Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 5, No. 3, 2023

An Overview of Hyperspectral Image Classification by Data-driven Deep Learning

Xiaochuan Yu [1,*], Mary B. Ozdemir [2], M. K. Joshie [2]

[1] College of Electrical and Information Engineering, Hunan University, Changsha, 410082, China
[2] School of Geosciences, The University of Sydney, Sydney, NSW 2006, Australia
*** Corresponding author:** Xiaochuan Yu (Email: xiaochuan_you@163.com)

Abstract: Hyperspectral imaging (HSI) in remote sensing is gaining significant attention due to its complexity, posing challenges for conventional machine learning in achieving accurate classification. The inherent nonlinear relationship between captured spectral information and materials further complicates hyperspectral imaging. Deep learning has emerged as an effective tool for feature extraction, finding widespread applications in image processing tasks. Motivated by its success, this survey integrates deep learning into hyperspectral imaging (HSI) classification, demonstrating commendable performance. The paper systematically reviews existing literature, providing a comparative analysis of strategies. Primary challenges in HSI classification for traditional methods are outlined, emphasizing the advantages of deep learning. Our framework categorizes works into three types: spectral-feature networks, spatial-feature networks, and spectral-spatial-feature networks, offering a comprehensive review of recent achievements and diverse approaches. Considering limited training samples in remote sensing and substantial data requirements for deep networks, strategies to enhance classification performance are presented, offering valuable insights for future studies. Experiments apply representative deep learning-based classification methods to real HSIs, providing practical validation. The survey contributes to understanding the current landscape in deep learning-based HSI classification and lays a foundation for future research in this evolving field.

Keywords: Hyperspectral; Remote Sensing; Classification; Imaging; Deep Networks.

1. Introduction

Hyperspectral imaging (HSI), vital in remote sensing, captures electromagnetic spectrum data across visible to near-infrared wavelengths. These images consist of numerous narrow spectral bands, each pixel representing a high-dimensional vector denoting spectral reflectance[1, 2]. With the ability to discern subtle spectral differences, hyperspectral imaging finds broad applications. Recent studies highlight HSI classification as a dynamic research area, facing challenges of large spatial variability in spectral signatures and limited training samples against high data dimensionality[3, 4]. Factors like illumination changes and environmental conditions contribute to spatial variability, while limited samples impact classifier generalization. Early HSI classification methods focused on spectral signatures, employing pixel-wise classifiers such as neural networks, support vector machines, and logistic regression. Some methods incorporated feature extraction techniques like PCA, ICA, and LDA. However, pixel-wise classifiers yielded unsatisfactory results due to neglecting spatial contexts.

Spatial features recently proved crucial for enhancing hyperspectral data representation and classification accuracies[5, 6]. Increasingly, spectral-spatial features-based frameworks are developed, integrating spatial contextual information into pixel-wise classifiers. Notably, techniques like extended morphological profiles (EMPs)[7], multiple kernel learning, and edge-preserving filtering[8] have been employed. In[9], spatial information from neighboring regions was incorporated into sparse representation models, recognizing that hyperspectral pixels often form linear combinations from the same class. Spatial consistency was further explored by segmenting HSIs into super pixels based on intensity or texture similarity[10]. While spectral-spatial methods demonstrated good performance, reliance on hand-crafted or shallow descriptors limits their applicability in challenging scenarios. Hand-crafted features, tailored for specific tasks, may lack the ability to discriminate subtle class variations or handle large intra-class variations. Extracting more discriminative features is crucial for enhancing HSI classification.

Data-driven method, specifically the Deep learning, is a growing trend in big data analysis, demonstrating breakthroughs in various computer vision tasks such as image classification[11-14], object detection, natural language processing[15], agricultural crop management[6] and financial risk prevention[14]. Motivated by these successes, deep learning has been applied to hyperspectral image classification with notable performance. Fig 1 illustrates the increasing publication trend in deep learning-based HSI classification, indicating continued exploration and research in this area. Unlike traditional hand-crafted methods, deep learning extracts informative features automatically through hierarchical layers, starting with simple features like texture and edge information in earlier layers and progressing to more complex features in deeper layers. This automatic learning process makes deep learning well-suited for handling diverse situations. Overall, deep learning is recognized as an effective feature extraction approach for HSI classification, with different networks focusing on extracting various feature types.

2. Data-driven Approach and Models

This section provides a brief overview of widely used deep network models in HSI classification, like stacked auto-encoders (SAEs), deep belief networks (DBNs), and generative adversarial networks (GANs). The auto-encoder

(AE) is a fundamental component of the stacked AE (SAE) [16-18]. An AE comprises one visible layer with d inputs, one hidden layer with L units, and one reconstruction layer with d units. Training involves mapping $x \in \mathbb{R}^d$ to $h \in \mathbb{R}^L$ in the hidden layer. SAEs are formed by stacking multiple AEs, connecting the output of one layer to the input of the next.

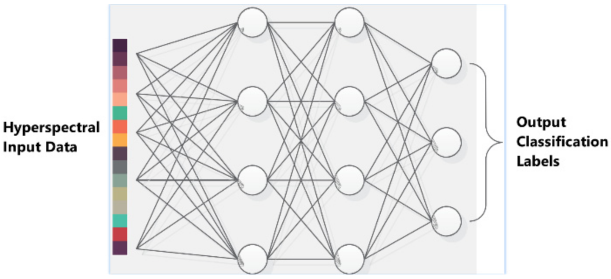

Fig 1. A general illustration of an SAE connected with a subsequent logistic regression classifier

The Restricted Boltzmann Machine (RBM) serves as the fundamental unit in a layer-wise training model and is integral to constructing a Deep Belief Network (DBN). To enhance the feature representation capability, multiple RBMs can be stacked to form a DBN, creating a deep hierarchical representation of the training data. For classification purposes, a logistic regression layer can be added to the end of the DBN, forming a spectral classifier. In machine learning, two main model classes exist: generative approaches, which learn distribution parameters to generate new samples, and discriminative approaches, which model label dependence on training data (i.e., x) for predicting labels from data samples.

GANs, a recent model, consist of a generative model (G) and a discriminative model (D) [19]. These models are trained adversarially, with G generating fake inputs to resemble real ones, while D distinguishes between real and fake inputs. This adversarial competition ensures continuous and effective training of the discriminator. GAN-based HSI classification framework means the generator G accepts HSI class labels (c) along with noise (z), and its output is defined accordingly. The discriminator D is trained using both real samples with class labels and fake data generated by G.

3. Deep Networks-Based Hsi Classification

Deep learning, highly successful in computer vision[20], has been applied to hyperspectral image (HSI) classification in remote sensing[21], offering advantages over traditional hand-crafted features by automatically learning discriminative features from complex hyperspectral data. This capability addresses the challenge of large spatial variability in spectral signatures mentioned in Section I. Numerous deep networks have emerged, extracting diverse features such as spectral, spatial, and spectral-spatial features. This section systematically reviews deep learning-based HSI classification methods within a framework, categorizing deep networks into spectral-feature networks, spatial-feature networks, and spectral-spatial-feature networks. Each network type is designed to extract specific features for subsequent classification, detailed in the following subsections.

Hyperspectral remote sensors typically provide hundreds of spectral bands, introducing redundancy in the data. Directly exploring original spectral vectors results in high computational costs and reduced classification performance. While traditional spectral feature extraction methods like PCA, ICA, and LDA [22] are effective, their simple linear processing struggles with the complex spectral properties of HSIs. This section introduces the spectral-feature networks, a deep learning-based framework for extracting profound spectral features. Previous research shows that incorporating spatial features enhances classification accuracies in HIS [23]. Such research covers spatial-feature networks that leverage deep networks for extracting spatial features. To achieve accurate HSI classification, these learned spatial features are fused with spectral features extracted by other techniques.

In[24], PCA initially reduced the dimensionality of hyperspectral data, followed by the exploitation of spatial information using a 2-D CNN in the neighborhood region of the input pixel. This combination of PCA and CNN not only extracted discriminative spatial features but also reduced computational costs. In a more advanced approach, Liang et al. [68] introduced sparse representation to encode deep spatial features extracted by a CNN into low-dimensional sparse features, enhancing feature representation and classification accuracies. Off-the-shelf CNNs is used like AlexNet and GoogLeNet for extracting deep spatial features [25]. These features were then fused to train a multiple-feature-based classifier. A novel HSI classification framework, the Deep Multiscale Spatial-Spectral Feature Extraction Algorithm, was presented in [26]. This approach utilized the pretrained FCN-8 to explore deep multiscale spatial structural information, followed by a weighted fusion mechanism to combine original spectral features and deep multiscale spatial features. The fused features were fed into a classifier for the final classification operation.

Rather than focusing solely on spectral or spatial features, spectral-spatial-feature networks aim to jointly extract deep spectral-spatial features for HSI classification. These features can be obtained through three approaches: 1) mapping low-level spectral-spatial features to high-level ones using deep networks; 2) directly extracting deep features from original data or principal components; 3) fusing separate deep spectral and spatial features. Spectral-spatial-feature networks are categorized into preprocessing-based, integrated, and postprocessing-based networks. Training deep networks requires a large number of samples, but remote sensing often has limited labeled data. This imbalance can lead to poor classification. Strategies to address this issue include data augmentation, an intuitive method to generate new training samples from known ones. Two main strategies for creating additional virtual samples are transformation-based and mixture-based sample generation. These methods will be discussed in detail in the following section.

(1) Transformation-based sample generation:

Due to the complex lighting conditions in HSIs, objects of the same class in different locations may be affected by varying radiations. Addressing this, virtual samples can be generated by transforming known samples, a method widely employed in [27]. Specifically, given a known training sample xi, a new virtual sample y can be obtained through $y = f(xi) + \gamma n$, where f is a transforming function (e.g., rotation, flipping, or mirroring), γ controls the weight of random Gaussian noise (n) introduced through factors like neighboring pixel interactions or imaging errors. The newly generated virtual sample y shares the same class as xi and can be used for deep network training.

(2) *Mixture-based sample generation*

In general, the objects of the same class usually show similar spectral characteristics in a certain range. This phenomenon makes it possible to generate a virtual sample from two given samples of the same class. In [56], [78], the virtual sample y is the linear combination of two training samples xi and xj.

While supervised feature learning has made significant strides in HSI classification, there is a pressing need for unsupervised or semi-supervised feature learning[28]. The primary goal of unsupervised/semi-supervised feature learning is to extract valuable features from unlabeled data. Recent research efforts [29] have focused on developing robust and effective unsupervised/semi-supervised feature learning frameworks based on deep learning for HSI classification. In unsupervised/semi-supervised deep feature learning, the top flowchart extracts informative features from unlabeled data in an unsupervised manner. The deep network, structured as an encoder-decoder paradigm, learns without using label information. Additionally, the classification performance can be enhanced by transferring the trained network and fine-tuning it on the labeled dataset, as shown at the bottom of the flowchart in Fig. 2.

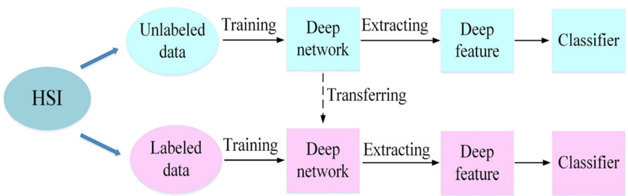

Fig 2. Illustration of unsupervised/semi-supervised feature learning associated with deep learning. The top part represents the unsupervised deep feature learning, and the bottom part illustrates the semi-supervised way for feature learning.

4. Experimental Strategy

This section comprises comprehensive experiments focusing on some key aspects. Firstly, it is the advantages of deep learning over traditional methods in HSI classification through a series of designed experiments. Secondly, we systematically compare the classification performance of recent state-of-the-art deep learning approaches. The visualization of learned deep features and network weights is conducted to gain insights into the "black box" nature of the models. Deep learning is often considered a black box, with internal network details remaining unclear in many applications. However, understanding these internal features is crucial for analyzing network performance and refining deep architectures. In this section, we use the Salinas dataset to visualize deep features. The weights of various convolutional kernels in the first layer are initially random. After training, the weight distribution becomes more regular, showcasing distinct textural features. For instance, the weights in the first kernel exhibit a gradual decrease from the left to the right side, a pattern divergent from randomly initialized weights[30]. Early convolution layers extract simple features like texture and edges, which are then composed into more complex high-level features through deeper layers. The automatic learning process makes CNNs well-suited for handling diverse situations.

5. Conclusion

Data-driven deep learning-based hyperspectral image (HSI) classification has gained significant attention in remote sensing, exhibiting commendable performance compared to traditional hand-crafted feature-based methods. This survey introduces key deep models like SAE, DBN, CNN, RNN, and GAN commonly used for HSI classification. It offers a unified framework for a comprehensive overview of existing deep learning-based methodologies, categorizing them into spectral-feature networks, spatial-feature networks, and spectral-spatial-feature networks. The comparison of traditional and deep learning-based HSI classification methods reveals the superior performance of deep learning methods, with the DFFN, incorporating residual learning and feature fusion, achieving the best results. Visualization of deep features and network weights aids in performance analysis and architectural design. Considering the limitation of available training samples in remote sensing, strategies to enhance classification performance are explored, with experiments confirming the efficacy of these approaches. Residual learning stands out as the most effective strategy, providing valuable insights for future studies in this field.

References

[1] S. Gui, R. Qin, Y. Tang, Sat2lod2: A Software For Automated Lod-2 Modeling From Satellite-Derived Orthophoto And Digital Surface Model, arXiv preprint arXiv:2204.04139 (2022).

[2] D. Huang, Y. Tang, R. Qin, An evaluation of PlanetScope images for 3D reconstruction and change detection–experimental validations with case studies, GIScience & Remote Sensing 59(1) (2022) 744-761.

[3] X. Lai, Y. Tang, L. Li, P. Chan, Q. Zeng, Study on microclimate observation network for urban unit: A case study in a campus of Shenzhen, China, Physics and Chemistry of the Earth, Parts A/B/C 110 (2019) 117-124.

[4] Y. Tang, S. Song, S. Gui, W. Chao, C. Cheng, R. Qin, Active and Low-Cost Hyperspectral Imaging for the Spectral Analysis of a Low-Light Environment, Sensors 23(3) (2023) 1437.

[5] P. Ghamisi, E. Maggiori, S. Li, R. Souza, Y. Tarablaka, G. Moser, A. De Giorgi, L. Fang, Y. Chen, M. Chi, New frontiers in spectral-spatial hyperspectral image classification: The latest advances based on mathematical morphology, Markov random fields, segmentation, sparse representation, and deep learning, IEEE geoscience and remote sensing magazine 6(3) (2018) 10-43.

[6] R. Tao, P. Zhao, J. Wu, N.F. Martin, M.T. Harrison, C. Ferreira, Z. Kalantari, N. Hovakimyan, Optimizing crop management with reinforcement learning and imitation learning, arXiv preprint arXiv:2209.09991 (2022).

[7] J. Li, P.R. Marpu, A. Plaza, J.M. Bioucas-Dias, J.A. Benediktsson, Generalized composite kernel framework for hyperspectral image classification, IEEE transactions on geoscience and remote sensing 51(9) (2013) 4816-4829.

[8] X. Kang, S. Li, J.A. Benediktsson, Spectral–spatial hyperspectral image classification with edge-preserving filtering, IEEE transactions on geoscience and remote sensing 52(5) (2013) 2666-2677.

[9] L. Fang, C. Wang, S. Li, J.A. Benediktsson, Hyperspectral image classification via multiple-feature-based adaptive sparse representation, IEEE Transactions on Instrumentation and Measurement 66(7) (2017) 1646-1657.

[10] L. Fang, N. He, S. Li, P. Ghamisi, J.A. Benediktsson, Extinction profiles fusion for hyperspectral images classification, IEEE Transactions on Geoscience and Remote Sensing 56(3) (2017) 1803-1815.

[11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with

convolutions, Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1-9.

[12] J. Wu, R. Tao, P. Zhao, N.F. Martin, N. Hovakimyan, Optimizing nitrogen management with deep reinforcement learning and crop simulations, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 1712-1720.

[13] J. Wu, D. Pichler, D. Marley, D. Wilson, N. Hovakimyan, J. Hobbs, Extended Agriculture-Vision: An Extension of a Large Aerial Image Dataset for Agricultural Pattern Analysis, arXiv preprint arXiv:2303.02460 (2023).

[14] Z. Xu, Analysis and research on financial risk prevention based on artificial intelligence algorithms, Proceedings of the 2nd International Conference on Bigdata Blockchain and Economy Management, ICBBEM 2023, May 19–21, 2023, Hangzhou, China, 2023.

[15] A. Bordes, X. Glorot, J. Weston, Y. Bengio, Joint learning of words and meaning representations for open-text semantic parsing, Artificial intelligence and statistics, PMLR, 2012, pp. 127-135.

[16] Y. Chen, Z. Lin, X. Zhao, G. Wang, Y. Gu, Deep learning-based classification of hyperspectral data, IEEE Journal of Selected topics in applied earth observations and remote sensing 7(6) (2014) 2094-2107.

[17] J. Wu, N. Hovakimyan, J. Hobbs, Genco: An auxiliary generator from contrastive learning for enhanced few-shot learning in remote sensing, arXiv preprint arXiv:2307.14612 (2023).

[18] Y. Wang, J. Wu, N. Hovakimyan, R. Sun, Double Dynamic Sparse Training for GANs, arXiv preprint arXiv:2302.14670 (2023).

[19] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Advances in neural information processing systems 27 (2014).

[20] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521(7553) (2015) 436-444.

[21] Y. Tang, J. Yan, Y. Wu, J. Hong, L. Xu, Z. Lin, Design of Remote Sensing Image Data Analysis and Processing Platform Based on Environmental Monitoring, Journal of Physics: Conference Series, IOP Publishing, 2021, p. 012056.

[22] T.V. Bandos, L. Bruzzone, G. Camps-Valls, Classification of hyperspectral images with regularized linear discriminant analysis, IEEE Transactions on Geoscience and Remote Sensing 47(3) (2009) 862-873.

[23] J.A. Benediktsson, P. Ghamisi, Spectral-spatial classification of hyperspectral remote sensing images, Artech House2015.

[24] J.M. Haut, M.E. Paoletti, J. Plaza, J. Li, A. Plaza, Active learning with convolutional neural networks for hyperspectral image classification using a new Bayesian approach, IEEE Transactions on Geoscience and Remote Sensing 56(11) (2018) 6440-6461.

[25] G. Cheng, Z. Li, J. Han, X. Yao, L. Guo, Exploring hierarchical convolutional features for hyperspectral image classification, IEEE Transactions on Geoscience and Remote Sensing 56(11) (2018) 6712-6722.

[26] L. Jiao, M. Liang, H. Chen, S. Yang, H. Liu, X. Cao, Deep fully convolutional network-based spatial distribution prediction for hyperspectral image classification, IEEE Transactions on Geoscience and Remote Sensing 55(10) (2017) 5585-5599.

[27] H. Lee, H. Kwon, Going deeper with contextual CNN for hyperspectral image classification, IEEE Transactions on Image Processing 26(10) (2017) 4843-4855.

[28] J. Wu, J. Hobbs, N. Hovakimyan, Hallucination improves the performance of unsupervised visual representation learning, Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 16132-16143.

[29] X. Zhang, Y. Liang, C. Li, N. Huyan, L. Jiao, H. Zhou, Recursive autoencoders-based unsupervised feature learning for hyperspectral image classification, IEEE Geoscience and Remote Sensing Letters 14(11) (2017) 1928-1932.

[30] Y. Tang, K. Yoshihiro, A. Tohru, B. Danushka, BS-1-12 Analyzing Patterns from Twitter Happiness referring Stock Market and American Unemployment Rate Announcement, Proceedings of the Electronic Information and Communication Society Comprehensive System Conference 2013(2) (2013) " S-23"-" S-24".