International Journal of Applied Earth Observations and Geoinformation 105 (2021) 102603

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Hyperspectral image classification on insufficient-sample and feature learning using deep neural networks: A review

Naftaly Wambugu [a], Yiping Chen [a,*], Zhenlong Xiao [a], Kun Tan [b], Mingqiang Wei [c], Xiaoxue Liu [a], Jonathan Li [a,d,*]

[a] *Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, Xiamen, FJ 361005, China*
[b] *Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China*
[c] *School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China*
[d] *Department of Geography & Environmental Management and Department of Systems Design Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

ARTICLE INFO

Keywords:
Deep learning
Image classification
Remote sensing
Hyperspectral imagery
Supervised learning

ABSTRACT

Over the years, advances in sensor technologies have enhanced spatial, temporal, spectral, and radiometric resolutions, thus significantly improving the size, resolution, and quality of imagery. These vast developments have inspired improvement in various hyperspectral images (HSI) classification applications such as land cover mapping, vegetation classification, urban monitoring, and understanding which are essential for better utilization of Earth's resources. HSI classification requires superior algorithms with greater accuracy, less computational complexity, and robustness to extract rich, spectral-spatial information. Deep convolution neural networks (DCCNs) have revolutionized image classification experience, with robust architectures being proposed from time to time. However, insufficient training samples have been earmarked as a significant bottleneck for supervised HSI classification and have not been fully explored in literature. To stimulate further research, this paper reviews current methods that handle labeled data insufficiency and the current feature learning methods for HSI classification using DCNNs. It also presents various methods' results on the three most popular public HSI datasets, together with intuitive observations motivating future research by the hyperspectral community.

1. Introduction

Advances in image acquisition techniques have upscaled spectral-spatial image resolutions, improved image processing approaches, and spurred a continuous generation of high volumes of quality data (Fu et al., 2017). Subsequently, high quality and low-cost data obtained from the sensors, coupled with the availability of advanced computing resources such as graphics processing units (GPUs) and parallel computing, has led to superior computer algorithms continuously enabling researchers to understand the ground surface, morphological changes, and human processes with greater precision and detail.

Both image classification and semantic segmentation are widely researched sub-domains of computer vision and have been applied in RS image analysis, face recognition, medical image segmentation, among others. In image classification, all objects within an image are grouped and categorized into a single class, while in semantic segmentation, each pixel in an image is assigned to a set of predefined classes/labels, where the same labels share certain characteristics (Kemker et al., 2018b). Hyperspectral image (HSI) classification has several applications such as land cover mapping and change detection (Crowson et al., 2019; Xu et al., 2019), soil organic carbon prediction (Meng et al., 2020), vegetation classification (Laliberte et al., 2011), forest biomass understanding and tree species identification and mapping (Modzelewska et al., 2020), urban monitoring and understanding (Chen et al., 2018a), volcanic activity investigation and monitoring (Modzelewska et al., 2020) among others, courtesy of the rich spectral-spatial information obtained from HSI data. However, it is costly and labor-intense to label HSI data due to the different sensors used to capture the data and the domain expertise involved. This explains why the existing labeled HSI benchmark datasets are few, making HSI classification lag behind other vision-based and image processing domains due to the limited annotated labels and the complicated nature of HSI data (Gao et al., 2018). Traditional classifiers such as random forest and support vector machines have had great success in HSI classification (Belgiu and Drăguţ, 2016), but only

* Corresponding authors at: Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, Xiamen, FJ 361005, China (J. Li).

E-mail addresses: chenyiping@xmu.edu.cn (Y. Chen), junli@uwaterloo.ca (J. Li).

https://doi.org/10.1016/j.jag.2021.102603
Received 11 May 2021; Received in revised form 19 October 2021; Accepted 20 October 2021
Available online 6 November 2021

use spectral information which limits their optimal performance. Moreover, the large dimensionality of spectral-spatial data, coupled with limited training samples impedes the improvement of classification performance (Zhong et al., 2018).

Deep Learning (DL) is a data-driven paradigm that follows end-to-end machine learning to make decisions without human-designed algorithms. DL has recently been applied extensively in various research domains such as computer vision, speech recognition, and natural language processing. In the past decades, several methods exploited spectral-spatial features from HSI data using stacked autoencoders (SAE) and deep belief networks (DBN) (Zou et al., 2020). Features were mainly extracted through a flattening layer that collapsed spatial dimensions into a 1-D vector. Several approaches such as principal component analysis (PCA), independent component analysis (ICA), and Fisher's linear discriminant analysis (LDA) have been used to achieve dimensionality reduction and feature extraction on the input data (Jiang et al., 2021). These approaches, however, often lead to the loss of essential spatial information, which is critical for image classification.

Various deep learning networks have been proposed for HSI classification; Restricted Boltzmann Machines (RBMs) (Midhun et al., 2014), autoencoder (AE) (Tao et al., 2015), deep belief network (DBN) (Li et al., 2014a), recurrent neural network (RNN) (Mou et al., 2017), and deep convolutional neural networks (CNNs). The advent of deep convolution neural networks (DCNNs) as subsets of DL architectures has transformed the image processing experience and is credited for handling image classification tasks (Hu et al., 2015). Broad adoption and popularity of DCNNs in handling image classification tasks can be associated with their ability to discriminately learn and extract hidden, complex, and underlying non-linear features from raw images, flexibility in the type of data they handle (such as spatial, spectral, point cloud, etc.), and customizable building blocks (number of layers, depth) (Krizhevsky et al., 2017; Wang et al., 2019c). Moreover, DCNNs are easily implemented in high-end processing units such as distributed systems and GPUs and have demonstrated outstanding performance in several application domains such as image compressive sensing (Zhou et al., 2021), visual object tracking (Zhang et al., 2019a; Zhang et al., 2021), and hyperspectral image classification (Zhong et al., 2018), among others. Feature extraction in DCNNs is achieved hierarchically. Lower layers extract generic features, and higher layers learn abstract and more task-specific representations through non-linear mapping relationships of stacked layers (convolutional layers, activation, and pooling layers) and a classifier (Chen et al., 2016).

Over the last few years, intense research has proposed a plethora of algorithms aimed at capturing complex features and abstract semantics from HSI data for in-depth understanding and exploitation of the rich spectral-spatial features (Li et al., 2019). However, large-scale HSI training data is limited due to the cost, complexity, and labeling constraints in developing such datasets (Shen et al., 2019), which cause sub-optimal learning of DCNN with large numbers of parameters. On this basis, the hyperspectral community has been devoted to proposing methods to generate more labeled samples and developing methods that can perform effectively with limited labeled samples.

While this review paper is not meant to delve into the technical aspects of HSI data, such as acquisition mechanisms, applications, and data analysis, it draws insights on methods and advances relating to insufficient labeled data for HSI image classification. Highly informative reviews on HSI data analysis using ML models, pre-processing, datasets, change detection, and applications have recently been presented (Audebert et al., 2019; Li et al., 2019; Signoroni et al., 2019). Moreover, Paoletti et al. (2019c) presented a comprehensive review on current HSI airborne and satellite data acquisition, pre-processing measures, classifiers, and learning process optimization (drop out, regularization, normalization, etc.). In this report, different from the other existing pieces of literature, our core contribution can be summarized as follows:

(1) We present an updated review of current methods handling training samples insufficiency in HSI classification.
(2) The most recent methods advancing feature learning in spectral-spatial-based DCNNs for HSI image classification are presented.
(3) A report on the performance of existing methods on select public datasets is provided, with short summaries and insightful discussions.

The remainder of the paper is organized as follows: Section 2 highlights new training samples generation. Methods for handling limited training samples are discussed in Section 3. Spatial-spectral learning-based DCNNs are highlighted in Section 4. Section 5 presents a performance report of the existing methods and discussions of various methods, while the conclusion is provided in Section 6.

1.1. *Existing public hyperspectral datasets*

Training samples are increasingly critical for all computer vision tasks that use the DL paradigm. The existing training sets are insufficient for training deep networks for HSI tasks. Using them can significantly cause overfitting (Castelluccio et al., 2015), while gathering and constructing satisfactory datasets for HSI tasks remains costly, time-consuming, and requires domain expertise (Mnih et al., 2015).

We present the currently available real public HSI datasets available to the scientific community with their respective details in Table 1. Notably, the existing datasets are much fewer compared to existing optical imagery datasets.

The taxonomy of our work is presented in Fig. 1. Methods of generating new training samples to supplement the available samples are highlighted in the branch labeled (a), methods designed to work with few or no training samples are presented in the branch (b), while branch (c) shows the current spectral-spatial based DCNNs advancing improved feature learning.

2. Generation of new training sample

Various cost-effective and more straightforward methods have been proposed to help generate new training samples to mitigate the cost and effort required for a real HSI benchmark dataset. In this section, we first highlight the issue of training samples insufficiency in HSI classification and later present intuitive and cost-effective methods used to create new training samples to supplement the insufficient training samples.

2.1. *Insufficient training samples in HSI classification in a nutshell*

Supervised classification methods use labeled samples to draw inferences from associations and relationships within the data using spectral, spatial, or combination of spectral-spatial features and later uses the inference to categorize unlabeled data during the testing process (Romero et al., 2016). Conversely, unsupervised methods classify input data without prior knowledge but rely on similarities and patterns within the data. Supervised DCNN models have been credited for achieving superior performance and are preferred over unsupervised models (Paoletti et al., 2017). However, they suffer limited labeled samples constraints and thus may not effectively learn the massive spatial-spectral parameters required for accurate HSI classification (Makantasis et al., 2015).

This bottleneck significantly inhibits their practical application in HSI data analysis. The construction of new HSI benchmark datasets to solve DL models' training needs is expensive, time-consuming, and presents a significant challenge (Mnih et al., 2015). Although several HSI datasets have been presented over a couple of years for an array of HSI applications, the currently labeled HSI samples are still far below the demands of the current DL models and for verifying the performance of evolving algorithms. Limited training samples cause model overfitting during training which significantly affects the model's performance.

Table 1
Existing real public HSI datasets.

Dataset Name	Sensor	Classes	Wavelength range	No of bands	Spatial resolution (mpp)
Salinas Valley (Graña et al., 2013)	AVIRIS	16	0.4–2.5 µm.	204	3.7
Indian Pines (Graña et al., 2013)	AVIRIS	16	0.4–2.5 µm.	224	20
Botswana (Graña et al., 2013)	NASA EO-1	14	0.400–2.500 µm	242	30
University of Pavia (Graña et al., 2013)	ROSIS	9	0.43–0.86 µm	103	1.3
Washington DC Mall (Zhu, 2017)	HYDICE	6	0.4–2.4 µm	210	–
San Diego Airport (Dong et al., 2018)	AVIRIS	7	0.37–2.51 µm	224	3.5
Jasper Ridge (Zhu, 2017)	AVIRIS	4	0.38–2.5 µm	224	–
Lunar Crater Volcanic Field (LCVF) (Dong et al., 2018)	AVIRIS		0.37–2.510 µm	224	20
HYDICE Urban (Dong et al., 2018)	HYDICE	6	0.4–2.5 µm	210	3
Baffin Bay (Han et al., 2019)	HYPERION SENSOR	3	0.35–2.58 µm	242	30
University of Houston (Xu et al., 2016)	CASI	31	0.38–1.05 µm	144	2.5
Kennedy Space Center (KSC) dataset (Zhou et al., 2019)	AVIRIS	12	400–2500 nm	176	18
Pavia City (Zhao et al., 2016)	ROSIS-03	4	0.43–0.86 µm	115	1.3
Umatilla County (Liu et al., 2017b)	Hyperion	7	0.35–2.58 µm	242	30
Chikusei (Yokoya and Iwasaki, 2016)	HYPERSPEC-VNIR	19	0.363 –1.018 µm	128	2.5
Samson (Zhu, 2017)	–	3	0.401–0.889 µm	156	–

Moreover, HSI images present a high burden on computational resources when processing the complex high dimensional data caused by the infinite number of features present in hyperspectral data (Alizadeh Moghaddam et al., 2020). Other challenges associated with HSI classification include the same class objects possessing different spectrums and different class objects possessing the same spectrum, redundant information caused by sensor calibrations, and noise further complicating class-specific features' separation.

In general, problems in HSI image classification can be summarized as follows: (1) Insufficient training samples. (2) Handling high dimensionality data. (3) Handling redundant information in HSI data. (4). High-class variance in HSI data.

2.2. Data augmentation

Data augmentation creates new training data from already existing sample data by applying transformations, thus generating realistic-like data through augmentations and geometric transformation such as shear, random rotations, shifts, flips on copies of existing images as a way of improving the model's robustness (Wong et al., 2016). This intuitive method is helpful in training DL networks and boosting network performance (Chen et al., 2018).

Data augmentation has widely and promisingly been applied in various studies (Chen et al., 2016; Xu et al., 2017) and achieved compelling results, proving to be a simple and yet effective way to create virtual samples. In hyperspectral data analysis, data augmentation has demonstrated notable improvements in DCNN generalization capabilities. Recently, Nalepa et al. (2020); Wang et al., 2019a proposed various data augmentation methods to increase the number of training samples in HSI classification as a cost-effective and simple way of handling labeled data insufficiency. Moreover, the albumentations library (Buslaev et al., 2020), provides fast and flexible image augmentations processes with many various image transform operations relating to color, contrast, brightness, and other geometric transformations and have been applied in image classification tasks (Liu et al., 2020a). Using data augmentation technique have substantively alleviated limited training samples challenge in HSI classification.

2.2.1. Transformation-based sample generation

Given that HSI imagery suffers complex and varied lighting conditions, objects belonging to the same class tend to vary in appearance at varying distances and perspectives due to different radiations and illuminations and possess similar spectral features in a given range (Aptoula et al., 2016). Using this basis, new virtual data can be generated by altering known samples. This data augmentation intuition has widely and promisingly been used in various studies (Lee and Kwon, 2017; Xu et al., 2017).

2.2.2. Mixture-based sample generation.

Naturally, objects belonging to the same class in hyperspectral images tend to possess analogous spectral features in a given range. Guided by this principle, virtual samples can be generated from two given samples of the same class by linearly combining the samples to generate virtual samples (Chen et al., 2016; Kang et al., 2018). These two methods have proved effective in improving the vitality of the model's performance in several tasks and have demonstrated that data augmentation is a promising technique for improving the generalization capabilities of DL networks. Nalepa et al., 2019a explored the power of data augmentation on hyperspectral data at inference time for deep networks and demonstrated improvement in the generalization capabilities of the DL networks.

2.3. Synthetic data generation

Gathering hundreds of thousands or millions of training samples required to train supervised networks can take a lot of man-hours to collect and label, thus making it impractical for many applications (Mnih et al., 2015). Although labeled data is difficult to obtain in the real world, it is easier to generate using simulation, making the dataset inherently less costly, faster to create, well-annotated, and not constrained by the availability of time or the physics of the natural world

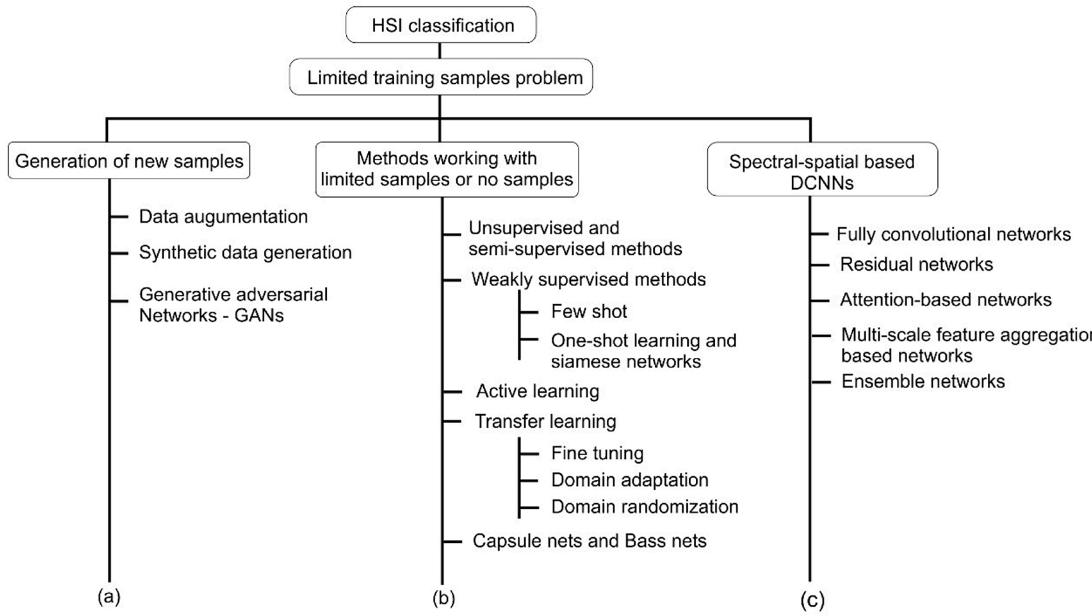

Fig. 1. A taxonomy of limited training samples and improved feature learning for HSI classification.

(Patki et al., 2016). Moreover, there is insignificant statistical evidence between the accuracy scores of experiments on synthesized data (Tremblay et al., 2018).

Unlike augmented data which is largely derived from real images with realistic transformation (such as translation, flipping, rotation, or the addition of noise) to increase the diversity of the training set, synthetic data is partly or completely artificially generated by a synthesizer network using generative adversarial networks (GANs) following methods such as compositing, styled transformation, and foreground and background augmentation, and have successfully been applied on images, text, audio and video data (Tripathi et al., 2019).

Over time, several DL models have used synthetic datasets to perform computer vision tasks such as object detection (Pepik et al., 2012) and scene understanding (Satkin et al., 2012). By effectively utilizing domain randomization (Tobin et al., 2017), the model interprets synthetic data as just part of the training dataset indistinguishable from physical information, thus exposing the model to a wide range of environments at training time (Peng et al., 2015). In this sense, synthetic datasets significantly improve the model's performance and, in some experiments, outperform photorealistic datasets (Davari et al., 2018; Patki et al., 2016). This has contributed to the improvements in computer vision and machine learning methods in semantic segmentation, object detection, object recognition, and image classification. Moreover, synthetic HSI data have been used to increase the training set through synthetically generated multispectral images (Kemker et al., 2018), HSI classification (Zhu et al., 2017), and anomaly detection (Zhao et al., 2017). Notably, synthetic datasets provide cheap and simple labeled data thus reducing the demand for massive training samples.

2.4. Generative adversarial networks - GANs

Generative adversarial networks (GANs) are an effective technique for generating new samples for training networks using the min–max strategy where one neural net successively generates fake samples from the original data. GANs consist of two parallel parts that are both parameterized as deep neural networks that can learn how to produce data from a dataset indistinguishable from the original data (Luo et al., 2019b). For each input image, a style image from a subset of different styles is selected, and a styled transformation of the original image is generated. Both original and styled images are fed to train the net. A generator G produces synthetic data given a noise variable input Z while a discriminator D identifies whether a sample is coming from the real data distribution Xr or the generated data distribution Xg as shown in Fig. 2. The discriminator D is trained to estimate the probability of a given sample coming from the real data distribution, whereas the generator G is optimized to "fool" the discriminator to offer a high probability for the generated data.

GANs have been reported to perform well even in instances with limited training samples and have shown excellent functionality in increasing the image resolution of input images (Marchesi, 2017). CycleGAN (Zhen et al., 2019), is powerful in style transfer from one image set to another. This variant of the generative adversarial network has been successfully used to convert visible images to generate synthetic InfraRed images training samples (Yun et al., 2019), and in combining visible band and infrared data to significantly increase the segmentation accuracy of RS datasets (Benjdira et al., 2019). In each of these works, real data were used to augment synthetic data for vision tasks during model training. Other variants of adversarial nets are Siamese-GAN (Huang and Chen, 2021), Sifting-GAN (Ma et al., 2019), and Attention-GAN, (Yu et al., 2020). Several other GAN-based frameworks have been proposed to alleviate the shortage of training samples through adversarial training for HSI classification (Zhong et al., 2020). GANs help to generate additional training data and have proven an intuitive and inexpensive method currently explored by many researchers in vision-based tasks.

3. Methods proposed to deal with limited samples or no samples

This section presents some intuitive methods proposed to train neural networks and produce competitive results despite the limitation of training samples, especially for HSI image classification.

3.1. Unsupervised methods and semi-supervised

Unlike supervised classification methods that use vast training labels to train the models, unsupervised methods learn relationships and associations from data directly without labels and estimate the class labels of the unlabeled samples (Zhou and Prasad, 2020). Using this idea, semi-supervised methods harness the benefits of both supervised and unsupervised methods. Ideally, in semi-supervised classification, the pre-

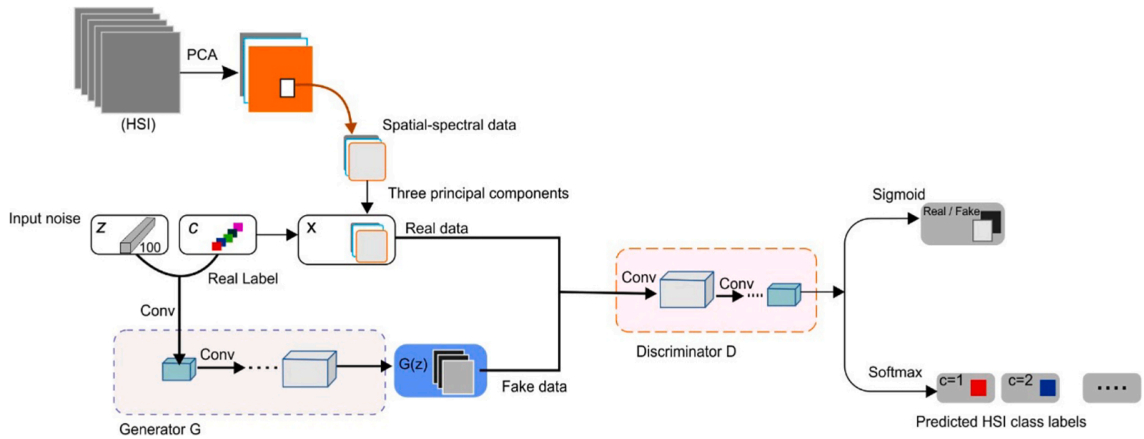

Fig. 2. An illustration of the GAN-based hyperspectral imagery (HSI) classification (Lin et al., 2018).

training is done using unlabeled data, while fine-tuning uses small labeled data sets. Semi-supervised methods notably offer better performance than unsupervised methods, are less costly in terms of labeling costs on data, and have been explored in various HSI classification tasks (Fang et al., 2018;).

Several fully connected network architectures have been proposed using a semi-supervised approach to classify HSI imagery (Liu et al., 2017a; Ma et al., 2016). Wu and Prasad (2018a), employed a non-parametric Bayesian clustering algorithm to generate pseudo labels that help in pre-training Convolution Recurrent Neural Networks (CRNN) for the HSI classification task. Mou et al. (2018) proposed an end-to-end unsupervised feature learning using convolution and deconvolution networks in place of encoder and decoder, while Zhan et al. (2018) employed GAN to design a semi-supervised feature learning framework for HSI classification, where the generator created counterfeit hyperspectral sample images that were similar to the real data to train the GAN.

This shows that semi-supervised learning mechanisms have notable capabilities in extracting good spectral and spatial features suitable for satisfactory results. Current semi-supervised feature learning assumes the encoder-decoder concept without using labels data.

3.2. *Weakly supervised methods*

Deep learning (DL) frameworks demand massive labeled samples to train and test network models accurately. However, the required labeled samples are unavailable for various applications due to the annotation costs and time required. Moreover, training DL models with few labeled samples results in overfitting, thus poor generalization performance on unseen data.

To overcome this, weakly supervised networks are trained with partially annotated images that are based on simple annotations. These annotations are easier to generate and less demanding as opposed to fully labeled images. This alleviates the burden of obtaining hand-labeled datasets, which can be costly or impractical (Torresani, 2014). Usually, weakly-supervised methods handle the limited ground truth annotation for HSI data using various approaches and for applications such as; annotating small sets of regions of interest for HSI classification (Signoroni et al., 2019), use of region growing method on the annotated areas (the annotates regions are referred as seeds) with pixel-level supervision for semantic segmentation (Moliner et al., 2020), and hyperspectral unmixing (Hong et al., 2019). In the subsequent subsections, we briefly discuss weakly supervised-based methods, namely few-shot and one-shot, and how they address the deficiency of training samples.

3.2.1. *One-shot learning and siamese network*

One-shot learning attempts to solve the limited training samples by classifying images given only a single training example for each category, using fast nearest-neighbor algorithms for efficient memory usage (Kaiser et al., 2017). The initial concept was inspired by the visual ability of human beings to learn a lot of information from just a single category or few images. With great success, one-shot learning has extensively been explored in facial recognition systems (Lake et al., 2013). The one-shot learning concept helps to alleviate the challenge of massive training samples required to train DL models. Other variants are zero-shot learning, where the model does not learn from any examples from the target class; and k-Shot learning, where the model observes k-examples from the target class during training (Palatucci et al., 2009).

Siamese network (Bromley et al., 1993), follows the one-shot learning concept where input pairs are fed to a pair of identical parallel networks that share the same configurations and parameters. The network is trained with a pairwise loss that minimizes the distance between image patches of the same class and maximizes the distance between image patches of different classes. By so doing, the network learns similarity scores on the input pairs instead of how neural networks learn to classify images for specific output classes (Koch et al., 2015). In this sense, few class instances are enough to train a network model and show competitive results. Siamese networks have been applied in HSI image analysis to extract non-linear, highly complex spectral-spatial features from the limited images even when class variance is wide (Liu et al., 2018a). Siamese networks learn hidden representations and semantic similarities from limited training samples by combining inputs into a single network. Other works have used the Siamese network in HSI classification (Cao et al., 2020), change detection (Tang et al., 2021a), object tracking (Abdelpakey et al., 2018), among other tasks.

3.2.2. *Few-shot learning*

Few-shot learning is a self-supervision paradigm that seeks to train models with little or no labeled data by optimizing models to recognize patterns within the data labels. This ensures that the resultant models can efficiently learn and recognize a set of classes even under limited training samples (Doersch et al., 2015). This transfer learning paradigm is used to overcome overfitting caused by data scarcity and is usually achieved in two steps. In the first step, the model is trained using base classes associated with a large set of training annotations such that the trained model gains visual analysis abilities in the form of learned representation. In the second stage, the model learns from a new set of classes using only a few samples from each class. Some variants of few-shot learning approaches that have been proposed include; gradient descent-based approaches (Ravi and Larochelle, 2017) which learn recognition tasks using less gradient descent iterations; metric-based approaches (Sung et al., 2018), which uses distance metrics between a test image and a set of training images in a few-shot task; among others. Few-shot learning has been explored in the HSI classification task to train a classifier from a given source domain and applied directly to a target domain with limited training labels without additional ground truth labels and in other instances, class labels guided in the generation of more samples (Qu et al., 2019). Lately, other works have explored few-shot learning to overcome domain shift in HSI classification (Li et al., 2021); explored relationships between samples using attention weighed graphs (Tong et al., 2020b); and handled unknown classes in HSI landcover classification (Liu et al., 2021).

3.3. Active learning

Active learning (AL) helps perform optimal selection instance sampling on a given target object to extract information using a given selection criterion while highlighting samples possessing the highest representation and low redundancy (Samat et al., 2015). Instead of generating random labels, the most informative instances are considered and selected for annotation during the learning process (Yanik and Sezgin, 2019). Given that HSI images contain high abstract and semantic features with high interclass similarities and low intra-class variance, accurate identification of the high-level semantic representations can be challenging even to a human-labeler. AL uses concepts like object-level learning to statistically learn key semantics of the labeled objects within the images to improve and simplify the classification task (Li and Guo, 2014). Besides, AL aims at selecting optimal discriminative features during the domain-expert annotation process using minimal labeled data. Recently AL has been explored in HSI image classification to improve classification performance and reduce labeling overheads (Mu et al., 2020b). A comprehensive survey on active learning algorithms for supervised spatial-spectral image classification is presented (Tuia et al., 2011) with comprehensive details.

3.4. Transfer learning

Training DL models from scratch can be time-consuming and require massive training samples. Transfer learning (TL) uses techniques to provide DL models with experiences in the form of training data to make inferences and later use for target tasks (Yosinski et al., 2014).

TL is mainly applied in tasks with little data to train a full-scale model from scratch. After the parameters transfer, the new network can use the learned features for classification tasks using supervised or unsupervised methods. To handle new tasks, the parameters of the top layers are randomly initialized. Usually, during pre-training, the network uses features extracted from the network to perform classification, while in fine-tuning, the network is tweaked with a small number of training samples for the target task. A network pre-trained on source data possesses knowledge on low-level features that can be transferred for other unseen tasks. Besides, since low-level features can best fit for most generic DL classification problems, pre-training helps the network learn certain generic features from the lower layers, such as color blobs, edges, as well as other low-level features that can be generalized for other image classification tasks (Castelluccio et al., 2015). Fig. 3. illustrates how features learned from low and middle network layers are transferred to another network with the same architecture as the learned one.

TL has successfully been applied in object recognition, image classification, scene recognition, fine-grained recognition, attribute detection, and image retrieval and has proved to be a great baseline concept for image analysis tasks. Using TL significantly reduces the demand for training samples, and the model's learning process is greatly accelerated, thus decreasing the model's training time compared to training a model from scratch. Moreover, TL improves network performance by transferring knowledge obtained from the source domain to the target domain (Deng et al., 2019) and has successfully been explored in RS applications such as image classification (Sumbul et al., 2018; Zhou and Prasad, 2018), crop type mapping (Nowakowski et al., 2021), tree species classification (Briechle et al., 2021), HSI classification (Liu et al., 2020b), to overcome the limitation of training samples.

However, due to many varying factors such as pose, illumination, image quality, and spectral disparities, there is always some impending domain variance between two application areas, causing degrading of the performance (Wang and Deng, 2018). This has influenced many researchers to work on new domain adaptation and randomization concepts to address domain variance in TL (Othman et al., 2017). These methods are discussed in the later sections.

Fine-tuning is an approach in transfer learning that involves tweaking the network with a small number of training samples for the target task. To achieve this, some initial network layers are frozen, and few top layers are adapted to learn features of the target to produce accurate predictions using an optimization algorithm task (Long et al., 2015b). In so doing, DL models can apply learned expertise to deliver accurate predictions for unseen data for various target tasks (Lundervold and Lundervold, 2019). However, training the network from scratch is recommended for situations where a wide domain shift exists between the original dataset and the target dataset while the network is initialized using pre-trained weights.

3.4.1. Domain adaptation

Domain adaptation (DA) is a transfer learning paradigm that exploits labeled information in one relevant application to execute new tasks in another application domain (Kouw and Loog, 2019). The goal of domain adaptation is to minimize the domain gap (where data from the source and target domains may either be considered homogeneous or

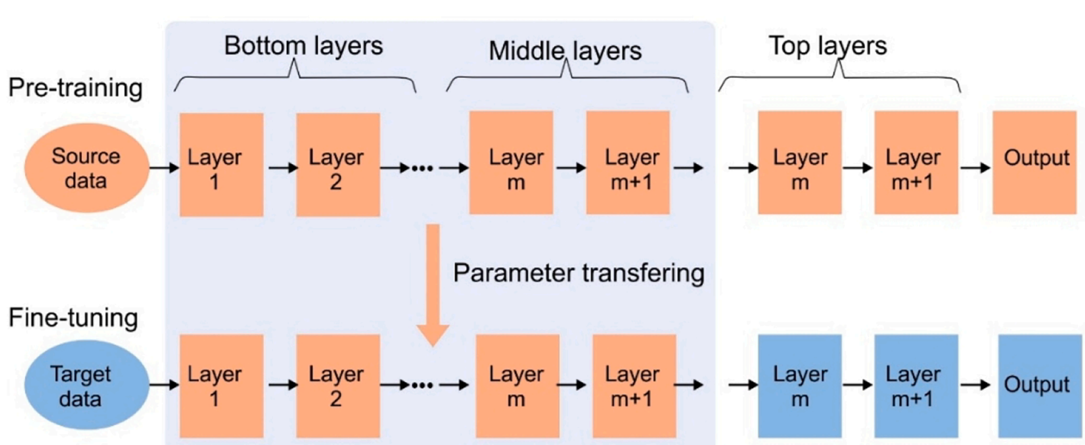

Fig. 3. Illustration of transfer learning (Pan and Yang, 2009).

heterogeneous) and to transfer knowledge in cases where a wide domain gap exists; while attempting to reduce dataset bias caused by the difference in the statistical distributions between training and test domains (Hoffman et al., 2014). In various cases, transfer learning may suffer "domain shift". A change in data distribution causes this condition as a result of several factors such as platform inconsistencies due to diverse sensor settings, image acquisition platforms, image calibration settings, and device mechanical configuration variations resulting in a situation referred to as "data shift problem" (Othman et al., 2017). Consequently, models trained on one platform may fail to perform well in another platform thus requiring an adaptive system that seeks to bridge the difference between two environments (Tuia et al., 2016).

A typical example is training a model on SAR imagery and testing it on an optical dataset. The two datasets fall on different domains, and thus a model that has been trained on one platform may not generalize well in another target task. DA in RS imagery on Hyperion, National Center for Airborne Laser Mapping (NCALM), and WorldView-2 datasets demonstrates the power of DA for HSI imagery (Song and Ma, 2017) and HSI image classification (Fan et al., 2019b) using the unsupervised approach on how it effectively addresses the issue of limited or unlabeled datasets. To solve this, vision-based models have been proposed to investigate fitting specific-domains models to target domains (Long et al., 2015b); re-training models in the target environment (Yosinski et al., 2014); pre-trained models adaptation (Li et al., 2016); using pre-trained weights for feature extraction (Gupta et al., 2016); and learning similar features between domains (Tzeng et al., 2014). DA can improve classification performance in situations where training data is scarce.

3.4.2. Domain randomization

The concept of domain randomization (Tobin et al., 2017) helps the model to interpret unreal (or synthetic) data as part of the training dataset indistinguishable from physical information, thus exposing the model to a wide range of environments at training time (Peng et al., 2015). Despite the significant contribution of synthetic datasets, some studies observe that they may suffer from a reality gap – a situation caused by quality differences between original data and generated synthetic data due to the natural richness, non-rigidity, and noise evident in real-world data and thus, models trained purely on synthetic data may not generalize well on real-world (Bousmalis et al., 2018). Moreover, synthetic datasets require validation against real-world data which complicates the processing tasks. Other challenges that synthetic datasets suffer include; lack of acceptance from some users, dependency on the quality of the data model, and the difficulty in keeping track of all necessary features required during replication. Besides, real HSI RS datasets possess high-class variance and redundant data, which presents a domain gap between synthetically generated HSI RS data and real HSI RS data (Rajpura et al., 2017). Several researchers have proposed domain randomization as one of the approaches to bridge the reality gap (Tobin et al., 2017), an active research area that can enable synthetic data to work well in practical and real-life applications.

3.5. Capsule networks and bass adaptive spectral-spatial networks

Capsule networks use a group of neurons (referred to as a capsule) to encode spatial information and the probability of an object being present in an image being identified. Specifically, capsule networks focus on position, rotation, shape, and scale details in a high dimension space (Sabour et al., 2017). Capsule networks were proposed to address the drawbacks caused by pooling layers in convolutional neural networks (CNNs) that lead to the loss of essential spatial information. Besides, general CNN-based network architecture has difficulties in accurately exploiting and learning complex relationships in high dimensional HSI data directly, which significantly affects their performance (Paoletti et al., 2019a). To deal with this, capsule networks use spectral-spatial capsules to learn and represent spatial and positional details, spectral information, and other essential transformational details regarding a scene or image's spatial features and positional orientation. The ability of capsule networks to learn and represent the complex relationship of features from an image or parts of an object have seen it applied in image classification tasks (Deng et al., 2018), object tracking (Abdelpakey et al., 2018), and generation of synthetic training samples (Jaiswal et al., 2018). Moreover, since RS possesses hidden representation and complex features, capsule networks have been explored in HSI classification (Ding et al., 2021; Xue, 2020) to tackle the limited training samples.

Additionally, HSI data contains high dimensional data, and some part of it is considered redundant. Processing the high dimensional data using machine learning models poses a great challenge due to the curse of dimensionality. Some former methods such as local Fisher's Discriminant Analysis and Principal Component Analysis have been proposed to deal with dimensionality reduction. However, these former methods result in the loss of some essential information, thus affecting classification performance. Given that not all spectral-band information from the HSI data is used during training, and since some spectral band information is application-specific, band-adaptive spectral-spatial (BASS) networks (Santara et al., 2017)has been proposed to address the redundant information and high dimensionality issue by extracting band-specific spectral-spatial features to train its network, resulting to lesser network training time and better model performance. Besides, since BASS network architecture has fewer independent connection weights, it performs well with few training samples.

3.6. Summary

The challenge in HSI classification based on the lack of sufficient training samples draws significant interest in the hyperspectral community. Many researchers have proposed methods to generate new samples from existing data or artificially to intuitively use the limited samples to train the models, while others have focused on developing network methods that can efficiently utilize the limited labeled samples. Limited training samples problem in HSI classification can be summarized as follows:

- Data augmentation and synthetic data generation through GANs and synthesizers are robust and cost-effective methods to increase training samples.
- Transfer learning can mitigate the demand for training samples by transferring learned weights from source data to initialize network weights, thus reducing training time and improving accuracy. Moreover, low-level features learned from source data can reduce training time compared to training the network from scratch.
- Weakly supervised methods utilize useful features learned from unlabeled data and using minimal labeled data to tune the classification network models, while active learning intuitively learns from the domain-expert annotation process to optimize classification using minimal training data.
- Data augmentation and transfer learning can significantly improve HSI classification accuracies even with limited training samples (Li et al., 2019).
- There is a need to develop new, larger, and more complex (e.g., with more classes and varied spectral range) reference hyperspectral datasets for comparative research in different HSI application domains.
- Since deep generative models can now synthesize hyperspectral pixels from scratch, more experiments on these data samples are necessary to validate their performance in classification and varied HSI tasks.

4. Spectral-spatial based DCNNs

DCNNs have been credited for their superiority in handling image classification problems and have gained popularity for their ability to learn discriminately, extract and represent hidden, complex, and non-linear features from raw images (Wambugu et al., 2021). Based on this fact, DCNNs have been applied in HSI image classification in the RS field. Over the last few years, intense research has proposed a plethora of networks aimed at capturing more complex features and descriptors necessary to capture non-linear and abstract semantics of HSI data for understanding using rich spatial and spectral features for an array of applications.

To achieve this, DCNNs architectures either use spectral, spatial, or spectral-spatial features in HSI data representation. This section first highlights joint spectral-spatial feature learning and later discusses the current DCNNs using joint spectral-spatial features for robust HSI feature representation. Specifically, the networks have been categorized based on the intuition of how they learn features based on the network architecture.

4.1. Joint spectral-spatial feature networks overview

The Hyperspectral data can be analyzed from either spatial, spectral, or joint spatial-spectral perspective. Most of the early DL methods only exploit data pixel-wise (1-dimensional approaches), working in the spectral direction by extracting spectral signatures from single pixels or groups of pixels either surrounding a central pixel or belonging to a given point of interest. This approach requires some prior knowledge and a pre-processing step to detect and map the regions of interest (usually done through segmentation). On the other hand, spatial-feature networks focus on extracting spatial features of the HSI data and later fuse the extracted features with spectral features extracted through other techniques (Jon Atli and Pedram, 2015). Different from others, spectral-spatial classifiers integrate both spectral and rich spatial features to boost classification performance (Fauvel et al., 2007). Instead of extracting the spectral and spatial features separately and later processing them together, the joint spectral-spatial 2-D CNNs extracts features from the original data directly.

4.2. Fully convolutional based networks

The success of deep learning in image classification traces many years back. Great advances can be related to contributions made by the full convolution network (FCN) (Long et al., 2015a), which allowed feature extraction through end-to-end training without the fixed size constrain of fully connected layers using convolution and deconvolution layers. In FCN, all output nodes are linked to preceding layers (regional input nodes) and used multiple convolutional layers followed by down-sampling layers that help the network achieve large receptive field coverage. However, down-sampling through pooling layers reduces the spatial dimensions of feature maps, leading to the loss of essential positional information of objects and greatly affecting image classification results. Since HSI data is complex and rich in multidimensional spectral-spatial features, it is difficult to process it like other generic images.

Most 2D and 3D CNN-based methods perform image classification by extracting robust and deep spectral-spatial features from raw HSI input images. The learning process can be achieved through feedforward and backpropagation processes simultaneously. In so doing, complex spectral-spatial features are extracted from HSI data. Several FCN-based frameworks have recently been proposed for RS analysis. For example, Zhang et al. (2019b) developed a fully convolutional network for remote sensing scene classification based on DenseNet, Zhao and Du (2016) extracted spatial features from the first three components bands of HSI data using a 2-D CNN and then combined the spatial information with spectral features for HSI classification, while Chen et al. (2016) designed big networks with strong constraints which can utilize virtual samples to improve classification results using 3D-CNN models to address the HSI feature extraction and classification problem with limited training samples. Moreover, FCNs have successfully been used for HSI classification (Hang et al., 2019; Li et al., 2019), hyperspectral image analysis (Jiao et al., 2017), and HSI scene parsing (Wang et al., 2021b) and have proven powerful in extracting useful discriminative features from hyperspectral data.

4.3. Residual based networks

He et al. (2016) introduced residual connections allowing the design of very deep networks that do not suffer gradient degradation problems using residual building blocks. Residual paths (also called skip connections) link up low-level and high-level layers to ensure efficient gradient flow. Densely connected network (DenseNets) (Huang et al., 2017) extends the ResNet concept by proposing dense skip connections from previous layers to immediate layers. By using dense blocks, low-level, mid-level, and high-level features are concatenated, ensuring efficient information sharing and gradient flow between all layers. Both ResNet and DenseNets have proved efficient in handling redundant information in spatial-spectral data (Yang et al., 2018), and have been explored in image classification tasks using super-resolution data (Wang et al., 2017), and HSI data (Kang et al., 2019; Paoletti et al., 2019b). Spectral and spatial residual blocks have been used to discriminatively learn and extract rich spectral-spatial information in hyperspectral imagery where residual blocks connect 3-D convolutional layers through identity mapping to facilitate better backpropagation of gradients (Zhong et al., 2018). Residual networks have become the predominant feature extraction architecture for many HSI data analysis architectures (Cao and Guo, 2020b; Zhong et al., 2017b).

4.4. Attention-based networks

The attention mechanism (Vaswani et al., 2017), inspired by the human cortex concept, aims at learning a weight map that represents the relative importance of activations within a layer or a channel. Since HSI data contains highly redundant information which creates bottlenecks in HSI classification, the attention mechanism helps the model to focus selectively on discriminative channels and ignore redundant information (Li et al., 2020). Moreover, attention mechanisms have shown improvements in the network's capability in discriminating the essential spatial and spectral channels and achieves better dimensionality reduction on HSI data than traditional pre-processing methods such as principal component analysis (PCA) without loss of essential information (Luo et al., 2019a). Attention mechanism intuition has been used in guided feature extraction on spatial-spectral data (Mei et al., 2019) using dual-branch spectral and spatial attention mechanisms implemented by combining CNNs and ResNet. Dong et al. (2019) used a band-attention mechanism called attention-GAN to improve feature learning in HSI data, while Hang et al. (2021) explored joint feature classification using two spectral and spatial subnets. In addition, attention mechanism has been explored in change detection (Wang et al., 2021a) using remotely sensed data and HSI image classification (Hang et al., 2021; Tang et al., 2021b) and have greatly improved feature exploitation and effective use of spectral-spatial information.

4.5. Multi-scale based networks

In multi-scale feature fusion, features are extracted at different spatial and spectral dimensions using convolution kernels of different sizes in a multi-layered or multi-branch structured network. Since HSIs exhibit a complex hierarchical distribution of spectral-spatial features, convolution kernels of the classical CNNs with a fixed size cannot effectively handle HSI feature extraction requirements. The outputs of the preceding layers are fed into the input of the successive layers, thus fusing sufficient spectral and spatial information extracted at different scales (Feilong and Wenhui, 2019). In this case, extraction of hierarchical features from different spatial dimensions coupled by skip connection between the layers alleviates vanishing gradient and guarantees feature reuse (Meng et al., 2019). Besides, fusing spectral-spatial features from multiple scales from all convolutional layers can extract more discriminative features. In addition, multiscale feature upsampling blocks increase the size of combined feature maps with different resolutions to utilize the information from different sizes and locations (Liu et al., 2018b). This concept has been explored in hyperspectral image classification (Cao et al., 2018), spectral fusion, and scene classification (Mu et al., 2020a; Zhang et al., 2020), using HSI data and has effectively improved feature extraction robustness and registered promising results.

4.6. Multi-level feature aggregation networks

The multi-level fusion concept is inspired by the understanding that shallow layers are sensitive to low-level features while deeper features can capture high-level semantics (Zeiler and Fergus, 2014). To efficiently strengthen feature propagation and improve the accuracy of downstream tasks, extracting and fusing multi-layer features from different CNN layers is necessary, rather than stacking the layers to form deeper complex networks with large-scale parameters that are harder to train and easier for overfitting (Chen et al., 2021). Moreover, fusing features from different levels (multi-branch feature fusion) enables the network to learn more discriminative information for HSI classification, which extensively requires low-level, mid-level, and high-level features (Jiao et al., 2017). Multi-branch fusion is notably an efficient method for obtaining finer features and combining features from different layers in feature extraction of HSI classification (Fang et al., 2019; Ge et al., 2020) and significantly reduces the network complexity and computational cost in complex and deep networks. This method has been proposed for HSI classification (Sun et al., 2020). Shen et al. (2019) extracted spectral and spatial information from hyperspectral data using a two-branch CNN network where compelling results were achieved. For example, (Li et al., 2020) developed a 2-D CNN to capture and fuse spatial-spectral features based on squeeze and excitation network, while Zhou et al. (2019) exploited spatial and spectral long-short term memory (LSTM) coupled with decision fusion mechanism to capture rich representation from HSI data. Multi-level and multi-path fusion have been combined with attention mechanisms to carry out change detection (Wang et al., 2021a) where a deep supervision network with different branches is used to reconstruct the change map producing superior results.

4.7. Ensemble networks

To better optimize feature learning, ensemble models with different structures have been explored to combine the benefits of different architectures and harness their strengths. Using this intuition, several methods have been combined and have made significant improvements in HSI classification. Siamese-GAN (Bashmal et al., 2018) combines Siamese network and adversarial network for aerial image classification. Wang et al. (2019b) combined GAN and attention mechanisms to approximate real HSI images' distribution using collaborative learning, while (Li et al., 2020) combined two different CNNs with cascaded attention mechanisms to extract HSI discriminate features. In other related works, Chen et al. (2017) and Minetto et al. (2018) combined ResNet and DenseNet architectures to form an ensemble CNN network for image classification, while Chen et al. (2019) performed transfer learning using an ensemble of ResNet and CNNs. Whereas some ensemble networks can work with a small training set, complex ensemble networks have demonstrated poor performance in HSI classification due to small training data (Roy et al., 2020). Recently ensemble networks have also been explored for HSI classification (Chen et al., 2019; Zheng et al., 2020). Its worth noting that merging networks to form complex ensemble networks may not guarantee superior performance, and thus the need to perform models' combination evaluation. Performance comparison on deep ensemble methods (snapshot and model combination) is available (Dede et al., 2019).

4.8. Summary

The challenge of dealing with abstract, complex, multidimensional, high-resolution spectral-spatial data, coupled with the limited availability of training samples, continues to draw attention to the hyperspectral community. More effort is dedicated to developing superior and robust DCNNs that can adequately learn both spectral and spatial representation to meet hyperspectral-data analysis needs. Several observations can be drawn from the discussed methods:

- Feature reuse methods have significantly enabled the design of deeper networks that can learn complex and non-linear relationships from the complex HSI data through residual connections.
- Attention-based networks address the issue of handled redundant data, which offers great challenges in processing HSI data by guiding extraction of relevant spectral-spatial information and has gained increased attention.
- Multi-branch / multi-layered methods learn abstract and high-level spectral-spatial representation using different branches and later fusing the features. Moreover, branching helps to scale the network's depth, thus reducing the computation burden associated with very deep networks.
- Most HSI data have a low spatial resolution making deep learning techniques designed for computer vision perform poorly since the spectral dimension prevails over the spatial neighborhood features in most cases. More approaches are required to improve the representation of spatial structure information. (Such as superpixel Correlation Coefficient)
- More experiments on ensemble methods (such as combining band adaptive spectral-spatial networks with existing DCNNs) need further exploration.

5. Performance of various methods

In this section, we present the performance of various methods on the 3 most popular HSI datasets, namely: Salinas Valley (SV), University of Pavia (UP), and Indian Pines (IP) as reported in the source publications. The three datasets can be obtained from the website (http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes).

5.1. Evaluation metrics

Various evaluation metrics are used to determine the quality of the classification results. Precision, Recall, mean intersection over union (mIoU), confusion metrics, pixel accuracy (PA), Overall accuracy (AO), Average Accuracy (AA), and Kappa Coefficiency (Kappa) are the most preferred performance metrics. (The highlighted metrics could not be discussed in detail for space considerations). More details on these evaluation metrics used can be accessed in (Lv and Wang, 2020). We use OA and Kappa as the performance evaluation metric in our report.

5.2. Performance of various methods on the 3 most popular datasets

The performance of various network architectures based on the discussed categories is presented in Table 2. The reported performance is ordered based on the year the methods were proposed.

Since HSI feature learning requires relatively deep networks to learn complex spectral information, integrating residual-based networks to other DCNNs can be powerful in mitigating gradient degradation problems that result from deeper networks. 3D-2D SSHDR, a residual learning-based network, attained OA 99.46% and Kappa of 99.38% on IP dataset and OA of 99.81%, and Kappa of 99.74% in UP datasets, respectively. Fully convolutional spatial propagation network (FCSPN) attained an OA of 99.63% and 99.61% on SV and IP datasets respectively. The results show that 3D-FCN can effectively capture spatial

Table 2
OA and Kappa (%) performance of various methods on the SV, IP, and UP datasets.

Method		Year	Datasets					
			SV		IP		UP	
			OA	*Kappa × 100*	*OA*	*Kappa × 100*	*OA*	*Kappa × 100*
Fully Convolutional Networks	CNN (Lee and Kwon, 2017)	2017	95.42	–	94.24	–	96.73	–
	Spectral–spatial 3-D fully convolutional network (SS3FCN) (Zou et al., 2020)	2020	81.32	–	71.47	–	79.89	–
	Fully convolutional spatial propagation network (FCSPN) (Jiang et al., 2021)	2021	99.63	99.78	99.61	99.56	–	–
Residual based Networks	Dual-Channel Densenet (Yang et al., 2018)	2018	98.89	–	98.28	–	–	–
	Spatial Residual Network (SSRN) (Zhong et al., 2018)	2018	–	–	99.19	99.07	99.79	99.72
	3D-2D SSHDR (Cao and Guo, 2020b)	2020	–	–	99.46	99.38	99.81	99.74
Multi-scale based fusion	Fully dense multiscale fusion network (FDMFN) (Meng et al., 2019)	2019	96.72	96.25	–	–	–	–
	Multiscale Spectral-Spatial Unified Networks (MSSN) (Wu et al., 2019)	2019	90.89	89.11	–	–	89.52	86.62
	Dual-scale crossover network (DSCN) (Cao and Guo, 2020a)	2020	–	–	99.62	99.57	99.84	99.78
Multi-branch based network	Deep multilayer fusion dense network. (MFDN) (Li et al., 2020)	2020	–	–	96.08	95.26	98.89	98.1
	2D–3D CNN (Ge et al., 2020)	2020	99.94	99.93	96.07	95.51	99.52	99.41
	Features adaptive fusion network (FAFNet) (Sun et al., 2020)	2020	–	–	99.24	–	99.54	–
Attention-based networks	Spectral-Spatial Attention Networks (SSAN) (Mei et al., 2019)	2019	–	–	99.67	98.37	99.24	98.17
	Double branch dual attention (DBDA) (Li et al., 2020)	2020	97.51	97.23	95.38	94.74	96.00	94.67
	Center attention module (CAM) (Hang et al., 2021)	2021	98.18	97.97	98.10	97.84	98.97	98.64
Bass, Siamese, and Capsule Networks	BassNet (Santara et al., 2017)	2017	95.36	94.80	96.77	96.12	97.48	96.62
	Siamese-CNN (Liu et al., 2018a)	2018			99.04	98.87	99.68	99.55
	Conv-Capsule (Zhu et al., 2019)	2019	99.17	99.07	–	–	–	–
Ensemble Networks	Deep CNN Ensemble (Chen et al., 2019)	2019	96.05	95.97	92.54	91.36	93.19	91.26
	HybridSN (Roy et al., 2020)	2019	99.98	99.98	98.39	98.16	99.72	99.64
	Stacked sparse autoencoder (SSAE) (Deng et al., 2019)	2019	99.26	99.18	–	–	99.85	99.79
Active Learning methods	Fast Patch-Free Global Learning Framework (Zheng et al., 2020)	2020	99.92	99.91	–	–	99.81	99.74
	HT-CNN-Attention (He et al., 2020)	2020	94.70	93.62	90.86	89.05	94.25	92.36
	Spectral-Spatial Feature Fusion using Spatial Coordinates (SSFFSC-AL) (Mu et al., 2020b)	2020	–	–	100	100	98.43	97.90

information from raw HSI data and refine spatial features using an attention mechanism.

As observed from Table 2, SSFFSC-AL and SSAE networks based on active learning and active transfer learning attained OA performance of 100% and 99.85% on the IP dataset and UP dataset, respectively. This demonstrates that the IP dataset may not be complex enough to test the full robustness and efficiency of the network. Siamese network that uses limited training samples to train the network attained an OA of 99.04% and 99.68% on the IP and UP datasets, respectively. Since HSI data contains immense redundant information, some objects on the HSI images do not belong to any class membership and are unnecessary during the classification process. Attention mechanisms and gated networks can be used to filter redundant information and help in projecting salient and discriminative spectral channels and spatial features required for HSI data classification. This can improve the convergence time and the classification accuracy scores while solving the redundancy data issue. Spectral-Spatial Attention Networks (SSAN) belonging to the category of guided networks attained an OA of 99.67% and 99.24% on IP and UP datasets, respectively. Through the use of the attention mechanism, the feature learning process has been optimized, and more critical information is obtained while discarding irrelevant information.

Multi-level-based networks optimize the feature learning process by striking a balance between network depth and width. While deeper networks learn more complex features, efficient HSI classification requires leverage between low-level, mid-level, and high-level features. This method significantly reduces the computation burden of deeper networks and still yields acceptable performance. 2D–3D multi-level-based CNN attained a competitive OA score of 99.94%, and 99.52%, and Kappa of 99.93%, and 99.41% in the SV and UP dataset, respectively. Besides, by fusing features extracted at different scales using different kernel sizes, multi-scale methods can extract and fuse sufficient spectral-spatial information necessary for HSI classification. Using this method, Dual-Scale Crossover Network (DSCN) achieved an OA score of 99.62% and 99.57%; and Kappa of 99.84%, and 99.78%; on IP and UP datasets, respectively. Multi-scale networks improve feature learning robustness by utilizing features obtained from different scales.

Additionally, given that spectral-spatial features from HSI data are complex, highly non-linear, and greatly challenging, combining different deep learning networks to form ensemble networks can be effective in exploiting the salient spectral features from HSI data as observed in the results. HybridSN network based on the ensemble method posted an OA of 99.98% and Kappa of 99.98% on the SV dataset. In this case, ensemble networks seek to harness performance improvements from different architectures.

Notably, since the experimental setup and environment for various methods are dissimilar, comparing the performance of various methods based on results posted in the source papers is impossible. Our work reported the performance of various methods without comparing them against each other. Future research can compare and evaluate various discussed methods under similar experimental conditions to provide precise performance comparison for various methods.

6. Conclusion

HSI image classification continues to draw interest in the blooming field of hyperspectral remote sensing due to its accrued benefits in many applications. Labeled data insufficiency has been reported as a significant bottleneck in training superior supervised DCNNs and has hampered the practical application of supervised DCNNs.

This review reported various methods of generating new training samples, discussed methods working with limited training samples, and highlighted the current methods advancing feature learning in DCNNs using joint spectral-spatial features. Additionally, we have reported the performance of some select methods on three popular HSI datasets and shared intuitive summaries. While researchers continue to dedicate more work to develop robust architectures for HSI applications, the lack of largescale HSI benchmark datasets with larger classes and wider spectral coverage remains a major bottleneck. Current methods have attained very high classification accuracy on popular datasets, making the real comparison of new and superior approaches almost impossible. The promising path presented by new samples creation through artificial synthesis can be exploited further to support new methods and unlock the full potential of deep learning in the hyperspectral domain.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 41871380 and U1605254, and the Natural Sciences and Engineering Research Council of Canada under Grant 50503-10284.

References

Abdelpakey, M., Shehata, M., Mohamed, M., 2018. DensSiam: end-to-end densely-siamese network with self-attention model for object tracking. Springer 11241, 463–473. https://doi.org/10.1007/978-3-030-03801-4_41.

Alizadeh Moghaddam, S.H., Mokhtarzade, M., Beirami, B.A., 2020. A feature extraction method based on spectral segmentation and integration of hyperspectral images. Int. J. Appl. Earth Obs. Geoinf. 89, 102097. https://doi.org/10.1016/j.jag.2020.102097.

Aptoula, E., Ozdemir, M.C., Yanikoglu, B., 2016. Deep learning with attribute profiles for hyperspectral image classification. IEEE Geosci. Remote Sens. Lett. 13 (12), 1970–1974. https://doi.org/10.1109/LGRS.2016.2619354.

Audebert, N., Le Saux, B., Lefevre, S., 2019. Deep learning for classification of hyperspectral data: A comparative review. IEEE Geosci. Remote Sens. Mag. 7 (2), 159–173. https://doi.org/10.1109/MGRS.624551810.1109/MGRS.2019.2912563.

Bashmal, L., Bazi, Y., Alhichri, H.S., Alrahhal, M.M., Ammour, N., Alajlan, N., 2018. Siamese-GAN: Learning invariant representations for aerial vehicle image categorization. Remote Sens. 10, 351. https://doi.org/10.3390/rs10020351.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Benjdira, B., Bazi, Y., Koubaa, A., Ouni, K., 2019. Unsupervised domain adaptation using generative adversarial networks for semantic segmentation of aerial images. Remote Sens. 11, 1369. https://doi.org/10.3390/rs11111369.

Briechle, S., Krzystek, P., Vosselman, G., 2021. Silvi-Net – A dual-CNN approach for combined classification of tree species and standing dead trees from remote sensing data. Int. J. Appl. Earth Obs. Geoinf. 98, 102292. https://doi.org/10.1016/j.jag.2020.102292.

Bromley, J., Bentz, J.W., Bottou, L., Guyon, I., LeCun, Y., Moore, C., Säckinger, E., Shah, R., 1993. Signature verification using a "siamese" time delay neural network. Int. J. Pattern Recognit. Artif. Intell. 7, 669–688. https://doi.org/10.1142/S0218001493000339.

Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A., 2020. Albumentations: Fast and flexible image augmentations. Inf. 11, 125. https://doi.org/10.3390/info11020125.

Cao, F., Guo, W., 2020. Cascaded dual-scale crossover network for hyperspectral image classification. Knowledge-Based Systems. 189, 105122. https://doi.org/10.1016/j.knosys.2019.105122.

Cao, F., Guo, W., 2020b. Deep hybrid dilated residual networks for hyperspectral image classification. Neurocomputing. 384, 170–181. https://doi.org/10.1016/j.neucom.2019.11.092.

Cao, X., Li, R., Wen, L.i., Feng, J., Jiao, L., 2018. Deep multiple feature fusion for hyperspectral image classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 (10), 3880–3891. https://doi.org/10.1109/JSTARS.460944310.1109/JSTARS.2018.2866595.

Cao, Z., Li, X., Jiang, J., Zhao, L., 2020. 3D convolutional siamese network for few-shot hyperspectral classification. J. Appl. Remote Sens. 14, 1–20. https://doi.org/10.1117/1.JRS.14.048504.

Bousmalis, K., Irpan, A., Wohlhart, P., Bai, Y., Kelcey, M., Kalakrishnan, M., Downs, L., Ibarz, J., Pastor, P., Konolige, K., Levine, S., Vanhoucke, V., 2018. Using simulation and domain adaptation to improve efficiency of deep robotic grasping. In: Proc. of the 2018 IEEE International Conference on Robotics and Automation (ICRA). Brisbane, QLD, Australia. pp. 4243-4250. doi: 10.1109/ICRA.2018.8460875.

Castelluccio, M., Poggi, G., Sansone, C., Verdoliva, L., 2015. Land use classification in remote sensing images by convolutional neural networks. arXiv 1. abs/1508.00092.

Chen, F., Jiang, H., Van de Voorde, T., Lu, S., Xu, W., Zhou, Y., 2018a. Land cover mapping in urban environments using hyperspectral APEX data: A study case in Baden, Switzerland. Int. J. Appl. Earth Obs. Geoinf. 71, 70–82. https://doi.org/10.1016/j.jag.2018.04.011.

Chen, Y., Jiang, H., Li, C., Jia, X., Ghamisi, P., 2016. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. IEEE Trans. Geosci. Remote Sens. 54 (10), 6232–6251. https://doi.org/10.1109/TGRS.2016.2584107.

Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2018b. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. 40 (4), 834–848. https://doi.org/10.1109/TPAMI.2017.2699184.

Chen, Y., Wang, Y., Gu, Y., He, X., Ghamisi, P., Jia, X., 2019. Deep learning ensemble for hyperspectral image classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12 (6), 1882–1897. https://doi.org/10.1109/JSTARS.460944310.1109/JSTARS.2019.2915259.

Chen, Z., Wang, C., Li, J., Fan, W., Du, J., Zhong, B., 2021. Adaboost-like End-to-End multiple lightweight U-nets for road extraction from optical remote sensing images. Int. J. Appl. Earth Obs. Geoinf. 100, 102341. https://doi.org/10.1016/j.jag.2021.102341.

Crowson, M., Hagensieker, R., Waske, B., 2019. Mapping land cover change in northern Brazil with limited training data. Int. J. Appl. Earth Obs. Geoinf. 78, 202–214. https://doi.org/10.1016/j.jag.2018.10.004.

Davari, A., Aptoula, E., Yanikoglu, B., Maier, A., Riess, C., 2018. GMM-based synthetic samples for classification of hyperspectral images with limited training data. IEEE Geosci. Remote Sens. Lett. 15 (6), 942–946. https://doi.org/10.1109/LGRS.2018.2817361.

Dede, M.A., Aptoula, E., Genc, Y., 2019. Deep network ensembles for aerial scene classification. IEEE Geosci. Remote Sens. Lett. 16 (5), 732–735. https://doi.org/10.1109/LGRS.885910.1109/LGRS.2018.2880136.

Deng, C., Xue, Y., Liu, X., Li, C., Tao, D., 2019. Active transfer learning network: A unified deep joint spectral-spatial feature learning model for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 57 (3), 1741–1754. https://doi.org/10.1109/TGRS.2018.2868851.

Deng, F., Pu, S., Chen, X., Shi, Y., Yuan, T., Pu, S., 2018. Hyperspectral image classification with capsule network using limited training samples. Sensors. 18, 3153. https://doi.org/10.3390/s18093153.

Ding, X., Li, Y., Yang, J., Li, H., Liu, L., Liu, Y., Zhang, C., 2021. An adaptive capsule network for hyperspectral remote sensing classification. Remote Sens. 13, 2445. https://doi.org/10.3390/rs13132445.

Chen, Y., Li, J., Xiao, H., Jin, X., Yan, S., Feng, J., 2017. Dual path networks. NIPS, 4467-4475. doi:10.5555/3294996.3295200.

Doersch, C., Gupta, A., Efros, A.A., 2015. Unsupervised visual representation learning by context prediction. IEEE, Santiago, Chile, pp. 1422–1430. https://doi.org/10.1109/ICCV.2015.167.

Dong, H., Zhang, L., Zou, B., 2019. Band Attention Convolutional Networks For Hyperspectral Image Classification. CoRR abs/1906.04379.

Dong, Y., Du, B., Zhang, L., Hu, X., 2018. Hyperspectral target detection via adaptive information - theoretic metric learning with local constraints. Remote Sens. 10, 1415. https://doi.org/10.3390/rs10091415.

Fan, W., Wei, T., Peng, J., Sun, W., 2019b. Ideal regularized kernel subspace alignment for unsupervised domain adaptation in hyperspectral image classification. IEEE, pp. 1–4. https://doi.org/10.1109/Multi-Temp.2019.8866985.

Fang, B., Li, Y., Zhang, H., Chan, J.-C.-W., 2018. Semi-supervised deep learning classification for hyperspectral image based on dual-strategy sample selection. Remote Sens. 10, 574. https://doi.org/10.3390/rs10040574.

Fang, J., Yuan, Y., Lu, X., Feng, Y., 2019. Robust space-frequency joint representation for remote sensing image scene classification. IEEE Trans. Geosci. Remote Sens. 57 (10), 7492–7502. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2019.2913816.

Fauvel, M., Chanussot, J., Benediktsson, J.A., Sveinsson, J.R., 2007. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. In: Proc. of the 2007 IEEE International Geoscience and Remote Sensing Symposium (IGRSS), pp. 4834-4837. doi:10.1109/TGRS.2008.922034.

Fu, G., Liu, C., Zhou, R., Sun, T., Zhang, Q., 2017. Classification for high resolution remote sensing imagery using a fully convolutional network. Remote Sens. 9, 498. https://doi.org/10.3390/rs9050498.

Feilong, C., Wenhui, G., et al., 2019. Cascaded dual-scale crossover network for hyperspectral image classification. Knowledge-Based Systems 189, 105122. https://doi.org/10.1016/j.knosys.2019.105122.

Gao, Q., Lim, S., Jia, X., 2018. Hyperspectral image classification using convolutional neural networks and multiple feature learning. Remote Sens. 10, 299. https://doi.org/10.3390/rs10020299.

Ge, Z., Cao, G., Li, X., Fu, P., 2020. Hyperspectral image classification method based on 2D–3D CNN and multibranch feature fusion. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 13, 5776–5788. https://doi.org/10.1109/JSTARS.460944310.1109/JSTARS.2020.3024841.

Gupta, A., Vedaldi, A., Zisserman, A., 2016. Synthetic data for text localisation in natural images. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 2315-2324. doi: 10.1109/CVPR.2016.254.

Graña, M., Veganzons, M., Ayerdi, B., 2013-2018. Hyperspectral remote sensing scenes. In: group, C.I. (Ed.), Faculty of Informatics of San Sebastián. Donostia (Gipuzkoa), Spain.

Han, Y., Gao, Y., Zhang, Y., Wang, J., Yang, S., 2019. Hyperspectral sea ice image classification based on the spectral-spatial-joint feature with deep learning. Remote Sens. 11, 2170. https://doi.org/10.3390/rs11182170.

Hang, R., Li, Z., Liu, Q., Ghamisi, P., Bhattacharyya, S.S., 2021. Hyperspectral image classification with attention-aided CNNs. IEEE Trans. Geosci. Remote Sens. 59 (3), 2281–2293. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2020.3007921.

Hang, R., Liu, Q., Hong, D., Ghamisi, P., 2019. Cascaded recurrent neural networks for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 57 (8), 5384–5394. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2019.2899129.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 770-778. doi:10.1109/CVPR.2016.90.

He, X., Chen, Y., Ghamisi, P., 2020. Heterogeneous transfer learning for hyperspectral image classification based on convolutional neural network. IEEE Trans. Geosci. Remote Sens. 58 (5), 3246–3263. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2019.2951445.

Hoffman, J., Guadarrama, S., Tzeng, E., Donahue, J., Girshick, R., Darrell, T., Saenko, K., 2014. LSDA: Large scale detection through adaptation. NIPS 4, 3536–3544. https://doi.org/10.5555/2969033.2969221.

Hong, D., Chanussot, J., Yokoya, N., Heiden, U., Heldens, W., Zhu, X.X., 2019. WU-Net: A weakly-supervised unmixing network for remotely sensed hyperspectral imagery. In Proc. of the 2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 373-376. doi:10.1109/IGARSS.2019.8899865.

Hu, W., Huang, Y., Li, W., Zhang, F., Li, H.-C., 2015. Deep convolutional neural networks for hyperspectral image classification. J. Sensors. 2015, 258619:1-258619:12. doi: 10.1155/2015/258619.

Huang, G., Liu, Z., Maaten, L.v.d., Weinberger, K.Q., 2017. Densely connected convolutional networks.In: Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 2261-2269. doi:10.1109/CVPR.2017.243.

Huang, L., Chen, Y., 2021. Dual-Path siamese CNN for hyperspectral image classification with limited training samples. IEEE Geosci. Remote Sens. Lett. 18 (3), 518–522. https://doi.org/10.1109/LGRS.885910.1109/LGRS.2020.2979604.

Jaiswal, A., AbdAlmageed, W., Wu, Y., Natarajan, P., 2018. CapsuleGAN: Generative adversarial capsule network 11131, 526–535. https://doi.org/10.1007/978-3-030-11015-4_38.

Jiang, Y., Li, Y., Zou, S., Zhang, H., Bai, Y., 2021. Hyperspectral image classification with spatial consistence using fully convolutional spatial propagation network. IEEE Trans. Geosci. Remote Sens. 1–13 https://doi.org/10.1109/TGRS.2021.3049282.

Jiao, L., Liang, M., Chen, H., Yang, S., Liu, H., Cao, X., 2017. Deep fully convolutional network-based spatial distribution prediction for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 55, 5585–5599. https://doi.org/10.1109/TGRS.2017.2710079.

Jon Atli, B., Pedram, G., 2015. Spectral-Spatial classification of hyperspectral remote sensing images. Artech.

Kaiser, L., Nachum, O., Roy, A., Bengio, S., 2017. Learning to Remember Rare Events. In: Proc. of International Conference on Learning Representations, (ICLR). https://openreview.net/pdf?id=SJTQLdqlg.

Kang, X., Li, C., Li, S., Lin, H., 2018. Classification of hyperspectral images by gabor filtering based deep network. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11, 1166–1178. https://doi.org/10.1109/JSTARS.2017.2767185.

Kang, X., Zhuo, B., Duan, P., 2019. Dual-path network-based hyperspectral image classification. IEEE Geosci. Remote Sens. Lett. 16, 447–451. https://doi.org/10.1109/LGRS.2018.2873476.

Kemker, R., Luu, R., Kanan, C., 2018a. Low-Shot Learning for the Semantic Segmentation of Remote Sensing Imagery. IEEE Trans. Geosci. Remote Sens. 56, 6214–6223. https://doi.org/10.1109/TGRS.2018.2833808.

Kemker, R., Salvaggio, C., Kanan, C., 2018b. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. ISPRS J. Photogramm. Remote Sens. 145, 60–77. https://doi.org/10.1016/j.isprsjprs.2018.04.014.

Koch, G., Richard Z., Ruslan S.. 2015. Siamese neural networks for one-shot image recognition.In: Proc. of ICML deep learning workshop. 2.

Kouw, W.M., Loog, M., 2019. TECHNICAL REPORT : An introduction to domain adaptation and transfer learning. arXiv:1812.11806. Delft University of Technology, p. 42. https://doi.org/10.13140/RG.2.2.33906.56004.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM. 60 (6), 84–90. https://doi.org/10.1145/3065386.

Laliberte, A.S., Goforth, M.A., Steele, C.M., Rango, A., 2011. Multispectral remote sensing from unmanned aircraft: image processing workflows and applications for rangeland environments. Remote Sens. 3, 2529–2551. https://doi.org/10.3390/rs3112529.

Lee, Hyungtae, Kwon, Heesung, 2017. Going deeper with contextual CNN for hyperspectral image classification. IEEE Trans. Image Process. 26 (10), 4843–4855. https://doi.org/10.1109/TIP.2017.2725580.

Li, S., Song, W., Fang, L., Chen, Y., Ghamisi, P., Benediktsson, J.A., 2019. Deep learning for hyperspectral image classification: An overview. IEEE Trans. Geosci. Remote Sens. 57 (9), 6690–6709. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2019.2907932.

Lake, B.M., Salakhutdinov, R., Tenenbaum, J.B., 2013. One-shot learning by inverting a compositional causal process. In: Proc. Neural Information Processing Systems (NIPS). pp. 2526-2534.

Li, X., Guo, Y., 2014. Multi-level adaptive active learning for scene classification. In : Proc. of the 2018 Eur. Conf. on Comput. Vis. (ECCV), pp. 234-249. doi:10.1007/978-3-319-10584-0_16.

Li, Z., Wang, T., Li, W., Du, Q., Wang, C., Liu, C., Shi, X., 2020e. Deep multilayer fusion dense network for hyperspectral image classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 13, 1258–1270. https://doi.org/10.1109/JSTARS.460944310.1109/JSTARS.2020.2982614.

Li, Y., Wang, N., Shi, J., Liu, J., Hou, X., 2016. Revisiting batch normalization for practical domain adaptation. Pattern Recognit. 80 https://doi.org/10.1016/j.patcog.2018.03.005.

Li, X., Ding, M., Pizurica, A., 2020. Deep feature fusion via two-stream convolutional neural network for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 58 (4), 2615–2629. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2019.2952758.

Li, Z., Liu, M., Chen, Y., Xu, Y., Li, W., Du, Q., 2021. Deep cross-domain few-shot learning for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 1–18 https://doi.org/10.1109/TGRS.2021.3057066.

Li, F., Wang, L., Han, W., 2020. Ensemble model with cascade attention mechanism for high-resolution image scene classification. Opt. Express. 28 https://doi.org/10.1364/oe.395866.

Li , T. , Zhang , J. , Zhang , Y. , 2014a . Classification of hyperspectral image based on deep belief networks. In: Proc. of the 2014 IEEE International Conference on Image Processing (ICIP), pp. 5132 - 5136. doi: 10.1109/ICIP.2014.7026039.

Li, R., Zheng, S., Duan, C., Yang, Y., Wang, X, 2020a. Classification of hyperspectral image based on double-branch dual-attention mechanism network. Remote Sens. 12, 582. https://doi.org/10.3390/rs12030582.

Lin, Z., Chen, Y., Ghamisi, P., Benediktsson, J., 2018. Generative adversarial networks for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 5046–5063. https://doi.org/10.1109/TGRS.2018.2805286.

Liu, B., Yu, X., Zhang, P., Yu, A., Fu, Q., Wei, X., 2018a. Supervised Deep Feature Extraction for Hyperspectral Image Classification. IEEE Trans. Geosci. Remote Sens. 56 (4), 1909–1921. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2017.2769673.

Liu, P., Zhang, H., Eom, K.B., 2017a. Active Deep Learning for Classification of Hyperspectral Images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10 (2), 712–724. https://doi.org/10.1109/JSTARS.460944310.1109/JSTARS.2016.2598859.

Liu, Q., Kampffmeyer, M., Jenssen, R., Salberg, A., 2020a. Dense Dilated Convolutions' Merging Network for Land Cover Classification. IEEE Trans. Geosci. Remote Sens. 58 (9), 6309–6320. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2020.2976658.

Liu, S., Du, Q., Tong, X., Samat, A., Pan, H., Ma, X., 2017b. Band Selection-Based Dimensionality Reduction for Change Detection in Multi-Temporal Hyperspectral Images. Remote Sens. 9 (10), 1008. https://doi.org/10.3390/rs9101008.

Liu, S., Shi, Q., Zhang, L., 2021. Few-shot hyperspectral image classification with unknown classes using multitask deep learning. IEEE Trans. Geosci. Remote Sens. 59 (6), 5085–5102. https://doi.org/10.1109/TGRS.2020.3018879.

Liu, Yao, Gao, Lianru, Xiao, Chenchao, Qu, Ying, Zheng, Ke, Marinoni, Andrea, 2020b. Hyperspectral Image Classification Based on a Shuffled Group Convolutional Neural Network with Transfer Learning. Remote Sens. 12 (11), 1780. https://doi.org/10.3390/rs12111780.

Liu, Y., Zhong, Y., Qin, Qianqing, 2018b. Scene Classification Based on Multiscale Convolutional Neural Network. IEEE Trans. Geosci. Remote Sens. 56 (12), 7109–7121. https://doi.org/10.1109/TGRS.2018.2848473.

Long, J., Shelhamer, E., Darrell, T., 2015a. Fully convolutional networks for semantic segmentation. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 3431-3440. https://doi.ieeecomputersociety.org/10.1109/CVPR.2015.7298965.

Long, M., Cao, Y., Wang, J., Jordan, M.I., 2015b. Learning Transferable Features with Deep Adaptation Networks. In : Proc. of the 2015 International Conference on Machine Learning (ICML), pp. 97-105. https://dl.acm.org/doi/10.5555/3045118.3045130.

Lundervold, A.S., Lundervold, A., 2019. An overview of deep learning in medical imaging focusing on MRI. Zeitschrift für Medizinische Physik 29 (2), 102–127. https://doi.org/10.1016/j.zemedi.2018.11.002.

Luo, X., Shu, B., Yu, H., Wang, H., Mu, J., 2019a. A Multi-scale Network based on Attention Mechanism for Hyperspectral Image Classification. In: Proc. of the 8th International Conference on Computing and Pattern Recognition (ICCPR), pp. 260-265. https://dl.acm.org/doi/10.1145/3373509.3373560.

YC This reference cannot be edited. Kindly help to update the reference to read as follows: Luo, Y., Zhu, L.-Z., Lu, B.-L., 2019b. A GAN-based Data augmentation method for multimodal emotion recognition: Advances in neural networks. In : Proc. of the International Symposium on Neural Networks (ISNN), pp. 141-150. doi: 10.1007/978-3-030-22796-8_16.

Lv, W., Wang, X., 2020. Overview of hyperspectral image classification. J. Sens. Sens. Syst. 2020, 1–13. https://doi.org/10.1155/2020/4817234.

Ma, D., Tang, P., Zhao, L., 2019. SiftingGAN: Generating and sifting labeled samples to improve the remote sensing image scene classification baseline in vitro. IEEE Geosci. Remote Sens. Lett. 16 (7), 1046–1050. https://doi.org/10.1109/LGRS.885910.1109/LGRS.2018.2890413.

Ma, X., Wang, H., Geng, J., 2016. Spectral-spatial classification of hyperspectral image based on deep auto-encoder. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (9), 4073–4085. https://doi.org/10.1109/JSTARS.460944310.1109/JSTARS.2016.2517204.

Makantasis, K., Karantzalos, K., Doulamis, A.D., Doulamis, N.D., 2015. Deep supervised learning for hyperspectral data classification through convolutional neural networks. In : Proc. of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 4959-4962. doi:10.1109/IGARSS.2015.7326945.

Mei, X., Pan, E., Ma, Y., Dai, X., Huang, J., Fan, F., Du, Q., Zheng, H., Ma, J., 2019. Spectral-spatial attention networks for hyperspectral image classification. Remote Sens. 11, 963. https://doi.org/10.3390/rs11080963.

Meng, X., Bao, Y., Liu, J., Liu, H., Zhang, X., Zhang, Y., Wang, P., Tang, H., Kong, F., 2020. Regional soil organic carbon prediction model based on a discrete wavelet analysis of hyperspectral satellite data. Int. J. Appl. Earth Obs. Geoinf. 89, 102111. https://doi.org/10.1016/j.jag.2020.102111.

Meng, Z., Li, L., Jiao, L., Feng, Z., Tang, X., Liang, M., 2019. Fully dense multiscale fusion network for hyperspectral image classification. Remote Sens. 11 (22), 2718. https://doi.org/10.3390/rs11222718.

Midhun, M., Nair, S.R., Prabhakar, V.N., Kumar, S.S., 2014. Deep model for classification of hyperspectral image using restricted boltzmann machine. In : Proc. of the 2014 International Conference on Interdisciplinary Advances in Applied Computing (ICONIAAC), pp. 1-7. doi:10.1109/ICIP.2014.7026039.

Minetto, R., Pamplona Segundo, M., Sarkar, S., 2018. Hydra: An ensemble of convolutional neural networks for geospatial land classification. IEEE Trans. Geosci. Remote Sens. 57 (9), 6530–6541. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2019.2906883.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A., Veness, J., Bellemare, M., Graves, A., Riedmiller, M., Fidjeland, A., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D., 2015. Human-level control through deep reinforcement learning. Nature. 518 (7540), 529–533. https://doi.org/10.1038/nature14236.

Modzelewska, A., Fassnacht, F.E., Stereńczak, K., 2020. Tree species identification within an extensive forest area with diverse management regimes using airborne hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. 84, 101960. https://doi.org/10.1016/j.jag.2019.101960.

Moliner, E., Romero, L.S., Vilaplana, V., 2020. Weakly supervised semantic segmentation for remote sensing hyperspectral imaging. Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP) 2273–2277. https://doi.org/10.1109/ICASSP40776.2020.9053384.

Mou, L., Ghamisi, P., Zhu, X., 2018. Unsupervised spectral-spatial feature learning via deep residual conv–deconv network for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 56 (1), 391–406. https://doi.org/10.1109/TGRS.2017.2748160.

Mou, L., Ghamisi, P., Zhu, X., Sensing, R., 2017. Deep recurrent neural networks for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 55, 3639–3655. https://doi.org/10.1109/TGRS.2016.2636241.

Mu, C., Guo, Z., Liu, Y., 2020a. A Multi-Scale and Multi-Level Spectral-Spatial Feature Fusion Network for Hyperspectral Image Classification. Remote Sens. 12, 125. https://doi.org/10.3390/rs12010125.

Mu, Caihong, Liu, Jian, Liu, Yi, Liu, Yijin, 2020b. Hyperspectral Image Classification Based on Active Learning and Spectral-Spatial Feature Fusion Using Spatial Coordinates. IEEE Access. 8, 6768–6781. https://doi.org/10.1109/Access.628763910.1109/ACCESS.2019.2963624.

Nalepa, J., Myller, M., Kawulok, M., 2019a. Hyperspectral Data Augmentation. CoRR abs/1903.05580.

Nalepa, J., Myller, M., Kawulok, M., 2020. Training and test-time data augmentation for hyperspectral image segmentation. IEEE Geosci. Remote Sens. Lett. 17 (2), 292–296. https://doi.org/10.1109/LGRS.885910.1109/LGRS.2019.2921011.

Nowakowski, A., Mrziglod, J., Spiller, D., Bonifacio, R., Ferrari, I., Mathieu, P., Garcia-Herranz, M., Kim, D., 2021. Crop type mapping by using transfer learning. Int. J. Appl. Earth Obs. Geoinf. 98, 102313. https://doi.org/10.1016/j.jag.2021.102313.

Othman, E., Bazi, Y., Melgani, F., Alhichri, H., Alajlan, N., Zuair, M., 2017. Domain Adaptation Network for Cross-Scene Classification. IEEE Trans. Geosci. Remote Sens. 55 (8), 4441–4456. https://doi.org/10.1109/TGRS.2017.2692281.

Pan, S.J., Yang, Q., 2009. A Survey on transfer learning. IEEE Trans. on Knowl. and Data En. 22 (10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191.

Paoletti, M.E., Haut, J.M., Plaza, J., Plaza, A., 2017. A new deep convolutional neural network for fast hyperspectral image classification. ISPRS J. Photogramm. Remote Sens. 145, 120–147. https://doi.org/10.1016/j.isprsjprs.2017.11.021.

Paoletti, Mercedes E., Haut, Juan Mario, Fernandez-Beltran, Ruben, Plaza, Javier, Plaza, Antonio, Li, Jun, Pla, Filiberto, 2019a. Capsule Networks for Hyperspectral Image Classification. IEEE Trans. Geosci. Remote Sens. 57 (4), 2145–2160. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2018.2871782.

Paoletti, Mercedes E., Haut, Juan Mario, Fernandez-Beltran, Ruben, Plaza, Javier, Plaza, Antonio J., Pla, Filiberto, 2019b. Deep Pyramidal Residual Networks for Spectral-Spatial Hyperspectral Image Classification. IEEE Trans. Geosci. Remote Sens. 57 (2), 740–754. https://doi.org/10.1109/TGRS.2018.2860125.

Paoletti, M.E., Haut, J.M., Plaza, J., Plaza, A., 2019c. Deep learning classifiers for hyperspectral imaging: A review. ISPRS J. Photogramm. Remote Sens. 158, 279–317. https://doi.org/10.1016/j.isprsjprs.2019.09.006.

Palatucci, M., Pomerleau, D., Hinton, G.E., Mitchell, T.M., 2009. Zero-shot Learning with Semantic Output Codes. 1410-1418. https://dl.acm.org/doi/10.5555/2984093.2984252.

Patki, N., Wedge, R., Veeramachaneni, K., 2016. The synthetic data vault. In . Proc. IEEE Int. Conf. on Data Sci. and Advanced Analytics (DSAA). doi: 10.1109/DSAA.2016.49.

Peng, X., Sun, B., Ali, K., Saenko, K., 2015. Learning deep object detectors from 3D models. Proc. Int. Conf. on Comput. Vis. (ICCV) 1278–1286. https://doi.org/10.1109/ICCV.2015.151.

Pepik, B., Stark, M., Gehler, P., Schiele, B., 2012. Teaching 3D geometry to deformable part models. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 3362-3369. doi: 10.1109/CVPR.2012.6248075.

Qu, Y., Baghbaderani, R.K., Qi, H., 2019. Few-shot hyperspectral image classification through multitask transfer learning. In : Paper presented in the 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), pp. 1-5. doi:10.1109/WHISPERS.2019.8920992.

Rajpura, P.S., Bojinov, H., Hegde, R.S., 2017. Object detection using deep CNNs trained on synthetic images. arXiv:1706.06782 2.

Ravi, S., Larochelle, H., 2017. Optimization as a model for few-shot learning. ICLR.

Romero, A., Gatta, C., Camps-Valls, G., 2016. Unsupervised deep feature extraction for remote sensing image classification. IEEE Trans. Geosci. Remote Sens. 54 (3), 1349–1362. https://doi.org/10.1109/TGRS.2015.2478379.

Roy, S.K., Krishna, G., Dubey, S.R., Chaudhuri, B.B., 2020. HybridSN: Exploring 3-D-2-D CNN feature hierarchy for hyperspectral image classification. IEEE Geosci. Remote Sens. Lett. 17 (2), 277–281. https://doi.org/10.1109/LGRS.885910.1109/LGRS.2019.2918719.

Sabour, S., Frosst, N., Hinton, G.E., 2017. Dynamic routing between capsules. In: Proc. Conf. Neural Inform. Process. Syst., Long Beach, USA, pp. 3856-3866.

Samat, A., Gamba, P., Du, P., Luo, J., 2015. Active extreme learning machines for quad-polarimetric SAR imagery classification. Int. J. Appl. Earth Obs. Geoinf. 35, 305–319. https://doi.org/10.1016/j.jag.2014.09.019.

Santara, A., Mani, K., Hatwar, P., Singh, A., Garg, A., Padia, K., Mitra, P., 2017. BASS Net: Band-adaptive spectral-spatial feature learning neural network for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 55 (9), 5293–5301. https://doi.org/10.1109/TGRS.2017.2705073.

Satkin, S., Lin, J., Hebert, M., 2012. Data-driven scene understanding from 3D models. In: Proc. of the British Machine Vis. Conf. (BMVC) 1–11. https://doi.org/10.5244/C.26.128.

Shen, Y., Xiao, L., Chen, J., Pan, D., 2019. A spectral-spatial domain-specific convolutional deep extreme learning machine for supervised hyperspectral image classification. IEEE Access. 7, 132240–132252. https://doi.org/10.1109/Access.628763910.1109/ACCESS.2019.2940697.

Signoroni, A., Savardi, M., Baronio, A., Benini, S., 2019. Deep learning meets hyperspectral image analysis: A multidisciplinary review. J. Imaging. 5, 52. https://doi.org/10.3390/jimaging5050052.

Song, J., Ma, L., 2017. Deep neural network-based domain adaptation for classification of remote sensing images. J. Appl. Remote Sens. 11 https://doi.org/10.1117/1.JRS.11.042612.

Sumbul, G., Cinbis, R.G., Aksoy, S., 2018. Fine-grained object recognition and zero-shot learning in remote sensing imagery. IEEE Trans. Geosci. Remote Sens. 56 (2), 770–779. https://doi.org/10.1109/TGRS.2017.2754648.

Sun, Z., Xu, Q., Li, F., Mei, Y., Bin, L., 2020. Hyperspectral image classification via hierarchical features adaptive fusion network. In: Proc. of the Advances in Brain Inspired Cognitive Systems Conf., (BICS) 2019. Springer, pp. 302–310. https://doi.org/10.1007/978-3-030-39431-8_29.

Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P.H.S., Hospedales, T.M., 2018. Learning to compare: Relation network for few-shot learning. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 1199-1208. doi:10.1109/CVPR.2018.00131.

Tang, C., Zhang, Z., He, G., Long, T., Wang, G., Wei, M., She, Wenqing, 2021a. An improved fully convolution network model for change detection in mining areas using sentinel-2 images. Remote Sens. Lett. 12 (7), 684–694. https://doi.org/10.1080/2150704X.2021.1925372.

Tang, Xu, Ma, Qiushuo, Zhang, Xiangrong, Liu, Fang, Ma, Jingjing, Jiao, Licheng, 2021b. Attention Consistent Network for Remote Sensing Scene Classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 2030–2045. https://doi.org/10.1109/JSTARS.460944310.1109/JSTARS.2021.3051569.

Tao, C., Pan, H., Li, Y., Zou, Z., 2015. Unsupervised spectral–spatial feature learning with stacked sparse autoencoder for hyperspectral imagery classification 12, 2438–2442. https://doi.org/10.1109/LGRS.2015.2482520.

Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., Abbeel, P., 2017. Domain randomization for transferring deep neural networks from simulation to the real world. In : Proc. of the IEEE/RSJ Int. Conf. on Intell. Robots and Systems (IROS), pp. 23-30. doi:10.1109/IROS.2017.8202133.

Torresani, L., 2014. In: Computer vision: A reference guide. Springer US, Boston, MA, pp. 883–885.

Tong, X., Yin, J., Han, B., Qv, H., 2020b. Few-shot learning with attention-weighted graph convolutional networks for hyperspectral image classification. In : Proc. of the 2020 IEEE International Conference on Image Processing (ICIP), pp. 1686-1690. doi: 10.1109/ICIP40778.2020.9190752.

Tremblay, J., Prakash, A., Acuna, D., Brophy, M., Jampani, V., Anil, C., To, T., Cameracci, E., Boochoon, S., Birchfield, S., 2018. Training deep networks with synthetic data: bridging the reality gap by domain randomization. IEEE, pp. 1082–10828. https://doi.org/10.1109/CVPRW.2018.00143.

Tripathi, S., Chandra, S., Agrawal, A., Tyagi, A., Rehg, J.M., Chari, V., 2019. Learning to generate synthetic data via compositing. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog., Long Beach, USA, pp. 461-470. doi:10.1109/CVPR.2019.00055.

Tuia, D., Persello, C., Bruzzone, L., 2016. Domain adaptation for the classification of remote sensing data: an overview of recent advances. IEEE Geosci. Remote Sens. Mag. 4 (2), 41–57. https://doi.org/10.1109/MGRS.2016.2548504.

Tuia, D., Volpi, M., Copa, L., Kanevski, M., Munoz-Mari, J., 2011. A survey of active learning algorithms for supervised remote sensing image classification. IEEE J. Sel. Top. Signal Process. 5 (3), 606–617. https://doi.org/10.1109/JSTSP.2011.2139193.

Tzeng, E., Hoffman, J., Zhang, N., Saenko, K., Darrell, T., 2014. Deep domain confusion: Maximizing for domain invariance. arXiv:1412.3474 1.

Wambugu, Naftaly, Chen, Yiping, Xiao, Zhenlong, Wei, Mingqiang, Aminu Bello, Saifullahi, Marcato Junior, José, Li, Jonathan, 2021. A hybrid deep convolutional neural network for accurate land cover classification. Int. J. Appl. Earth Obs. Geoinf. 103, 102515. https://doi.org/10.1016/j.jag.2021.102515.

Wang, C., Liu, Y., Bai, X., Tang, W., Lei, P., Zhou, J., 2017. Deep residual convolutional neural network for hyperspectral image super-resolution. In: Image and Graphics. Springer, pp. 370–380. https://doi.org/10.1007/978-3-319-71598-8_33.

Wang, Cong, Zhang, Lei, Wei, Wei, Zhang, Yanning, 2020. Hyperspectral Image Classification With Data Augmentation and Classifier Fusion. IEEE Geosci. Remote Sens. Lett. 17 (8), 1420–1424. https://doi.org/10.1109/LGRS.885910.1109/LGRS.2019.2945848.

Wang, Decheng, Chen, Xiangning, Jiang, Mingyong, Du, Shuhan, Xu, Bijie, Wang, Junda, 2021a. ADS-Net: An Attention-Based deeply supervised network for remote sensing image change detection. Int. J. Appl. Earth Obs. Geoinf. 101, 102348. https://doi.org/10.1016/j.jag.2021.102348.

Wang, D., Du, B., Zhang, L., 2021b. Fully Contextual Network for Hyperspectral Scene Parsing. IIEEE Trans. Geosci. Remote Sens 1–16. https://doi.org/10.1109/TGRS.2021.3050491.
Wang, M., Deng, W., 2018. Deep visual domain adaptation: A survey. Neurocomputing. 312, 135–153. https://doi.org/10.1016/j.neucom.2018.05.083.
Wang, Q., Liu, S., Chanussot, J., Li, X., 2019b. Scene Classification With Recurrent Attention of VHR Remote Sensing Images. IEEE Trans. Geosci. Remote Sens. 57 (2), 1155–1167. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2018.2864987.
Wang, Q., Yuan, Z., Du, Q., Li, X., 2019c. GETNET: A General End-to-End 2-D CNN Framework for Hyperspectral Image Change Detection. IEEE Trans. Geosci. Remote Sens. 57, 3–13. https://doi.org/10.1109/TGRS.2018.2849692.
Wong, S.C., Gatt, A., Stamatescu, V., McDonnell, M.D., 2016. Understanding Data Augmentation for Classification: When to Warp? In: Proc. Int. Conf. on Digital Image Computing: Techniques and Applications (DICTA). DICTA, pp. 1–6. https://doi.org/10.1109/DICTA.2016.7797091.
Wu, H., Prasad, S., 2018. Semi-supervised deep learning using pseudo labels for hyperspectral image classification. IEEE Trans. Image Process. 27 (3), 1259–1270. https://doi.org/10.1109/TIP.2017.2772836.
Wu, S., Zhang, J., Zhong, C., 2019. Multiscale spectral-spatial unified networks for hyperspectral image classification. In : Proc. of the 2019 IEEE International Geoscience and Remote Sensing Symposium (IGRASS), pp. 2706-2709. doi:10.1109/TGRS.2018.2827407.
Xu, X., Li, J., Plaza, A.J., 2016. Fusion of hyperspectral and LiDAR data using morphological component analysis. Proc. IEEE Int. Geosci. and Remote Sens. Symposium (IGRASS) 3575–3578. https://doi.org/10.1109/IGARSS.2016.7729926.
Xu, X., Li, W., Ran, Q., Du, Q., Gao, L., Zhang, B., 2017. Multisource remote sensing data classification based on convolutional neural network. IEEE Trans. Geosci. Remote Sens. 56 (2), 937–949. https://doi.org/10.1109/TGRS.2017.2756851.
Xu, Y., Du, B., Zhang, L., Cerra, D., Pato, M., Carmona, E., Prasad, S., Yokoya, N., Hansch, R., Le Saux, B., 2019. Advanced multi-sensor optical remote sensing for urban land use and land cover classification. in outcome of the 2018 IEEE GRSS data fusion contest. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12, 1709–1724. https://doi.org/10.1109/JSTARS.2019.2911113.
Xue, Z., 2020. A general generative adversarial capsule network for hyperspectral image spectral-spatial classification. Remote Sens. Lett. 11 (1), 19–28. https://doi.org/10.1080/2150704X.2019.1681598.
Yang, G., Gewali, U.B., Ientilucci, E., Gartley, M., Monteiro, S.T., 2018. Dual-channel densenet for hyperspectral image classification. Proc. IEEE Int. Geosci. and Remote Sens. Symposium (IGARSS) 2595–2598. https://doi.org/10.1109/IGARSS.2018.8517520.
Yanik, E., Sezgin, T.M., 2019. Active scene learning. CoRR abs/1903.02832.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention is All you need. Advances in Neural Information Processing Systems 30. In: Proc. Conf. Neural Inform. Process. Syst., 2017, pp. 5998-6008.
Marchesi, M., 2017. Megapixel size image creation using generative adversarial networks. arXiv:1706.00082 1.
Yokoya, N., Iwasaki, A., 2016. Airborne hyperspectral data over Chikusei.
Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep neural networks?, Advances in Neural Information Processing Systems 27: In: Proc. Conf. Neural Inform. Process. Syst., pp. 3320-3328.
Yu, Y., Li, X., Liu, F., 2020. Attention GANs: Unsupervised deep feature learning for aerial scene classification. IEEE Trans. Geosci. Remote Sens. 58 (1), 519–531. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2019.2937830.
Yun, K., Yu, K., Osborne, J., Eldin, S., Nguyen, L., Huyen, A., Lu, T., 2019. Improved visible to IR image transformation using synthetic data augmentation with cycle-consistent adversarial networks. CoRR abs/1904.11620. doi: 10.1117/12.2519121.
Zeiler, M.D., Fergus, R., 2014. Visualizing and Understanding Convolutional Networks. Proc. Eur. Conf. on Comput. Vis. (ECCV) 818–833. https://doi.org/10.1007/978-3-319-10590-1_53.
Zhan, Y., Hu, D., Wang, Y., Yu, X., 2018. Semisupervised hyperspectral image classification based on generative adversarial networks. IEEE Geosci. Remote Sens. Lett. 15 (2), 212–216. https://doi.org/10.1109/LGRS.2017.2780890.
Zhang, Jianming, Jin, Xiaokang, Sun, Juan, Wang, Jin, Li, Keqin, 2019a. Dual Model Learning Combined With Multiple Feature Selection for Accurate Visual Tracking. IEEE Access. 7, 43956–43969. https://doi.org/10.1109/Access.628763910.1109/ACCESS.2019.2908668.
Zhang, J., Lu, C., Li, X., Kim, H.-J., Wang, J., 2019b. A full convolutional network based on DenseNet for remote sensing scene classification. Math. Biosci. Eng. 16, 3345–3367. https://doi.org/10.3934/mbe.2019167.
Zhang, J., Lu, C., Wang, J., Yue, X.-G., Lim, S.-J., Al-Makhadmeh, Z., Tolba, A., 2020. Training convolutional neural networks with multi-size images and triplet loss for remote sensing scene classification. Sensors. 20, 1188. https://doi.org/10.3390/s20041188.
Zhang, J., Sun, J., Wang, J., Yue, X., 2021. Visual object tracking based on residual network and cascaded correlation filters. J. Ambient Intell. Humanized Comput. 12 (8), 8427–8440. https://doi.org/10.1007/s12652-020-02572-0.
Zhao, C., Deng, W., Yan, Y., Yao, X., 2017. Progressive line processing of kernel RX anomaly detection algorithm for hyperspectral imagery. Sensors. (Basel) 17 (8), 1815. https://doi.org/10.3390/s17081815.
Zhao, C., Yao, X., Huang, B., 2016. Real-time anomaly detection based on a fast recursive kernel RX algorithm. Remote Sens. 8 (12), 1011. https://doi.org/10.3390/rs8121011.
Zhao, W., Du, S., 2016. Spectral-spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach. IEEE Trans. Geosci. Remote Sens. 54 (8), 4544–4554. https://doi.org/10.1109/TGRS.2016.2543748.
Zhen, L., Hai, Y., Guo, M., Xiao, L., et al., 2019. Image style transfer based on improved CycleGAN. Optics and Precision Engineering 27, 1836–1844. https://doi.org/10.3788/OPE.20192708.1836.
Zheng, Z., Zhong, Y., Ma, A., Zhang, L., 2020. FPGA: Fast patch-free global learning framework for fully end-to-end hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 58 (8), 5612–5626. https://doi.org/10.1109/TGRS.3610.1109/TGRS.2020.2967821.
Zhong, Z., Li, J., Clausi, D.A., Wong, A., 2020. Generative adversarial networks and conditional random fields for hyperspectral image classification. IEEE Trans. Cybern. 50 (7), 3318–3329. https://doi.org/10.1109/TCYB.622103610.1109/TCYB.2019.2915094.
Zhong, Z., Li, J., Luo, Z., Chapman, M., 2018. Spectral-spatial residual network for hyperspectral image classification: a 3D deep learning framework. IEEE Trans. Geosci. Remote Sens. 56 (2), 847–858. https://doi.org/10.1109/TGRS.2017.2755542.
Zhou, F., Hang, R., Liu, Q., Yuan, X., 2019. Hyperspectral image classification using spectral-spatial LSTMs. Neurocomputing. 328, 39–47. https://doi.org/10.1016/j.neucom.2018.02.105.
Zhou, S., He, Y., Liu, Y., Li, C., Zhang, J., 2021. Multi-channel deep networks for block-based image compressive sensing. IEEE Trans. Multimedia. 23, 2627–2640. https://doi.org/10.1109/TMM.2020.3014561.
Zhou, X., Prasad, S., 2018. Deep feature alignment neural networks for domain adaptation of hyperspectral data. IEEE Trans. Geosci. Remote Sens. 56 (10), 5863–5872. https://doi.org/10.1109/TGRS.2018.2827308.
Zhong, Z., Li, J., Ma, L., Jiang, H., Zhao, H., 2017b. Deep residual networks for hyperspectral image classification, 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 1824-1827. doi:10.1109/IGARSS.2017.8127330.
Zhu, F., 2017. hyperspectral unmixing: Ground truth labeling, datasets, benchmark performances and survey. arXiv: Computer Vision and Pattern Recognition.
Zhu, K., Chen, Y., Ghamisi, P., Jia, X., Benediktsson, J.A., 2019. Deep convolutional capsule network for hyperspectral image spectral and spectral-spatial classification. Remote Sens. 11, 223. https://doi.org/10.3390/rs11030223.
Zhou, X., Prasad, S., 2020. Advances in deep learning for hyperspectral image analysis—addressing challenges arising in practical imaging scenarios. In: Prasad, S., Chanussot, J. (Eds.), Hyperspectral Image Analysis: Advances in Machine Learning and Signal Processing. Springer, pp. 117–140.
Zhu, W., Chayes, V., Tiard, A., Sanchez, S., Dahlberg, D., Bertozzi, A.L., Osher, S., Zosso, D., Kuang, D., 2017. Unsupervised classification in hyperspectral imagery with nonlocal total variation and primal-dual hybrid gradient algorithm. IEEE Trans. Geosci. Remote Sens. 55 (5), 2786–2798. https://doi.org/10.1109/TGRS.2017.2654486.
Zou, L., Zhu, X., Wu, C., Liu, Y., Qu, L., 2020. Spectral-spatial exploration for hyperspectral image classification via the fusion of fully convolutional networks. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 13, 659–674. https://doi.org/10.1109/JSTARS.460944310.1109/JSTARS.2020.2968179.