7 × 7 domain range

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \tag{1}$$

Equation (1) represents the generating function of the Gaussian convolution kernel, where x, y denotes the distance from the centroid and σ is the variable variance. A higher σ means a flatter distribution, while a lower σ means a more prominent center. We experimentally tried different σ values, and finally set sigma to the empirical value 7.

It can be found that the present model makes extensive use of (1 × 1 × 1) 3-D convolution kernel, and the spectral features of each pixel in the 7 × 7 neighborhood are extracted separately, while the use of Gaussian convolution kernel can calculate the influence on the central image element by the spatial distance of the neighboring spectral features. As shown in Fig. 2, this Gaussian kernel convolves and merges the 128 feature subchannels extracted by the network separately to finally obtain the depth spectral feature of the central image element. The spectral features are flattened, and the output is activated using the soft-max function with a fully connected network as the output layer. Each convolutional layer of this model is connected with a batch normal layer and a Relu activation layer, and the input hyperspectral data is normalized by demeaning in a band. The parameters of the network are shown in Table I.

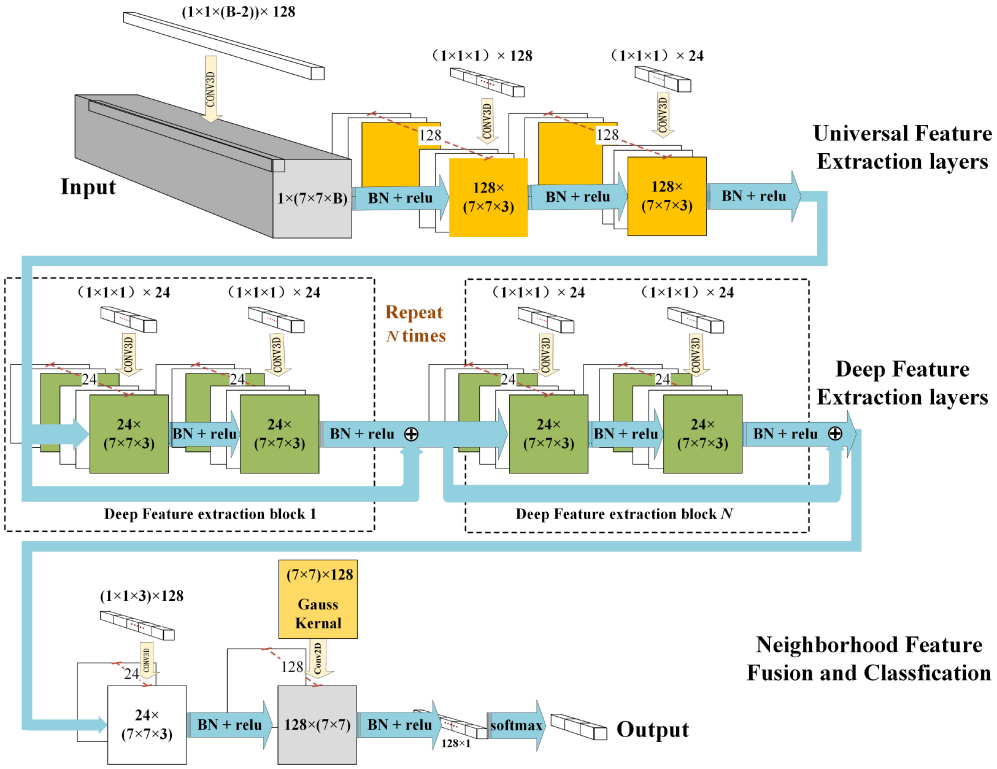

Fig. 2. Schematic diagram of GFEN structure.

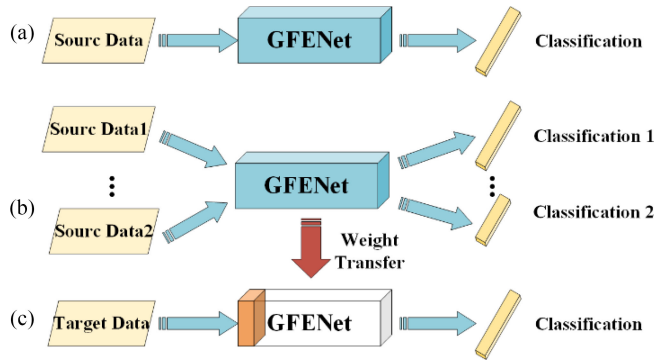

Fig. 3. (a) Single-source domain learning. (b) Multisource domain learning. (c) Parameter transfer with target domain learning.

C. Network Transfer Approach

The transfer learning using this model aims to train a robust classification network using hyperspectral data in the source domain containing a large number of training samples, and migrate the network model to the target domain with only a small number of samples for relearning and classification. The shallow layer of the network acquires generic spectral features, and the deep layer contains object-specific depth features in the target domain.

1) *Training the source domain classification network*: The training can be performed with a single source domain or, with multiple source domains of data sharing the middle layer of the network for joint training of multiple tasks, as shown in Fig. 3(a) and (b).

2) *Target domain migration learning*: We will compare two migration approaches. First, the network parameters obtained from the source domain training are migrated to the target domain model, while the shallow layer of the network is frozen (orange part in Fig. 3 and Table I), and the deeper network is fine-tuned with the target domain samples only to reacquire the features exclusive to the target task. Second, based on the network obtained from the previous training step, the shallow network is unfrozen and fine-tuned again to classify the required. The second step is to thaw the shallow network based on the network trained in the previous step, and repeatedly fine-tune the shallow features required for classification to improve classification accuracy.

III. Experiment Results and Analysis

A. Experimental Data

In this article, three popular datasets collected by two hyperspectral sensors are used for the experiments.

1) Indian pine dataset: This dataset was collected by the AVIRIS sensor at the Indian pine test site in northwestern Indiana, it consists of 145 × 145 pixels and 224 spectral reflection bands in the wavelength range of 400–2500 nm. Due to data quality, bands 1, 33, 97, 161, and 20 water vapor absorption bands ([104–108], [150–163], 220) were removed, and 200 bands were finally retained. The dataset contains 10249 samples with labels, which are classified into 16 feature categories, and the specific sample distribution is shown in Table II.

2) Salinas dataset: Also collected by the AVIRIS sensor, 204 bands were retained after excluding the water vapor absorption

TABLE II
INDIAN PINE DATASET SAMPLE DISTRIBUTION

	Class	Samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93
	Total:	**10249**

TABLE III
SALINAS DATASET SAMPLE DISTRIBUTION

	Class	Samples
1	Brocoli-green-weeds-1	2009
2	Brocoli-green-weeds-2	3726
3	Fallow	1976
4	Fallow-rough-plow	1394
5	Fallow-smooth	2678
6	Stubble	3959
7	Celery	3579
8	Grapes-untrained	11271
9	Soil-vinyard-develop	6203
10	Corn-senesced-green-weeds	3278
11	Lettuce-romaine-4wk	1068
12	Lettuce-romaine-5wk	1927
13	Lettuce-romaine-6wk	916
14	Lettuce-romaine-7wk	1070
15	Vinyard-untrained	7268
16	Vinyard-vertical-trellis	1807
	Total:	**54129**

bands ([108–112], [154–167], 224). This data contains 512 × 217 pixels with a spatial resolution of 3.7 m. It has 54 129 samples, divided into 16 types of features, as shown in Table III.

3) Botswana dataset: This data was acquired by the Hyperion sensor on NASA EO-1 satellite. The raw data has 242 bands, again covering the spectral range of 400–2500 nm with a spectral resolution of less than 10 nm and a spatial resolution of 30 m. To mitigate the effects of sensor calibration errors and equipment anomalies, only 145 of these bands ([10–55], [82–97], [102–119], [134–164], [187–220]) were used as candidate features in the dataset. The dataset contains 1476 × 256 pixels with 3248 tagged samples divided into 14 categories as shown in Table IV.

In the abovementioned three datasets, the first two datasets are collected from the AVIRIS sensor, but contain slightly different bands. The last dataset is collected from the Hyperion sensor, and the number of bands and the spatial resolution are significantly different, which can be used for transfer verification of data

TABLE IV
BOTSWANA DATASET SAMPLE DISTRIBUTION

	Class	Samples
1	Water	270
2	Hippo grass	101
3	Floodplain grasses1	251
4	Floodplain grasses2	215
5	Reeds	269
6	Riparian	269
7	Firescar	259
8	Island interior	203
9	Acacia woodlands	314
10	Acacia shrublands	248
11	Acacia grasslands	305
12	Short mopane	181
13	Mixed mopane	268
14	Long mopane	95
	Total:	**3248**

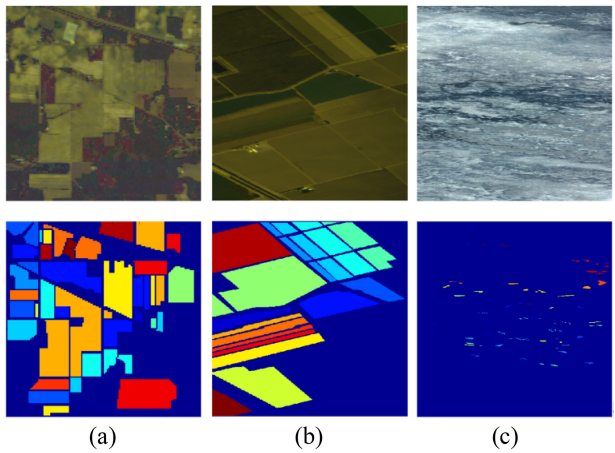

Fig. 4. Hyperspectral image classification dataset with pseudocolor images of some regions and classification maps.

from different sensors. Fig. 4 shows the pseudocolor images and classification maps for some areas of the three datasets.

B. Experimental Setup and Validation

The experiments first validated the effectiveness of the GFEN model proposed in this article for training and classification on a single dataset, then illustrated the role brought by spectral alignment in the feature migration process, and then validated the classification accuracy of migration learning for each of the three datasets, finally made a design and outlook on the construction of a multitask classification model. The experiments used overall accuracy (OA), average accuracy (AA), and Kappa coefficient as performance evaluation metrics. Except for the first experiment, which used 200 training samples in order to facilitate performance comparison with other models, we focused on testing the classification performance in the case of very small number of samples.

Experiment 1: Single-dataset classification performance test

The single dataset classification experiments were selected to test the Indian pine and Salinas datasets, with reference to the four advanced algorithms compared in [21], which are CNN, SSRN, VGG, HT-CNN-Attention. In this article, the GFEN

TABLE V
RESULTS OF CLASSIFICATION EXPERIMENTS ON INDIAN PINE DATASET

Class No.	CNN	SSRN	VGG	HT-CNN-Att.	GFEN
OA	86.81	83.21	87.49	90.86	**95.42**
AA	68.30	62.88	65.69	74.68	**83.82**
kappa(%)	84.20	80.86	84.94	89.05	**94.77**
1	38.70	25.00	36.69	**39.06**	0.00
2	84.64	87.16	88.58	**93.81**	91.43
3	63.15	95.61	41.56	60.52	**97.67**
4	55.37	20.60	50.74	41.96	**92.67**
5	56.41	84.93	51.77	68.53	**97.68**
6	96.90	99.71	92.58	95.86	**99.16**
7	42.54	28.70	15.56	74.07	**96.30**
8	85.84	97.65	86.20	86.38	**100.00**
9	12.75	9.28	8.89	**48.42**	0.00
10	90.33	80.77	91.50	**93.56**	93.07
11	94.21	83.09	96.28	95.69	**97.38**
12	58.07	70.77	62.57	87.93	**97.25**
13	85.99	98.48	88.29	96.49	**98.51**
14	96.63	96.99	97.63	95.07	**99.60**
15	67.38	38.76	81.58	66.50	82.59
16	63.96	53.85	60.62	51.08	**97.80**
Traning Time (sec.)	23.5	132.51	32.28	26.92	96.16
Test Time(sec.)	9.23	3.68	10.32	5.33	3.12

TABLE VI
RESULTS OF CLASSIFICATION EXPERIMENTS ON SALINAS DATASET

Class No.	CNN	SSRN	VGG	HT-CNN-Att.	GFEN
OA	88.40	88.73	88.97	**94.70**	89.04
AA	92.48	92.63	92.83	92.37	**95.65**
Kappa(%)	88.27	87.51	87.29	**93.62**	87.88
1	81.75	**100.00**	85.25	97.42	99.57
2	88.89	**100.00**	96.00	98.85	93.35
3	89.35	72.85	89.24	75.55	**99.79**
4	80.01	99.78	**100.00**	**100.00**	99.23
5	88.40	98.68	98.46	98.35	**99.75**
6	90.45	99.85	100.00	92.50	**100.00**
7	97.96	99.43	79.68	100.00	**100.00**
8	69.63	62.90	74.50	**99.22**	55.35
9	89.33	97.87	99.67	**100.00**	99.71
10	85.75	88.08	96.70	95.28	**98.65**
11	88.92	86.54	83.82	98.75	**98.90**
12	82.07	99.41	98.91	**99.58**	99.54
13	82.65	84.74	93.57	66.05	**98.21**
14	85.41	**99.33**	79.13	74.12	97.38
15	76.80	93.40	87.21	84.25	**94.12**
16	56.06	98.71	75.14	57.26	**98.11**
Traning Time (sec.)	28.99	141.57	63.62	22.44	102.35
Test Time(sec.)	82.52	66.69	88.43	88.48	42.88

algorithm is compared with the same condition. In total, 200 randomly selected labeled samples were trained for each dataset, and the remaining samples were used for testing. The test results are shown in Tables V and VI. The experimental results show that in the test of Indian Pine data, the OA, AA, and Kappa coefficients obtained by the GFEN are better than the other four types of algorithms. In the test of Salinas data, the AA of GFEN is higher than the other algorithms, but the OA and Kappa are lower than the HT-CNN with attention mechanism.

We conduct our experiments on a computer with an Intel Core i7-8700K CPU with 3.7 GHz, 16 GB of DDR4 RAM, an NVIDIA GeForce GTX 1080Ti GPU. As shown in Tables V–VI, the training time of GFEN is shorter than SSRN and longer than other compared models, but the test time of GFEN is the shortest in five algorithms.

It can be concluded that the GFEN model designed in this article has no less than the other excellent networks in hyperspectral image classification performance.

TABLE VII
OA TABLE OF THE INDIAN PINE DATASET CLASS A/B TRANSFER LEARNING TEST

Samples per class		A、Consistent-band transfer		B、Inconsistent-band transfer	
	Self-training	Frozen	Fine-tune	Frozen	Fine-tune
1	69.87±1.53	**75.85**±0.03	72.75±0.12	66.20±0.14	69.40±0.69
2	70.34±1.70	**77.60**±0.01	76.12±0.30	70.01±0.22	75.54±0.26
3	75.01±1.20	79.47±0.27	**80.68**±4.27	73.39±0.50	80.51±2.57
4	75.54±1.98	79.79±0.20	**81.76**±0.68	73.36±0.66	79.31±1.21
5	77.07±2.50	80.36±0.35	**81.12**±3.18	75.63±0.45	78.66±4.24

Experiment 2: The Need for Spectral Alignment

Usually, when using hyperspectral classification datasets for model training, the spectral alignment of the training set and the test set is the same regardless of which bands are selected and what alignment is taken by the data. Thus, the positions of the feature spectral bands learned by the model can be effectively used for feature extraction in the test set. However, inaccurate spectral positions in transfer learning of different datasets will greatly reduce the effectiveness of using the transfer model. We split the same dataset into two copies containing different sample categories. One is for the source domain learning and the other is for the target domain. For example, all labeled samples in the first 1–8 classes in the Indian Pine dataset are used as the source domain for network training, and the remaining 9–16 classes are used as the target domain to test the effect of migration learning. One to five samples of each class in the target domain are selected as the training set, and the rest are used as the test set. Two approaches are adopted for training and testing: a) keeping the source domain data unchanged, i.e., the source domain band arrangement is highly consistent with the target domain. b) Randomly disrupting the source domain band order to represent the case of inconsistent source and target domain band arrangement.

The experimental results are presented in Table VII. The experiments show that in the case of insufficient samples in the target domain, the spectrally aligned source domain model A can provide effective shallow features, e.g., the frozen shallow parameters can significantly improve the overall classification accuracy (>6%) in the case of 1-2 training samples per class, while the shallow parameters B with different band ordering have no effect on the performance improvement or even have the opposite effect (negative transfer). Fig. 5 shows the curves of classification accuracy with the number of samples. Freezing the shallow parameters directly provides the best classification effect when there are only 1-2 training samples in the target domain, and as the number of own samples increases, unfreezing the shallow parameters for retraining can further improve the classification performance. In either case, transfer based on spectral alignment can achieve a better migration effect.

Experiment 3: Same/different sensor data transfer learning experiment

This experiment focuses on examining the classification accuracy improvement brought by transfer learning. Given that the SSRN model has stable performance and has been used as a comparison algorithm in several papers [19]–[21], the

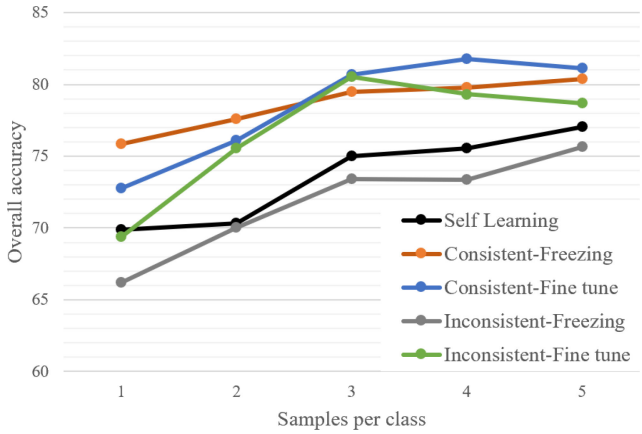

Fig. 5. Classification accuracy curve of Indian Pine data with 1–5 samples per class.

SSRN is selected as the performance evaluation criterion for the GFEN model in this experiment. The SSRN, GFEN, frozen shallow GFEN transfer learning (F-GFEN), and fine-tuned shallow GFEN transfer learning (T-GFEN) classification experiments were performed for each dataset separately.

The shared spectral space of AVIRIS and Hyperion sensors with a total of 421 bands was constructed, and the Indian Pine, Salinas, and Botswana datasets were each spectrally projected into this space. For each dataset, the remaining two datasets were used as the source domain, using 90% of the samples for training, and the rest samples were used for the validation set for the source domain model. For the target domain training, 1, 5, 10, and 15 training samples were randomly selected for each class of data, and the remaining samples were used as tests. Tables VIII–X show the overall classification accuracies of the experiments on the three datasets, with 10 repetitions of each set, and the mean values are presented in the tables.

The experimental results show that transfer learning can obtain higher classification accuracy regardless of the datasets from the same sensor or different sensors. For the Indian Pine dataset, when transfer learning with frozen shallow layer is performed, Salinas data obtained from the same sensor as the source domain can obtain better results, but the fine-tuning after thawing the shallow layer, the Botswana model brings more significant accuracy improvement. Similarly, for the Salinas data, the Indian pine data from the same sensor brings better migration results. As for the Botswana dataset, although the

TABLE VIII
OA OF CLASSIFICATION OF INDIAN PINE DATASET

Samples per class	SSRN	GFEN	From Salinas		From Botswana	
			F-GFEN	T-GFEN	F-GFEN	T-GFEN
1	36.21	47.84	47.99	50.20	47.65	**50.53**
5	70.47	78.18	72.17	75.90	72.25	**77.60**
10	82.22	86.71	78.68	**87.49**	76.20	86.99
15	85.20	88.36	82.62	89.17	80.25	**89.39**

TABLE IX
OA OF CLASSIFICATION OF SALINAS DATASET

Samples per class	SSRN	GFEN	From Indian Pine		From Botswana	
			F-GFEN	T-GFEN	F-GFEN	T-GFEN
1	74.66	76.63	77.71	**79.44**	76.12	78.25
5	89.75	90.05	88.52	**90.97**	87.24	90.42
10	92.28	91.27	92.43	**93.31**	90.67	93.02
15	92.93	91.84	93.55	**94.13**	92.47	92.01

TABLE X
OA OF CLASSIFICATION OF BOTSWANA DATASET

Samples per class	SSRN	GFEN	From Indian Pine		From Salinas	
			F-GFEN	T-GFEN	F-GFEN	T-GFEN
1	87.41	91.42	90.13	93.05	93.21	**93.76**
5	95.75	95.61	96.21	96.35	97.03	**97.08**
10	97.15	96.79	97.91	97.27	**98.15**	97.56
15	97.43	97.07	98.45	97.67	**98.62**	97.78

TABLE XI
OA OF MULTISOURCE DOMAIN MIGRATION CLASSIFICATION FOR EACH OF THE THREE DATASETS

Datasets	Indian Pine		Salinas		Botswana	
Samples per class	F-GFEN	T-GFEN	F-GFEN	T-GFEN	F-GFEN	T-GFEN
1	45.53	50.23	74.40	78.79	89.26	91.32
5	64.76	76.23	85.98	90.68	94.88	96.50
10	68.86	87.64	90.66	93.20	97.65	97.28
15	73.94	89.14	90.44	94.30	98.18	97.42

data used for migration were from different sensors, both had positive effects on the classification performance improvement.

Experiment 4: Multisource domain transfer learning Multisource domain transfer learning refers to the use of multiple datasets to coconstruct source domain networks to train generic classification features for hyperspectral images. As shown in Fig. 3(b), a source domain network is constructed in the form of multitask classification using dual datasets, and for the three datasets used in this article, two datasets are used as the source domain and the remaining one as the target domain for migration experiments. For the two source-domain datasets, 3000 samples are randomly selected for training each; for the target-domain dataset, only 1, 5, 10, and 15 samples are selected for each class, and the remaining samples are used for testing.

Table XI shows the classification results of multisource domain migration learning for the three target domains. In this experiment, the transfer approach of freezing the shallow layer first and fine-tuning, it later outperforms the classification effect of freezing only the shallow layer, and its final classification accuracy is between the accuracy of migration using two source domains alone, but avoids the problem of source domain data selection.

IV. CONCLUSION

In this article, we propose a spectral projection strategy for spectral alignment and feature extraction for heterogenous hyperspectral data of different scenes, and design a new hyperspectral hierarchical feature extraction network GFEN by combining this data preprocessing method. Four groups of experiments are designed to verify 1) the classification capability of the model, 2) the significance and effect of spectral projection, 3) the performance improvement brought by migration learning,

and 4) the significance and results of coconstruction of feature networks in multiple source domains.

It should be noted that the feature extraction and classification network designed in this article focuses on solving the problem of spectral feature migration, and has not yet considered the migration of spatial features, and therefore, does not consider the differences in spatial resolution of different sensors. In the multitask training of multisource domain data, only the cross entropy of 1:1 for different classification tasks is currently used as the loss function. In our future work, we will consider more intelligent processing and analysis methods of hyperspectral images [29] and add more datasets from different sensors, including a large number of UAV hyperspectral data, to establish a more robust hyperspectral classification model by modeling the spectral variability and complex noises [30].

ACKNOWLEDGMENT

The authors would like to thank the School of Surveying and Geospatial Engineering, University of Tehran, Iran, for providing hyperspectral data online.

REFERENCES

[1] D. Hong *et al.*, "Interpretable hyperspectral artificial intelligence: When nonconvex modeling meets hyperspectral remote sensing," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 2, pp. 52–87, Feb. 2021.

[2] M. Ahmad *et al.*, "Hyperspectral image classification—Traditional to deep models: A survey for future prospects," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 968–999, 2022.

[3] J. Yao, D. Hong, L. Xu, D. Meng, J. Chanussot, and Z. Xu, "Sparsity-enhanced convolutional decomposition: A novel tensor-based paradigm for blind hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022.

[4] Y. Zheng, S. Liu, Q. Du, H. Zhao, X. Tong, and M. Dalponte, "A novel multitemporal deep fusion network (MDFN) for short-term multitemporal HR images classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 10691–10704, 2021.

[5] D. Hong *et al.*, "Spectralformer: Rethinking hyperspectral image classification with transformers," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–15, 2022, doi: 10.1109/TGRS.2021.3130716.

[6] S. Liu, D. Marinelli, L. Bruzzone, and F. Bovolo, "A review of change detection in multitemporal hyperspectral images: Current techniques, applications, and challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 140–158, Feb. 2019.

[7] B. Rasti *et al.*, "Feature extraction for hyperspectral imagery: The evolution from shallow to deep: Overview and toolbox," *IEEE Geosci. Remote Sens. Mag.*, vol. 8, no. 4, pp. 60–88, Apr. 2020.

[8] R. Hang, Q. Liu, D. Hong, and P. Ghamisi, "Cascaded recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5384–5394, Aug. 2019.

[9] S. K. Roy *et al.*, "Revisiting deep hyperspectral feature extraction networks via gradient centralized convolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–19, 2022.

[10] A. B. Hamida, A. Benoit, P. Lambert, and C. B. Amar, "3-D deep learning approach for remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4420–4434, Aug. 2018.

[11] Y. Li, H. Zhang, and Q. Shen, "Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network," *Remote Sens.*, vol. 9, no. 1, 2017, Art. no. 67.

[12] Z. Zhong, J. Li, Z. Luo, and M. Chapman, "Spectral–spatial residual network for hyperspectral image classification: A 3-D deep learning framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 847–858, Feb. 2018.

[13] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Jul. 2021.

[14] 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," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 277–281, Feb. 2020.

[15] Y. Luo, J. Zou, C. Yao, X. Zhao, T. Li, and G. Bai, "HSI-CNN: A novel convolution neural network for hyperspectral image," in *Proc. Int. Conf. Audio Lang. Image Process.*, 2018, pp. 464–469.

[16] J. Yue, W. Zhao, S. Mao, and H. Liu, "Spectral–spatial classification of hyperspectral images using deep convolutional neural networks," *Remote Sens. Lett.*, vol. 6, no. 6, pp. 468–477, 2015.

[17] R. Mdrafi, Q. Du, A. C. Gurbuz, B. Tang, L. Ma, and N. H. Younan, "Attention-based domain adaptation using residual network for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 6424–6433, 2020.

[18] Z. Xue *et al.*, "HResNetAM: Hierarchical residual network with attention mechanism for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3566–3580, 2021, doi: 10.1109/JSTARS.2021.3065987.

[19] R. Hang, Z. Li, Q. Liu, P. Ghamisi, and S. S. Bhattacharyya, "Hyperspectral image classification with attention aided CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 2281–2293, Mar. 2021.

[20] H. Zhang, Y. Li, Y. Jiang, P. Wang, Q. Shen, and C. Shen, "Hyperspectral classification based on lightweight 3-D-CNN with transfer learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5813–5828, Aug. 2019.

[21] X. He, Y. Chen, and P. Ghamisi, "Heterogeneous transfer learning for hyperspectral image classification based on convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3246–3263, May 2020.

[22] J. Lin, L. Zhao, S. Li, R. Ward, and Z. J. Wang, "Active-learning-incorporated deep transfer learning for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 11, pp. 4048–4062, Nov. 2018.

[23] C. Deng, Y. Xue, X. Liu, C. Li, and D. Tao, "Active transfer learning network: A unified deep joint spectral–spatial feature learning model for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1741–1754, Mar. 2019.

[24] Y. Qu *et al.*, "Physically constrained transfer learning through shared abundance space for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10455–10472, Dec. 2021.

[25] D. Hong *et al.*, "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4340–4354, May 2021.

[26] R. Hang, F. Zhou, Q. Liu, and P. Ghamisi, "Classification of hyperspectral images via multitask generative adversarial networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1424–1436, Feb. 2021.

[27] R. Hang *et al.*, "Classification of hyperspectral and LiDAR data using coupled CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4939–4950, Jul. 2020.

[28] S. K. Roy, A. Deria, D. Hong, B. Rasti, A. Plaza, and J. Chanussot, "Multimodal fusion transformer for remote sensing image classification," 2022, *arXiv:2203.16952*.

[29] B. Zhang *et al.*, "Progress and challenges in intelligent remote sensing satellite systems," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 1814–1822, 2022.

[30] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.

Bin Yang received the M.S. degree in optical engineering from the HuaZhong University of Science and Technology, Wuhan, China, in 2005, and the Ph.D. degree in resource and environment remote sensing from the School of Geosciences and Info-Physics, Central South University, Changsha, China, in 2015.

He was a Postdoctoral Fellow with the National University of Defense Technology, China, and he is currently with the School of Geography Science, Hunan Normal University, Changsha, China. His research interests include hyperspectral image processing, target detection and classification, deep learning, and high-performance computing.

Shunshi Hu received the Ph.D. degree in cartography and geography information system from the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China, in 2013.

He is currently a Full Faculty with the Key Laboratory of Geospatial Big Data Mining and Application, Hunan Normal University, Changsha, China. He has authored and coauthored some papers and conducted some scientific research projects on the topics of his research interests, which include remote sensing mechanism and applications, remote sensing digital image processing, and hyperspectral remote sensing applications.

Qiandong Guo received the B.E. degree in remote sensing science and technology from Wuhan University, Wuhan, China, in 2011, the M.S. degree in cartography and geographic information system from the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China, in 2014, and the Ph.D. degree in geography from the University of South Florida, Tampa, FL, USA, in 2018.

His research interests include hyperspectral image processing, target detection, and land use and land cover changes.

Danfeng Hong (Senior Member, IEEE) received the M.Sc. degree (*summa cum laude*) in computer vision from the College of Information Engineering, Qingdao University, Qingdao, China, in 2015, and the Dr.-Ing degree (*summa cum laude*) from the Signal Processing in Earth Observation, Technical University of Munich, Munich, Germany, in 2019.

He is currently a Professor with the Key Laboratory of Computational Optical Imaging Technology, Aerospace Information Research Institute, Chinese Academy of Sciences (CAS). Before joining CAS, he has been a Research Scientist and led a Spectral Vision Working Group at the Remote Sensing Technology Institute, German Aerospace Center, Oberpfaffenhofen, Germany. He was also an Adjunct Scientist at GIPSA-lab, Grenoble INP, CNRS, Université Grenoble Alpes, Grenoble, France. His research interests include signal/image processing, hyperspectral remote sensing, machine/deep learning, artificial intelligence, and their applications in Earth Vision.

Dr. Hong is a Topical Associate Editor for IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS), an Editorial Board Member of Remote Sensing, and an Editorial Advisory Board Member of *ISPRS Journal of Photogrammetry and Remote Sensing*. He was a recipient of the Best Reviewer Award of the IEEE TGRS in 2021 and 2022, and the Best Reviewer Award of the IEEE JSTARS in 2022, the Jose Bioucas Dias Award for recognizing the outstanding paper at WHISPERS in 2021, the Remote Sensing Young Investigator Award in 2022, and the IEEE GRSS Early Career Award in 2022.