Spectral–Spatial Features Exploitation Using Lightweight HResNeXt Model for Hyperspectral Image Classification

Dhirendra Prasad Yadav, Deepak Kumar, Anand Singh Jalal, Ankit Kumar, Surbhi Bhatia Khan, Thippa Reddy Gadekallu, Arwa Mashat & Areej A. Malibari

Published online: 04 Sep 2023.

Submit your article to this journal ⬈

Article views: 107

View related articles ⬈

View Crossmark data ⬈

Taylor & Francis
Taylor & Francis Group

Spectral–Spatial Features Exploitation Using Lightweight HResNeXt Model for Hyperspectral Image Classification

Exploitation des caractéristiques spectrales-spatiales à l'aide d'un modèle HResNeXt léger pour la classification d'images hyperspectrales

Dhirendra Prasad Yadav[a,b], Deepak Kumar[b], Anand Singh Jalal[a], Ankit Kumar[a], Surbhi Bhatia Khan[c,d], Thippa Reddy Gadekallu[c,e,f,g,h], Arwa Mashat[i], and Areej A. Malibari[j]

[a]Department of Computer Engineering & Applications, G.L.A. University, Mathura, Uttar Pradesh, India; [b]Department of Computer Engineering, NIT Meghalaya, Shillong, Meghalaya, India; [c]Department of Electrical and Computer Engineering, Lebanese American University, Byblos, Lebanon; [d]Department of Data Science, School of Science, Engineering and Environment, University of Salford, Salford, UK; [e]Zhongda Group, Haiyan County, Jiaxing City, Zhejiang Province 314312, China; [f]School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India; [g]College of Information Science and Engineering, Jiaxing University, Jiaxing 314001, China; [h]Division of Research and Development, Lovely Professional University, Phagwara, India; [i]Department of Information System, Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh, Saudi Arabia; [j]Department of Industrial and Systems Engineering, College of Engineering, Princess Nourah Bint Abdulrahman University, PO Box 84428, Riyadh 11671, Saudi Arabia

ABSTRACT

Hyperspectral image classification is vital for various remote sensing applications; however, it remains challenging due to the complex and high-dimensional nature of hyperspectral data. This paper introduces a novel approach to address this challenge by leveraging spectral and spatial features through a lightweight HResNeXt model. The proposed model is designed to overcome the limitations of traditional methods by combining residual connections and cardinality to enable efficient and effective feature extraction from hyperspectral images, capturing both spectral and spatial information simultaneously. Furthermore, the paper includes an in-depth analysis of the learned spectral–spatial features, providing valuable insights into the discriminative power of the proposed approach. The extracted features exhibit strong discriminative capabilities, enabling accurate classification even in challenging scenarios with limited training samples and complex spectral variations. Extensive experimental evaluations are conducted on four benchmark hyperspectral data sets, the Pavia university (PU), Kennedy Space Center (KSC), Salinas scene (SA), and Indian Pines (IP). The performance of the proposed method is compared with the state-of-the-art methods. The quantitative and visual results demonstrate the proposed approach's high classification accuracy, noise robustness, and computational efficiency superiority. The HResNeXt obtained an overall accuracy on PU, KSC, SA, and IP, 99.46%, 81.46%, 99.75%, and 98.64%, respectively. Notably, the lightweight HResNeXt model achieves competitive results while requiring fewer computational resources, making it well-suited for real-time applications.

RÉSUMÉ

La classification d'images hyperspectrales est vitale pour diverses applications de télédétection. Cependant, cela reste difficile en raison de la nature complexe et de la haute dimensionnalité des données hyperspectrales. Cet article présente une nouvelle approche pour relever ce défi en tirant parti des caractéristiques spectrales et spatiales grâce à un modèle HResNeXt léger. Le modèle proposé est conçu pour surmonter les limites des méthodes traditionnelles en combinant les connexions résiduelles et la cardinalité pour permettre une extraction efficace des caractéristiques des images hyperspectrales, capturant simultanément les informations spectrales et spatiales. En outre, l'article comprend une analyse approfondie des caractéristiques spectrales et spatiales apprises, fournissant des informations précieuses sur le pouvoir discriminatif de l'approche proposée. Les caractéristiques extraites présentent de fortes capacités discriminantes, permettant une classification précise même dans des scénarios difficiles avec de petits échantillons d'entraînement et des variations spectrales complexes. Les évaluations expérimentales ont été menées sur quatre

CONTACT Ankit Kumar ✉ iiita.ankit@gmail.com

ensembles de données hyperspectrales de référence: PU, KSC, SA et IP. La performance de la méthode proposée est comparée aux méthodes de pointe. Les résultats quantitatifs et visuels démontrent une haute précision des classifications pour l'appropre proposée, sa robustesse au bruit et sa supériorité dans l'efficacité de calcul. Le HResNeXt a obtenu un OA sur PU, KSC, SA et IP, de 99,46%, 81,46%, 99,75% et 98,64%, respectivement. Notamment, le modèle HResNeXt léger permet d'obtenir des résultats compétitifs tout en nécessitant moins de ressources de calcul, ce qui le rend bien adapté aux applications en temps réel.

Introduction

Hyperspectral pictures contain hundreds of continuous spectral bands that can be utilized to distinguish between various substances. As a result, hyperspectral pictures are now widely recognized as a crucial data source in remote sensing for object recognition and classification. Numerous classification strategies, notably supervised models, have been developed for labeling hyperspectral information. Supervised classification methods have been used for many classification tasks using random forest (Sun et al. 2019; Joelsson et al. 2005; Gadekallu et al. 2023) and support vector machine (SVM; Ravi et al. 2022; Saab et al. 2022). A random forest is an algorithm that averages out a set of data. The final classes of test samples are chosen either by a majority vote or the maximum posterior (MAP) rule, and a collection of decision trees is generated from a set of randomly selected subsamples of the training data. In contrast, an SVM seeks a hyperplane to prioritize differences across classes. However, "shallow" models like the random forest and SVM (Melgani and Bruzzone 2004; Waske et al. 2010) are considered inferior to "deep" networks that can obtain hierarchical, deep representations of features (Guo et al. 2022; Mou et al. 2020).

Many supervised approaches for HSI categorization (Audebert et al. 2019) have been proposed over the past 20 years. When HSI classification first began, spectral data were less available. Standard spectral classification using the SVM was reported in He and Chen (2021). In addition, several SVM-based classifiers (Deng et al. 2018; Ghamisi et al. 2017) have been proposed to manage the land cover classification of HSI due to SVM (Chen et al. 2022). SVM-based methods have poor sensitivity to huge dimensionality. The spatial features of HSI have been extracted using a variety of morphological operations, including morphological profiles (MPs; Chen et al. 2022), extended morphological profiles (EMPs; Benediktsson et al. 2005), extended multi-attribute profiles (EMAPs; Dalla Mura et al. 2011), and extinction profiles (EPs; Fang et al. 2018).

The use of deep learning algorithms in processing remote sensing images, particularly in HSI categorization (Wang et al. 2023; Xu et al. 2022; Ji et al. 2023), has the potential to revolutionize the sector radically. Depending on the different features utilized in the classification process, it is feasible to categorize deep learning-based HSI classification strategies into three primary categories. These three types of networks have been utilized based on geographical information, spectral properties, and hybrid networks. Both spectral–spatial feature-based networks and spatial feature-based networks have received more attention in recent years (Zhuo et al. 2022; Fu et al. 2023), which is possible because the HSI categorization is dependent not only on geographical information but also on spectral information.

The R-VCANet (Pan et al. 2017), Bayesian 2D convolutional neural networks (CNNs; Cao et al. 2018), and the squeeze multi-bias network (SMBN; Fang et al. 2019) are a few examples that have been used for land cover classification using spatial features. On the other hand, an HSI has an excessive number of channels, which frequently results in two-dimensional convolution kernels that are overly deep. The number of parameters has also increased significantly. Consequently, HSI classification methods could be based on three-dimensional CNNs. A deep contextual CNN (Bashir et al. 2023), as defined by (Bashir et al. 2023), employs several three-dimensional local convolutional filters of various sizes. A deep contextual CNN enables simultaneous utilization of an HSI's spatial and spectral components. To enhance the extraction of the essential spectral–spatial aspects of HSIs, Chen et al. (2016) created a three-dimensional CNN-based feature extraction model with regularization.

Ben Hamida et al. (2018) created a new three-dimensional deep learning technique to process spectral and spatial data concurrently using less computer power (i.e., floating point operations, FLOPs). Even though the overall number of parameters for a 3D CNN may be less, it still needs more processing resources than a 2D CNN. This is due to the depths it must penetrate and the absence of a bird's-eye perspective of the spectral data. The most current approach developed by Roy et al. (2020) utilized 3D and 2D CNN layers to develop a deep learning model called HybridSN. The HybridSN model improved the

classification accuracy through joint exploitation of the spectral and spatial features.

However, deep learning models are difficult to train despite their impressive HSI classification performance because of the challenges of obtaining tagged pixels and the high cost of labeling them. In addition, the pixels among these labeled data are not distributed equitably. When dealing with irregularly dispersed data and a limited number of samples, it is incredibly challenging to construct reliable deep-learning models (Hang et al. 2019; Mou and Zhu 2020) that have excellent performance and need fewer processing resources. A novel lightweight 3D convolutional neural network is an asymmetric inception network (AINet ;Fang et al. 2022). This concentrated on spectral characteristics rather than geographical settings and utilized a data fusion transfer learning approach to speed up training and improve the initialization of the model. However, its performance could be more optimal when used on smaller samples. The network was created using the double-branch dual-attention (DBDA) method described in Li et al. (2020), which simultaneously utilized spectral and spatial properties and achieved high accuracy across a broad range of HSI data sets using channel and spatial attention mechanisms that enhance the feature maps.

While state-of-the-art performance in CNN models for HSI classification has developed, certain limitations remain. For example, while using a CNN-based technique, some aspects of the input HSI are ignored and must be thoroughly explored. The CNN technique is vector-based; therefore, it reads the inputs as a set of pixel vectors (Linzen et al. 2016). HSI's data structure in the spectral domain is fundamentally built on a sequence. Because of this, CNN might cause data loss while processing hyperspectral pixel vectors (Vallathan et al. 2021). Second, the long-range sequential reliance needed to switch between band locations can be challenging to model. As the size of the kernel and the number of layers restrict the receptive field of CNNs, they could be better at gathering long-range relationships of input data (Peng et al. 2022). The convolutional processes focus on a small area around the input point. Because HIS (Vaswani et al. 2017) often consists of hundreds of spectral bands, understanding its long-range correlations is challenging.

The Transformer (Glorot and Bengio 2010; Jiang and Chen 2022; Hong et al. 2022) paradigm was recently proposed for use in natural language processing. The concept of self-attention serves as the foundation of this approach. By paying close attention, the Transformer (El-Assal et al. 2022; Xie et al. 2017; Yadav et al. 2022) may infer a worldwide dependence among a set of inputs. While training,

deep learning models (Saab et al. 2022; Arikumar et al. 2022) like Transformers frequently experience the vanishing-gradient problem, which hinders or even prohibits convergence. Even while these backbone networks and their modifications (Garg et al. 2022; Grupo de Inteligencia Computacional (GIC) 2023) have shown promise in classification accuracy, they still need to adequately characterize spectral series information (Sharma and Biswas 2018; Zhao et al. 2022) – particularly regarding collecting minor spectral disparities along spectral dimensions. Several recent methods are discussed in Table 1.

1. To reduce computational resources, spectral features are obtained through a single layer of three-dimensional convolution.
2. We have utilized a modified ResNeXt network with fewer trainable parameters to improve the classification accuracy and reduce computation resources.
3. Four distinct data sets and six different state-of-the-art methodologies are used to assess the model's performance.

Proposed method

We assume the spectral–spatial hyperspectral data cubes, where I^m is the input, W, H, D are the spectral bands' total number, H is their width, and D is their combined height. A single HSI pixel in I^m has D spectral band and forms a one-hot label vector $L = (L_1, L_2, ...L_c) \in R^{1 \times 1 \times c}$, where c stands for types of land cover. However, the mixed land cover classes shown in the hyperspectral pixels give a lot of variety within each class and numerous similarities across classes. The proposed model architecture is shown in Figure 1.

The classification of the class model would need to be reliable and effective for it to work. First, to eliminate the spectral redundancy, the principal component analysis (PCA) is used along the spectral bands of the initial HSI data (Im). The number of spectral bands is reduced from D to A via the PCA, but the spatial dimensions remain the same (W, width and H, height). We minimized the number of spectral bands while preserving the essential spatial information. Following PCA, the reduced data cube will be S number of bands after principal component analysis. The data cube is broken up into smaller, overlapping patches, and the pixel in the center of the cube is used to determine the truth label. The 3D data are then sent to a 3D CNN. Within this network, convolution is carried out with a 3D kernel (El-Assal et al. 2022), which records spectral characteristics from contiguous bands.

Table 1. Summary of the recent methods used for HSI classification.

Method	Model	Data Set	Overall Accuracy (OA)
Ding et al. [52]	Graph attention neural network	Pavia University,	71.76%
		Salinas	82.61%
		Houston	63.82%
Yao et al. [53]	Hybrid multi-graph neural network (DHMG)	Pavia University,	97.81%
		Salinas	98.33%
		Houston	93.31%
Zhang et al. [54]	Single-source Domain Expansion Network (SDEnet)	Pavia center,	81.94%
		Houston	79.96%
		GID Dataset	77.95%
Liu et al. [55]	Deep contrastive learning network (DCLN)	Pavia University,	89.70%
		Salinas	94.88%
		Houston	82.19%
		Indiana Pines	83.74%
Zhaou et al. [56]	Multiscale convolutional fusion network (AMGCFN)	Pavia University,	97.31%
		Salinas	97.59
		Indiana Pines	93.94%
Wang et al. [57]	Transfer fusion network (TFNet)	Pavia University,	98.64%
		Salinas	98.56%
		Houston	90.50%
Dang et al. [58]	Double-Branch Feature Fusion Transformer	Kennedy Space Center Salinas	98.64%
		University of Pavia Houston	98.50%
			98.76%
			90.08%
Wang et al. [59]	Hybrid network model (HNM)	Indian Pines	96.59%
		Houston	97.50%
		WHU-Hi-HanChuan	97.58%
Xie et al. [60]	Consistency-based prototype network (FCPN)	Pavia University,	95.15%
		Salinas	94.61%
		Indiana Pines	84.71%
Xu et al. [61]	Multiscale and cross-level attention learning (MCAL) network	Indian Pines	94.925
		Pavia Center	99.30%
		HyRANK-Loukia	82.16%
		WHU-Hi-HanChuan	95.30%

The main contribution of the methods is as follows.
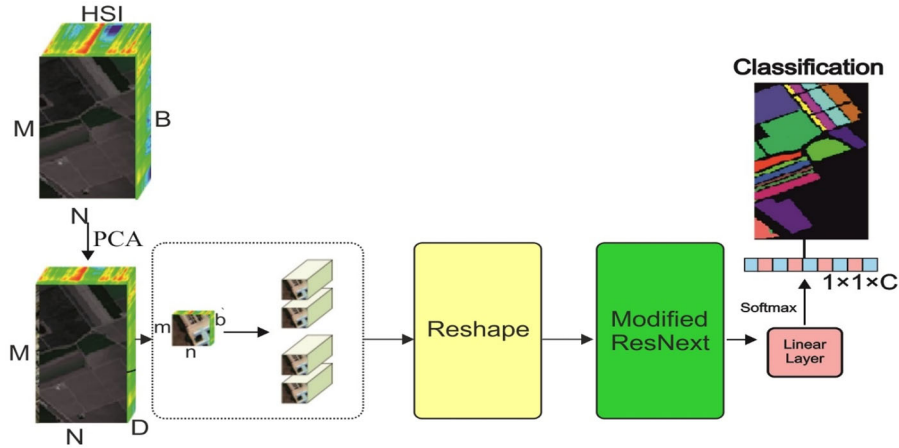
Figure 1. The model architecture of the proposed method.

For the ith feature map of the jth layer's spatial position's activation value (x, y, z) is calculated as follows.

$$u_{j,i}^{x,y,z} = \Phi\Big(b_{j,i} + \sum_{\tau=1}^{d_{l-1}} \sum_{\lambda=-\eta}^{\eta} \sum_{\alpha=-\gamma}^{\gamma} \sum_{\beta=-\delta}^{\delta} w_{j,i,\tau}^{\beta,\alpha,\lambda} \times u_{j-1}^{x+\beta,y+\alpha+z+\lambda}\Big)$$

(1)

where Φ = activation function, $b_{j,i}$ = bias parameter, d_{l-1} = feature map of the (l–1)th layer, $2\eta + 1, 2\gamma +$

$1, 2\delta + 1$ is the kernel's width, height, depth and $w_{j,i}$ = weight parameter of the ith feature map of the jth layer.

Modified ResNeXt

The features obtained from the 3D CNN layers are first reshaped and fed into the modified ResNeXt for spatial feature extraction. The original ResNeXt included 23×10^6 parameters, resulting in substantial calculation costs (Xie et al. 2017). We reduced the

number of trainable parameters down to 9×10^6 by lowering the convolution size of the first layer filter from 7×7 to 5×5 because a larger filter size dampened the intensity of the edge pixels. Because of this, there was an increase in the number of false negative values. The sizes of conv2, conv3, conv4, and conv5 are all decreased similarly, as shown in Table 2. The modified ResNeXt model architecture is shown in Figure 2.

The modified ResNeXt contains grouped convolution, ReLU activation and residual blocks. The pooling layer maps the features of the CNN block obtained from the cluster of adjacent neurons. Pixels are separated, and surrounding pooling units rarely overlap, reducing overfitting (Saab et al. 2022; Yadav et al. 2022). Further, convolution is performed by the inner dot product of neurons in the convolution layer to generate aggregate transform as follows.

$$\sum_{i=1}^{N} w_i x_i \qquad (2)$$

where, $x = (x_1, x_2, \ldots x_N) = $ input vector to the N channel and $w_i = $ filter weight of the ith neuron. To reduce the dimension depth, elementary transform $w_i x_i$ is replaced by a more generic function called aggregated transformation, as shown in Eq. (3).

Table 2. Parameters comparison of the original ResNeXt and Modified ResNeXt.

$$F(x) = \sum_{i=1}^{C} T_i(x) \qquad (3)$$

where $T_i(x)$ can be an arbitrary function. Analogous to a simple neuron, T_i should project x into an (optionally low-dimensional) embedding and then transform it.

Local response normalization (LRN)

The normalization of the input does not affect the saturation of the deep learning algorithm using ReLU. ReLU enables the neurons in the model to learn with fewer positive training examples (Arikumar et al. 2022). Despite this, we determine the action of neurons at a given location (x, y) by employing a kernel (k) to facilitate generalization. Afterward, ReLU nonlinearity is applied in the HResNeXt. The LRN of the neurons with N layers is determined as follows.

$$LRN_{x,y}^{k} = \alpha_{x,y}^{k} / \left(t + \alpha \sum_{i-\max(0,k,n/2)}^{\min(N,1,k+n/2)} (\alpha_{u,v}^{i})^2 \right)^{\beta} \qquad (4)$$

where t, α, β are hyperparameter constants, and n is the adjacent kernel feature map. The efficiency of deep CNN models is dependent on their architecture. The major component of the deep CNN model is hyperparameters. Through the proper use of hyperparameters, CNN classification accuracy can be improved. The division by zero is avoided by setting $t = 2$. The consecutive pixels of input layers that undergo normalization are defined by n. The normalization process α is set to 10e-4, and the contrasting constant β is set to 0.65. The feature map generated

Figure 2. Modified ResNeXt block diagram for HSI classification.

by the flatten layer is passed to the Softmax layer. Finally, the Softmax layer converts probabilities into their corresponding class.

Experimental result and discussion

Data set

The proposed method is evaluated on four standard open access data sets Salinas scene (SA), Indian Pines (IP), Pavia University (PU), and Kennedy Space Center (KSC) (Grupo de Inteligencia Computacional (GIC) 2023). The 224-band AVIRIS sensor in orbit took this picture of the Salinas Valley in California, and it has an impressive level of detail (3.7-m pixels). In the area that was studied, there are a total of 512 lines and 217–112, 154–167, and 224). Radiance data were the only way to access this image at the sensor level. It is made up of both wilderness and agricultural terrain, namely grapes. The underlying reality of Salinas may be broken down into 16 different levels. The Indian Pines test site in northwest Indiana was the location of the AVIRIS sensor images that were taken and included in the IP collection. It has a 145×145 pixel resolution and can pick up wavelengths between 0.4 and 2.5×10^{-3} m. Cropland makes up two-thirds of IP, while forests and other types of perennial natural vegetation comprise the other third of the territory.

Along with the low-density residences, other buildings, and narrower roadways, the region has a rail line, two large dual-lane highways, two big dual-lane motorways, and two significant dual-lane highways. As the shot was taken in June, many crops are still in their early phases of development, with coverage of less than 5%. This is because of the date of the photo acquisition. There are 16 different kinds of ground truth, all related to one another. The number of spectral bands utilized in this investigation was cut down to 200 after bands in 220-nm ranges were removed from consideration. The ROSIS sensor gathered the information for the PU when it flew over Pavia in northern Italy. It has a resolution of 610×610 pixels and contains 103 spectral bands; however, none of the samples in either picture is meaningful. The resolution of the device is 610×610 pixels. It is possible to obtain a geometric resolution of 1.3 m. The data from the sample were divided into 99 different groups according to the underlying facts.

On March 23, 1996, the Florida-based Kennedy Space Center (KSC) was photographed by the NASA-operated AVIRIS satellite. The 224 bands from which AVIRIS may collect data have a 10-nm bandwidth and a center wavelength ranging from 400 nm to 2500 nm. The data from KSC have an 18-m spatial resolution. After considering water absorption and low SNR bands, the study utilized 176 bands. The comprehensive land cover maps were created using color infrared images captured by the Kennedy Space Center. Because many species in this area have similar spectral signatures, it is challenging to identify the vegetation in this area. The diverse land applications in this place have been organized into thirteen distinct categories. A detailed description of the data set is shown in Tables 3–6.

Experimental setup

In the proposed study, the experiment is conducted using Python 3.8 on NVIDIA Quadro RTX4000 GPU, having 128 GB RAM and a dual graphics card of 8 GB. For each data set, the initial learning rate was set to 0.0001 and trained for 100 epochs using the Adam optimizer with a mini-batch size of 64.

Quantitative result analysis

SVM, 1D CNN, 2D CNN, 3D CNN, HybridSN, and Spectral Former (SF) are all machine learning and deep learning-based approaches that are compared to gauge the method's performance [42]. All parameters are kept at their literature-referenced values for consistency's sake. Because there are few instances in several classes of the IP data set, the experiment is conducted by splitting the PU, KSC, and SA data sets

Table 3. IP data set description with a land cover color map (Sharma and Biswas 2018; Zhao et al. 2022).

Class ID	Land cover	Train	Test	Color Code
1	Alfalfa_C1	5	41	C1
2	Com-notill_C2	143	1285	C2
3	Com-mintill_C3	83	747	C3
4	Com_C4	23	214	C4
5	Grass-pasture_C5	48	435	C5
6	Grass-trees_C6	73	657	C6
7	Grass-pasture-mowed_C7	3	25	C7
8	Hay-windrowed_C8	48	430	C8
9	Oats_C9	2	18	C9
10	Soybean-notill_C10	97	875	C10
11	Soybean-mintill_C11	245	2210	C11
12	Soybean-clean_C12	59	534	C12
13	Wheat_C13	20	185	C13
14	Woods_C14	126	1139	C14
15	Buildings-Grass-Trees-Drives_C15	39	347	C15
16	Stone-Steel-Towers_C16	9	84	C16

*Color Code Details

Table 4. PU data set description with land cover color map.

Class ID	Land cove	Train	Test	Color Code
1	Asphalt_C01	332	6299	C01
2	Meadows_C02	932	17117	C02
3	Gravel_C03	105	1994	C03
4	Trees_C04	153	2911	C04
5	Sheets_C05	67	1278	C05
6	Bare Soil_C06	251	4778	C06
7	Bitumen_C07	66	1264	C07
8	Bricks_C08	184	3498	C08
9	Shadows_C09	47	900	C09

*Color Code Details for table 4.

Table 5. KSC data set description with land cover color map.

Class ID	Land cover	Train	Test	Color Code
1	Scrub_C01	76	685	C01
2	Willow swamp_C02	24	219	C02
3	CP Hammock_C03	25	231	C03
4	Slash Pine_C04	25	227	C04
5	OakBroadleaf_C05	16	145	C05
6	Hardwood_C06	22	207	C06
7	Swamp_C07	10	95	C07
8	Graminoid marsh_C08	43	401	C08
9	Spartinamarsh_C09	52	468	C09
10	Cattailmarsh_C010	40	364	C10
11	Saltmarsh_C011	41	378	C11
12	Mud flat_C012	50	453	C12
13	Water_C013	92	835	C13

*Color Code Details for table 5.

Table 6. SA data set description with land cover color map.

Class ID	Land cover	Train	Test	Color Code
1	Brocoligreenweeds_A_C01	100	1909	C01
2	Brocoli_green_weeds_B_C02	161	3565	C02
3	Fallow_C03	99	1877	C03
4	Fallowroughplow_C04	70	1324	C04
5	Fallowsmooth_C05	134	2544	C05
6	Stubble__C06	198	3761	C06
7	Celery_C07	179	3400	C07
8	Grapesuntrained_C08	564	10707	C08
9	Soilvinyarddevelop_C09	310	5893	C09
10	Comsenescedgreenweeds_C010	164	3114	C10
11	Lettuceromaine4wk_C011	53	1015	C11
12	Lettuceromaine5wk_C012	96	1831	C12
13	Lettuceromaine6wk_C013	46	870	C13
14	Lettuceromaine7wk_C014	54	1016	C14
15	Vinyarduntrained_C015	363	6905	C15
16	Vinyardverticaltrellis_C016	90	1717	C16

*Color Code Details for table 6.

into 5% for training and 95% for validation. We split the data set in half, using the first 10% for training and the second 90% for validation. For HSI

Table 7. Performance evaluation on the PU data set.

ID	SVM	1D CNN	2D CNN	3D CNN	HybridSN	SF	HResNeXt
1	91.26	96.47	97.24	98.32	**99.64**	97.72	98.78
2	96.67	97.72	98.32	**99.27**	97.38	98.52	99.21
3	64.25	87.36	87.92	91.47	86.75	**99.53**	98.62
4	93.65	94.83	95.28	96.28	92.71	98.27	**99.37**
5	97.29	97.42	**99.64**	98.62	99.12	97.34	99.42
6	87.23	97.32	99.42	97.54	97.98	99.16	**99.68**
7	88.62	92.54	90.23	87.65	88.37	97.87	**98.72**
8	82.34	94.84	93.45	92.27	91.28	**99.27**	98.85
9	87.54	85.68	94.86	96.32	94.45	96.15	**98.56**
AA	87.65	93.91	95.15	95.30	94.19	98.20	**99.02**
OA	88.62	94.73	97.68	97.42	96.52	98.85	**99.46**
Kappa	84.78	92.74	94.65	95.78	93.26	97.18	**98.62**

Table 8. Performance evaluation on the KSC datas et.

ID	SVM	1D CNN	2D CNN	3D CNN	HybridSN	SF	HResNeXt
1	74.34	68.24	78.46	72.28	83.16	**86.52**	85.27
2	51.26	64.13	58.89	63.13	71.24	76.63	**85.64**
3	45.26	47.38	51.44	64.72	68.81	78.81	**80.12**
4	52.15	56.27	46.37	49.64	56.83	75.27	**78.12**
5	62.24	42.44	64.84	68.86	**94.28**	87.78	94.05
6	46.62	43.45	49.88	42.43	53.76	68.46	**72.78**
7	32.18	36.94	38.26	45.34	46.82	68.35	**72.14**
8	44.38	46.37	54.78	**75.22**	73.48	65.28	74.64
9	56.67	65.32	48.56	57.78	56.87	68.92	**78.16**
10	47.24	43.85	45.28	48.54	45.86	48.84	**57.43**
11	49.24	48.76	55.63	67.26	69.72	78.65	**81.72**
12	53.23	62.75	65.87	63.72	70.13	81.52	**84.36**
13	70.82	66.82	67.82	68.27	73.52	**86.42**	85.20
AA	52.74	53.42	55.85	60.55	66.50	74.73	**79.21**
OA	53.24	54.84	56.83	61.87	67.68	76.28	**81.46**
Kappa	52.28	51.84	54.68	59.28	65.82	73.23	**78.87**

classification, we have used the libsvm toolbox3's support for SVMs by adjusting the RBF's two parameters. The 2D CNN consists of a softmax layer and three 2D convolutional blocks. The convolutional blocks of 2D CNNs use the same 1D conventional layer, BN layer, max-pooling layer, and ReLU activation function as their 1D counterparts. Separate spatial and spectral features extractors of size 3332, 3364, and 11,128 are included in each 2D convolutional layer. The 3D CNN has two convolutional layers that use 3D max-pooling and batch normalization to provide optimal results. HybridSN combines three 3D convolutional layers with one 2D convolutional layer. The spectral former used a cross-layer skip connection to extract features in both patch- and pixel-based ways. Local and global attention methods also improve the HSI's classifying precision. We have summarized the classification performance of each method and proposed method from Tables 7–10.

The SVM classification performance is lower in several classes due to a lack of high-dimensional features, whereas 1D CNN improves the classification accuracy through one-directional convolution. Further, 2D CNN calculates spatial features via

convolution in both directions. 3D CNN computation cost is high but capable of extracting high-dimensional spectral features. To make use of spectral and spatial characteristics and enhance classification performance, HybridSN used 3D and 2D CNN layers. The global and local attention of the feature is provided using the SF network through a transformer, which improves the accuracy. However, this requires high computation costs and large volumes of data.

The proposed HResNeXt utilized spectral and spatial features via one 3D CNN layer and modified the 2D convolutional block of the ResNeXt network to improve the classification. In addition, the computation cost is less due to fewer parameters.

### Performance evaluation on different patch sizes

Path size plays an essential role in the computation model for HSI data. We can see in Figure 3 that the

Table 9. Performance evaluation on the SA data set.

ID	SVM	1D CNN	2D CNN	3D CNN	HybridSN	SF	HResNeXt
1	95.82	96.16	97.52	98.32	99.24	**100**	**100**
2	94.84	95.25	96.14	97.35	98.92	99.98	**100**
3	97.62	98.47	98.54	97.24	**100**	99.06	99.28
4	98.21	97.24	98.27	98.52	97.82	98.92	**99.18**
5	96.28	98.52	99.86	**100**	98.74	97.84	98.28
6	95.32	96.38	**98.78**	98.12	98.23	99.16	**100**
7	94.25	95.78	97.28	97.32	98.25	98.63	**99.62**
8	84.26	86.35	93.76	91.82	96.12	97.16	**98.94**
9	96.34	97.35	98.74	97.24	98.52	98.82	**99.35**
10	91.26	92.45	94.26	91.72	94.74	**99.88**	98.92
11	92.54	94.34	95.34	94.84	97.42	97.37	**98.48**
12	93.26	94.26	97.67	98.15	98.85	99.64	**100**
13	90.15	78.17	92.42	88.18	96.29	97.87	**98.65**
14	97.24	96.22	98.84	99.05	97.28	98.82	**99.96**
15	85.42	87.72	92.66	95.64	98.76	98.25	**99.84**
16	85.28	88.76	**97.76**	92.82	93.65	**100**	98.54
AA	93.01	93.34	96.74	96.04	97.68	98.83	**99.32**
OA	95.18	96.46	97.85	97.56	98.25	98.87	**99.75**
Kappa	92.65	93.28	95.14	95.87	96.84	97.78	**98.86**

Table 10. Performance evaluation on the IP data set.

ID	SVM	1D CNN	2D CNN	3D CNN	HybridSN	SF	HResNeXt
1	48.64	54.25	72.36	84.32	89.25	96.72	**98.24**
2	74.16	76.46	82.42	85.48	92.53	95.42	**97.13**
3	69.75	78.32	84.25	**96.37**	94.46	95.84	95.18
4	64.42	74.86	78.12	98.56	96.34	98.38	**98.85**
5	78.36	81.24	75.78	95.62	94.54	97.56	**98.38**
6	81.24	87.12	86.92	**98.46**	97.73	97.94	96.28
7	75.36	78.38	88.64	97.82	97.64	98.52	**99.34**
8	85.72	89.53	92.36	95.25	99.28	99.38	**100**
9	46.38	55.34	64.52	88.28	86.62	88.92	**91.53**
10	77.53	82.28	82.38	87.58	93.32	98.65	**99.48**
11	84.18	87.68	92.32	96.74	**97.82**	96.16	95.19
12	75.32	72.27	**98.12**	97.36	96.32	96.25	96.87
13	81.96	84.54	87.54	96.32	98.92	97.98	**99.57**
14	90.28	93.26	96.87	98.66	99.25	**98.84**	96.45
15	46.25	65.78	78.42	92.26	94.87	95.28	**97.32**
16	75.86	86.38	91.87	**97.82**	97.13	97.48	96.10
AA	72.21	77.98	84.56	94.18	95.38	96.83	**97.24**
OA	74.45	79.36	92.14	95.74	96.14	96.78	**98.64**
Kapa	71.82	75.84	83.43	92.87	94.82	95.72	**96.86**
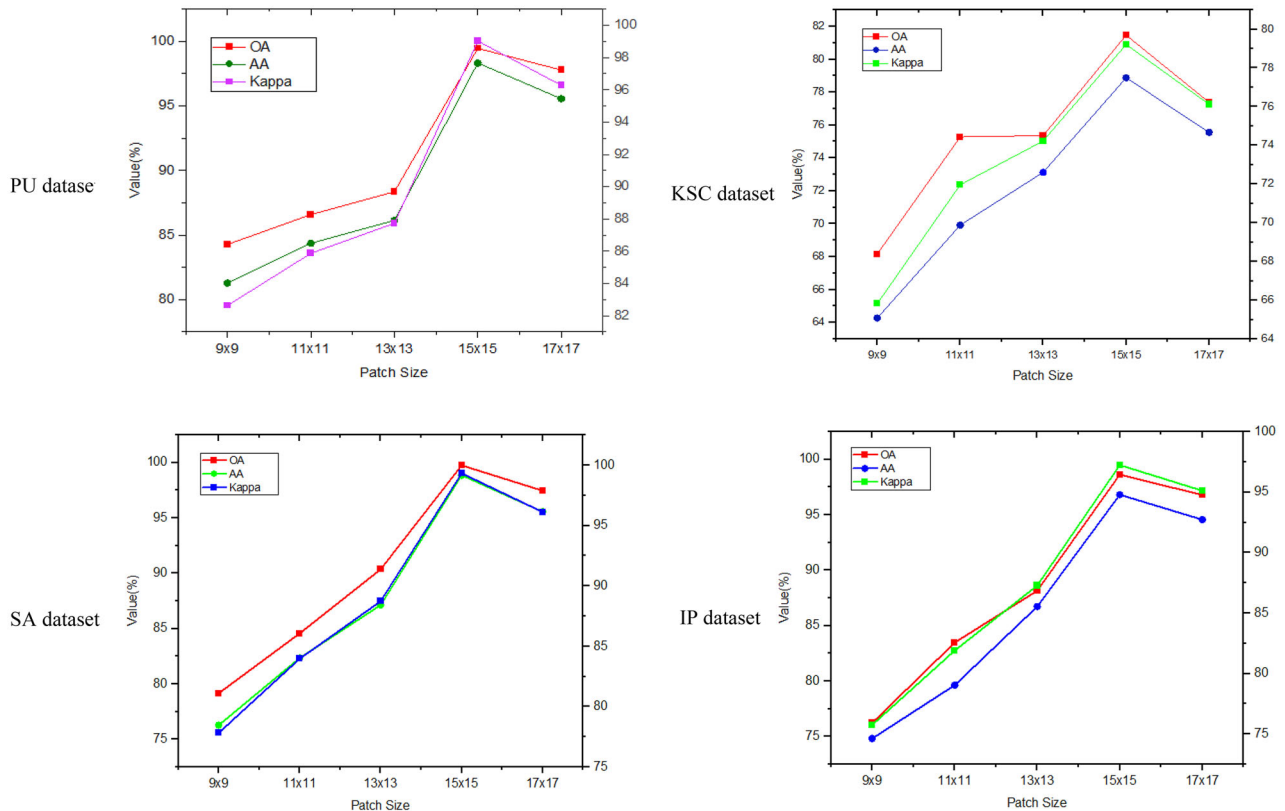
Figure 3. Effect of patch size on classification performance.
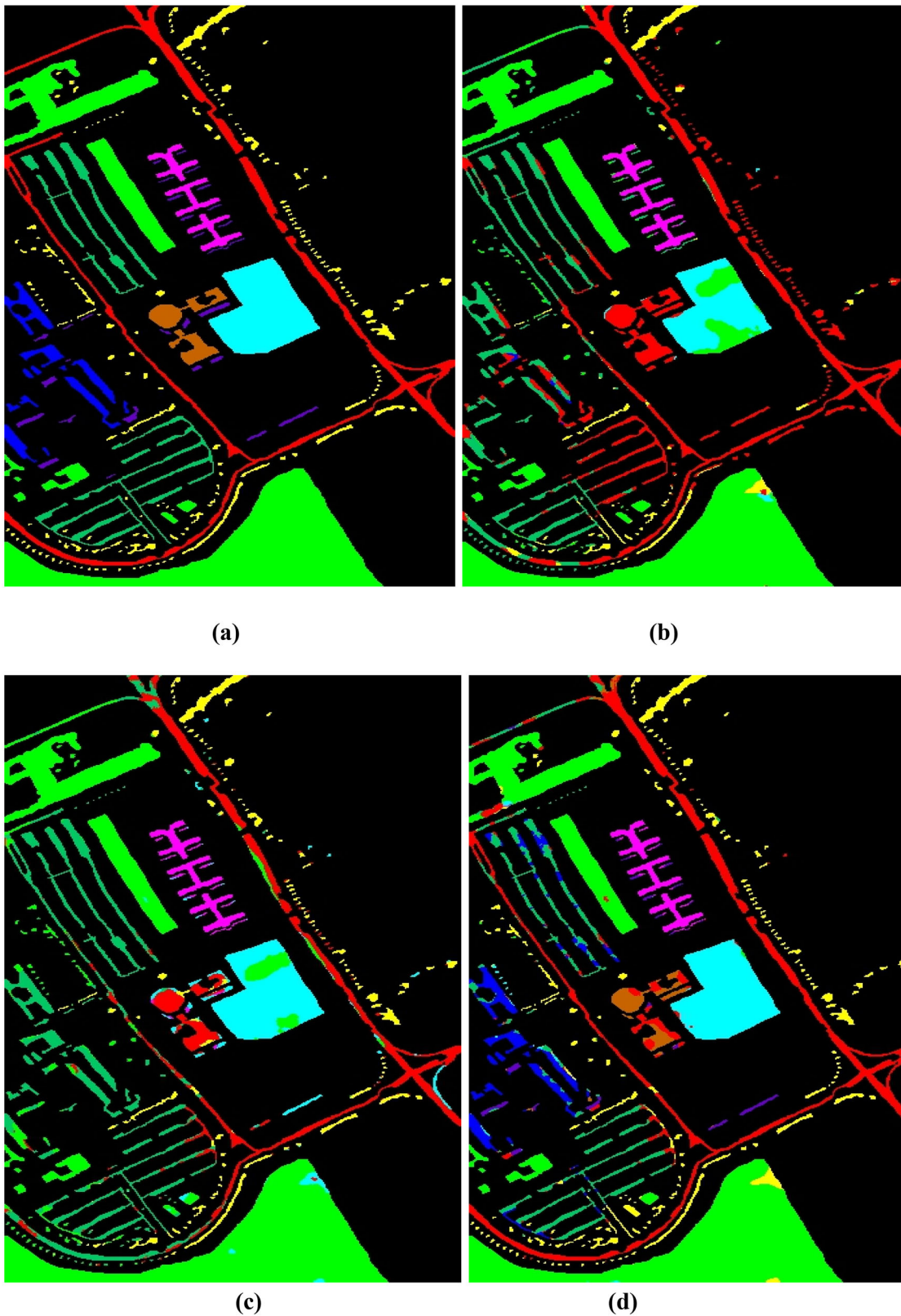
Figure 4. The class visual map of PU data set and GT, 3(a) GT, 3(b) SVM, 3(c) 1D CNN, 3(d) 2D CNN, 3(e) 3D CNN, 3(f) HybridSN, 3(g) SF, and 3(h) HResNeXt, respectively.

(e)

(f)

(g)

(h)
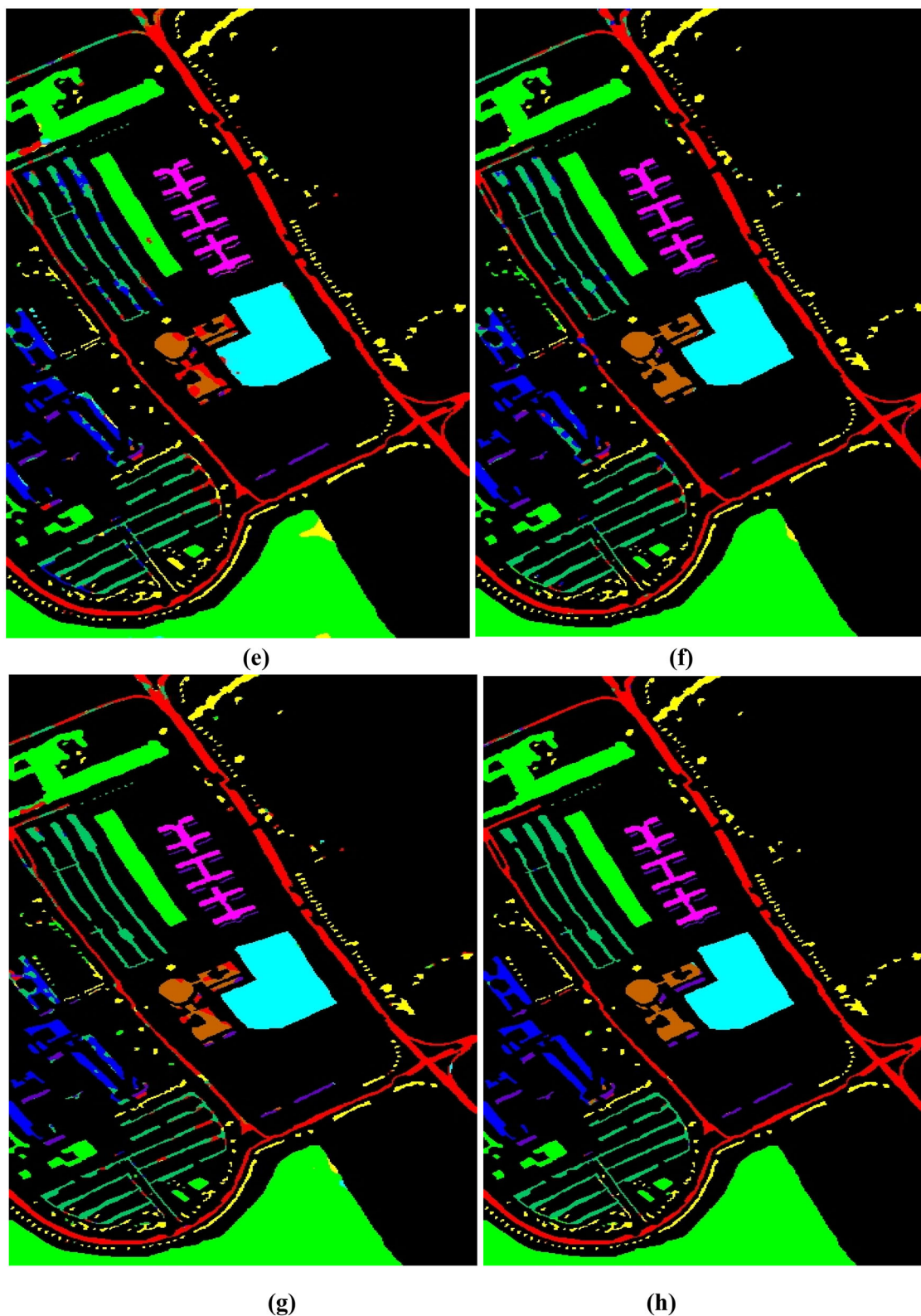
Figure 4. Continued.

HResNeXt accuracy is less in 9×9 and 11×11, whereas the highest classification accuracy is achieved on a 15×15 patch size. Further, increasing the patch size reduces classification accuracy.

Visual analysis

We present a class visual map of the PU, KSC, SA, and IP data sets in Figures 4–7. In Figure 4, we can see that the land cover classification map using SVM is less close

(a) GT 3
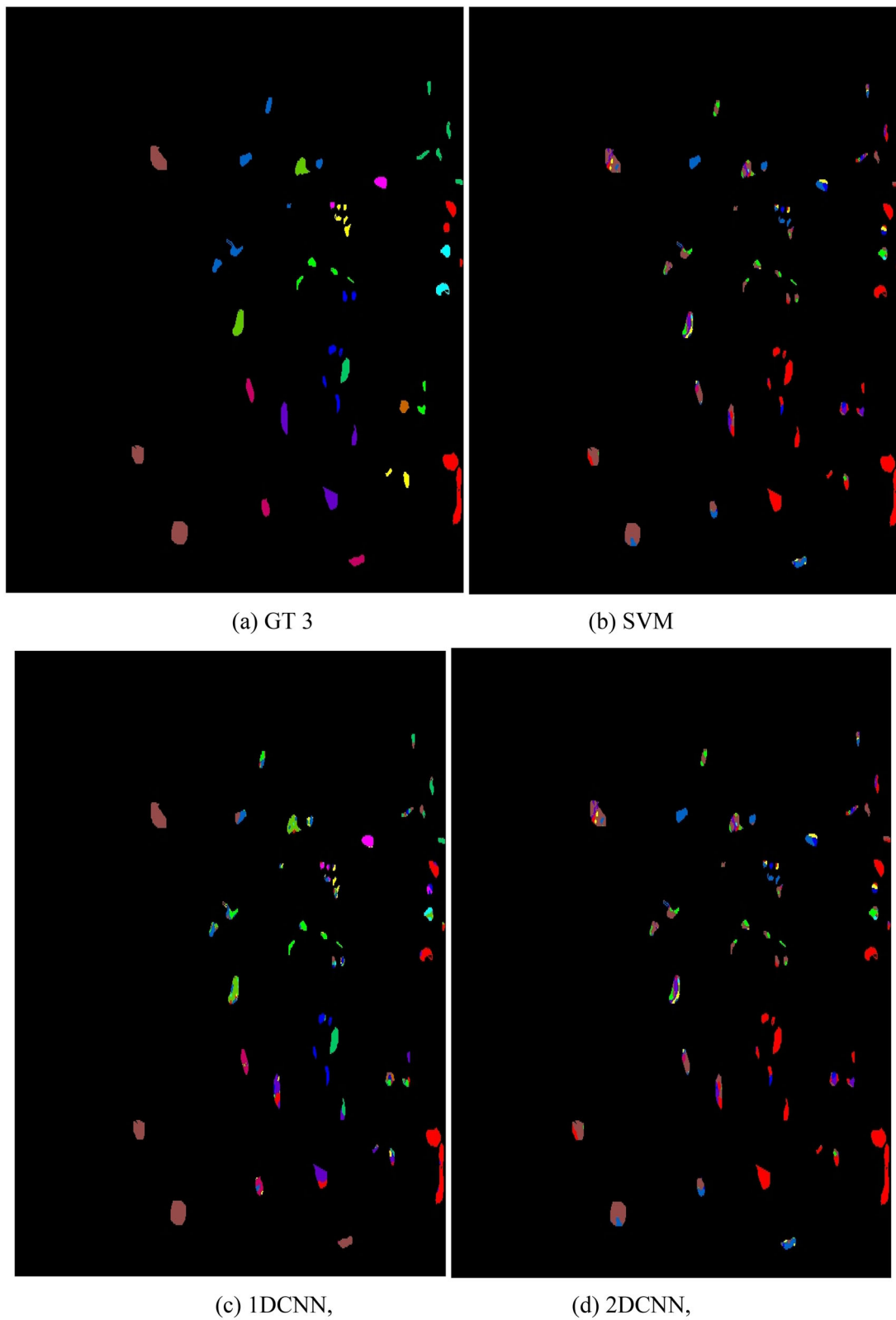
(b) SVM

(c) 1DCNN,

(d) 2DCNN,

Figure 5. The class visual map of KSC dataset and GT, 3(a) GT, 3(b) SVM, 3(c) 1D CNN, 3(d) 2D CNN, 3(e) 3D CNN, 3(f) HybridSN, 3(g) SF, and 3(h) HResNeXt, respectively.

(e) 3DCNN

(f) HybridSN

(g) SF and

(h)HResNeXt
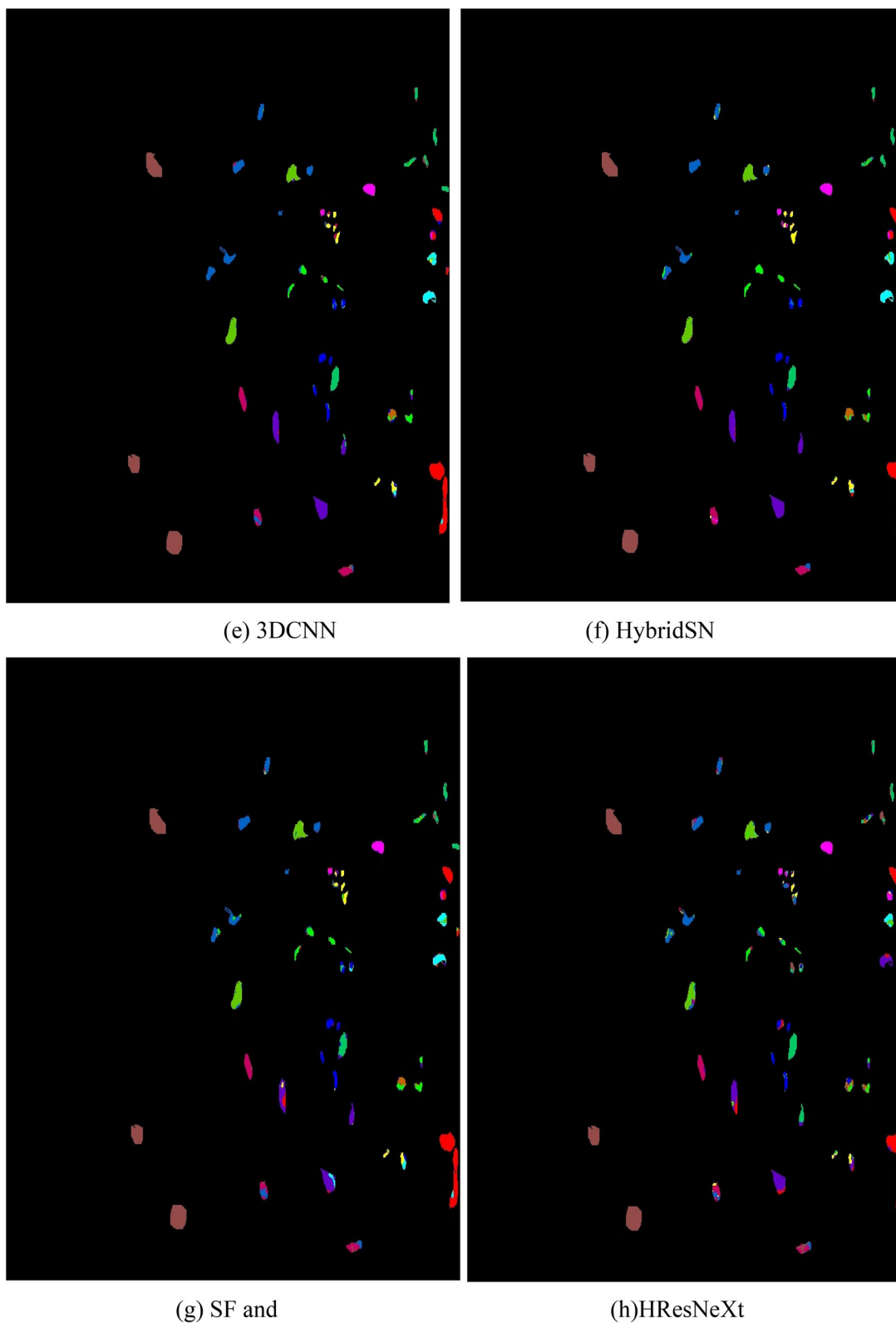
to the ground truth (GT) in several classes, especially in Asphalt, Bitumen, Self-Blocking Bricks, and Shadows classes. In contrast, the 1D CNN has improved visual maps in several classes. Much better object visualization can be seen in the 2D CNN approach, which achieved a very close map in the Painted metal sheets class compared to GT. The visual map of the Meadows class using 3D CNN is much better than other methods.
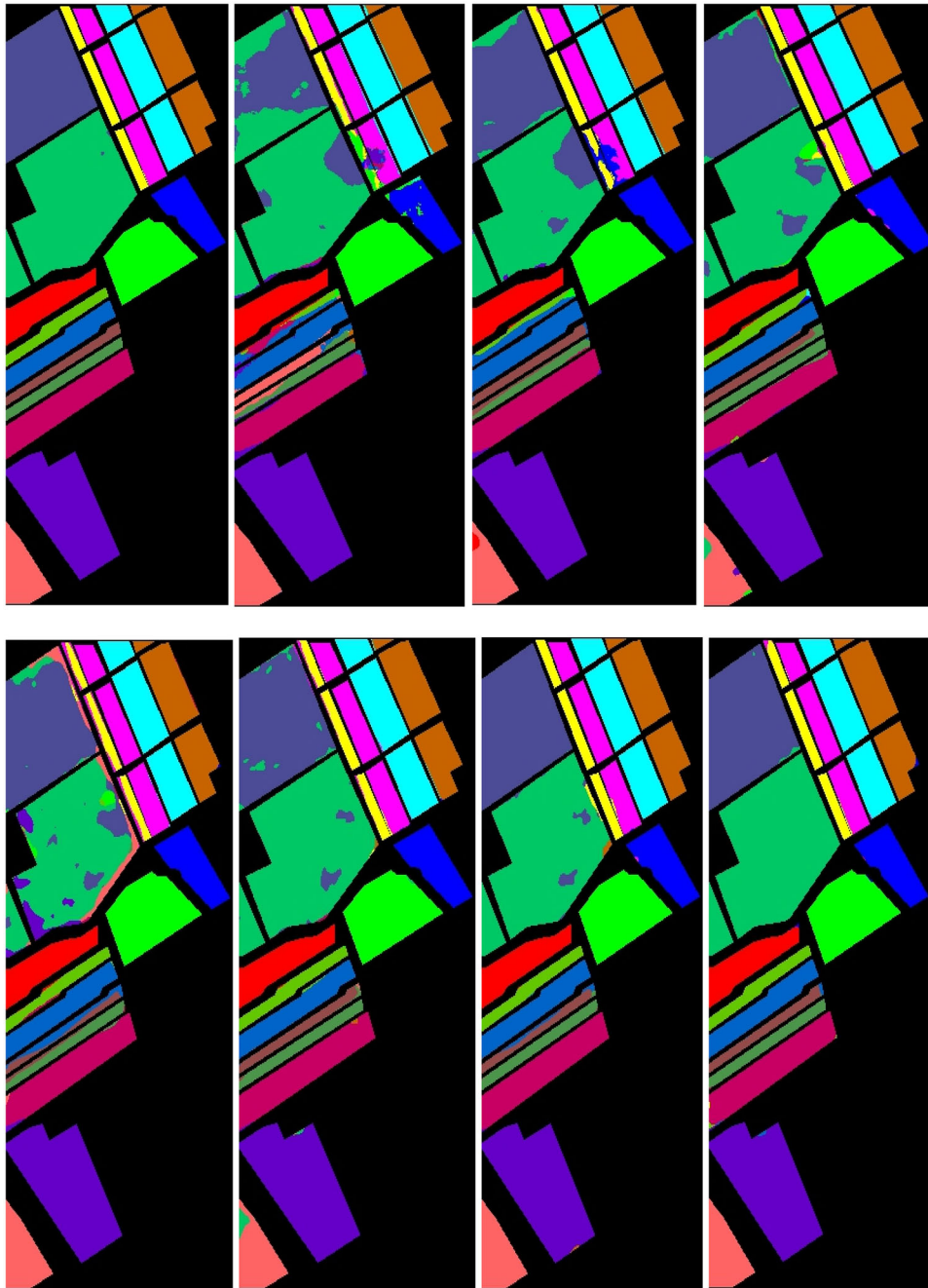
Figure 6. The class visual map of SA data set and GT, 3(a) GT, 3(b) SVM, 3(c) 1D CNN, 3(d) 2D CNN, 3(e) 3D CNN, 3(f) HybridSN, 3(g) SF, and 3(h) HResNeXt, respectively.

In contrast, the HybridSN visual map of the Asphalt class is similar to GT. The SF utilized global feature attention to improve the land cover classification map. The HResNeXt visual map and GT are very close in the Trees, Bare Soil, Bitumen, and Shadows classes. Similarly, in Figure 5, we can observe that the classification map of SVM, 1D CNN, and 2D CNN in several land covers suffered from noise. However, 3D CNN has a better Graminoid marsh class classification map than other methods. The Oak and Scrub classes' HybridSN and SF land cover map is much better.

Furthermore, the proposed method classification maps are very close to GT in several other classes.

In Figure 6, we can see that the classification map of the land covers using SVM is further from the ground truth (GT) in several classes. In contrast, the 1D CNN has improved visual maps in several classes. Much better object visualization can be seen in the 2D CNN approach, which achieved a very close map in the Celery class compared to GT. The visual map of the Stubble class using 3D CNN is much better than other methods. In contrast, the HybridSN visual
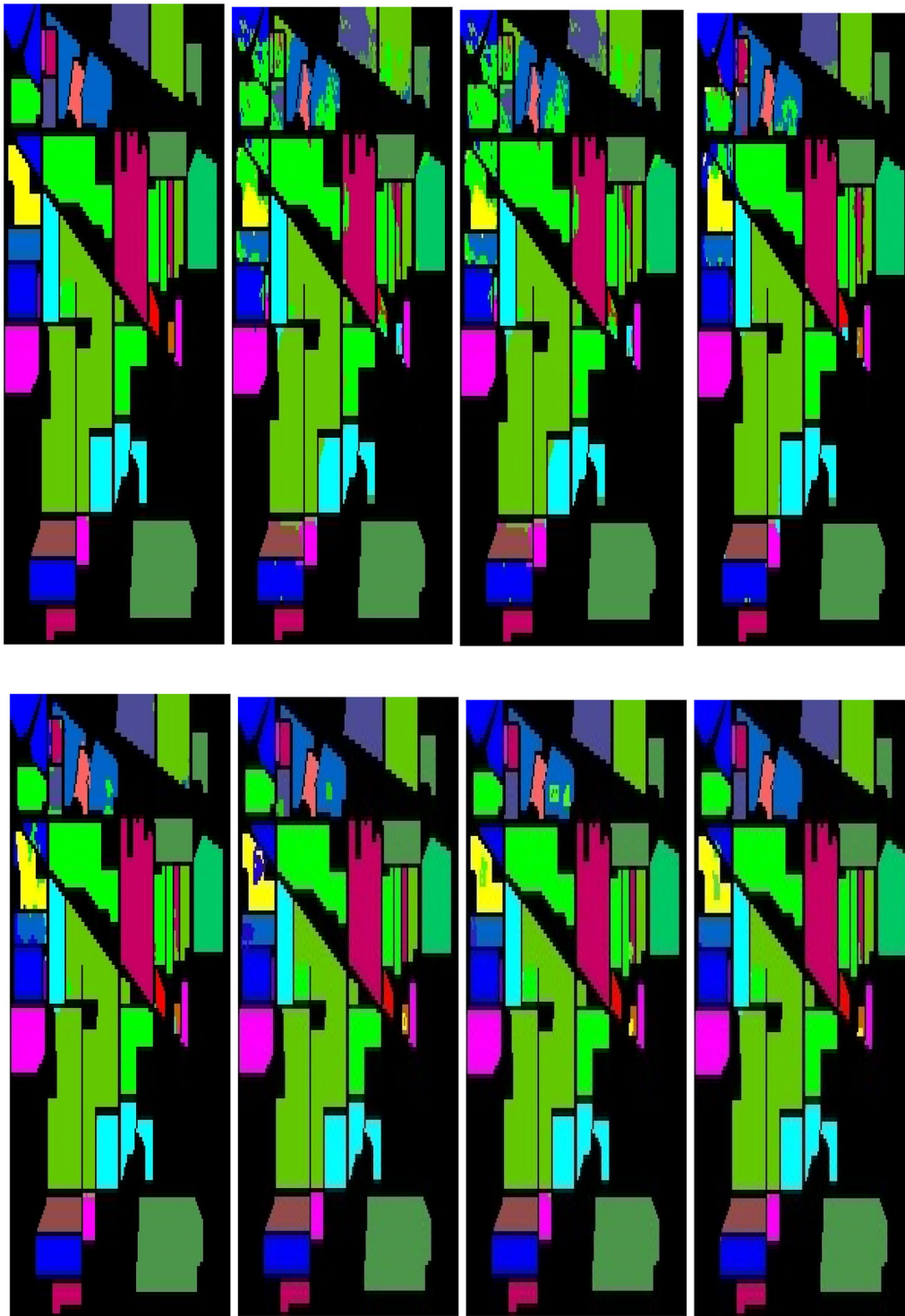
Figure 7. The class visual map of IP data set and GT, 3(a) GT, 3(b) SVM, 3(c) 1D CNN, 3(d) 2D CNN, 3(e) 3D CNN, 3(f) HybridSN, 3(g) SF, and 3(h) HResNeXt, respectively.

map of the Fallow classes is similar to GT. The SF utilized global feature attention to improve the land cover classification map observed in Brocoli_green_weeds_1 class. The HResNeXt visual map and GT are very GT in several other classes.

Similarly, in Figure 7, we can observe that the classification map of SVM and 1D CNN suffered from noise.

The 2D CNN has improved the visual map in the Soybean-clean class. However, 3D CNN has a better classification map in Corn-mintill and Grass-trees classes than other methods. The HybridSN and SF land cover map of the Soybean-min-till and Woods class is much better. Furthermore, the proposed method classification maps are very close to GT in several other classes.
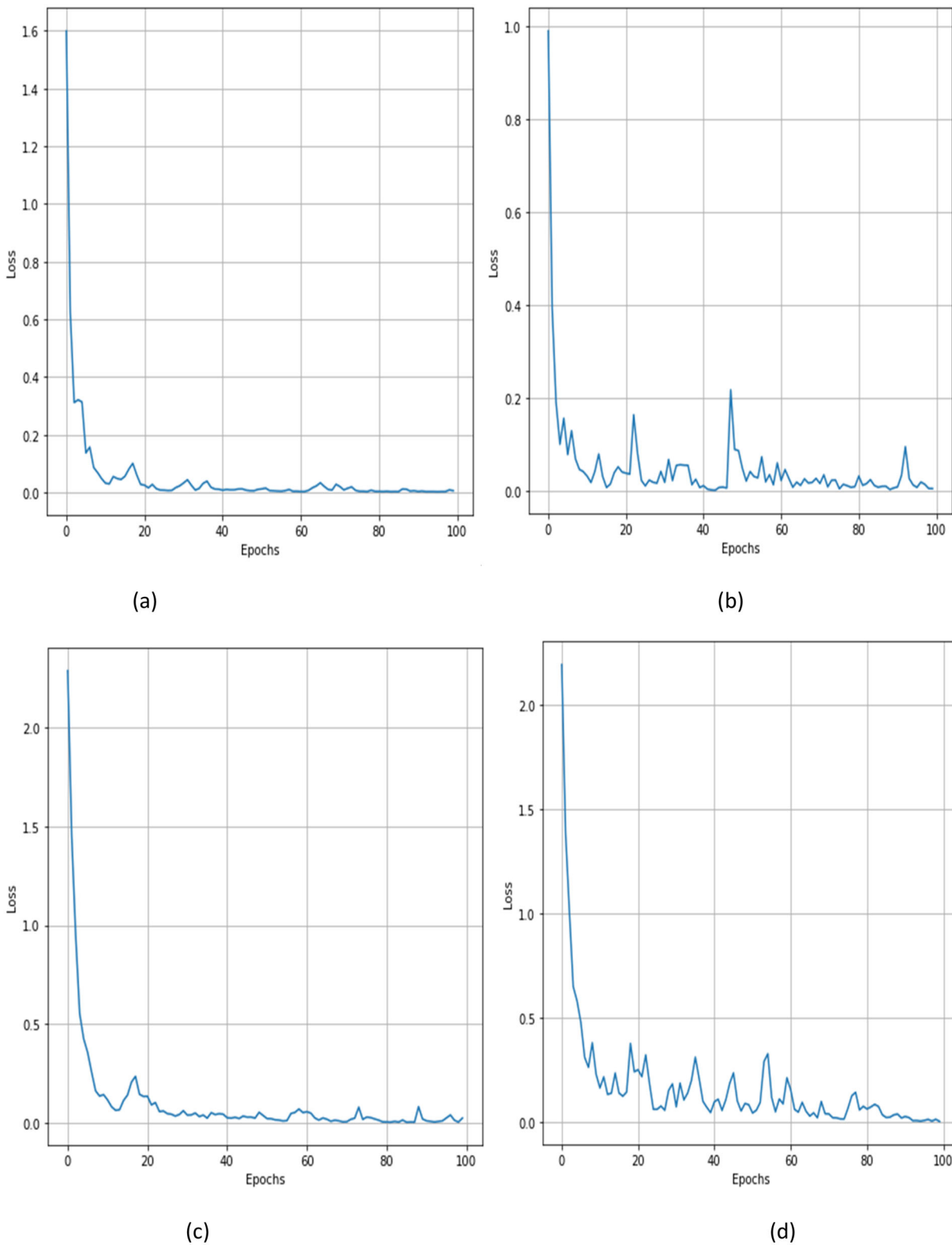
Figure 8. The training loss of the proposed model on (a) PU, (b) KSC, (c) SA, and (d) IP.

The classification accuracy was significantly improved by using a variety of machine learning and deep learning techniques. Nevertheless, computation costs and highly efficient method is less that can be used for real-time applications. We compared these methodologies with the proposed HResNext syste g m by usinquantitative and visual maps. The SVM-based approach cannot extract high-dimensional features

Table 11. Computation time on training and validation data.

Data set	Train Time (min)	Test Time (s)
PU	1.63	7.86
KSC	1.02	2.54
SA	1.14	3.86
IP	1.85	6.82

due to its design limitations. The feature extraction process used by 1D CNN is called directional convolution. In addition, 2D CNN is capable of calculating spatial features in both directions, but it does not have spectral features. 3D CNN computation cost is high but capable of extracting high-dimensional spectral features.

The training loss of the proposed method on different data sets

We have performed several experiments on data sets. We notice no significant changes in the training accuracy after 100 epochs. Therefore, the model was trained for 100 epochs only. By doing this, computation costs can be reduced. The training loss curve proposed method on PU, KSC, SA, and IP is shown in Figure 8.

The computation time of the proposed method on training and validation data

We have summarized the training and validation time of the proposed method on PU, KSC, SA, and IP data sets. The detailed summary of the training time in minutes and test time in seconds is shown in Table 11.

Conclusion

Hyperspectral image classification is challenging and requires a sophisticated method to better utilize the rich spatial and spectral features. Many machine learning and deep learning techniques enhance classification accuracy. Nevertheless, computation costs and highly efficient method is less that can be used for real-time applications. The lightweight HResNeXt model is specifically designed to overcome traditional methods' limitations. The HResNeXt successfully captures spectral and spatial information concurrently. In the proposed study, we utilized only one 3D convolution block for spectral features and a modified 2D residual block to capture spatial features. The original ResNeXt has many trainable parameters, which can increase the computation cost. Hence, first, we reduced the trainable parameters that reduce the costs.

After that, we jointly extracted spectral and spatial features to improve the quantitative and visual performance. Subsequently, it enables efficient and effective feature extraction from hyperspectral images, resulting in competitive classification accuracy. The HResNeXt obtained an OA on PU, KSC, SA, and IP, 99.46%, 81.46%, 99.75%, and 98.64%, respectively. In future study, we will explore more advanced and lightweight graph CNNs and vision transformers. In addition, the integration of handcrafted features and deep features can be used to improve classification accuracy. Further, high-dimensional features extracted by the model can be optimized using the nature-inspired algorithm to enhance the classification performance. The computation cost of the algorithm is still a big challenge that needs further reduction. In addition, the dimension reduction algorithm PCA has been applied in the proposed study. There may be slight differences in the performance of other dimension reduction algorithms. Further, the model can be implemented on a real-time data set.

Acknowledgement

References

Arikumar, K.S., Deepak Kumar, A., Gadekallu, T.R., Prathiba, S.B., and Tamilarasi, K. 2022. "Real-time 3D object detection and classification in autonomous driving environment using 3D LiDAR and camera sensors." *Electronics*, Vol. 11(No. 24): p. 4203. doi:10.3390/electronics11244203.

Audebert, N., Le Saux, B., and Lefevre, S. 2019. "Deep learning for classification of hyperspectral data: A comparative review." *IEEE Geoscience and Remote Sensing Magazine*, Vol. 7(No. 2): pp. 159–173. doi:10.1109/MGRS.2019.2912563.

Bashir, M.F., Javed, A.R., Arshad, M.U., Gadekallu, T.R., Shahzad, W., and Beg, M.O. 2023. "Context aware emotion detection from low resource Urdu language using deep neural network." *ACM Transactions on Asian and Low-Resource Language Information Processing*, Vol. 22 (No. 5): pp. 1–30. doi:10.1145/3528576.

Ben Hamida, A., Benoit, A., Lambert, P., and Ben Amar, C. 2018. "3-D deep learning approach for remote sensing image classification." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 56(No. 8): pp. 4420–4434. doi:10.1109/TGRS.2018.2818945.

Benediktsson, J.A., Palmason, J.A., and Sveinsson, J.R. 2005. "Classification of hyperspectral data from urban areas based on extended morphological profiles." *IEEE*

Transactions on Geoscience and Remote Sensing, Vol. 43(No. 3): pp. 480–491. doi:10.1109/TGRS.2004.842478.

Cao, X., Zhou, F., Xu, L., Meng, D., Xu, Z., and Paisley, J. 2018. "Hyperspectral image classification with Markov random fields and a convolutional neural network." *IEEE Transactions on Image Processing : a Publication of the IEEE Signal Processing Society*, Vol. 27(No. 5): pp. 2354–2367. doi:10.1109/TIP.2018.2799324.

Chen, Y., Chen, Z., Guo, D., Zhao, Z., Lin, T., and Zhang, C. 2022. "Underground space use of urban built-up areas in the central city of Nanjing: Insight based on a dynamic population distribution." *Underground Space.*, Vol. 7(No. 5): pp. 748–766. doi:10.1016/j.undsp.2021.12.006.

Chen, Y., Jiang, H., Li, C., Jia, X., and Ghamisi, P. 2016. "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 54(No. 10): pp. 6232–6251. doi:10.1109/TGRS.2016.2584107.

Dalla Mura, M., Villa, A., Benediktsson, J.A., Chanussot, J., and Bruzzone, L. 2011. "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis." *IEEE Geoscience and Remote Sensing Letters*, Vol. 8(No. 3): pp. 542–546. doi:10.1109/LGRS.2010.2091253.

Deng, F., Pu, S., Chen, X., Shi, Y., Yuan, T., and Pu, S. 2018. "Hyperspectral image classification with capsule network using limited training samples." *Sensors (Basel, Switzerland)*, Vol. 18(No. 9): p. 3153. doi:10.3390/s18093153.

El-Assal, M., Tirilly, P., and Bilasco, I.M. 2022. 2D versus 3D convolutional spiking neural networks trained with unsupervised STDP for human action recognition. http://arxiv.org/abs/2205.13474

Fang, B., Liu, Y., Zhang, H., and He, J. 2022. "Hyperspectral image classification based on 3D asymmetric inception network with data fusion transfer learning." *Remote Sensing*, Vol. 14(No. 7): p. 1711. doi:10.3390/rs14071711.

Fang, L., He, N., Li, S., Ghamisi, P., and Benediktsson, J.A. 2018. "Extinction Profiles Fusion for Hyperspectral Images Classification." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 56(No. 3): pp. 1803–1815. doi:10.1109/TGRS.2017.2768479.

Fang, L., Liu, G., Li, S., Ghamisi, P., and Benediktsson, J.A. 2019. "Hyperspectral image classification with squeeze multibias network." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 57(No. 3): pp. 1291–1301. doi:10.1109/TGRS.2018.2865953.

Fu, C., Yuan, H., Xu, H., Zhang, H., and Shen, L. 2023. "TMSO-Net: Texture adaptive multiscale observation for light field image depth estimation." *Journal of Visual Communication and Image Representation*, Vol. 90: p. 103731. doi:10.1016/j.jvcir.2022.103731.

Gadekallu, T.R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P.K.R., and Srivastava, G. 2023. "Deep neural networks to predict diabetic retinopathy." *Journal of Ambient Intelligence and Humanized Computing*, Vol. 14(No. 5): pp. 5407–5420. doi:10.1007/s12652-020-01963-7.

Garg, H., Gupta, N., Agrawal, R., Shivani, S., and Sharma, B. 2022. "A real time cloud-based framework for glaucoma screening using EfficientNet." *Multimedia Tools and Applications*, Vol. 81 (No. 24): pp. 34737–34758. doi:10.1007/s11042-021-11559-8.

Ghamisi, P., Yokoya, N., Li, J., Liao, W., Liu, S., Plaza, J., Rasti, B., and Plaza, A. 2017. "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art." *IEEE Geoscience and Remote Sensing Magazine*, Vol. 5(No. 4): pp. 37–78. doi:10.1109/MGRS.2017.2762087.

Glorot, X., and Bengio, Y. 2010. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (pp. 249–256).

Grupo de Inteligencia Computacional (GIC). 2023. Ehu.eus website: https://www.ehu.eus/ccwintco/index.php

Guo, Z., Yu, K., Li, Y., Srivastava, G., and Lin, J.C.-W. 2022. "Deep learning-embedded social internet of things for ambiguity-aware social recommendations." *IEEE Transactions on Network Science and Engineering*, Vol. 9(No. 3): pp. 1067–1081. doi:10.1109/TNSE.2021.3049262.

Hang, R., Liu, Q., Hong, D., and Ghamisi, P. 2019. "Cascaded recurrent neural networks for hyperspectral image classification." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 57(No. 8): pp. 5384–5394. doi:10.1109/TGRS.2019.2899129.

He, X., and Chen, Y. 2021. "Modifications of the multi-layer perceptron for hyperspectral image classification." *Remote Sensing*, Vol. 13(No. 17): p. 3547. doi:10.3390/rs13173547.

Hong, D., Han, Z., Yao, J., Gao, L., Zhang, B., Plaza, A., and Chanussot, J. 2022. "SpectralFormer: Rethinking hyperspectral image classification with transformers." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 60: pp. 1–16. doi:10.1109/TGRS.2021.3130716.

Ji, J., Liu, S., Zhang, F., Liao, X., Wang, S., and Liao, J. 2023. "Hyperspectral Image Classification Based on Unsupervised Regularization." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 16: pp. 1871–1882. doi:10.1109/JSTARS.2023.3241662.

Jiang, Z., and Chen, L. 2022. "Multisemantic Level Patch Merger Vision Transformer for diagnosis of pneumonia." *Computational and Mathematical Methods in Medicine*, Vol. 2022: pp. 7852958. doi:10.1155/2022/7852958.

Joelsson, S.R., Benediktsson, J.A., and Sveinsson, J.R. 2005. Random forest classifiers for hyperspectral data. In *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05*. IEEE. doi:10.1109/IGARSS.2005.1526129.

Li, R., Zheng, S., Duan, C., Yang, Y., and Wang, X. 2020. "Classification of hyperspectral image based on double-branch dual-attention mechanism network." *Remote Sensing*, Vol. 12(No. 3): pp. 582. doi:10.3390/rs12030582.

Linzen, T., Dupoux, E., and Goldberg, Y. 2016. "Assessing the ability of LSTMs to learn syntax-sensitive dependencies." *Transactions of the Association for Computational Linguistics*, Vol. 4: pp. 521–535. doi:10.1162/tacl_a_00115.

Melgani, F., and Bruzzone, L. 2004. "Classification of hyperspectral remote sensing images with support vector machines." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 42(No. 8): pp. 1778–1790. doi:10.1109/TGRS.2004.831865.

Mou, L., and Zhu, X.X. 2020. "Learning to pay attention on spectral domain: A spectral attention module-based

convolutional network for hyperspectral image classification." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 58(No. 1): pp. 110–122. doi:10.1109/TGRS.2019.2933609.

Mou, L., Lu, X., Li, X., and Zhu, X.X. 2020. "Nonlocal Graph Convolutional Networks for Hyperspectral Image Classification." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 58(No. 12): pp. 8246–8257. doi:10.1109/TGRS.2020.2973363.

Pan, B., Shi, Z., and Xu, X. 2017. "R-VCANet: A new deep-learning-based hyperspectral image classification method." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 10(No. 5): pp. 1975–1986. doi:10.1109/JSTARS.2017.2655516.

Peng, Y., Zhang, Y., Tu, B., Li, Q., and Li, W. 2022. "Spatial–Spectral Transformer With Cross-Attention for Hyperspectral Image Classification." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 60: pp. 1–15. doi:10.1109/TGRS.2022.3203476.

Ravi, C., Tigga, A., Reddy, G.T., Hakak, S., and Alazab, M. 2022. "Driver identification using optimized deep learning model in smart transportation." *ACM Transactions on Internet Technology*, Vol. 22(No. 4): pp. 1–17. doi:10.1145/3412353.

Roy, S.K., Krishna, G., Dubey, S.R., and Chaudhuri, B.B. 2020. "HybridSN: Exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification." *IEEE Geoscience and Remote Sensing Letters*, Vol. 17(No. 2): pp. 277–281. doi:10.1109/LGRS.2019.2918719.

Saab, S., Jr, Phoha, S., Zhu, M., and Ray, A. 2022. "An adaptive polyak heavy-ball method." *Machine Learning*, Vol. 111 (No. 9): pp. 3245–3277. doi:10.1007/s10994-022-06215-7.

Sharma, M., and Biswas, M. 2018. "CRCOED: Collaborative representation-based classification using odd even decomposition for hyperspectral remote sensing imagery." *Procedia Computer Science*, Vol. 143: pp. 458–465. doi:10.1016/j.procs.2018.10.418.

Sun, Y., Fu, Z., and Fan, L. 2019. "A novel hyperspectral image classification pattern using random patches convolution and local covariance." *Remote Sensing*, Vol. 11(No. 16): p. 1954. doi:10.3390/rs11161954.

Vallathan, G., John, A., Thirumalai, C., Mohan, S., Srivastava, G., and Lin, J.C.-W. 2021. "Suspicious activity detection using deep learning in secure assisted living IoT environments." *The Journal of Supercomputing*, Vol. 77(No. 4): pp. 3242–3260. doi:10.1007/s11227-020-03387-8.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., and Polosukhin, I. 2017. Attention is all you need. http://arxiv.org/abs/1706.03762

Wang, S., Hu, X., Sun, J., and Liu, J. 2023. "Hyperspectral anomaly detection using ensemble and robust collaborative representation." *Information Sciences*, Vol. 624: pp. 748–760. doi:10.1016/j.ins.2022.12.096.

Waske, B., van der Linden, S., Benediktsson, J.A., Rabe, A., and Hostert, P. 2010. "Sensitivity of support vector machines to random feature selection in classification of hyperspectral data." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48(No. 7): pp. 2880–2889. doi:10.1109/TGRS.2010.2041784.

Xie, S., Girshick, R., Dollar, P., Tu, Z., and He, K. 2017. Aggregated residual transformations for deep neural networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. doi:10.1109/CVPR.2017.634.

Xu, X., Shen, B., Ding, S., Srivastava, G., Bilal, M., Khosravi, M.R., Menon, V.G., Jan, M.A., and Wang, M. 2022. "Service offloading with deep Q-network for digital twinning-empowered internet of vehicles in edge computing." *IEEE Transactions on Industrial Informatics*, Vol. 18(No. 2): pp. 1414–1423. doi:10.1109/TII.2020.3040180.

Yadav, D.P., Jalal, A.S., and Prakash, V. 2022. "Human burn depth and grafting prognosis using ResNeXt topology based deep learning network." *Multimedia Tools and Applications*, Vol. 81(No. 13): pp. 18897–18914. doi:10.1007/s11042-022-12555-2.

Zhao, B., Ragnarsson, H.I., Ulfarsson, M.O., Cavallaro, G., and Benediktsson, J.A. 2022. "Predicting Classification Performance for Benchmark Hyperspectral Datasets." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 15: pp. 4180–4193. doi:10.1109/JSTARS.2022.3173893.

Zhuo, Z., Du, L., Lu, X., Chen, J., and Cao, Z. 2022. "Smoothed lv distribution based three-dimensional imaging for spinning space debris." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 60: pp. 1–13. doi:10.1109/TGRS.2022.3174677.