Deep Learning Patch-Based Approach for Hyperspectral Image Classification

Papia F. Rozario,
Dept. of Geography and Anthropology
University of Wisconsin-Eau Claire
Eau Claire, WI, USA
rozaripf@uwec.edu

Eleana Ruehmann
Dept. of Geography and Anthropology
University of Wisconsin-Eau Claire
Eau Claire, WI, USA
ruehmael2663@uwec.edu

Tyler Pham
Dept. of Computer Science
University of Minnesota-Twin Cities
Minneapolis, MN, USA
pham0443@umn.edu

Tianqi Sun
Dept. of Computer Science
University of Wisconsin-Eau Claire
Eau Claire, WI, USA
sunt0034@uwec.edu

Jacob Jensen
Dept. of Math and Computer Science
Minot State University
Minot, ND, USA
jacob.d.jensen@minotstateu.edu

Hengrui Jia
Dept. of Math, Statistics,
and Computer Science
Macalester College
Saint Paul, MN, USA
hjia@macalester.edu

Zhongyue Yu
Dept. of Computer Science
University of Wisconsin-Eau Claire
Eau Claire, WI, USA
yuz1472@uwec.edu

Rahul Gomes,
Dept. of Computer Science
University of Wisconsin-Eau Claire
Eau Claire, WI, USA
gomesr@uwec.edu

***Abstract*—Classification of hyperspectral images is an important step of image interpretation from high spatial resolution imagery. Different studies demonstrate that spatial features can provide complementary information for increasing the accuracy of hyperspectral image classification. In this study, we evaluate different methods of spectral-spatial classification of hyperspectral images that are based on denoising methods using convolutional autoencoders. The resulting high-dimensional vectors of spectral features are classified by supervised algorithms such as support vector machine (SVM), maximum likelihood (ML), and random forest (RF). The experiments are performed on several widely known hyperspectral images that reveal a patch-based 3D convolutional autoencoder is more effective in reducing noise in the dataset and retaining spectral-spatial information. Random Forest classifier provides the highest classification accuracy across all the models.**

***Index Terms*—Autoencoders, UNet, Convolution Neural Networks, Hyperspectral Imagery.**

I. INTRODUCTION

HYPERSPECTRAL imagery (HSI) is used across many different fields including agriculture [1], biology [2], and forestry [3]. Hyperspectral imaging is the application of analyzing a wide spectrum of light rather than just assigning primary colors to each individual pixel. The light in each of these pixels is broken down into many individual spectral bands. While more bands are able to provide information pertaining to a wide array of objects, processing such a large dataset is challenging. Neural Networks have greatly improved the spectral spatial classification of hyperspectral images in remote sensing for the betterment of knowledge discovery [4], [5]. However, with the growing complexity of these networks, these solutions have necessitated the requirement of sophisticated computational hardware, thereby limiting their application. Traditional machine learning models are a suitable alternative, but different algorithms have limitations to express non-linearity thereby increasing the need for human intervention.

Imagery derived from hyperspectral satellites collects information in the form of very fine spectral bands covering a wavelength range in the electromagnetic spectrum. This is far more compared to multispectral images used for a variety of factors in remote sensing like change detection [6]–[8], crop analysis [9]–[11], and monitoring natural disasters [12]. This provides a huge volume of data and the possibility of distinguishing materials spectrally making the above-mentioned tasks far more accurate. Particularly, crop analysis and vegetation identification amongst others can be classified accurately. In traditional hyperspectral classification systems, classifiers are only able to consider spectral signatures without considering the correlations between the pixel of interest and its neighboring pixels [13]. Numerous classification techniques have been developed such as the K-nearest-neighbor classifier [14], maximum-likelihood estimation [14], [15], artificial neural networks [16], and kernel-based techniques [16]. In particular, support vector machines [15], [16] have demonstrated excellent performance for such classification. However, it is a very challenging task due to the tiny distinction among spectral signatures of various types in the same families, such as tillage in the corn fields. There are some spectral-spatial classifiers developed for Features-level fusion. For example, Generalized Composite Kernel (GCK) for a combination of both spectral and spatial information were employed by multinomial logistic regression, and support vector machine are introduced in [15], [17]–[19]. In addition to the composite classifier framework, much research focuses on spatial feature extraction.

Redundancy can be all throughout an HSI which can make feature extraction less effective. Thus, reducing the redundancy can be beneficial to do before extracting the features from the image. Plotting images for exploring the average relative brightness can be useful in selecting important bands from the HSIs, but this does not solve the redundancy problem of the HSI. The highest variant bands can be very similar to each other or correlated to one another. This would be redundant to choose bands that share very similar values or features. To counteract this, getting rid of correlated bands by taking bands that correlate a lot and getting rid of the least important one out of the two is advisable. This can be continued until the desired number of bands is reached or the specified correlation is met. The way to determine if a pair of bands are correlated is to use the correlation coefficient. The correlation coefficient gives a value between 0 and 1 with 0 not being correlated at all and 1 being completely correlated. There are a couple of methods for comparing the bands to one another and reducing the band set. The first method is adjacent correlation comparison which takes adjacent bands and compares them to get a correlation coefficient. The second method is all correlation comparison. This method compares one band with every other band minus itself and produces a list of correlation coefficients.

While single-band comparisons can reveal significant patterns, they are not able to depict non-linearity as extensively as neural networks. In this research, we explore how different variations of neural networks perform on the classification of HSI. Neural networks have been used as an autoencoder to reduce noise in the dataset followed by the application of machine learning for classification.

II. PREVIOUS WORK

A. *Denoising methods used in HSI*

HSI is often processed to reduced dimensionality and researchers have explored different pathways to address this issue. The following are examples of denoising without application of deep learning. These are mostly specific to the dataset making transfer learning challenging. The authors in [20] proposed a HSI denoising method based on obvious redundancy and correlation. This is a noise reduction method based on joint spectral-spatially distributed sparse representation. According to the experimental results and analysis, this method can greatly improve the denoising performance. The highest Overall Accuracy (OA) and kappa coefficient values rose by 12.24% and 0.1433 respectively after the application of this method. The method exploits the intra-band structure and inter-band correlation during joint sparse representation and joint dictionary learning. In joint spectral-spatial sparse coding, inter-band correlations are exploited to capture similar structures and preserve spectral continuity. The authors in [21] proposed a multi-linear algebra approach to jointly achieve denoising and dimensionality reduction. This method is mainly based on spatial low-rank approximation and spectral dimensionality reduction. From the definition of matrix rank, we know that if the image is regarded as a matrix, the less the number of its bases, the less the number of linearly independent vectors corresponding to the bases, and the smaller the rank of the matrix. We can usually use low rank for image denoising. Principal Component Analysis (PCA) is the most popular dimensionality reduction method. According to the experimental results, the OA nearly increased 5% by using this method. Zhang et. al. [22] proposed a new HSI Restoration method using Low-Rank Matrix Restoration (LRMR). An HSI denoising method was proposed by jointly exploiting local/global redundancy and correlation (RAC) of HSI in spatial and spectral domains. This method removes noise by modeling RAC and by learning a sparse approximation of the dictionary data. But using only the local RAC in the spectral domain might cause spectral distortion so a low-rank constraint was used to deal with global RAC in the spectral domain. The experimental results had no noise and visually clearer than the traditional method of video denoising by sparse 3D(VBM3D). The OA was over 85% [23]. In [24] the authors proposed a low-rank tensor recovery (LRTR) method to remove mixed noise in HSI data. Comparing to the traditional low-rank tensor decomposition methods, the LRTR method can preserve the global structure of HSI while removing Gaussian noise and sparse noise. The experimental results also show that the proposed LRTR method can achieve promising performance only in removing Gaussian noise. However, LRTR still could make high OA and Class-Specific Accuracy (CA) which were most over 80%.

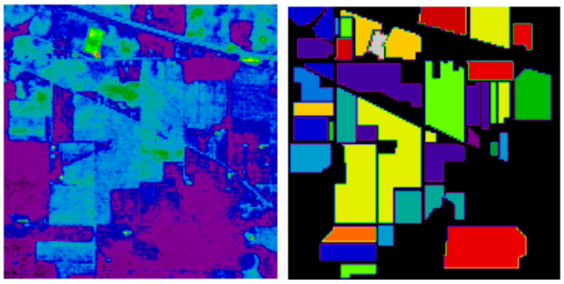

Fig. 1: Indian Pines dataset (**a**) Output from a single band and ,(**b**) classes in the dataset

III. METHODOLOGY

A. *Dataset*

For this research, public access datasets of Indian Pines, Pavia Center, and Pavia University were used. Indian Pines dataset has a size of (145×145) and has 199 useful bands ranging from a wavelength of 400nm to 2500nm. The dataset consists of two-thirds agriculture and one-third forest land and other natural vegetation. There are some but minimal manmade structures present in the dataset's scene. The 16

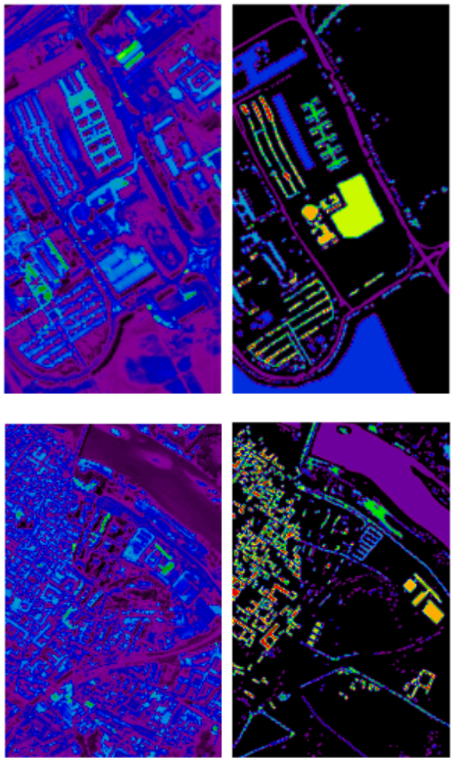

Fig. 2: Pavia University dataset (**a**) Output from a single band and (**b**) classes in the dataset. Pavia Center dataset (**c**) Output from a single band and (**d**) classes in the dataset

classes of the ground truth consist of crops such as corn, alfalfa, soybean, wheat, etc. Most of the crops listed in the ground truth classes include variations of how the crops were planted or tended to. Figure 1a shows the Indian Pines dataset and Figure 1b shows its corresponding classes.

Classes	Indian Pines	Pavia University	Pavia Center
0	Alfalfa	Asphalt	Water
1	Corn-notill	Meadows	Trees
2	Corn-mintill	Gravel	Asphalt
3	Corn	Trees	Self-Blocking Bricks
4	Grass-pasture	Painted metal sheets	Bitumen
5	Grass-trees	Bare Soil	Tiles
6	Grass-pasture-mowed	Bitumen	Shadows
7	Hay-windrowed	Self-Blocking Bricks	Meadows
8	Oats	Shadows	Bare Soil
9	Soybean-notill		
10	Soybean-mintill		
11	Soybean-clean		
12	Wheat		
13	Woods		
14	Buildings-Grass-Trees-Drives		
15	Stone-Steel-Towers		

TABLE I: Classes in the three datasets

Pavia Center and Pavia University, located in northern Italy, both have a groundtruth of nine classes such as asphalt, trees, painted metal sheets, etc. Pavia Center dataset has a size of 1096 × 715 with 102 bands ranging from a wavelength of 430 to 860 nm. Pavia University has a dataset size of 610 × 340 with 103 bands that range from a wavelength of 430 to 860 nm. Similar to Pavia Center, Pavia University has a ground truth size of nine classes. Figure 2a and 2c shows both Pavia University and Pavia Center datasets. Figure 2b and 2d show their corresponding classes. Table I shows all the classes in these three datasets.

Figure 3 shows the wavelength of these bands after the removal of noisy data. Information about classes can be found in this repository [25]. To produce this image each pixel is split into different classification categories. Each pixel has its own set of brightness values for each band or channel, so the brightness value can be averaged among all of the pixels in a given classification category. This is done to each band until getting a set of average relative brightness values for each classification category is. The graphs in Fig 1 can show which bands are the most variant.

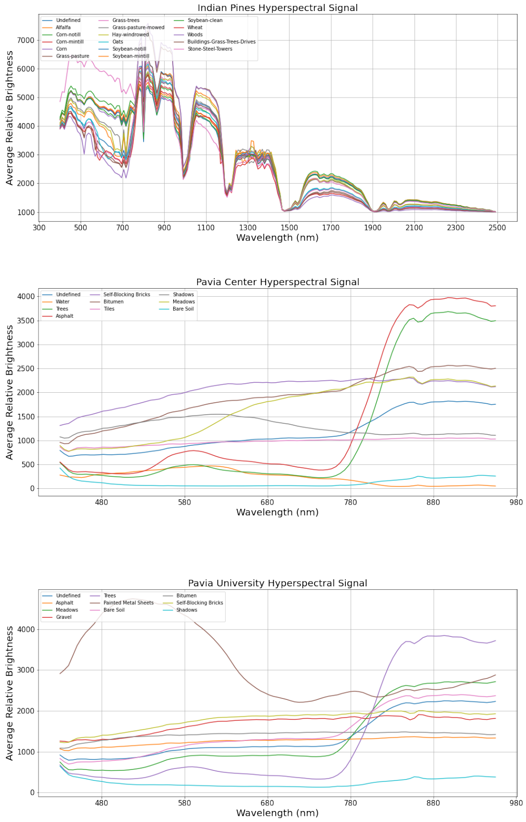

Fig. 3: Wavelength distribution of HSI bands (**a**) Indian Pines (**b**) Pavia Center and (**c**) Pavia University

B. Convolutional Neural Networks

Convolutional Neural Network (CNN) is an important part of deep learning focusing on visual images [26], [27]. It takes an input image and then analyzes, or more explicitly segments, the image according to the whole image or parts of the image's features, then gives an output that classifies the whole image or each part of the image into a specific category. CNN includes three layers: the convolutional layer, the pooling layer, and the fully connected layer. The convolutional layer is the layer that can extract local features from images. Networks with more layers can iteratively extract more complex features from lower-layer features [28].

CNNs can be used as Autoencoders (AE) which are unsupervised artificial neural network (ANN) that takes an input vector and reduces it to a lower dimensional space then reconstructs it back to the original input vector. An image can be processed by AE by taking the image and passing it through layers of downsampling then upsampling to regenerate the same image. This can be useful for extracting features from HSI. The features that are most important about the HSI would hopefully be preserved by the AE, and then can be used for classification. It would reduce the noise in the HSI. For this research, a special type of CNN called UNet [29] was used as the AE.

C. Machine learning

1) Maximum Likelihood: Maximum likelihood estimation is a mathematical method to determine the parameter values for a certain mathematical model so that the model can produce results that are closer to the collected and observed data, or in other words, the model can be as accurate as possible. Suppose we have collected some data points which obey the Gaussian (normal) distribution, the way to calculate the maximum likelihood estimates is given in the equation below.

$$P(x;\ \mu,\ \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

Here 'x' is the data points we have observed. The semicolon after 'x' is used to denote that 'x' is the data points we have collected and the μ and σ are the parameters we need to determine. The goal is to find the value of μ and σ so that the expression above can produce its maximum value [30].

2) Support Vector Machines: A support vector machine is a supervised machine learning model that comes with the associated learning algorithms that analyze data for classification and regression analysis. The objective of the support vector machine algorithm is to find a hyperplane in N-dimensional space (where N is the number of features) that has the maximum margin between data points and both classes so that it can distinctly classify the data points [31].

3) Random Forest: In machine learning, a random forest is a classifier that contains multiple decision trees, and its output class is determined by the mode of the class output by individual trees. The Random Forest algorithm has three main hyperparameters that need to be set before training. These include node size, number of trees, and number of features sampled. Random Forest classifiers can be used to solve regression or classification problems in remote sensing [32]–[34]. Each decision tree is constructed in a random forest using a random subset of the training data and a random subset of the input features. This randomization helps to reduce overfitting and increase the accuracy of the model. When making a prediction, the random forest algorithm aggregates the predictions of all the decision trees in the forest, either by majority vote (for classification problems) or by averaging (for regression problems).

IV. EXPERIMENTS

For processing the Indian Pines data set with 199 hyperspectral bands, a UNet-based autoencoder was used with three convolution blocks each having two convolution layers. Batch normalization along with Relu as an activation function was used. Dropout was selected at 7% with a kernel size of (3×3). Since the UNet is being used to reduce noise in the data set, mean squared error was used as the loss function. A similar approach was used for both the Pavia center and Pavia university datasets. Since the image size was significantly larger, the number of kernels used was larger. For example, the Indian Pines used 64. 128, and 256 kernels in three convolution blocks while the Pavia Center and University utilized 256, 512, and 1024 kernels respectively. Details about training time as well as other parameters are shown in Table II. All three UNet models were trained for 100 epochs.

TABLE II: Training Parameters of UNet Autoencoders

	Indian Pines		**Pavia Center**		**Pavia University**	
Type	2D	3D	2D	3D	2D	3D
Patch Applied	No		No	(136, 88, 104)	No	(136, 88, 104)
Parameters (millions)	4.56	13.28	137.15	13.28	137.16	13.28
Training Time(s)	5.342	261.04	272.86	4898	82.6	1230.19
Optimizer	Adam					
Loss	Mean Squared Error					
Epoch	100					
Dropout	0.07					
GPU	Nvidia Tesla V100S (32 GB)					

Figure 4 shows the difference between convolution operation for 2D and 3D variants. For the 3D approach, kernels of size $(3\times3\times3)$ were used. The number of filters in the Indian Pines dataset was set to 64 like the 2D counterpart with the same number of convolution blocks. The number of training parameters was higher considering that the 3D approach uses both spectral and spatial features from HSI. It was also trained for 100 epochs. The HSI was preprocessed to convert its dimension from $(145\times145\times199)$ to $(144\times144\times200)$. This ensured that the image was divisible by a 3D convolution kernel. For the Pavia Center and Pavia University datasets, a different approach was taken due to their size. For example, Pavia Center had a resolution of $(1096\times715\times102)$ which was too large for a 3D-UNet model. The Pavia Center was resampled to $(1088\times704\times104)$. This was followed by creating 3D patches with a resolution of $(136\times88\times104)$. A total of 64 patches were generated. Pavia University was resampled from $(610\times340\times103)$ to $(544\times352\times104)$. 16 different 3D patches of $(136 \times 88 \times 104)$ were then used to train the UNet model. Since the patch sizes were quite similar to the Pavia Center, the number of kernels used for each downsampling layer was 64, 128, and 256. After the training process, these patches were stitched back and resampled to generate Pavia Center and Pavia University with the original dimensions.

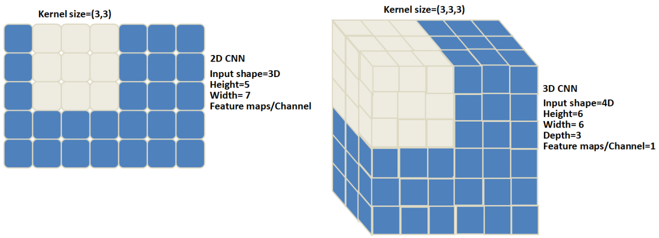

Fig. 4: Example 2D vs 3D convolution with a kernel of size 3. For Pavia Center and Pavia University, 3D patches were generated prior to AE application. Feature maps/Channel varied based on UNet layers.

TABLE III: Accuracy and Kappa values of the Machine Learning models after application of CNN Autoencoders

CNN Type	**Accuracy**			**Kappa**		
			Pavia Center			
	ML	**RF**	**SVM**	**ML**	**RF**	**SVM**
2D	79.9 %	97.34	96.26	0.73	0.96	0.95
3D	84.18	98.1	97.34	0.78	0.97	0.96
			Pavia University			
2D	51.74	82.83	77.39	0.42	0.77	0.68
3D	55.65	87.76	77.11	0.44	0.84	0.68
			Indian Pines			
2D	41.81	66.32	47.14	0.35	0.61	0.35
3D	54.18	84.6	66.01	0.48	0.82	0.59

Three machine learning algorithms were used to test the efficacy of the proposed approach. These algorithms reported significant differences as well. A total of 10% random samples were used as a training dataset to build these classification models. Results of the algorithms on the remaining 90% of the test data are shown in Table III. Our first observation is that the 3D approach to image denoising performed significantly well compared to the 2D approach across all algorithms. It was also observed that the Random Forest algorithm performed best when compared to Maximum Likelihood and Support Vector Machines using Radial Basis Kernel. The Random Forest model was built using 100 trees.

Tables IV, V, and VI display per-class precision, recall, and F1-score for all the three algorithms corresponding to the Random Forest output. Using the 3D approach significantly increased the per-class performance. For example, Corn-notill, Corn-mintill, and Corn almost doubled their precision metrics in the Indian Pines dataset. This shows that the Random Forest approach coupled with 3D spectral-spatial denoising is able to identify subtle differences in agriculture signature. Performance improvements were also observed for urban areas in Pavia University. For example, the precision for classifying Bitumen increased from 0.69 to 0.85. Pavia Center reported minimal improvements as well. The precision of Asphalt increased from 0.83 to 0.91. This shows that the patch-based approach has a great potential to extract information from HSI.

TABLE IV: Metrics from Random Forest output on Indian Pines dataset

Classes	**Precision**		**Recall**		**F1-Score**		**Support**
	2D	**3D**	**2D**	**3D**	**2D**	**3D**	
0	0.58	1	0.45	0.7	0.51	0.82	40
1	0.57	0.78	0.55	0.83	0.56	0.8	1285
2	0.47	0.77	0.44	0.66	0.46	0.71	738
3	0.48	0.84	0.29	0.66	0.36	0.74	207
4	0.85	0.92	0.7	0.86	0.77	0.89	429
5	0.74	0.9	0.89	0.9	0.81	0.9	663
6	0.58	0.94	0.46	0.71	0.51	0.81	24
7	0.92	0.97	0.96	1	0.94	0.98	423
8	1	0.9	0.28	0.5	0.43	0.64	18
9	0.58	0.84	0.5	0.8	0.54	0.82	873
10	0.61	0.8	0.77	0.87	0.68	0.84	2228
11	0.42	0.78	0.26	0.65	0.32	0.71	535
12	0.82	0.91	0.8	0.94	0.81	0.92	186
13	0.89	0.93	0.93	0.99	0.91	0.96	1141
14	0.71	0.92	0.46	0.85	0.56	0.88	350
15	0.92	0.98	0.27	0.98	0.42	0.98	85
						Total	9225

TABLE V: Metrics from Random Forest output on Pavia University dataset

Classes	**Precision**		**Recall**		**F1-Score**		**Support**
	2D	**3D**	**2D**	**3D**	**2D**	**3D**	
0	0.8	0.89	0.85	0.85	0.82	0.87	6015
1	0.86	0.89	0.94	0.97	0.9	0.93	16778
2	0.74	0.83	0.44	0.71	0.55	0.77	1882
3	0.96	0.95	0.84	0.82	0.89	0.88	2759
4	0.99	0.99	0.99	0.99	0.99	0.99	1197
5	0.69	0.85	0.53	0.67	0.6	0.75	4530
6	0.87	0.81	0.83	0.82	0.85	0.82	1196
7	0.67	0.74	0.72	0.84	0.69	0.78	3301
8	1	1	0.99	0.99	0.99	1	841
						Total	38499

TABLE VI: Metrics from Random Forest output on Pavia Center dataset

Classes	**Precision**		**Recall**		**F1-Score**		**Support**
	2D	**3D**	**2D**	**3D**	**2D**	**3D**	
0	1	1	1	1	1	1	59415
1	0.94	0.96	0.93	0.96	0.94	0.96	6824
2	0.83	0.91	0.85	0.9	0.84	0.9	2781
3	0.76	0.79	0.8	0.78	0.78	0.78	2397
4	0.92	0.93	0.87	0.96	0.9	0.94	5915
5	0.91	0.93	0.97	0.96	0.94	0.94	8350
6	0.96	0.95	0.87	0.9	0.91	0.93	6607
7	0.99	1	0.99	0.99	0.99	1	38465
8	1	1	0.99	1	0.99	1	2583
						Total	133337

V. CONCLUSION

The preliminary findings in this research support that a patch-based 3D convolution is both scalable and reduces noise in HSI. Random Forest has superior classification accuracy when compared to other machine-learning models. Future work will look into feature selection methods coupled with the impact of kernel dimensions on classification performance. Feature selection may further reduce training data prior to the application of deep learning thereby reducing resources hopefully without impacting classification accuracy.

VI. Acknowledgment

This research project was funded by the National Science Foundation (NSF)-Research Experience for Undergraduates OAC- 2150191. The computational resources of the study were provided by the Blugold Center for High-Performance Computing under NSF grant CNS-1920220.

References

[1] Bing Lu, Phuong D Dao, Jiangui Liu, Yuhong He, and Jiali Shang. Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sensing*, 12(16):2659, 2020.

[2] B Weber, C Olehowski, T Knerr, J Hill, K Deutschewitz, DCJ Wessels, B Eitel, and B Büdel. A new approach for mapping of biological soil crusts in semidesert areas with hyperspectral imagery. *Remote Sensing of Environment*, 112(5):2187–2201, 2008.

[3] Michele Dalponte, Lorenzo Bruzzone, Loris Vescovo, and Damiano Gianelle. The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas. *Remote Sensing of Environment*, 113(11):2345–2355, 2009.

[4] Frédéric Ratle, Gustavo Camps-Valls, and Jason Weston. Semisupervised neural networks for efficient hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 48(5):2271–2282, 2010.

[5] Rahul Gomes, Papia Rozario, and Nishan Adhikari. Deep learning optimization in remote sensing image segmentation using dilated convolutions and shufflenet. In *2021 IEEE International Conference on Electro Information Technology (EIT)*, pages 244–249. IEEE, 2021.

[6] Papia F Rozario, Peter Oduor, Larry Kotchman, Michael Kangas, et al. Quantifying spatiotemporal change in landuse and land cover and accessing water quality: A case study of missouri watershed james sub-region, north dakota. *Journal of Geographic Information System*, 8(06):663, 2016.

[7] Papia F Rozario, Peter Oduor, Larry Kotchman, Michael Kangas, et al. Transition modeling of land-use dynamics in the pipestem creek, north dakota, usa. *Journal of Geoscience and Environment Protection*, 5(03):182, 2017.

[8] Buddhika Madurapperuma, Papia Rozario, Peter Oduor, and Larry Kotchman. Land-use and land-cover change detection in pipestem creek watershed, north dakota. *International Journal of GEOMATICS and GEOSCIENCES*, 5(3):416–426, 2015.

[9] Anna Pacheco and Heather McNairn. Evaluating multispectral remote sensing and spectral unmixing analysis for crop residue mapping. *Remote Sensing of Environment*, 114(10):2219–2228, 2010.

[10] Paulina Lyubenova Raeva, Jaroslav Šedina, and Adam Dlesk. Monitoring of crop fields using multispectral and thermal imagery from uav. *European Journal of Remote Sensing*, 52(sup1):192–201, 2019.

[11] Juan Augusto Paredes, Jessenia González, Carlos Saito, and Andrés Flores. Multispectral imaging system with uav integration capabilities for crop analysis. In *2017 First IEEE International Symposium of Geoscience and Remote Sensing (GRSS-CHILE)*, pages 1–4. IEEE, 2017.

[12] Papia F Rozario, Buddhika D Madurapperuma, and Yijun Wang. Remote sensing approach to detect burn severity risk zones in palo verde national park, costa rica. *Remote Sensing*, 10(9):1427, 2018.

[13] Yonghao Xu, Bo Du, Fan Zhang, and Liangpei Zhang. Hyperspectral image classification via a random patches network. *ISPRS journal of photogrammetry and remote sensing*, 142:344–357, 2018.

[14] Le Sun, Zebin Wu, Jianjun Liu, Liang Xiao, and Zhihui Wei. Supervised spectral–spatial hyperspectral image classification with weighted markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 53(3):1490–1503, 2014.

[15] Le Sun, Chenyang Ma, Yunjie Chen, Hiuk Jae Shim, Zebin Wu, and Byeungwoo Jeon. Adjacent superpixel-based multiscale spatial-spectral kernel for hyperspectral classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(6):1905–1919, 2019.

[16] Le Sun, Chenyang Ma, Yunjie Chen, Yuhui Zheng, Hiuk Jae Shim, Zebin Wu, and Byeungwoo Jeon. Low rank component induced spatial-spectral kernel method for hyperspectral image classification. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(10):3829–3842, 2019.

[17] Le Sun, Feiyang Wu, Tianming Zhan, Wei Liu, Jin Wang, and Byeungwoo Jeon. Weighted nonlocal low-rank tensor decomposition method for sparse unmixing of hyperspectral images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:1174–1188, 2020.

[18] Yushi Chen, Xing Zhao, and Xiuping Jia. Spectral–spatial classification of hyperspectral data based on deep belief network. *IEEE journal of selected topics in applied earth observations and remote sensing*, 8(6):2381–2392, 2015.

[19] Hoo-Chang Shin, Matthew R Orton, David J Collins, Simon J Doran, and Martin O Leach. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4d patient data. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1930–1943, 2012.

[20] Jie Li, Qiangqiang Yuan, Huanfeng Shen, and Liangpei Zhang. Noise removal from hyperspectral image with joint spectral–spatial distributed sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*, 54(9):5425–5439, 2016.

[21] Nadine Renard, Salah Bourennane, and Jacques Blanc-Talon. Denoising and dimensionality reduction using multilinear tools for hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 5(2):138–142, 2008.

[22] Hongyan Zhang, Wei He, Liangpei Zhang, Huanfeng Shen, and Qiangqiang Yuan. Hyperspectral image restoration using low-rank matrix recovery. *IEEE transactions on geoscience and remote sensing*, 52(8):4729–4743, 2013.

[23] Yong-Qiang Zhao and Jingxiang Yang. Hyperspectral image denoising via sparse representation and low-rank constraint. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1):296–308, 2014.

[24] Haiyan Fan, Yunjin Chen, Yulan Guo, Hongyan Zhang, and Gangyao Kuang. Hyperspectral image restoration using low-rank tensor recovery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(10):4589–4604, 2017.

[25] Hyperspectral remote sensing scenes - grupo de inteligencia computacional (gic). https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes. Accessed: 2022-11-26.

[26] Mostofa Ahsan, Rahul Gomes, and Anne Denton. Application of a convolutional neural network using transfer learning for tuberculosis detection. In *2019 IEEE International Conference on Electro Information Technology (EIT)*, pages 427–433. IEEE, 2019.

[27] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.

[28] Swalpa Kumar Roy, Gopal Krishna, Shiv Ram Dubey, and Bidyut B Chaudhuri. Hybridsn: Exploring 3-d–2-d cnn feature hierarchy for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 17(2):277–281, 2019.

[29] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

[30] In Jae Myung. Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology*, 47(1):90–100, 2003.

[31] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20:273–297, 1995.

[32] Rahul Gomes, Mostofa Ahsan, and Anne Denton. Random forest classifier in sdn framework for user-based indoor localization. In *2018 IEEE International Conference on Electro/Information Technology (EIT)*, pages 0537–0542. IEEE, 2018.

[33] Mahesh Pal. Random forest classifier for remote sensing classification. *International journal of remote sensing*, 26(1):217–222, 2005.

[34] Rahul Gomes, Anne Denton, and David Franzen. Quantifying efficiency of sliding-window based aggregation technique by using predictive modeling on landform attributes derived from dem and ndvi. *ISPRS International Journal of Geo-Information*, 8(4):196, 2019.