

A SURVEY: DEEP LEARNING CLASSIFIERS FOR HYPERSPECTRAL IMAGE CLASSIFICATION

¹MURALI KANTHI¹, ²T. HITENDRA SARMA, ³C. SHOBHA BINDU

^{1, 3}Department of Computer Science & Engineering, JNTUCEA, Anantapur. A.P., India.

²Department of Information Technology, Vasavi College of Engineering, Hyderabad, Telangana, India.

E-mail: ¹murali.kanthi@gmail.com, ²hitendrasarma@iecee.org, ³shobabindhu@gmail.com

ABSTRACT

Hyperspectral imaging (HSI) is a popular subject in remote sensing data processing because of the huge quantity of data included in these images, which enables for improved description and utilization of the surface of the earth by integrating abundant spectral and spatial data. However, due to the high dimensionality of HSI data and the limited number of labelled examples available, conducting HSI image classification poses significant technical and pragmatic hurdles. Approaches to HSI classification with deep learning have gained major successes in recent years as new deep learning algorithms emerge, giving unique prospects for hyperspectral image classification research and development. Initially, a quick introduction to standard deep learning (DL) models is provided, followed by a comparison of the performance of common DL based HSI approaches. Finally, the difficulties and future research prospects are explored.

Keywords: *Remote Sensing Image, Deep Learning, Auto-Encoder, Convolutional Neural Network, Stacked, Deep Belief Network.*

1. INTRODUCTION

Remote hyperspectral sensing is an acquisition of digital images in numerous small spectral contiguous bands of earth material and creates complete spectral signatures. Hyperspectral imaging (HSI) gathers information and analyses it over the electromagnetic spectrum. The goal of HSI is to get the spectrum of each pixel in a scene image in order to identify objects, materials, or processes. [1]. At a nominal spectral resolution of 10 nanometers, a spectral band is defined as a distinct interval of the electromagnetic spectrum with wavelengths ranging from 0.4 micrometers to 0.5 micrometers in one spectral band. The resulting hyperspectral images not only include a lot of spectral data, but they also have a lot of spatial data on the ground features. Hence, many applications have made use of hyperspectral remote sensing., such as precision agriculture [2], crop monitoring [3], and land resources [4], HSI has been employed in environmental protection to detect oil spills [5], water quality analysis [6], gas [7] and vegetation coverage [8], health of forests [9]. In the medical field, HSI has been used for skin testing to assess the health of human skin [10], military and defense applications [11]. Other domains where the usage

of HSI had useful outcomes included security and so forth [12]. Classification is a basic activity that plays a significant part in land use and land cover applications, among the aforementioned applications. Several approaches for quickly processing and classifying hyperspectral images have been developed. [13]. In a variety of HSI data classification techniques, machine learning (ML) and deep learning (DL) methodologies are employed. This includes unsupervised methods like clustering. Often supervised classifiers are preferable because they may give high classification accuracies, however, while the limited availability of training data may influence these approaches, since they usually need several examples to get such good results.

As deep learning technology has evolved fast over the past decade, it has attracted a great deal of interest. Deep learning technology, in contrast to standard machine learning models, does not require the creation of artificial feature patterns and can automatically learn data patterns. As a result, it has achieved great results in the domains of speech recognition, autonomous driving, semantic segmentation, object identification, and natural language processing. It was recently brought

into the field of HSI categorization as well. A variety of novel deep learning HSI classification models have been developed by researchers. Deep Learning techniques can automatically generate meaningful representations of input data at distinct scales. These learnt features have been successful in a variety of machine vision tasks.

Deep belief networks were suggested using restricted Boltzmann machines, which represented a significant and early advance in DL. Following that, work based on Auto-encoder was developed, which trained the various intermediate levels of representation locally at each level. Convolutional neural networks, another DL architecture, have recently achieved significant results in computer vision, thanks to the model's deep structure, which allows it to capture and generalize filtering mechanisms by performing convolutions in the image domain, yielding abstract mathematical and efficient features. Despite its enormous promise, using Deep Learning for image classification in Remote Sensing (RS) introduces a slew of new problems. This is due to a number of factors: For first off, many RS data, particularly hyperspectral images (HSIs), contain hundreds of bands, requiring a high number of neurons in a DL network to retain a significant amount of data in a small region. The spectral curve vectors across bands, aside from the obvious geometrical motifs inside per band, may also give useful information. However, further study is needed to determine how to use this knowledge. Second, to obtain their often-remarkable results, DL techniques rely on a large number of labelled samples. Unfortunately, there are very few labelled samples in the RS data. Third, RS images are more sophisticated than traditional natural scene images. A wide variety of objects may be included in the high-resolution RS images. In the first place, HSIs may be measured using a variety of sensors. A DL network structures for RS image classification cannot be constructed directly because to the complexity of the data.

DL's use in RS image categorization is difficult because of the reasons listed above. Recognizing this, a number of techniques were presented in recent years to cope with such issues. An overview of such advancements is presented in this survey, with a particular focus on two essential aspects: pixel-wise classifying HSIs, and when it comes to the former, scene categorization for elevated airborne or satellite images is crucial., it's about figuring out what category each of the pixels in an RS scene image falls into, while the latter is

about automatically assigning semantic labels to each of the images.

The research in RS analysis is important due to its potential applications in real life. Hyperspectral imaging results in multiple bands of images that make the analysis challenging due to the increased volume of data. The spectral, as well as the spatial correlation between different bands, conveys useful information regarding the scene of interest. Recently, the convolutional neural network (CNN) is one of the most frequently used deep learning-based methods for HSI classification. These approaches are mostly based on 2-DCNN to extract the spatial and spectral features separately for HSI classification. On the other hand, the HSI classification performance is highly dependent on both spatial and spectral information. In the literature, there are several deep learning models proposed based on 3D-CNN proved to be more efficient than 2D-CNN models to extract joint spatial-spectral features for HSI classification. In this direction, we proposed different 3D-CNN based approaches to extract joint spatial-spectral features for HSI classification.

The following is how this survey is structured. The second part describes common DL models used in HSI image classification, such as stacked auto-encoders, DBNs, and CNNs. The third section compares the classification performances of common deep learning techniques for HSI images. The fourth section reviews the current work and the fifth section address future difficulties, identifying prospective areas for future study in HSI image categorization using deep learning methods.

2. TYPICAL DEEP NEURAL NETWORK MODELS

We'll examine at three typical deep neural network models that were used to classify RS images in this part.

2.1 Auto Encoders (AEs)

Effective feature extraction becomes a key preprocessing step for modelling the data's underlying structures and relationships, decreasing the Hughes effect and the curse of dimensionality while handling with HSI data classification problems. To conduct unsupervised coding from HSI data, autoencoders have been frequently utilized as deep models. In its operating mode, the AE design does not conduct any classification tasks, but instead reconstructs the input data by decreasing $\min \|X - X^*\|$ [14]. In reality, the fundamental

strength of these models is their capacity to reflect the exact input data into a new space, resulting in squeezed, stretched, or perhaps even equitably results with the lowest degree of deformation conceivable. As shown in Figure 1, this projection is carried out using a conventional architecture consisting of encoder and decoder networks connected by a bottleneck layer that depicts the hidden region. Due to the existing connection between neighboring bands, HSI-AEs evolved as generally pixel-wise techniques, which are commonly used to perform dimensionality reduction (DR) and greater extent spectral FE. In this case, the encoder takes the spectral pixel $X \in \mathbb{N}^{n_{bands}}$ as input and represents it in a new space $\mathbb{R}^{n_{new}}$ using a hierarchical collection of recognition weights or encoder components. The resulting code vector or code dictionary $c_i \in \mathbb{R}^{n_{new}}$ is then sent into the decoder, which applies a set of $L_{encoder}$ generating weights on the output matrix in order to restore and/or resemble the input matrix x'_i .

In the literature, there have been numerous AE models developed for HSI data analysis. Zhu et al. [15], proposed an unsupervised tied AE for spatial frequency FE based on the highest uncertainty ratio as a pre-processing DR process and perfect through Softmax categorization. Okan et al. [16], proposed a pixel-wise stacked AE, which uses a two-step training method that includes unsupervised and supervised learning before the final supervised classification using regression analysis. Mughees et al. [17], created an SAE for spectrum processing, while an evolutionary perimeter alteration segmentation technique is used for spatial analysis. As a consequence, a majority voting-based approach is used to merge the spectral-based categorization image with the spatial-based narrow band segmentation image. Paul et al. [18], propose a segmented-SAE for spectral-spatial HSI classification process as an advancement on the SAE, which uses consensual data to conduct spectral segmentation and syntactic statuses to comprehend the spatial data contained in the HSI cube, minimizing its sophistication and computational periods.

Zhou et al. [19], developed a two-stage AE termed concise and discrete SAE, in which the first stage trains a discriminative SAE to learn a physical infrastructure by reducing the training error, and the second phase conducts data classification while updating the DSAE's parameters. AEs have also been used with neural

networks like CNNs to derive spectral characteristics.

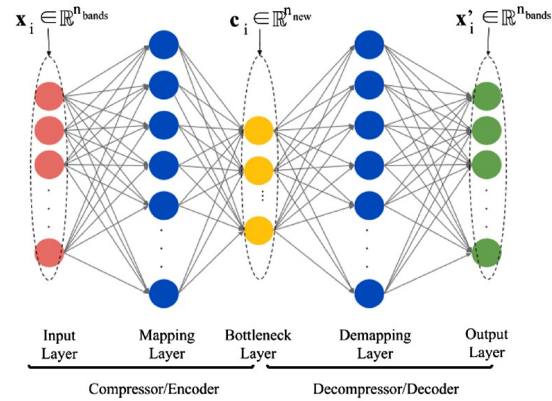


Figure 1: Typical Representation of Autoencoder, A Bottleneck Layer Connects the Two Major Parts: An Encoder and A Decoder.

2.2 Deep Belief Networks (DBNs)

DBNs construct a probabilistic graphical model in the form of a directed acyclic graph by combining probability and graph theory. Several papers in the literature discuss how to construct DBNs by stacking unsupervised networks like limited Boltzmann models with a pessimistic training algorithm as the optimizer.

DBNs have been used to conduct FE in HSI data processing as a version of the AE framework incorporating pessimistic layer-wise learning. In this regard, Li et al. [20] employ a DBN for extracting features, stacking spectral-spatial data and using logistic regression to classify. Chen et al. [21] also uses three DBNs to retrieve spatial, spectral, and spatial-spectral elevated characteristics from HSI data in a systematic way, and then use regression models to accomplish the final classification model. Several efforts have been made to improve the HSI classification efficiency of this type of DNN. For example, Le et al. [22], examine the hyper-parameters utilized by Chen et al. [21]'s spectral and spectral-spatial DBNs, whereas Zhong et al. [23], developed a diversification DBN for HSI classification that maintain standards the pre-training and fine-tuning processes by a factor of ten. Guofeng et al. [24] use kernel PCA and PCA to enhance conventional DBN learning approach and prevent the influence of gradient disappearance. Zhou et al. [25], created a group belief network based on DBNs that takes into account the properties of combined spatial-spectral characteristics from HSI by changing the lower

layer of every RBM that makes up the network model.

Despite DBNs are potential DL techniques for HSI information classification, however suffer from the same restriction as SAEs: such neural models are intended for evaluating 1-dimensional, thus the extensive spatial information included in HSI cubes should be high accuracy in order to access the data simultaneously. Finally, this sort of spectral-spatial processing can't fully include the spatial-contextual data contained in HSI cubes.

2.3 Convolutional Neural Networks (CNN)

CNN uses a multi-layer trained architecture consisting of a convolution, pooling, and fully connected layers. The input maps are concatenated using trainable kernels and then passed through from the activation function to produce the outcome image features as a consequence of convolution layers. Figure 2 depicts a typical CNN architecture. Every hidden layer entity is coupled by shared weights to the local responsive field over the input instead of being entirely connected to the input in the converting layer, which may be 2-dimensional size $m \times n$ feature maps in the convolution layer. The convolution layer produces a concentration of the X^i input feature maps with a K^i kernel of the size $k \times k \times m$ and a nonlinear element activation feature, which subsequently applies to the X^i output maps. To predict classification labels, fully connected with Softmax layers are applied after stacked layers have been completed.

A CNN's architecture is made up of two distinct components, which may be viewed as two networks. The FE-net, which is made up of a various leveled stack of extracting features and recognition phases that gets to know elevated interpretations of the inputs, and the classifier, which is made up of a stack of FC layers that accomplishes the very last classification process of

calculating the subscription of every input, are both prepared as a final method to enhance all of the weights in the CNN. According to Murugan et al. [26], the FE-net is made up of several hierarchically stratified separation and identification stages, with CONV, activation, and POOL layers being widely utilized in these sub mapping functions. The CNN approach can extract local stationarity properties of X in this way, revealing the properties that are common throughout the data source using regional kernels. In reality, the CNN's feature extraction is quite similar to that of other DNN models, in that the initial stages may identify identifiable characteristics, while the final phases integrate all of the characteristics identified by the preceding layers to detect further creative traits. However, as shown in Figure 2, the flexibility in kernel design helps in effective and environmental retrieval of spatial, spectral, and spectral-spatial features, whereas the regionally existence of convolutional kernels, combined with parameter-sharing throughout layers, reduces the value of variables that need to be fine-tuned by the approach, producing evaluations faster. The classifier net, in particular, conducts the final classification utilizing the data received by the FE-network. Typically, this component is done in phases made up of ReLU and FC layers, with Softmax placed on last FC layer.

Furthermore, a normal MLP structure or alternative classifiers like SVM or logistic regression can be used to create the classifier net. The classifier can also be ignored, with the first portion, the FE-net, being used for different applications like unsupervised FE. There are three kinds of CNN methods for HSI categorization in the present literature, regardless of whether they conduct spatial, spectral, or spectral-spatial selection of feature. We'll look at some of the works that are currently available in each category in the sections that follow.

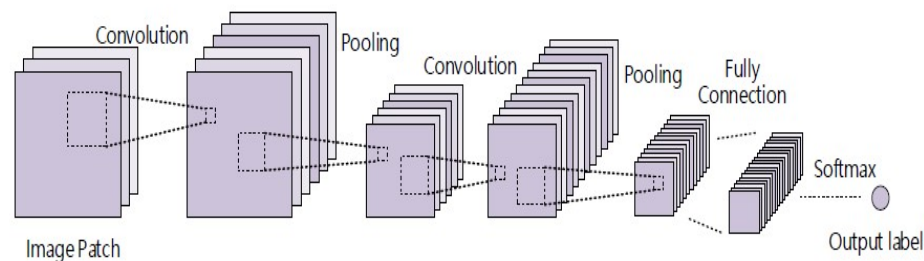


Figure 2: Typical CNN Structure Illustration.

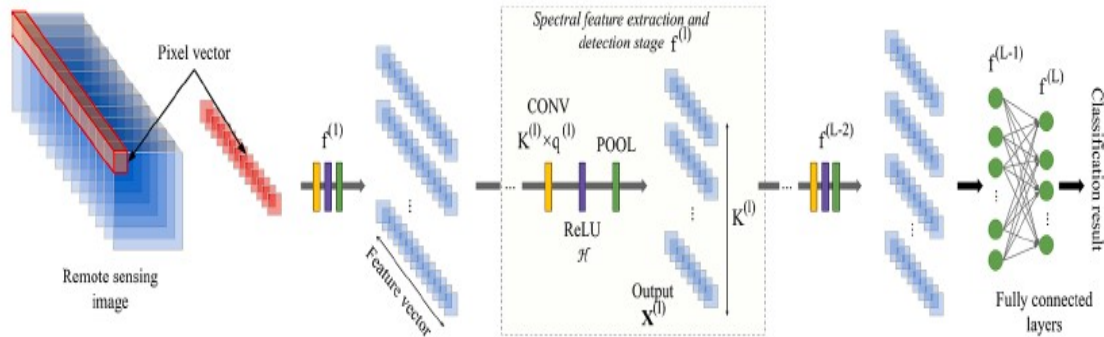


Figure 3: Traditional Spectral Convolutional Model Architecture Employed By 1D-CNN

2.3.1 Spectral CNN models

In spectral models like the one shown in Figure 3, the input data is spectral pixels $x_i \in \mathbb{N}^{n_{channels}}$, where $n_{channels}$ can be the amount of bands n_{bands} or an acceptable number of spectral channels n_{new} , retrieved using PCA, to which 1D kernels were implemented on every CONV layer, $K^{(l)} \times q^{(l)}$, resulting in an output $X^{(l)}$ composed of $K^{(l)}$ feature vectors. Li et al. [27] present a 1D-CNN framework for investigating spectral data associated between pixels by retrieving pixel pair features from the source information, with the central pixel and every one of its immediate neighbors as source. Likewise, Du and Li [28] build subtraction PPFs for HSI target detection, with the spectral difference between the central pixel and its surrounding pixels as the 1D-CNN model's input.

Mei et al. [29] add batch normalization layers, a dropout mechanism, and a novel nonlinear activation function to the 1D-CNN architecture, training the model using the pixel's wavelength, the spectral mean of nearby pixels, and the standard deviations per spectral band of adjacent pixels.

2.3.2 Spatial CNN models

When it comes to spatial models, they solely take into account spatial data from the HSI dataset. In this case, 2D-CNN architectures are commonly used to analyze spatial data, with each CONV layer applying $K^{(l)} \times k^{(l)} \times k^{(l)}$ kernels to the input data, yielding $K^{(l)}$ feature maps as shown in Figure 4. By decreasing the spectral dimension with a DR technique like PCA and reducing spatial regions of $d \times d$ pixel-centered neighbors, spatial data may be recovered from the primary HSI dataset.

Haut et al. [30] prepare a 2D-CNN with one PC, whereas Liang and Li [31] train a 2D-CNN with three PCs and use sparse coding to post-

process the recovered spatial features to produce a discrete lexicon of more reflective spatial characteristics for classification. The random patches network is a 2D-CNN framework presented by Xu et al. [32], in which input data is lightened using PCA and three PCs are considered. In addition, Zheng et al. [33] use a 2D-CNN to accomplish final classification with six PCs as input. Zhao et al. [34] present a 2D-CNN framework for retrieving deep spatial characteristics that uses a multiscale convolutional AE derived on the Laplacian pyramid, and a PCA to retrieve three PCs on the other. Then, using logistic regression as a classifier, the collected spatial characteristics are combined with the spectral data. Zhu et al. [35] have presented a 2D-CNN adaptive HSI classification network, made up of adaptive convolutions and down sampling that merge the surrounding metadata of each input sample adaptively.

2.3.3 Spectral-spatial CNN models

As shown in Figure 5, There are models that take into consideration both spectral and spatial characteristics from the HSI cube. Many different techniques and architectures may be devised to accomplish spectral-spatial processing because of the versatility of CNN models. 2D-CNN architectures with spectral-spatial handcrafted features can conduct the spectral-spatial processing.

He et al. [36], for example, use covariance matrices to train the 2D-CNN model, which store the spectral-spatial examples of various regions of 20 PCs, resulting in variationally correlation maps. Aptoula et al. [37] employ attribute profiles as input to the 2D-CNN model, using spatial-spectral information that APs may collect in an image at multiple scales. Yue et al. [38] create a 2D-CNN model to analyze spectral-spatial characteristics by

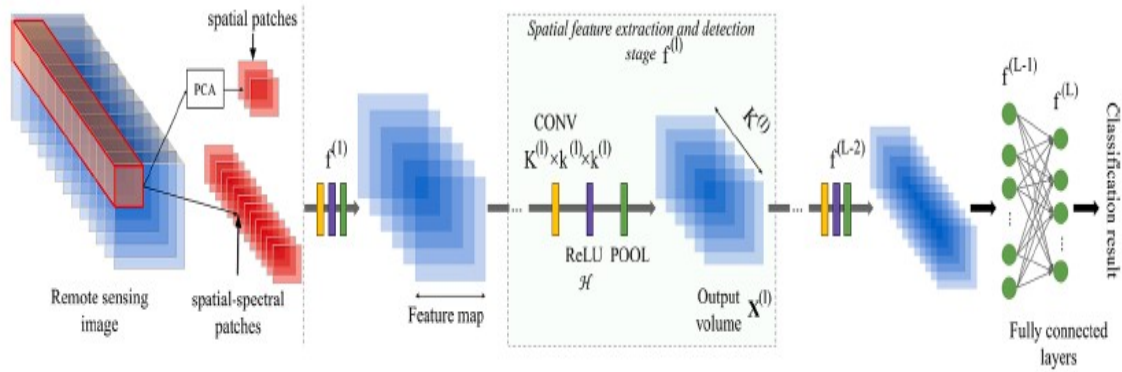


Figure 4: Traditional Spatial Convolutional Model Architecture Employed By 2D-CNN

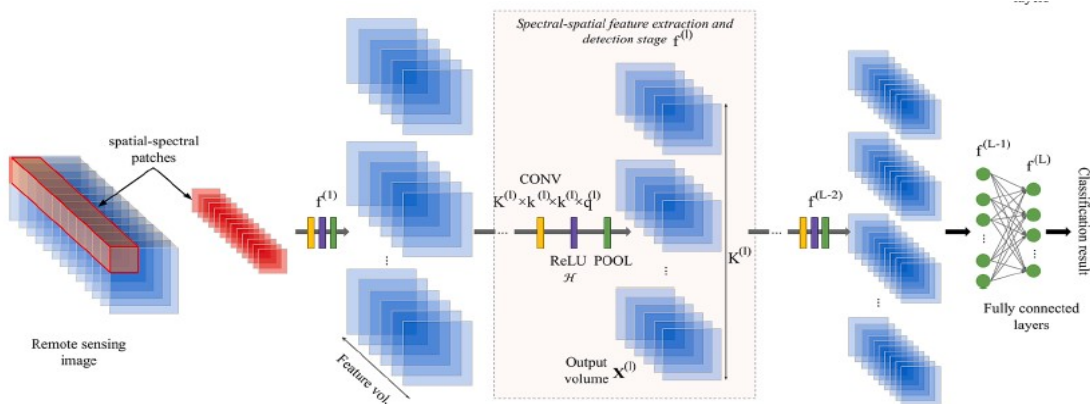


Figure 5: Traditional Spatial-Spectral Convolutional Model Architecture Employed By 3D-CNN

combining spectral information into three feature regions and appending them to spatial regions. As an example, Zhao and Du [39] presented a spectral-spatial feature-based classification framework that uses a 2D-CNN to find spatial information, while a symmetric local differential embed is used to retrieve spectral features. A SAE is used by Yue et al. [40] to retrieve spectral characteristics, while a 2D-CNN with spatial pyramid pooling performs a multiscale spatial FE. It has been shown that the hierarchical spatial-spectral features derived from the 1D-CNN1D or 2D-CNN may be used with Softmax regression classifiers to get the final classification.

For spectral-spatial classification, 3D-CNN is often used in addition to 1D-CNN and 2D-CNN. The 3-D filters of size $K^{(l)} \times k^{(l)} \times k^{(l)} \times q^{(l)}$ are capable of extracting high-level features in a natural fashion, resulting in $K^{(l)}$ feature volumes as output. Chen et al. [41] compare three types of convolutional models using HSI data as input blocks, while Li et al. [42]

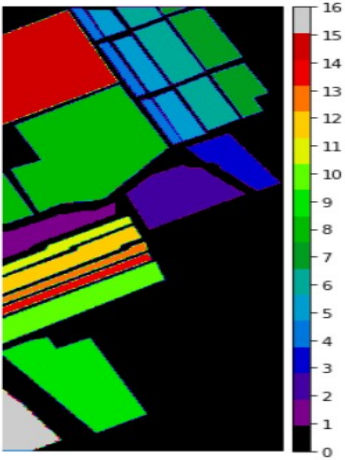
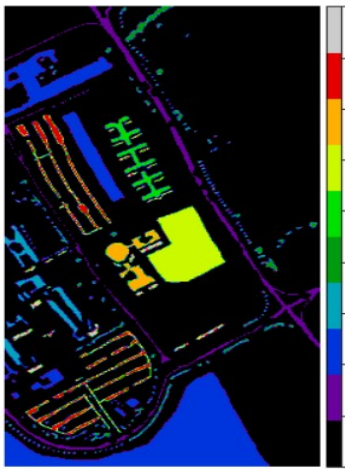
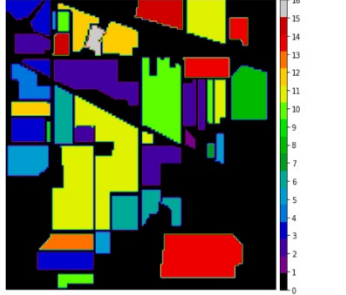
compare the spectral-spatial 3D-CNN framework, two spectral-based techniques, and the spatial 2D-CNN for HSI data categorization, highlighting the advantages of using convolutional models. The existing literature, like that for 1D-CNN and 2D-CNN models, offers increasingly complicated and advanced techniques for HSI processing utilizing 3D-CNN structures. Luo et al. [43], for example, create a hybrid CNN2D-3D architecture that can handle overfitting concerns by utilizing a three-dimensional kernel as the network's initial layer to retrieve relevant features from the original 3D inputs, which have a limited neighboring frame. The acquired features are subsequently reshaped into a single matrix, which is then sent to the next 2-D kernel, as well as the pooling and FC layers, resulting in a final classification.

As seen by recent articles, researchers are attempting to build deeper and more complicated networks by including novel configurations, additional nodes and links, finer optimizers and functionalities, as well as other innovations.

TABLE 1: Description of Popularly used Benchmark Hyperspectral Image Datasets

Parameters	Salinas (SA)	Pavia University (PU)	Indian Pines (IP)
Spatial Dimension	512 × 217	610 × 340	145 × 145
No. of Spectral Bands	200	115	200
No. of Classes	16	9	16
Wavelength Range	360–2500 μm	0.43–0.86 μm	0.4–2.5μm
Sensor	AVIRIS	ROSIS	AVIRIS

TABLE 2: Number of Available Samples in the Salinas Scene, Pavia University, and Indian Pines Datasets with Ground-Truths.

Salinas (SA)			Pavia University (PU)			Indian Pines (IP)		
								
S. No.	Land Cover Type	Samples	S. No.	Land Cover Type	Samples	S. No.	Land Cover Type	Samples
1	Broccoli-green-weeds-1	2009	1	Asphalt	6631	1	Alfalfa	46
2	Broccoli-green-weeds-2	3726	2	Meadows	18649	2	Corn-notill	1428
3	Fallow	1976	3	Gravel	2099	3	Corn-mintill	830
4	Fallow-rough-plow	1394	4	Trees	3064	4	Corn	237
5	Fallow-smooth	2678	5	Painted metal sheets	1345	5	Grass-pasture	483
6	Stubble	959	6	Bare Soil	5029	6	Grass-trees	730
7	Celery	3579	7	Bitumen	1330	7	Grass-pasture	28
8	Grapes-untrained	11271	8	Self-Blocking Brick	3682	8	Hay-windrowed	478
9	Soil-vinyard-develop	6203	9	Shadows	947	9	Oats	20
10	Corn-senesced-green	3278				10	Soybean-notill	972
11	Lettuce-romaine-4wk	1068				11	Soybean-mintill	2455
12	Lettuce-romaine-5wk	1927				12	Soybean-clean	593
13	Lettuce-romaine-6wk	916				13	Wheat	205
14	Lettuce-romaine-7wk	1070				14	Woods	1265
15	Vinyard-untrained	7268				15	Buildings-Grass	386
16	Vinyard-vertical-trellis	1807				16	Stone-Steel	93
Total Samples		54129	Total Samples		42776	Total Samples		10249

All category accuracies have a mean value of AA. The following equation (1) is for Kappa for more complete indicator:

$$\hat{K} = \frac{N \sum_{i=1}^r X_{ii} - \sum_{i=1}^r (X_{i+} \cdot X_{+i})}{N^2 - \sum_{i=1}^r (X_{i+} \cdot X_{+i})} \quad (1)$$

where N represents the total samples, r denotes the total number of categories, and X_{ij} is the $(i, j)^{th}$ value of the confusion matrix X, with X_{i+} and X_{+i} representing the summation of the i^{th} row and i^{th} column of X, respectively.

3. DATASETS DESCRIPTION AND PERFORMANCE INDICATORS

We compare the most common deep learning-based models to conventional machine learning based models after analyzing the key models and frameworks in order to evaluate the benefits and advantages that deep learning models may achieve in terms of reliability and classification accuracy. Towards this aim, the AVIRIS-collected Indian Pines (IP) and Salinas Valley (SV) sceneries, as well as the ROSIS-collected University of Pavia (UP) scene, were selected to produce the experimental component of the study. Table 1 illustrates the specifics of these data sets and Table 2 summarizes the number of labelled samples per category and the ground-truth information provided in distinct HSI datasets. Three evaluation measures, average accuracy (AA), overall accuracy (OA), and kappa coefficient (K), are used to assess classification results on benchmark data sets, as is typical in the literature. The overall sample count is equal to the number of correctly classified samples divided by the total number of samples.

4. EXPERIMENTAL RESULTS AND ANALYSIS

We look at three popular DL models for HSI classification in this section: CNN, SAE, and DBN, which employ spectral, spatial, and spectral-spatial data. Pavia University dataset is considered as an example for the comparisons. For CNN and

DBN-based classification, 10% and 50% labeled samples are randomly picked as training examples, respectively. For SAE-based classification, labeled samples are separated into three groups: learning, validating, and test set, using a 6:2:2 split ratio. Table 3 shows the classification findings. In terms of classification accuracy, spectral-spatial feature-based classification techniques showed the best performance, followed by spatial feature-based classification techniques in second place, and spectral feature-based classification techniques in third place, as shown in Table 3.

In HSI, some pixels belong to distinct objects and have the same spectral character, while others belong to the same item but have distinct spectral characteristics. It's challenging to identify those pixels only on the basis of spectral information. Furthermore, methods that simply consider spectral data frequently miss out on spatial pattern characteristics, having a negative impact. Spatial feature-based classification techniques use the spatial information in HSI to classify objects. However, during the dimensionality reduction process, some spectral information may be lost. Spatial-spectral feature-based categorization methods maximize HSI by combining spatial and spectral information. As a consequence, they produce the best results.

We evaluate five spatial-spectral feature-based classification models based on SAE, DBN, 2D-CNN, and 3D-CNN by separating the labelled samples 1:1 into train and test. Table 4 lists the experimental findings, whereas Figures 6, 7, and 8 illustrate the visual classification results for the IP, PU, and SA datasets, respectively. As demonstrated in Figures 6 through 8, CNN-based HSI spectral-spatial feature classification techniques, such as 3D-CNN and 2D-CNN, outperform SAE and DBN-based methods. It's worth noting that HSI classification techniques based on SAEs and DBNs are established before in history than those based on CNNs. However, statistically, the number of papers on the use of CNNs for HSI classification appears to be increasing at a quicker rate recently, and CNN performance is typically superior.

TABLE 3: Classification Accuracies (in %) on Pavia University Datasets.

Model	Spectral feature			Spatial feature			Spectral-spatial feature		
	OA	AA	Kappa ($\times 100$)	OA	AA	Kappa ($\times 100$)	OA	AA	Kappa ($\times 100$)
CNN [41]	92.28	92.55	90.37	94.04	97.52	92.43	99.54	99.77	99.56
SAE [44]	95.14	94.01	93.70	98.12	97.32	97.55	98.52	97.82	98.07
DBN [45]	96.42	95.09	95.30	98.62	97.95	98.19	99.05	98.48	98.75

TABLE 4: Classification Accuracies (in %) on PU, SA, and IP Datasets.

Models	Pavia University			Salinas			Indian Pines		
	OA	AA	Kappa	OA	AA	Kappa	OA	AA	Kappa
SAE [17]	89.35	88.28	0.8745	90.46	89.73	0.8924	98.64	98.92	97.56
DBN [21]	99.05	98.48	0.9875	99.27	98.92	0.9902	95.95	95.45	0.9539
2D-CNN [30]	99.09	99.53	0.9880	99.64	99.12	0.9942	97.50	98.66	0.9715
3D-CNN [42]	99.39	98.85	0.9920	99.82	99.36	0.9957	99.07	98.68	0.9893

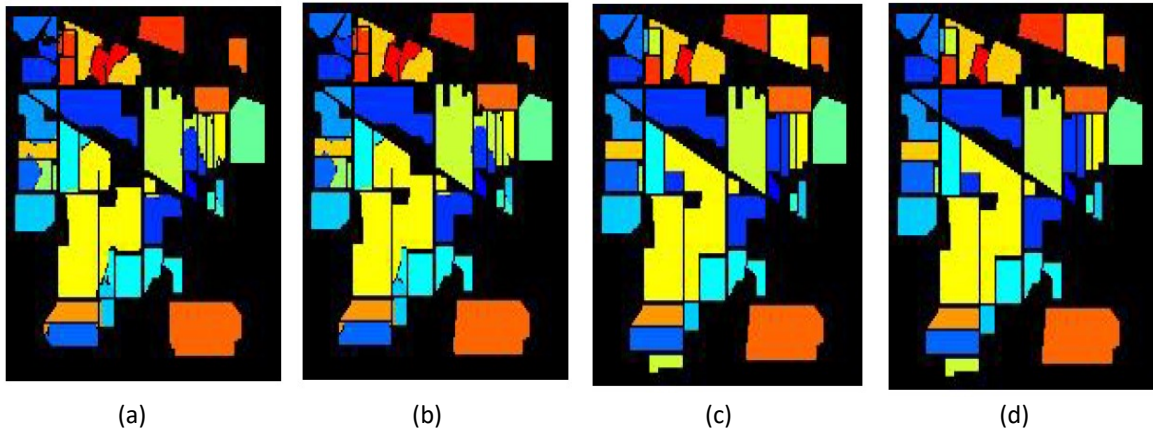


Figure 6: IP Dataset Classification Maps, (a) SAE, (b) DBN, (c) 2D-CNN, (d) 3D-CNN

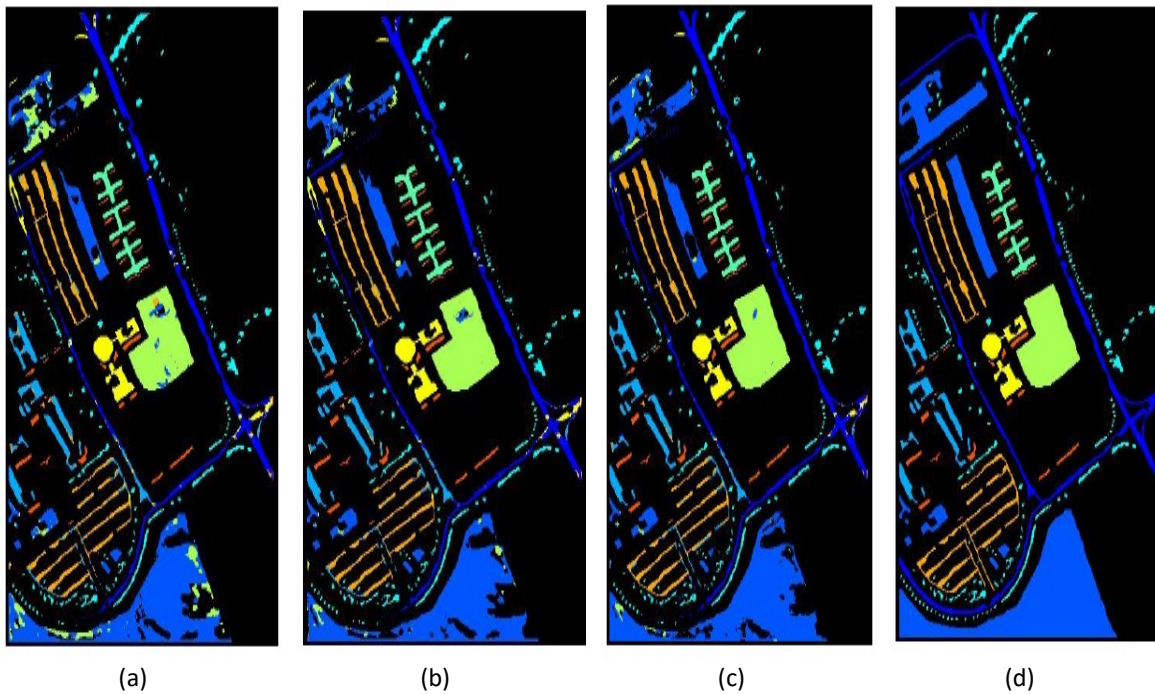


Figure 7: PU Dataset Classification Maps, (a) SAE, (b) DBN, (c) 2D-CNN, (d) 3D-CNN

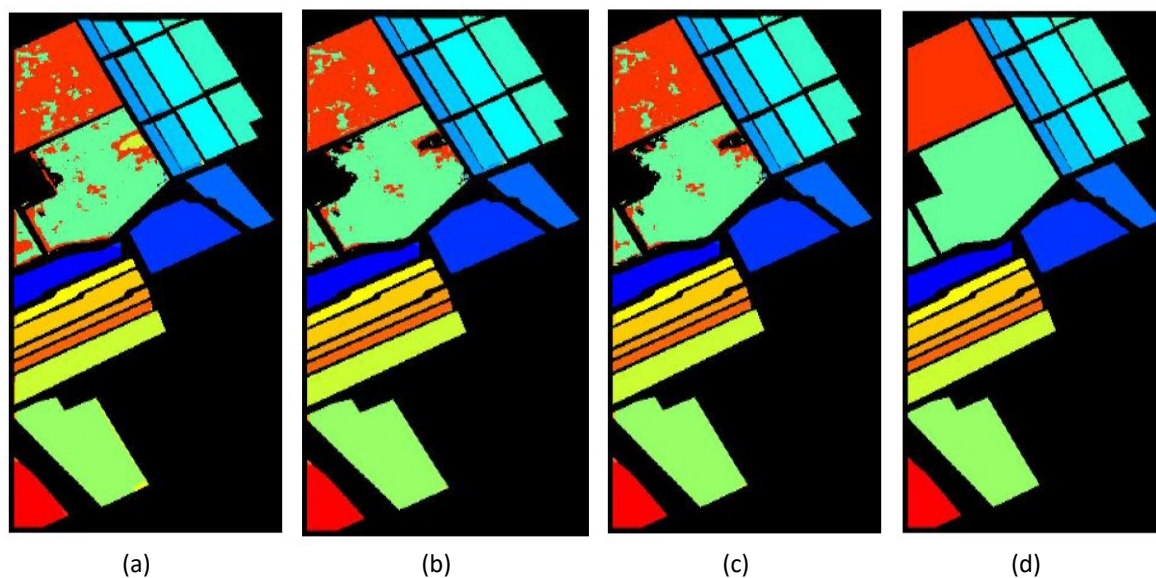


Figure 8: SA Dataset Classification Maps, (a) SAE, (b) DBN, (c) 2D-CNN, (d) 3D-CNN

5. CONCLUSION

We briefly discussed a variety of common deep learning approaches that can be utilized to conduct Hyperspectral image classification in this literature review, including CNNs, SAEs, and DBNs. We conducted a systematic assessment of present state-of-the-art deep learning methods for HSI classification. DL has been used to explore classification techniques based on spectral features, spatial features, and combined spectral and spatial features. We've also compared and examined the results of several common approaches. The usefulness of DL-based HSI classification approaches in tackling real-world issues has been demonstrated, however this accomplishment does not yet exhibit the maximum capabilities of DL. Because of the growing availability of Remotely sensed data and computing resources, fast progress of DL in remote sensing image classification is predicted in the future years. Nonetheless, there is indeed a still far to go in terms of achieving maximum potential when dealing with several challenging factors. Other HSI classification approaches have also been developed. These approaches have the potential to improve performance and should thus be considered. For example, Zhu et al. [46], suggested a multiple 3D feature fusion model to retrieve spectral-spatial characteristics, while Fang et al. [47], developed a unique fused network to utilize the data for HSI classification. Recent advancements aren't

addressed in depth in this paper, although they could be beneficial in the future.

REFERENCES

- [1] C.-I. Chang, Hyperspectral imaging: techniques for spectral detection and classification. Springer Science & Business Media, 2003, vol. 1.
- [2] M. Teke, H. S. Deveci, O. Haliloğlu, S. Z. Gürbüz, and U. Sakarya, "A short survey of hyperspectral remote sensing applications in agriculture," in 2013 6th International Conference on Recent Advances in Space Technologies (RAST). IEEE, 2013, pp. 171–176.
- [3] I. B. Strachan, E. Pattey, and J. B. Boisvert, "Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance," *Remote Sensing of environment*, vol. 80, no. 2, 2002, pp. 213–224.
- [4] S. Chabrilat, R. Milewski, T. Schmid, M. Rodriguez, P. Escribano, M. Pelayo, and A. Palacios-Orueta, "Potential of hyperspectral imagery for the spatial assessment of soil erosion stages in agricultural semi-arid Spain at different scales," in 2014 IEEE Geoscience and Remote Sensing Symposium. IEEE, 2014, pp. 2918–2921.
- [5] F. Salem, M. Kafatos, T. El-Ghazawi, R. Gomez, and R. Yang, "Hyperspectral image analysis for oil spill detection," in Summaries of

- NASA/JPL Airborne Earth Science Workshop, Pasadena, CA, 2001, pp.5–9.
- [6] S. Jay and M. Guillaume, "A novel maximum likely hood-based method for mapping depth and water quality from hyperspectral remote-sensing data," *Remote Sensing of Environment*, vol. 147, 2014, pp. 121–132.
- [7] P. Kuflik and S. R. Rotman, "Band selection for gas detection in hyperspectral images," in *2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel*. IEEE, 2012, pp. 1–4.
- [8] C. Jänicke, A. Okujeni, S. Cooper, M. Clark, P. Hostert, and S. van derLinden, "Brightness gradient-corrected hyperspectral image mosaics for fractional vegetation cover mapping in northern california," *Remote Sensing Letters*, vol. 11, no. 1, 2020, pp. 1–10.
- [9] X. Shang and L. A. Chisholm, "Classification of australian native forest species using hyperspectral remote sensing and machine-learning classification algorithms," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, 2013, pp. 2481–2489,
- [10] Z. Du, M. K. Jeong, and S. G. Kong, "Band selection of hyperspectral images for automatic detection of poultry skin tumors," *IEEE Transactions on Automation Science and Engineering*, vol. 4, no. 3, 2007, pp. 332–339.
- [11] Y. H. El-Sharkawy and S. Elbasuney, "Hyperspectral imaging: A new prospective for remote recognition of explosive materials," *Remote Sensing Applications: Society and Environment*, vol. 13, 2019, pp. 31–38,
- [12] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing dataanalysis and future challenges," *IEEE Geoscience and remote sensing magazine*, vol. 1, no. 2, 2013, pp. 6–36.
- [13] G. Cheng, J. Han, and X. Lu, "Remote sensing image scene classification: Benchmark and state of the art" *Proceedings of the IEEE*, vol.105, no. 10, 2017, pp. 1865–1883.
- [14] Plaut, Elad. "From principal subspaces to principal components with linear autoencoders." *arXiv preprint arXiv:1804.10253* (2018).
- [15] Zhu, Jiang, et al. "Auto-encoder based for high spectral dimensional data classification and visualization." *2017 IEEE Second International Conference on Data Science in Cyberspace (DSC)*. IEEE, 2017.
- [16] Özdemir, A. Okan Bilge, B. Ekin Gedik, and C. Yasemin Yardımcı Çetin. "Hyperspectral classification using stacked autoencoders with deep learning." *2014 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. IEEE, 2014.
- [17] Mughees, Atif, and Linmi Tao. "Efficient deep auto-encoder learning for the classification of hyperspectral images." *2016 International Conference on Virtual Reality and Visualization (ICVRV)*. IEEE, 2016.
- [18] Paul, S., Kumar, D.N., Spectral-spatial classification of hyperspectral data with mutual information based segmented stacked autoencoder approach. *ISPRS J. Photogramm. Remote Sens.* 138, 265–280, 2018.
- [19] Zhou, Peicheng, et al. "Learning compact and discriminative stacked autoencoder for hyperspectral image classification." *IEEE Transactions on Geoscience and Remote Sensing* 57.7 (2019): 4823-4833.
- [20] Li, T., Zhang, J., Zhang, Y., Classification of hyperspectral image based on deep belief networks. In: *Proc. IEEE Int. Conf. Image Proces.* 2014, pp. 5132–5136.
- [21] Chen, Y., Zhao, X., Jia, X., Spectral-Spatial Classification of Hyperspectral Data Based on Deep Belief Network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8 (6), 2015, 2381–2392.
- [22] Le, J.H., Yazdanpanah, A.P., Regentova, E.E., Muthukumar, V., A deep belief network for classifying remotely-sensed hyperspectral data. In: *Bebis, G., Boyle, R., Parvin, B., Koracin, D., Pavlidis, I., Feris, R., McGraw, T., Elenndt, M., Kopper, R., Ragan, E., Ye, Z., Weber, G. (Eds.), Advances in Visual Computing*. Springer International Publishing, Cham, 2015. pp. 682–692.
- [23] Zhong, P., Gong, Z., Li, S., Schönlieb, C.B., Learning to diversify deep belief networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* 55 (6), 2017, 3516–3530.
- [24] Guofeng, T., Yong, L., Lihao, C., Chen, J., June 2017. A dbn for hyperspectral remote sensing image classification. In: *2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. pp. 1757–1762.
- [25] Zhou, X., Li, S., Tang, F., Qin, K., Hu, S., Liu, S., Deep learning with grouped features for spatial spectral classification of hyperspectral images. *IEEE Geosci. Remote Sens. Lett.* 14 (1), 2017, 97–101.

- [26] Murugan, P., Feed forward and backward run-in deep convolution neural network. arXiv preprint arXiv:1711.03278, 2017.
- [27] Li, W., Wu, G., Zhang, F., Du, Q., Hyperspectral image classification using deep pixel-pair features. *IEEE Trans. Geosci. Remote Sens.* 55 (2), 2017, 844–853.
- [28] Du, J., Li, Z., A hyperspectral target detection framework with subtraction pixel pair features. *IEEE Access* 6, 2018, 45562–45577.
- [29] Mei, S., Ji, J., Bi, Q., Hou, J., Du, Q., Li, W., Integrating spectral and spatial information into deep convolutional neural networks for hyperspectral classification. In: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 5067–5070.
- [30] Haut, J., Paoletti, M., Plaza, J., Plaza, A., Hyperspectral image classification using random occlusion data augmentation. *IEEE Geosci. Remote Sens. Lett.* 2019.
- [31] Liang, H., Li, Q., Hyperspectral imagery classification using sparse representations of convolutional neural network features. *Remote Sens.* 8 (2), 2016, 99.
- [32] Xu, Y., Du, B., Zhang, F., Zhang, L., Hyperspectral image classification via a random patches network. *ISPRS J. Photogramm. Remote Sens.* 142, 2018, 344–357.
- [33] Zheng, Z., Zhang, Y., Li, L., Zhu, M., He, Y., Li, M., Guo, Z., He, Y., Yu, Z., Yang, X., Liu, X., Luo, J., Yang, T., Liu, Y., Li, J., Classification based on deep convolutional neural networks with hyperspectral image. In: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2017, pp. 1828–1831.
- [34] Zuo, Z., Shuai, B., Wang, G., Liu, X., Wang, X., Wang, B., Chen, Y., Convolutional recurrent neural networks: Learning spatial dependencies for image representation. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2015, pp. 18–26.
- [35] Zhu, J., Fang, L., Ghamisi, P., Deformable convolutional neural networks for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* 15 (8), 2018, 1254–1258.
- [36] He, N., Paoletti, M.E., Haut, J.n.M., Fang, L., Li, S., Plaza, A., Plaza, J., Feature extraction with multiscale covariance maps for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* 2018, 1–15.
- [37] Aptoula, E., Ozdemir, M.C., Yanikoglu, B., Deep learning with attribute profiles for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* 13 (12), 2016, 1970–1974.
- [38] Yue, J., Zhao, W., Mao, S., Liu, H., Spectral-spatial classification of hyperspectral images using deep convolutional neural networks. *Remote Sens. Lett.* 6 (6), 2015, 468–477.
- [39] Zhao, W., Du, S., Spectral-spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach. *IEEE Trans. Geosci. Remote Sens.* 54 (8), 2016, 4544–4554.
- [40] Yue, J., Mao, S., Li, M., A deep learning framework for hyperspectral image classification using spatial pyramid pooling. *Remote Sens. Lett.* 7 (9), 2016, 875–884.
- [41] Chen, Y., Jiang, H., Li, C., Jia, X., Ghamisi, P., Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* 54 (10), 2016, 6232–6251.
- [42] Li, Y., Zhang, H., Shen, Q., Spectral-spatial classification of hyperspectral imagery with 3d convolutional neural network. *Remote Sens.* 9 (1), 2017, 67.
- [43] Luo, Y., Zou, J., Yao, C., Zhao, X., Li, T., Bai, G., Hsi-cnn: A novel convolution neural network for hyperspectral image. In: 2018 International Conference on Audio, Language and Image Processing (ICALIP), 2018, pp. 464–469.
- [44] Chen, Y., Lin, Z., Zhao, X., Wang, G., & Gu, Y. Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6), 2017, 2094–2107.
- [45] Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. Deep learning in remote sensing: A review. arXiv:1710.03959, 2017.
- [46] Zhu, J., Hu, J., Jia, S., & Li, Q. Multiple 3-d feature fusion framework for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, PP (99), 2018, 1–14.
- [47] Fang, L., He, N., Li, S., Ghamisi, P., & Benediktsson, J. A. Extinction profiles fusion for hyperspectral images classification. *IEEE Transactions on Geoscience and Remote Sensing*, PP (99), 2018, 1–13.