

Semisupervised deep learning using consistency regularization and pseudolabels for hyperspectral image classification

Xiang Hu^{ORCID},^a Tong Zhou,^{a,b} and Yuanxi Peng^{a,*}

^aNational University of Defense Technology, College of Computer,
State Key Laboratory of High-Performance Computing, ChangSha, China

^bNational University of Defense Technology, Beijing Institute for Advanced Study,
ChangSha, China

Abstract. Hyperspectral image (HSI) classification is a focus area in remote sensing research, wherein redundant spectral information poses a significant challenge and deep-learning-based classifiers have achieved better performance than traditional methods have. Training a deep-learning-based classifier requires numerous labeled samples. However, collecting such a substantial amount of labeled hyperspectral data is difficult. Semisupervised classification of HSIs has thus received increasing attention, where semisupervised learning classifiers function based on labeled and unlabeled data. A new training method for semisupervised HSI classification is proposed. Specifically, consistency regularization and pseudolabeling are combined as a semisupervised training framework, without the introduction of a complex mechanism. Our proposed algorithm can work without the need to change the conventional convolutional neural network model architecture. Unlike previous deep-learning-based methods, our approach does not require data reconstruction to obtain unsupervised loss. This means that our model can be much less computationally intensive. From the results of experiments on three public hyperspectral datasets, our proposed method outperforms several state-of-the-art methods. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JRS.16.026513](https://doi.org/10.1117/1.JRS.16.026513)]

Keywords: semisupervised learning; consistency regularization; pseudolabels; hyperspectral image classification.

Paper 210455 received Jul. 28, 2021; accepted for publication May 10, 2022; published online May 31, 2022.

1 Introduction

Hyperspectral image (HSI) classification is a subject garnering very high research attention in hyperspectral remote sensing. HSIs can provide rich spectral and spatial information, which improves their utility in various applications. However, the abundant spectral information also leads to low classification accuracy, which is termed the Hughes phenomenon. On the contrary, limited labeled hyperspectral samples also lead to difficulty in HSI classification. In the real world, increasing quantities of hyperspectral data are becoming available with the development of information acquisition technology. However, most of these data are unlabeled, and labeling the data is extremely laborious and time-consuming. Nevertheless, semisupervised learning (SSL) methods can achieve good performance via usage of a large volume of unlabeled data and a small amount of labeled data. Thus, SSL methods are increasingly being proposed to address this challenge.

In the early stages of semisupervised HSI classification, several traditional methods have been applied. For instance, the transductive support-vector machine (SVM) was proposed for semisupervised HSI classification.¹ Gomez-Chova et al.² used the Laplacian SVM (LapSVM) by introducing the graph Laplacian matrix to process unlabeled data. Yang et al.³ designed the spatio-spectral Laplacian SVM by considering both spatial and spectral information.

*Address all correspondence to Yuanxi Peng, pyx@nudt.edu.cn

Recently, deep learning has shown great potential for use in computer vision^{4,5} and HSI classification. In this regard, Zhang et al.⁶ proposed a spectral–spatial residual network consisting of spectral and spatial residual blocks to extract spectral and spatial features. By replacing the residual network with the densely connected structure and using different convolution kernel sizes, Wang et al.⁷ proposed a fast dense spectral–spatial convolution network.

Although previous studies on deep learning for supervised HSI classification with a small number of training samples have achieved noteworthy success, these studies could not use unlabeled data to enhance the classification performance. Thus, designing a semisupervised method based on deep learning is of great value, and ladder networks have been introduced to deep learning for semisupervised HSI classification.⁸ Fang et al.⁹ employed the cotraining method with a model based on ResNets.¹⁰ However, such methods require the addition of new parts to the model or designing of multiple models, which increases the computational effort.

Most recently, a new semisupervised method called Fixmatch¹¹ has emerged as the state-of-the-art semisupervised framework in computer vision. This method works using existing convolutional neural network (CNN) models, and changes to the model structure are unnecessary. It comprises two components: consistency regularization and pseudolabeling. The consistency regularization component depends on data augmentation. However, many classical data augmentation methods in the computer vision field cannot be applied to HSIs directly. For example, color distortion, a classical image augmentation method, is not suitable for HSIs because it changes the spectral information. Consequently, HSI data augmentation has not been sufficiently researched. We therefore attempted to apply data augmentation for HSI semisupervised classification and designed a simple SSL framework for HSI classification.

In this paper, inspired by Fixmatch,¹¹ we propose a semisupervised method for HSI classification. Semisupervised methods for HSI classification can work with both labeled and unlabeled data, which is of high practical significance. Because labeling HSI data is time-consuming and labor-intensive. The contributions of this study are as follows:

1. We propose a method for semisupervised HSI classification based on consistency regularization and pseudolabeling. The pseudolabels are produced based on the original image which work as targets when the augmented images are fed into the model. In our method, both the original image and the augmented image are used when performing consistency regularization.
2. Our proposed algorithm couples the SSL with conventional CNN without the need to change the model structure.
3. We explore the use of data augmentation on HSI semisupervised classification, where the consistency regularization relies on data augmentation. We adjust data augmentation methods to make the consistency regularization available for HSI classification. We modify the classical Cutout¹² method to prevent it from removing the spectral information of the center pixel and combine it with horizontal or vertical flip as our augmentation method. Our experiments prove that the proposed method is superior to the state-of-the-art approach.

The remainder of this paper is organized as follows: Sec. 2 describes our proposed method in detail. The experimental results for the three datasets are shown in Sec. 3. Finally, Sec. 4 provides the conclusions.

2 Proposed Method

Our proposed method comprises two crucial elements, consistency regularization, and pseudolabeling, which will be introduced separately. Next, a summary of our proposed models is presented.

2.1 Consistency Regularization

Consistency regularization¹³ is a vital component of many state-of-the-art SSL algorithms. The working fundamentals of consistency regularization are based on the assumption that the model

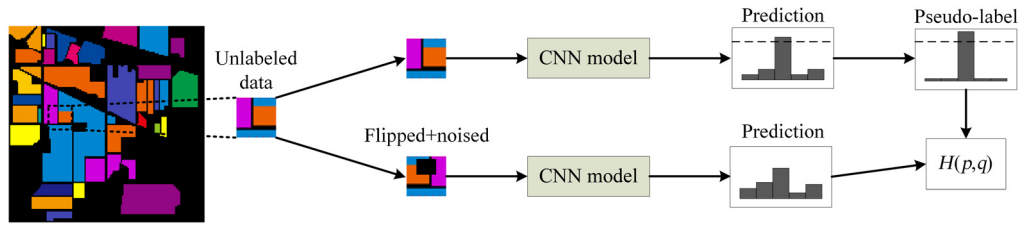


Fig. 1 The overall architecture of our proposed framework.

should yield consistent results when inputs are different visions of the same data. The model is trained with a supervised loss and an unsupervised loss. The unsupervised loss is formulated as

$$\|p_{\text{model}}(y|\text{Augment}(x); \theta) - p_{\text{model}}(y|\text{Augment}(x); \theta)\|_2^2. \quad (1)$$

The $\text{Augment}(x)$ is a stochastic transformation. Thus, the two terms in Eq. (1) are not equal.

2.2 Pseudolabeling

Pseudolabeling¹⁴ is also a typical semisupervised method. It relies on the model to give labels (the argmax of the model's output) for unlabeled data. Only the labels whose largest probability exceeds a predefined threshold are retained. The pseudolabeling process employs the following loss function:

$$\frac{1}{\mu B} \sum_{b=1}^{\mu B} 1(\max(q_b) \geq \tau) H(\hat{q}_b, q_b), \quad (2)$$

where $q_b = p_m(y|u_b)$, $\hat{q}_b = \arg \max(q_b)$, and τ is the threshold.

2.3 Summary of Our Proposed Model

The complete architecture is shown in Fig. 1. In our model, the loss function consists of two elements, supervised loss ℓ_s and unsupervised loss ℓ_u . ℓ_s refers to the standard cross-entropy loss, and ℓ_u is formulated as

$$\ell_u = \frac{1}{\mu B} \sum_{b=1}^{\mu B} 1(\max(q_b) \geq \tau) H(\hat{q}_b, p_m(y|\mathcal{A}(u_b))). \quad (3)$$

where $q_b = p_m(y|u_b)$ and $\hat{q}_b = \arg \max(q_b)$. u_b represents the unlabeled data, and $\mathcal{A}(u_b)$ represents the augmented unlabeled data. The loss function of our method is defined as

$$\ell = \ell_s + \lambda_u \ell_u, \quad (4)$$

where λ_u is the weight parameter.

All the above equations were gained from Ref. 11. The augmentation method is illustrated in Algorithm 1. The size of the submatrix in Algorithm 1 is not fixed.

3 Experimental Results

3.1 Experimental Datasets and Parameter Setup

We conducted experiments on three popular hyperspectral datasets: Indian Pines, Pavia University, and Salinas, as shown in Fig. 2.

The Indian Pines image was acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) in 1992 from northwestern Indiana. It contains 145×145 pixels with 224 spectral

Algorithm 1 Selecting random rectangular area to remove spectral information

```

1: Input: input image  $I$ ; image size  $w \times h \times c$ .
2: Output: augmented image  $I^*$ .
3: Randomly flip the input image  $I$  horizontally or vertically
4: Generate a matrix of the size  $(w \times h)$  using 1
5: Select a random submatrix in this matrix and change the elements inside to 0
6: if the center point of the matrix is in the submatrix then
7:     change the element of that point to 1
8: end if
9: for  $i = 1$  to  $c$  do
10:     multiply the image in the  $i$ 'th channel by this matrix to obtain the augmented image  $I^*$ 
11: end for
12: Return the augmented image  $I^*$ 

```



Fig. 2 (a) False-color Indian Pines image. (b) False-color Salinas image. (c) False-color University of Pavia image.

bands in the wavelength range of 0.4 to 2.5 μm . Considering the water absorption effect, the researchers removed 24 bands, and the 10,249 labeled pixels were grouped into 16 classes.

The University of Pavia scene was acquired by the Reflective Optics System Imaging Spectrometer in 2002 over Pavia, northern Italy. The image contains 610×340 pixels with 103 spectral bands. The spectral wavelength ranges from 0.43 to 0.86 μm . The 42,776 labeled pixels were classified into nine categories.

The Salinas dataset was acquired by the AVIRIS sensor over Salinas Valley, California. The data cube size is $512 \times 217 \times 224$ with the wavelength range from 0.4 to 2.5 μm . Twenty water absorption bands were discarded, and 200 bands were available for analysis. The 54,129 labeled pixels were divided into 16 categories.

All the experiments were performed using the RTX Titan GPU. The learning rate was 0.0001, and the minibatches had a size of 256. The model was trained using the Adam optimizer for 150 epochs. The unsupervised loss weight parameter was $\min(0.01 \times \text{epoch}, 1)$, and the threshold was 0.98. According to Fixmatch,¹¹ we think a high threshold value is necessary to ensure a high level of accuracy. We chose 0.9, 0.94, and 0.98 as the range of the threshold value. We chose hybridSN¹⁵ as the backbone network. The input formats were also the same as in Ref. 15, $25 \times 25 \times 30$ for the Indian Pines, and $25 \times 25 \times 15$ for the Pavia University and Salinas images. Principal component analysis was used to reduce the data dimensions. We compare our proposed

model with five methods, including four traditional methods and one deep-learning-based method. The traditional methods are mentioned in Ref. 16: transductive SVM,¹⁷ local and global consistency graph-based algorithm (LGC),¹⁸ label propagation algorithm (LPA),¹⁹ and label propagation algorithm combined with particle cooperation and competition (LPAPCC).¹⁶ The deep-learning-based method is SS-CNN.⁸ This model requires data reconstruction to obtain unsupervised loss. We trained this model for the same number of epochs as our model. The training and testing samples are the same as in Ref. 16, as shown in Tables 1–3. The percentages in parentheses are the corresponding ratios.

3.2 Analysis of Classification Results

The classification results over three HSI datasets are listed in Tables 4–6. The results of the traditional methods as shown in these tables are obtained from Ref. 16. The best values are marked in bold. We conducted the experiments five times consecutively. The numbers after the plus-minus signs are the standard deviations of the corresponding metrics. In this paper, the overall classification accuracy (OA), average classification accuracy (AA), and kappa coefficient (Kappa) are adopted as model performance evaluation metrics. The OA is the ratio of the correctly classified samples to all the test samples. The AA denotes the average classification accuracy in each category. Finally, Kappa is the measure of agreement between the ground truth and classification and is interpreted as follows: <0.2, poor agreement; 0.2 to 0.4, fair agreement; 0.4 to 0.6, moderate agreement; 0.6 to 0.8, good agreement; and 0.8 to 1, excellent agreement.

As shown in Tables 4–6, our proposed method achieved the best performance over three experimental datasets. The AA of our model is much better than that of the other contrast methods. This means that our method is applicable to unbalanced data. Considering that our model did not achieve the highest classification accuracy on some kinds of samples, we think this

Table 1 Training and testing samples for Indian Pines.

No.	Name	Labeled	Unlabeled	Test	Total
1	Alfalfa	5 (10.87%)	8 (17.39%)	33 (71.74%)	46 (100%)
2	Corn-notill	42 (2.94%)	386 (27.03%)	1000 (70.03%)	1428 (100%)
3	Corn-mintill	24 (2.89%)	225 (27.11%)	581 (70.00%)	830 (100%)
4	Corn	7 (2.95%)	64 (27.00%)	166 (70.04%)	237 (100%)
5	Grass-pasture	14 (2.90%)	130 (26.92%)	339 (70.19%)	483 (100%)
6	Grass-trees	21 (2.88%)	198 (27.12%)	511 (70.00%)	730 (100%)
7	Grass-pasture-mowed	5 (17.86%)	3 (10.71%)	20 (71.43%)	28 (100%)
8	Hay-windrowed	14 (2.93%)	129 (26.99%)	335 (70.08%)	478 (100%)
9	Oats	5 (25.00%)	1 (5.00%)	14 (70.00%)	20 (100%)
10	Soybean-notill	29 (2.98%)	262 (26.95%)	681 (70.06%)	972 (100%)
11	Soybean-mintill	73 (2.97%)	663 (27.01%)	1719 (70.02%)	2455 (100%)
12	Soybean-clean	17 (2.87%)	160 (26.98%)	416 (70.15%)	593 (100%)
13	Wheat	6 (2.93%)	55 (26.83%)	144 (70.24%)	205 (100%)
14	Woods	37 (2.92%)	342 (27.04%)	886 (70.04%)	1265 (100%)
15	Buildings-grass-trees-drives	11 (2.85%)	104 (26.94%)	271 (70.21%)	386 (100%)
16	Stone-steel-towers	5 (5.38%)	22 (23.66%)	66 (70.97%)	93 (100%)
	Total	315 (3.07%)	2752 (26.85%)	7182 (70.08%)	10,249 (100%)

Table 2 Training and testing samples for Pavia University.

No.	Name	Labeled	Unlabeled	Test	Total
1	Asphalt	99 (1.49%)	1890 (28.50%)	4642 (70.00%)	6631 (100%)
2	Meadows	279 (1.50%)	5315 (28.50%)	13,055 (70.00%)	18,649 (100%)
3	Gravel	31 (1.48%)	598 (28.49%)	1470 (70.03%)	2099 (100%)
4	Trees	45 (1.47%)	874 (28.52%)	2145 (70.01%)	3064 (100%)
5	Painted metal sheets	20 (1.49%)	383 (28.48%)	942 (70.04%)	1345 (100%)
6	Bare soil	75 (1.49%)	1433 (28.49%)	3521 (70.01%)	5029 (100%)
7	Bitumen	19 (1.43%)	380 (28.57%)	931 (70.00%)	1330 (100%)
8	Self-blocking bricks	55 (1.49%)	1049 (28.49%)	2578 (70.02%)	3682 (100%)
9	Shadows	14 (1.48%)	270 (28.51%)	663 (70.01%)	947 (100%)
	Total	637 (1.49%)	12,192 (28.50%)	29,947 (70.01%)	42,776 (100%)

Table 3 Training and testing samples for Salinas.

No.	Name	Labeled	Unlabeled	Test	Total
1	Brocoli-green-weeds-1	12 (0.60%)	590 (29.37%)	1407 (70.03%)	2009 (100%)
2	Brocoli-green-weeds-2	22 (0.59%)	1095 (29.39%)	2609 (70.02%)	3726 (100%)
3	Fallow	11 (0.56%)	581 (29.40%)	1384 (70.04%)	1976 (100%)
4	Fallow-rough-plow	8 (0.57%)	410 (29.41%)	976 (70.01%)	1394 (100%)
5	Fallow-smooth	16 (0.60%)	787 (29.39%)	1875 (70.01%)	2678 (100%)
6	Stubble	23 (0.58%)	1164 (29.40%)	2772 (70.02%)	3959 (100%)
7	Celery	21 (0.59%)	1052 (29.39%)	2506 (70.02%)	3579 (100%)
8	Grapes-untrained	67 (0.59%)	3314 (29.40%)	7890 (70.00%)	11,271 (100%)
9	Soil-vineyard-develop	37 (0.60%)	1823 (29.39%)	4343 (70.01%)	6203 (100%)
10	Corn-senesced-green-weeds	19 (0.58%)	964 (29.41%)	2295 (70.01%)	3278 (100%)
11	Lettuce-romaine-4wk	6 (0.56%)	314 (29.40%)	748 (70.04%)	1068 (100%)
12	Lettuce-romaine-5wk	11 (0.57%)	567 (29.42%)	1349 (70.01%)	1927 (100%)
13	Lettuce-romaine-6wk	5 (0.55%)	269 (29.37%)	642 (70.09%)	916 (100%)
14	Lettuce-romaine-7wk	6 (0.56%)	315 (29.44%)	749 (70.00%)	1070 (100%)
15	Vineyard-untrained	43 (0.59%)	2137 (29.40%)	5088 (70.01%)	7268 (100%)
16	Vineyard-vertical-trellis	10 (0.55%)	532 (29.44%)	1265 (70.01%)	1807 (100%)
	Total	317 (0.59%)	15,914 (29.40%)	37,898 (70.01%)	54,129 (100%)

additionally reflects the high stability of our method. According to the Kappa results, our proposed method performs well on the three datasets. This method achieved the highest classification accuracy on six of the nine classes in the Pavia University dataset. These results further confirm the superiority of our proposed method.

From the above results shown in Tables 4–6, we can reach several conclusions. First, our proposed method shows the most accurate and stable performance. Second, from a comparison

Table 4 Classification results on Indian Pines.

Class names	TSVM	LGC	LPA	LPAPCC	SS-CNN	Proposed
Alfalfa	0.7251	0.4371	0.0478	0.0547	0.94 ± 0.06	0.97 ± 0.03
Corn-notill	0.8698	0.514	0.9421	0.9587	0.68 ± 0.04	0.86 ± 0.03
Corn-mintill	0.7214	0.5599	0.9326	0.9228	0.71 ± 0.01	0.89 ± 0.02
Corn	0.8806	0.4205	0.931	0.9574	0.95 ± 0.03	0.87 ± 0.05
Grass-pasture	0.9737	0.8059	0.9955	0.9851	0.93 ± 0.05	0.90 ± 0.06
Grass-trees	0.944	0.7727	0.9758	0.987	0.87 ± 0.01	0.96 ± 0.02
Grass-pasture-mowed	0.3597	0.4289	0.6609	0.5325	0.61 ± 0.17	0.89 ± 0.05
Hay-windrowed	0.9983	0.9413	1	1	0.94 ± 0.01	0.98 ± 0.02
Oats	0.5356	0.1723	0.818	0.8862	0.60 ± 0.09	0.68 ± 0.11
Soybean-notill	0.8255	0.5348	0.9337	0.936	0.70 ± 0.06	0.90 ± 0.01
Soybean-mintill	0.8613	0.5228	0.9599	0.9634	0.70 ± 0.03	0.95 ± 0.00
Soybean-clean	0.6773	0.4165	0.9214	0.9344	0.56 ± 0.08	0.90 ± 0.10
Wheat	0.9806	0.805	0.9965	0.997	0.95 ± 0.02	0.97 ± 0.03
Woods	0.9867	0.8388	0.9977	0.9936	0.89 ± 0.02	0.96 ± 0.04
Buildings-grass-trees-drives	0.9345	0.4765	0.9511	0.9522	0.81 ± 0.07	0.89 ± 0.11
Stone-steel-towers	0.9757	0.9906	0.9871	0.978	0.97 ± 0.02	0.88 ± 0.05
OA	0.8648	0.6188	0.8768	0.8886	0.76 ± 0.01	0.92 ± 0.02
AA	0.8281	0.6023	0.8682	0.8776	0.80 ± 0.01	0.90 ± 0.01
Kappa	0.8458	0.5534	0.8607	0.8741	0.72 ± 0.01	0.91 ± 0.02

Table 5 Classification results on Pavia University.

Class names	TSVM	LGC	LPA	LPAPCC	SS-CNN	Proposed
Asphalt	0.7126	0.5135	0.5332	0.5335	0.91 ± 0.02	0.97 ± 0.01
Meadows	0.9123	0.9198	0.9917	0.9915	0.96 ± 0.01	0.99 ± 0.00
Gravel	0.6837	0.6265	0.8954	0.9297	0.80 ± 0.06	0.94 ± 0.06
Trees	0.8198	0.4756	0.9695	0.9777	0.99 ± 0.00	0.99 ± 0.00
Painted metal sheets	0.9572	0.8932	0.996	0.9977	1 ± 0	0.9851 ± 0.01
Bare soil	0.8869	0.9445	0.9949	0.996	0.95 ± 0.02	0.99 ± 0.00
Bitumen	0.708	0.6068	0.9251	0.9059	0.92 ± 0.03	0.95 ± 0.05
Self-blocking bricks	0.6626	0.5431	0.8738	0.8851	0.80 ± 0.03	0.91 ± 0.03
Shadows	0.4501	0.2614	0.8894	0.9443	0.99 ± 0.00	0.96 ± 0.01
OA	0.8227	0.7295	0.8488	0.8499	0.93 ± 0.00	0.98 ± 0.00
AA	0.7548	0.6427	0.8966	0.9068	0.92 ± 0.00	0.97 ± 0.01
Kappa	0.7626	0.6502	0.8005	0.802	0.91 ± 0.01	0.97 ± 0.00

Table 6 Classification results on Salinas.

Class names	TSVM	LGC	LPA	LPAPCC	SS-CNN	Proposed
Brocoli-green-weeds-1	0.9378	0.1983	0.2134	0.2174	0.87 ± 0.23	0.98 ± 0.02
Brocoli-green-weeds-2	0.975	1	1	1	0.96 ± 0.02	0.99 ± 0.00
Fallow	0.9747	0.9877	0.9971	0.9979	0.92 ± 0.04	0.99 ± 0.00
Fallow-rough-plow	0.8457	0.9349	0.9922	0.99	0.97 ± 0.02	0.98 ± 0.01
Fallow-smooth	0.9696	0.9384	0.9989	0.9978	0.93 ± 0.05	0.99 ± 0.00
Stubble	0.9947	1	1	1	1 ± 0	0.99 ± 0.00
Celery	0.9649	0.9783	0.9933	0.9956	0.98 ± 0.01	0.98 ± 0.03
Grapes-untrained	0.925	0.9944	0.9994	0.9996	0.72 ± 0.07	0.97 ± 0.03
Soil-vineyard-develop	0.994	0.6112	0.5211	0.5459	0.98 ± 0.01	0.99 ± 0.00
Corn-senesced-green-weeds	0.9827	0.9916	0.9974	0.9948	0.91 ± 0.02	0.99 ± 0.00
Lettuce-romaine-4wk	0.9859	1	1	1	0.95 ± 0.01	0.94 ± 0.07
Lettuce-romaine-5wk	0.9731	0.9761	0.9937	0.9972	0.95 ± 0.02	0.99 ± 0.00
Lettuce-romaine-6wk	0.8525	0.9619	0.997	0.9927	0.98 ± 0.01	0.90 ± 0.10
Lettuce-romaine-7wk	0.9383	0.9082	0.9684	0.9815	0.99 ± 0.00	0.97 ± 0.02
Vineyard-untrained	0.8584	0.9978	0.9992	0.9991	0.75 ± 0.15	0.98 ± 0.00
Vineyard-vertical-trellis	0.9796	0.9606	1	1	0.99 ± 0.00	0.99 ± 0.00
OA	0.938	0.7646	0.7555	0.7681	0.85 ± 0.03	0.98 ± 0.01
AA	0.947	0.9025	0.9169	0.9193	0.93 ± 0.01	0.98 ± 0.01
Kappa	0.931	0.7418	0.7269	0.7433	0.84 ± 0.04	0.98 ± 0.01

with the TSVM results, the deep-learning-based framework proves superior to the traditional machine learning-based algorithm. Third, the Indian Pines dataset is much more difficult to classify than the Salinas scene. Moreover, the LGC and LPA results indicate that the single mechanism does not perform well in HSI semisupervised classification. Therefore, methods based on multiple mechanisms need to be investigated further. Besides, our proposed method gets better performance than SS-CNN. It indicates that deep-learning-based methods can achieve good results in HSI semisupervised classification without changing the model architecture. Finally, the LPAPCC result suggests a significant advantage in using the graph mechanism. In future work, we will attempt to introduce the graph mechanism to design a semisupervised framework.

3.3 Influence of Data Removal

Data removal is an important part of our augmentation method. Here, to evaluate its effects, we compare the performance of our proposed method and our method without data removal. The results are shown in Table 7. As shown in Table 7, our proposed method surpasses the performance of our method with image flip only.

Although the enhancement is negligible on the Indian Pines dataset and University of Pavia dataset, we think it is because that the proportion of unlabeled data in the training data is small. This can be verified by a larger enhancement in Salinas dataset whose unlabeled data occupy a relatively larger proportion. However, in the real world, the unlabeled data are much more than labeled data. We suppose that data removal can have a greater impact on accuracy with a large proportion of unlabeled data.

Table 7 Classification results of our method without data removal and our method with data removal.

Dataset	Metric	Without data removal	With data removal
Indian Pines	OA	0.91 ± 0.00	0.92 ± 0.02
	AA	0.91 ± 0.02	0.90 ± 0.01
	Kappa	0.90 ± 0.01	0.91 ± 0.02
University of Pavia	OA	0.97 ± 0.00	0.98 ± 0.00
	AA	0.96 ± 0.00	0.97 ± 0.01
	Kappa	0.97 ± 0.00	0.97 ± 0.00
Salinas	OA	0.96 ± 0.03	0.98 ± 0.01
	AA	0.96 ± 0.02	0.98 ± 0.01
	Kappa	0.95 ± 0.03	0.98 ± 0.01

3.4 Influence of Parameter

The threshold and the weight of unlabeled loss in our proposed method are important. To find the best parameters for our proposed method, we conducted several experiments. We compared the performance of different parameters. The results are presented in Tables 8 and 9. From the results about different thresholds, we can see our proposed method did not achieve the best accuracy over Indian Pines. However, the differences are very small. Moreover, our proposed method is superior to other methods over University of Pavia and Salinas. We can get that the threshold of 0.98 is strongly robust. From the results about different weights, we can easily conclude that our proposed method achieved the highest accuracy.

Besides, from the results, we can find that the sensitivity of our model to the parameters is different on different datasets. We suppose the reason is that the proportions of unlabeled data in the training data are different on different datasets. Because the proportion of unlabeled data in the training data is small on Indian Pines, the model is not sensitive to these parameters. The smaller the proportion of unlabeled data in the training data is, the less sensitive the model is to these parameters.

In summary, it has been shown from these results that the model accuracy is not significantly affected by the weight and threshold. Our proposed method is very robust and therefore practical in real-world applications.

Table 8 Classification results of our method with different thresholds.

Dataset	Metric	0.9	0.94	0.98
Indian Pines	OA	0.92 ± 0.01	0.91 ± 0.02	0.92 ± 0.02
	AA	0.89 ± 0.01	0.90 ± 0.03	0.90 ± 0.01
	Kappa	0.91 ± 0.02	0.90 ± 0.03	0.91 ± 0.02
University of Pavia	OA	0.97 ± 0.01	0.97 ± 0.00	0.98 ± 0.00
	AA	0.96 ± 0.01	0.96 ± 0.01	0.97 ± 0.01
	Kappa	0.96 ± 0.01	0.97 ± 0.00	0.97 ± 0.00
Salinas	OA	0.96 ± 0.01	0.96 ± 0.01	0.98 ± 0.01
	AA	0.95 ± 0.03	0.95 ± 0.03	0.98 ± 0.01
	Kappa	0.96 ± 0.01	0.95 ± 0.02	0.98 ± 0.01

Table 9 Classification results of our method with different weights.

Dataset	Metric	1	$\min(0.01 \times \text{epoch}, 1)$
Indian Pines	OA	0.92 ± 0.02	0.92 ± 0.02
	AA	0.89 ± 0.02	0.90 ± 0.01
	Kappa	0.91 ± 0.02	0.91 ± 0.02
University of Pavia	OA	0.95 ± 0.02	0.98 ± 0.00
	AA	0.94 ± 0.01	0.97 ± 0.01
	Kappa	0.94 ± 0.03	0.97 ± 0.00
Salinas	OA	0.97 ± 0.00	0.98 ± 0.01
	AA	0.97 ± 0.01	0.98 ± 0.01
	Kappa	0.97 ± 0.00	0.98 ± 0.01

4 Conclusions

A deep-learning-based framework for HSI semisupervised classification is proposed in this paper. The framework is composed of two components: consistency regularization and pseudo-labeling. It is used on existing CNN models to enable them to utilize unlabeled data. Unlike the case in previous research, a simpler SSL framework is proposed herein, and no complex mechanism is introduced. The proposed method's loss consists of supervised and unsupervised losses. The unsupervised loss is essentially the cross-entropy loss based on the augmented version of unlabeled data. The label is based on the output of the corresponding unlabeled data. Our proposed algorithm can work without the need to change the conventional CNN model architecture. Besides, unlike previous deep-learning-based methods, our approach does not require data reconstruction to obtain unsupervised loss. Therefore, our model can be much less computationally intensive. Our experiments demonstrate that the proposed method can achieve high classification accuracy and outperform state-of-the-art methods, while being robust to model parameters.

In future work, we will propose a better data augmentation method for HSI classification, and the generative adversarial network²⁰ is the focus of our subsequent study. An attempt will be made to introduce additional methods, such as the graph mechanism, to HSI semisupervised classification because the original pseudolabeling does not solve the problem of wrong labels. Moreover, an attempt is underway to design a new pseudolabeling method.

Acknowledgments

The authors declare no conflict of interest.

References

1. L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for semisupervised classification of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.* **44**(11), 3363–3373 (2006).
2. L. Gómez-Chova et al., "Semisupervised image classification with laplacian support vector machines," *IEEE Geosci. Remote Sens. Lett.* **5**(3), 336–340 (2008).
3. L. Yang et al., "Semi-supervised hyperspectral image classification using spatio-spectral laplacian support vector machine," *IEEE Geosci. Remote Sens. Lett.* **11**(3), 651–655 (2013).
4. Y. Liu et al., "Hierarchical learning with backtracking algorithm based on the visual confusion label tree for large-scale image classification," *Vis. Comput.* **38**, 897–917 (2021).

5. Y. Liu et al., "Visual tree convolutional neural network in image classification," in *24th Int. Conf. Pattern Recognit. (ICPR)*, IEEE, pp. 758–763 (2018).
6. Z. Zhong et al., "Spectral–spatial residual network for hyperspectral image classification: a 3-D deep learning framework," *IEEE Trans. Geosci. Remote Sens.* **56**(2), 847–858 (2017).
7. W. Wang et al., "A fast dense spectral–spatial convolution network framework for hyperspectral images classification," *Remote Sens.* **10**(7), 1068 (2018).
8. B. Liu et al., "A semi-supervised convolutional neural network for hyperspectral image classification," *Remote Sens. Lett.* **8**(9), 839–848 (2017).
9. B. Fang et al., "Semi-supervised deep learning classification for hyperspectral image based on dual-strategy sample selection," *Remote Sens.* **10**(4), 574 (2018).
10. K. He et al., "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 770–778 (2016).
11. K. Sohn et al., "Fixmatch: simplifying semi-supervised learning with consistency and confidence," *Adv. Neural Inf. Process. Syst.* **33**, 596–608 (2020).
12. T. DeVries and G. W. Taylor, "Improved regularization of convolutional neural networks with cutout," arXiv:1708.04552 (2017).
13. P. Bachman, O. Alsharif, and D. Precup, "Learning with pseudo-ensembles," in *27th Int. Conf. Adv. Neural Inf. Process. Syst.*, pp. 3365–3373 (2014).
14. D.-H. Lee et al., "Pseudo-label: the simple and efficient semi-supervised learning method for deep neural networks," in *Workshop Challenges in Represent. Learn., ICML*, Vol. 3, No. 2, p. 896 (2013).
15. S. K. Roy et al., "Hybridsn: exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.* **17**(2), 277–281 (2019).
16. Z. He et al., "A constrained graph-based semi-supervised algorithm combined with particle cooperation and competition for hyperspectral image classification," *Remote Sens.* **13**(2), 193 (2021).
17. T. Joachims et al., "Transductive inference for text classification using support vector machines," in *ICML*, Vol. 99, pp. 200–209 (1999).
18. D. Zhou et al., "Learning with local and global consistency," in *Adv. Neural Inf. Process. Syst.*, pp. 321–328 (2004).
19. Z. Xiaojin and G. Zoubin, "Learning from labeled and unlabeled data with label propagation," Tech. Rep., Technical Report CMU-CALD-02–107, Carnegie Mellon University (2002).
20. I. Goodfellow et al., "Generative adversarial networks," *Commun. ACM* **63**(11), 139–144 (2020).

Xiang Hu received his BS degree in computer science from the National University of Defense and Technology (NUDT), China, in 2019. He is now working toward a master's degree at NUDT. His research interests are in the areas of pattern recognition, image processing, and remote sensing.

Tong Zhou received his PhD from NUDT, China. He is currently an assistant researcher at NUDT. His main research is focused on deep learning and hyperspectral remote sensing.

Yuanxi Peng received his BS degree in computer science from Sichuan University, China, in 1988, and his MS and PhD degrees in computer science from NUDT, China, in 1998 and 2001, respectively. He was a visiting professor in the Department of Electronic and Computer Engineering, University of Toronto, Canada, during 2010 to 2011. He has been a professor in the Computer School at NUDT since 2011. His research interests are in the areas of high performance computing, deep learning, and remote sensing.