EARTH SCIENCES
RESEARCH JOURNAL
Earth Sci. Res. J. Vol. 28, No. 2 (June, 2024): 161 - 174

UNIVERSIDAD NACIONAL DE COLOMBIA

GEOPHYSICS

Investigation of the performances of Support Vector Machine, Random Forest, and 3D-2D Convolutional Neural Network for Hyperspectral Image Classification

Eren Can Seyrek[1]* and Murat Uysal[1,2]
1. Geomatics Engineering, Faculty of Engineering, Afyon Kocatepe University, TR-3200, Afyonkarahisar, Türkiye
2. Remote Sensing and GIS Application and Research Center, Afyon Kocatepe University, TR-3200, Afyonkarahisar, Türkiye
*Corresponding author: ecseyrek@aku.edu.tr

ABSTRACT

Classification of the hyperspectral images (HSIs) is one of the most challenging tasks hyperspectral remote sensing. Various Machine Learning classification algorithms have been implemented to HSI classification. In recent years, several Convolutional Neural Network (CNN) architectures were developed for HSI classification. The aim of this study is to test the performance of CNN, and well-known Support Vector Machine and Random Forest algorithms using the HyRANK Loukia, Houston 2013, and Salinas Scene datasets. The findings indicate that the Modified HybridSN CNN outperformed other algorithms across all datasets, as demonstrated by various performance evaluation metrics.

Keywords: Hyperspectral image classification; support vector machine; random forest; convolutional neural networks; Houston 2013; HyRANK; Salinas Scene;

Investigación sobre el desempeño de máquinas de vectores de soporte, bosque aleatorio y redes neuronales convolucionales 3D y 2D en la clasificación de imágenes hiperespectrales

RESUMEN

La clasificación de imágenes hiperespectrales (HSI, del inglés hyperspectral images) es una de las tareas más complejas de la detección remota hiperespectral. Varios algoritmos de aprendizaje de máquinas se han implementado en la clasificación de las HSI. Recientemente, varias arquitecturas basadas en redes neuronales convolucionales (CNN, del inglés Convolutional Neural Networks) se han desarrollado para esta clasificación de imágenes hiperespectrales. El objetivo de este estudio es evaluar el desempeño de las CNN y los algoritmos de máquinas de vectores de soporte y de bosque aleatorio con los conjuntos de datos HyRANK Loukia, Houston 2013 y Salinas Scene. Los resultados demuestran que el modelo Modified HybridSN CNN superó a otros algoritmos en todos los conjuntos de datos de acuerdo con lo que demuestran varias métricas de evaluación de desempeño.

Palabras clave: clasificación de imágenes hiperespectrales; máquinas de vectores de soporte; bosque aleatorio; redes neuronales convolucionales; Houston 2013; HyRANK; Salinas Scene

Record
Manuscript received: 15/10/2022
Accepted for publication: 17/07/2024

How to cite item:
Seyrek, E. C., & Uysal, M. (2024). Investigation of the performances of Support Vector Machine, Random Forest, and 3D-2D Convolutional Neural Network for Hyperspectral Image Classification. *Earth Sciences Research Journal, 28*(2), 161-174. https://doi.org/10.15446/esrj.v28n2.105296

ISSN 1794-6190 e-ISSN 2339-3459
https://doi.org/10.15446/esrj.v28n2.105296

1. Introduction

Hyperspectral imaging sensors acquire data in hundreds of narrow contiguous spectral bands along the electromagnetic spectrum. Hyperspectral images (HSIs) offer detailed spectral and spatial information about the surface of the Earth, enabling precise discrimination of the surface objects and land cover types. HSIs have a wide range of applications, such as land cover classification (Akar & Tunc Gormus, 2021), land cover change detection (Erturk et al., 2015), precision agriculture (Bhosle & Musande, 2020; Park & Lu, 2015; Teke et al., 2013), forestry (Adão et al., 2017), geology (Van der Meer et al., 2012), environmental monitoring (Stuart et al., 2019), urban planning (Heiden et al., 2012), medical imaging (Lu & Fei, 2014), food quality control (Basantia et al., 2018), and military defense (Ardouin et al., 2007). Consequently, research focusing on the processing and analysis of HSI data has experienced significant growth in recent years.

Classification is a fundamental and crucial technique in HSI processing, which is utilized to identify the land category of each pixel in the HSI (Meng et al., 2021). HSI classification is a more challenging process because images have a large number of spectral bands and have lower spatial resolution than multispectral images. The supervised classification approach generates decision boundaries by learning from training samples (i.e., ground truth pixels) with the aim of minimizing empirical risk and structural risk (Chen et al., 2020).

In the literature, a range of supervised Machine Learning (ML) techniques have been proposed for HSI classification. For instance, Support Vector Machine (SVM) (Gualtieri et al., 1999; Kavzoglu & Colkesen, 2009; Melgani & Bruzzone, 2004), Random Forest (RF) (Chan & Paelinckx, 2008; Waske et al., 2009), Adaboost (Chan & Paelinckx, 2008), Canonical Correlation Forest (Xia et al., 2016) Extreme Gradient Boosting (Loggenberg et al., 2018), Light Gradient Boosting Machines (Ustuner, 2024), etc. have been utilized for this purpose. However, many of these algorithms require a complex and precise user-defined parameter-tuning process. Moreover, they generally perform pixel-wise classification based solely on spectral information, neglecting the spatial relationship between adjacent pixels in the classification process (Chen et al., 2020).

In recent years, Deep Learning, an emerging field within ML, has grown in popularity as a research subject due to its ability to extract different levels of features from various types of datasets, including image, video, and voice. Convolutional Neural Networks (CNNs) are a widely used Deep Learning method for processing data in multiple arrays, such as signals, sequences, images, and videos. The CNNs have demonstrated remarkable performances in tasks including speech recognition, face detection, object detection, and image classification (Akın & Cömert, 2023; LeCun et al., 2015). As a consequence of their ability to achieve state-of-the-art performance, CNNs have become increasingly prominent in remote sensing applications (Abdikan et al., 2023; Cheng et al., 2022; Ghanbari et al., 2021). Notably, a variety of CNN architectures have been developed specifically for the classification of HSIs (Chen et al., 2020; Chen et al., 2016; Fırat et al., 2022; Hang et al., 2020; Luo et al., 2018; Roy et al., 2019). However, there is a lack of publicly available HSI datasets including ground truth data. Many studies within the literature have relied upon commonly used datasets such as Indian Pines, Salinas Scene, University of Pavia, and Kennedy Space Center HSI datasets to evaluate the performances of the classification algorithms. In recent years, several novel HSI datasets were introduced, such as Houston 2013, Houston 2018, HyRANK (Karantzalos et al. 2018), and WHU-Hi (Zhong et al., 2020). In light of these advancements, conducting algorithm evaluations on recently introduced benchmark HSI datasets allows for a more comprehensive and impartial examination of the tests within a broader context.

This study evaluated the performance of the various ML algorithms using a variety of HSI datasets. Traditional ML classifiers, such as SVM and RF, were compared with a Deep Learning approach represented by CNN. HybridSN CNN (Roy et al., 2019) was chosen as the CNN architecture, but in light of literature suggesting that employing various activation functions and optimizers can enhance classification performance (Agarwal et al., 2021; Bera & Shrivastava, 2020; Dubey et al., 2022; Hao et al., 2020; Seyrek & Uysal, 2024; Vani & Rao, 2019), a modified version called Modified HybridSN was developed, integrating Mish activation function (Misra, 2019) and Adamax optimizer (Kingma & Ba, 2014). The HyRANK Loukia, Houston 2013, and Salinas Scene datasets were employed for testing the classification performances of the algorithms to provide diversity in terms of spectral and spatial resolution of the data. To mitigate the Hughes Phenomena resulting from limited ground truth data and high spectral band correlation between within the HSI datasets, Principal Component Analysis (PCA) transformation was applied during the data preparation stage. Therefore, the classification performance of the algorithms was investigated without employing any preprocessing steps apart from PCA transformation for dimensionality reduction to HSI datasets. To ensure consistent evaluation conditions across all algorithms, the data were split into training and testing sets with a ratio of 10% for training and 90% for testing. The performance of the algorithms was assessed by comparing metrics such as Overall Accuracy (OA), Kappa Coefficient (κ), Producer Accuracy (PA), User Accuracy (UA), and f score, respectively. In addition, differences in performances between algorithms were analyzed by performing McNemar's Test.

2. Machine Learning Algorithms

2.1. Support Vector Machines (SVM)

SVM is a well-known non-parametric and supervised ML algorithm based on statistical learning theory, developed by Vapnik (Cortes & Vapnik, 1995). The algorithm is based on the principle of determining the hyperplane that optimally separates two classes. SVM constructs its model without making any assumptions regarding the data distribution. In the case of two-class classification problems where classes are linearly separable, SVM selects one decision boundary that minimizes generalization error from an infinite number of linear decision boundaries. As a result, the selected decision boundary provides the largest margin between classes on the feature space, where the margin is determined as the sum of the distances from the closest samples of classes (support vectors) to the hyperplane (Vapnik, 1995). Due to its consideration of support vectors located on the decision boundaries of the classification process, SVM is particularly useful for small training data sets (Mountrakis et al., 2011).

In the majority of classification scenarios, including the classification of remotely sensed images based on sample pixels, the separation of classes cannot be achieved through a linear hyperplane within the input space. To address this challenge, Boser et al. (1992) proposed a method of transforming the input data onto a higher dimensional input space by using kernel functions. While linear, polynomial, radial basis function (RBF), and sigmoid kernels are commonly used kernel functions on implementing the nonlinear SVM in the literature, polynomial and RBF kernels are the most suitable functions for remotely sensed image classification problems (Kavzoglu & Colkesen, 2009; Melgani & Bruzzone, 2004; Pal & Mather, 2005). Furthermore, Mountrakis et al. (2011) emphasized that the RBF kernel function is the most suitable kernel function for remote sensing data analysis. In this study, the RBF kernel was employed to classify HSIs using the SVM algorithm. Two user-defined parameters, namely the regularization parameter C and kernel width γ, must be specified for the RBF kernel. As the optimal parameter combinations for achieving the best classification performance are unknown for a given dataset, an exhaustive search for optimal parameters is necessary (Hsu et al., 2003). In these experiments, the C and γ parameters of the RBF kernel were determined using a grid search method coupled with cross-validation.

2.2. Random Forest (RF)

The RF algorithm is a commonly used ensemble learning algorithm based on decision trees, utilized for tasks such as classification, regression, and feature selection. Breiman (2001) developed the Bagging algorithm with the philosophy that an ensemble of classifiers can achieve superior classification performance compared to a single classifier. The RF algorithm combines multiple multivariate decision trees, with each tree trained on a distinct asset of different training data (Breiman, 2001; Rodriguez-Galiano et al., 2012). To construct decision trees within RF, the Classification and Regression Trees methodology is applied without pruning (Breiman, 2001; Breiman et al., 1984). The Gini Index serves as the attribute selection criterion, evaluating the impurity of attributes (e.g., spectral bands in hyperspectral data) with respect to classes during the constructing of the model (Pal, 2005). The Gini Index for a given training dataset T can be expressed as follows:

$$Gini = \sum_{j \neq i} \left(\frac{f(C_i,T)}{|T|} \right) \left(\frac{f(C_j,T)}{|T|} \right) \quad (1)$$

Where the $f(C_j,T)/|T|$ denotes the probability that a selected case (e.g., a pixel) belongs to the class C_i. During the training of the RF model, approximately 2/3 of the samples (referred to as "in-bag") within each randomly selected subset are utilized to construct each decision tree in the forest. The remaining 1/3 of the samples (referred to as "out-of-bag") are employed to validate the performance of the final ensemble using an internal cross-validation approach (Sahin et al., 2020). To determine the final label of an unknown sample (e.g., a pixel), the classification results of each decision tree in the forest are combined by performing majority voting, with each classifier vote carrying equal weight (Chan & Paelinckx, 2008). Implementing an RF model requires setting two user-defined parameters: the number of trees (*Ntree*) and the number of randomly selected features at each node (*Mtry*) (Sheykhmousa et al., 2020). The parameter *Mtry* is typically set the square root of the total number of input features, while the *Ntree* parameter is commonly determined as 500, as the error tends to stabilize prior to this number of classification trees (Belgiu & Drăguţ, 2016). Moreover, both theoretical and empirical studies have reported that the classification accuracy of the RFs is less sensitive to the parameter *Ntree* compared to *Mtry* (Kulkarni & Sinha, 2012). Nevertheless, the parameters *Ntree* and *Mrty* were determined through a grid search method utilizing cross-validation.

2.3. Convolutional Neural Networks (CNN)

CNN is a specialized neural network structure for processing data organized in a grid-like topology, such as a digital image formed by three 2-dimensional (2D) arrays. Unlike conventional neural networks that primarily utilize matrix multiplication, CNN contains at least one convolution (LeCun et al., 1998). CNN is designed to handle input data in multiple array formats, such as 1-dimensional data, including signals and sequences; 2D data, including images and audio spectrograms; and 3-dimensional (3D) data, including video and volumetric images (LeCun et al., 2015). CNN is typically consisting of three fundamental components: convolution layers, pooling layers, and fully connected layers, each serving different roles (Li et al., 2018). Within the convolution layer, input data subjects to convolution operation with learnable kernels, followed by an activation function to generate output feature maps. The 2D and 3D convolution operations can be mathematically represented as Equation (2) and Equation (3), respectively.

$$v_{ij}^{xy} = f\left(b_{ij} + \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} w_{ijp}^{qr} v_{(i-1)p}^{(x+q)(y+r)} \right) \quad (2)$$

$$v_{ij}^{xyz} = f\left(b_{ij} + \sum_{p=0}^{P_i-1} \sum_{s=0}^{S_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} w_{ijp}^{qrs} v_{(i-1)p}^{(x+q)(y+r)(z+s)} \right) \quad (3)$$

where v denotes the output variable in the feature map. S, Q, and R represent spectral and spatial dimensions of the kernel, respectively. The indices of the kernel are indicated by (s, q, r), while (z, x, y) represent the indices of the feature map, corresponding to the two spatial and one spectral dimension, respectively. The kernel parameters are represented by k, while the indexes of the input layer, output layer, and feature maps are denoted by i, j, and p, respectively. P indicates the number of feature maps. The bias term is denoted by b. The activation function is represented by f, was selected as a Rectified Linear Unit (ReLU) in this study due to its faster computing rate without a gradient diffusion problem. The ReLU function is defined as follows (Nair & Hinton, 2010; Wang et al., 2020):

$$f(x) = (0, x) \quad (4)$$

In most CNN architectures, a pooling layer follows a convolution layer, reducing the dimensionality of feature maps. Two common pooling operations are commonly used, named max pooling and average pooling. Following the pooling layer, the final fundamental CNN layer is the fully connected layer, similar to backpropagation in traditional neural networks. The learned features extracted from the input image by the CNN serve as the output of the fully connected layer. Connecting this output to a learning classifier, such as Softmax, the classification process is implemented (Krizhevsky et al., 2012).

In this study, the HybridSN CNN model, introduced by Roy et al. (2019), was employed for classifying HSI images. The structure of the HybridSN CNN model utilized in this study is shown in Figure 1. Upon reviewing some studies in the literature, it has been observed that activation functions and optimizers have an impact on the performance of CNN models. For example, some studies observed that the Adamax optimizer improved the classification performance in the CNN model (Ariff & Ismail, 2023; Seyrek & Uysal, 2024; Vani & Rao, 2019). Furthermore, there are research findings indicating that Mish activation function (Equation 5) also enhances classification performance (Misra, 2019; Seyrek & Uysal, 2024; Zhang et al., 2019). Based on this information, the HybridSN CNN architecture was modified to utilize the Adamax optimizer and Mish activation function. Consequently, the resulting CNN architecture is referred to as the Modified HybridSN CNN for the remainder of the study.

$$Mish(x) = x \times \tanh\, tanh\, (\log\, log\, (1 + e^x)) \quad (5)$$

The model comprises three 3D convolution layers, enabling simultaneous extraction of both spectral and spatial features. Subsequently, a 2D convolution layer performs exclusively spatial feature extraction. This ensures the effective extraction of both spectral and spatial features from a hyperspectral image. During the data preparation stage, the HSI image subjected to PCA was partitioned into small overlapping 3D patches, each associated with a ground truth label corresponding to the central pixel. Let's assume that $P \in R^{(S \times S \times B)}$, where S represents the size of the 3D patches and B denotes the number of spectral bands of the PCA-transformed HSI data. The total number of 3D patches generated is $(M\text{-}S+1) \times (N\text{-}S+1)$, where M is the width and N is the height of HSI data. During the data preparation for training and testing the CNN model, 3D patches with any ground truth data corresponding to the central pixel are excluded.

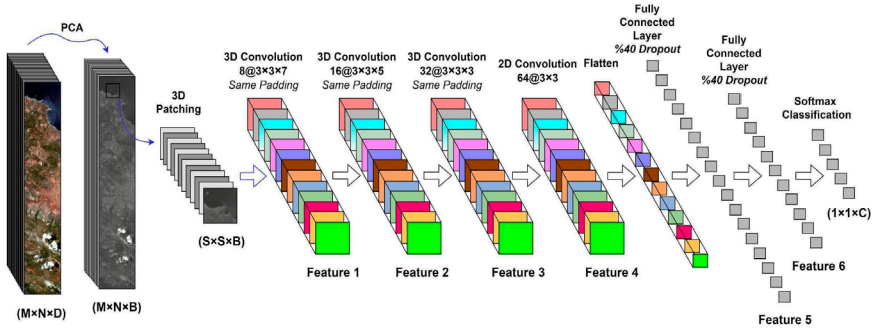

Figure 1. The modified HybridSN architecture

Table 1. Summary of the CNN models for the HyRANK Loukia, DFC 13, and Salinas Scene datasets.

CNN Architecture Model Summary				
Layer	**Output Shape**	**HyRANK Loukia: # of Parameters**	**DFC 13: # of Parameters**	**Salinas Scene: # of Parameters**
input_1 (Input Layer)	(13, 13, 15, 1)	0	0	0
conv3d (Conv3D)	(11, 11, 9, 8)	512	512	512
conv3d_1 (Conv3D)	(9, 9, 5, 16)	5776	5776	5776
conv3d_2 (Conv3D)	(7, 7, 3, 32)	13856	13856	13856
reshape (Reshape)	(7, 7, 96)	0	0	0
conv2d (Conv2D)	(5, 5, 64)	55360	55360	55360
flatten (Flatten)	(1600)	0	0	0
dense (Dense)	(256)	409856	409856	409856
dropout (Dropout)	(256)	0	0	0
dense_1 (Dense)	(128)	32896	32896	32896
dropout_1 (Dropout)	(128)	0	0	0
dense_2 (Dense)	HyRANK Loukia:(14); DFC 13: (15); Salinas Scene: (16)	1806	1935	2064
Total Parameters:	-	520062	520191	520320

Other Parameters	
Parameter	**Value**
Activation Function	HybridSN: ReLU; Modified HybridSN: Mish
Dropout Rate	%40
Number of Epochs	100
Batch Size	256
Optimizer	HybridSN: Adam (β_1: 0.9, β_2: 0.999, ϵ:1e-08, learning_rate: 0.001); Modified HybridSN: Adamax (β_1: 0.9, β_2: 0.999, ϵ:1e-08, learning_rate: 0.001)

3. Datasets

In order to evaluate the performance of the classification algorithms, three publicly available HSI datasets were employed. The first dataset, named HyRANK, was developed by the International Society for Photogrammetry and Remote Sensing (ISPRS), Working Group III/4 (Karantzalos et al., 2018). It comprises five distinct HSIs acquired from the National Aeronautics and Space Administration's (NASA) EO-1 Hyperion sensor. Among these, only two HSIs, namely Dioni and Loukia, are accompanied by ground-truth data. The Loukia has 30 meters spatial resolution and a hypercube size of 249×945×176. Ground truth data for Loukia contains 14 Coordination of Information on the Environment (CORINE) land cover classes, carefully annotated by interpretation experts (Karantzalos et al., 2018). Detailed information regarding the number of training and testing samples for this data set is provided in Table 2.

The second dataset, named Houston 2013, was acquired from the University of Houston campus and the surrounding urban area on June 23, 2012, by the National Science Foundation (NSF)-funded Center for Airborne Laser Mapping (NCALM). The dataset has a spatial resolution of 2.5 meters and comprises 144 spectral bands in the range between 380 and 1050 nanometers. The spatial dimensions of the HSI are 349×1905×144. Fifteen distinct ground classes were determined by the Institute of Electrical and Electronics Engineers - Geoscience and Remote Sensing Society (IEEE GRSS) Image Analysis and Data Fusion Technical Committee for organizing the 2013 Data Fusion Contest (DFC). The original dataset included both training and validation data. In this study, training and validation data were merged and employed as ground truth data. Detailed statistics regarding the numbering of train and testing samples of the dataset are provided in Table 3.

The third dataset, referred to as the Salinas Scene, was acquired on October 9, 1998, from Salinas Valley, California, United States of America (USA), utilizing NASA's airborne Airborne Visible / Infrared Imaging Spectrometer (AVIRIS) sensor. The dataset has a spatial resolution of 3.7 meters and comprises 224 spectral bands. Notably, the original dataset underwent preprocessing, during which water vapor absorption bands within the ranges of 108-112 and 154-167 were excluded. The dataset contains 16 ground classes in three main groups: vegetables, bare soils, and vineyard fields. Following processing, the dataset's dimensions stand at 512×217×204. Details regarding the number of training and testing samples for this dataset are provided in Table 4.

Table 2. Overview of training and testing samples and corresponding training ratios utilized in the HyRANK Loukia dataset.

Class Number	Class Name	# Train Samples	Train Ratio	# Test Samples
1	Dense urban fabric	29	2.15%	259
2	Mineral extraction sites	7	0.52%	60
3	Non irrigated arable land	54	4.00%	488
4	Fruit trees	8	0.59%	71
5	Olive groves	140	10.37%	1261
6	Broad leaved forest	22	1.63%	201
7	Coniferous forest	50	3.70%	450
8	Mixed forest	107	7.93%	965
9	Dense sclerophyllous vegetation	379	28.07%	3414
10	Sparce sclerophyllous vegetation	280	20.74%	2523
11	Sparsely vegetated areas	41	3.04%	363
12	Rocks and sand	49	3.63%	438
13	Water	139	10.30%	1254
14	Coastal water	45	3.33%	406
	Total	**1350**	**100.00%**	**12153**

Table 3. Overview of training and testing samples and corresponding training ratios utilized in the Houston 2013 dataset.

Class Number	Class Name	# Train Samples	Train Ratio	# Test Samples
1	Healthy grass	137	7.93%	1237
2	Stressed grass	145	8.40%	1309
3	Synthetic grass	80	4.63%	715
4	Trees	126	7.30%	1138
5	Soil	130	7.53%	1168
6	Water	34	1.97%	305
7	Residential	148	8.57%	1328
8	Commercial	135	7.82%	1219
9	Road	155	8.98%	1399
10	Highway	142	8.22%	1282
11	Railway	157	9.09%	1409
12	Parking lot 1	143	8.28%	1286
13	Parking lot 2	64	3.71%	571
14	Tennis court	51	2.95%	459
15	Running track	80	4.63%	718
	Total	**1727**	**100.00%**	**15543**

Table 4. Overview of training and testing samples and corresponding training ratios used in the Salinas Scene dataset.

Class Number	Class Name	# Train Samples	Train Ratio	# Test Samples
1	Brocoli_green_weeds_1	372	3.71%	2009
2	Brocoli_green_weeds_2	197	6.87%	3726
3	Fallow	139	3.64%	1976
4	Fallow_rough_plow	268	2.57%	1394
5	Fallow_smooth	396	4.95%	2678
6	Stubble	358	7.32%	3959
7	Celery	1127	6.61%	3579
8	Grapes_untrained	620	20.82%	11271
9	Soil_vinyard_develop	328	11.46%	6203
10	Corn_senesced_green_weeds	107	6.06%	3278
11	Lettuce_romaine_4wk	193	1.98%	1068
12	Lettuce_romaine_5wk	91	3.57%	1927
13	Lettuce_romaine_6wk	107	1.68%	916
14	Lettuce_romaine_7wk	727	1.98%	1070
15	Vinyard_untrained	181	13.43%	7268
16	Vinyard_vertical_trellis	372	3.34%	1807
	Total	**5211**	**3.71%**	**54129**

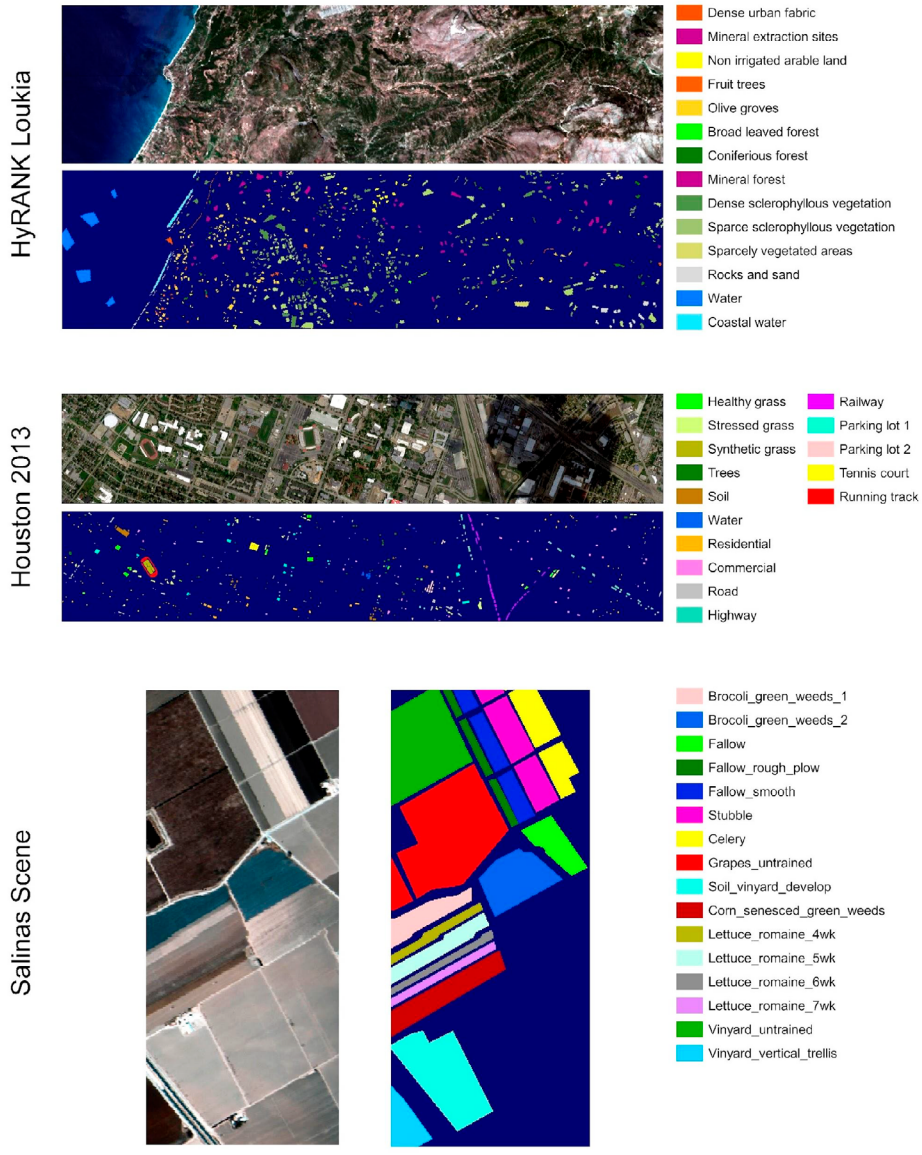

Figure 2. True color image compositions and corresponding ground truth data for the HyRANK Loukia, Houston 2013, and Salinas Scene datasets.

4. Experimental setup

Classification of HSIs was performed using a workstation with Intel® Xeon™ E-2136 processor and NVIDIA GeForce RTX-2070 SUPER graphic processor. The software setup comprised Windows 11 x64 as an operating system and Python 3.7.9 for programming tasks. Model development utilized Keras 2.3.1 (Chollet, 2015), Tensorflow 2.1.0 (Abadi et al., 2016) with CUDA support, and Scikit-learn 0.23.2 (Pedregosa et al., 2011) libraries.

5. Performance evaluation

To assess and compare the performance of the classification models, several evaluation metrics were computed from the confusion matrices. These metrics included OA, PA, UA, f score, and κ. OA, a commonly used performance metric, represents the ratio of correctly classified pixels to the total number of pixels. The f score is calculated as the harmonic mean of PA (precision) and UA (recall) for each class. The κ is a statistical measure indicating the agreements between ground truth and classification maps. These calculated range from 0 to 1, with higher values indicating better classification performance. Additionally, McNemar's test was employed to examine the statistical significance of prediction differences between the classification algorithms. McNemar's test is a non-parametric test based on Chi-square distribution and is utilized to determine dichotomous dependent variable between two related groups (i.e., pixel classes predicted by two different classification models). The test value (x^2) is calculated from a 2×2 sized contingency table by following equation (Foody, 2004).

$$\chi^2 = \frac{\left(|f_{12} - f_{21}| - 1\right)^2}{\left(f_{12} - f_{21}\right)} \tag{6}$$

where f_{12} represents the number of samples that are correctly classified by the first model but misclassified by the second model, and f_{21} represents the number of the samples misclassified by the first model, and correctly classified by the second model. If the derived x^2 value exceeds the critical value from x^2 table (3.841) at a 95% confidence level, the null hypothesis is rejected. The rejection indicates statistically significant difference in OA between the predictions of the two models.

6. Results

Initially, PCA transformation was implemented on the HSIs, retaining the first 15 principal components across all datasets. Subsequently, 10% of the ground truth data was split for training the models, with the remaining 90% reserved for testing the models. Grid Search was performed for SVM and RF algorithms on each HSI dataset to determine the optimal user-defined parameters. The search ranges for each parameter, along with the determined optimal parameters for each dataset, are detailed in Table 5.

Table 5. Hyperparameter tuning results with Grid Search method.

Algorithm	Parameter	Search Range	HyRANK Loukia	DFC13	Salinas Scene
SVM	C	0.001, 0.01, 0.1, 1 ,10, 100, 1000	10	100	100
	γ	1, 0.1, 0.01, 0.001, 0.0001, 0.00001	0.01	0.1	0.01
RF	Mrty	'sqrt', 'log2'	log2	sqrt	sqrt
	Ntree	100, 250, 500, 750, 1000, 1200	100	750	1200

The evaluation of classification performance using the HyRANK Loukia dataset is summarized in Table 6. Note that bold texts highlight the highest superior values within the same performance metric row. As can be seen from the table, the Modified HybridSN CNN algorithm achieves the highest OA and κ coefficient with 89.69% and 87.74%, respectively. In contrast, RF exhibits the lowest OA and κ coefficient with 80.86% and 76.20%, respectively, for the HyRANK Loukia dataset. Upon examining the class-based accuracy metrics, it is evident that the both CNN algorithms demonstrates superior performance compared to other algorithms across most classes. Of particular significance is its sufficiency in classes such as Dense urban fabric, Fruit trees, Broad leaved forest, Coniferous forest, and Mixed forest, where both SVM and RF underperformed, exhibiting lower UA scores. This observation underscores the significant contribution of the feature extraction process of CNN to the accuracy of classifying low-resolution HSI. When examining class-based accuracy metrics for CNN models, it was found that the Modified HybridSN CNN model slightly improved the classification accuracy for certain classes (Dense urban fabric, Fruit trees, Coniferous forest, Dense sclerophyllous vegetation, Sparse sclerophyllous vegetation, and Sparsely vegetated areas) compared to the original HybridSN CNN model. In terms of training processing times, SVM stands out for its rapid processing, whereas the RF requires more processing time. In contrast, RF had the shortest classification time for processing the entire image, whereas SVM took the longest time for classification.

The classification maps of the algorithms applied to the HyRANK Loukia dataset, alongside the true color composition of the scene, are given in Figure 3. Visual interpretation of the classification maps is very challenging owing to the discontinuous land patterns and low spatial resolution of the scene. Nonetheless, notable classification errors are discernible in certain pixels. For instance, the SVM and both CNN algorithms misclassifies some pixels located within the inner side of the sea (e.g., pixels incorrectly classified as land within the sea). Furthermore, certain pixels, which should have been classified as coastal water, are misclassified as sparse sclerophyllous vegetation. Additionally, the CNN algorithm eliminated certain details, such as roads. Despite the overall similarity in classification maps produced by CNN algorithms, the Modified HybridSN algorithm seems to provide a superior classification map compared to the original HybridSN model. This conclusion is drawn from the classification of the sea region in the lower right corner of the scene and the lower incidence of misclassified pixels in the sea.

Table 6. The performance analysis of the classification algorithms on HyRANK Loukia HSI dataset.

Class Name	SVM			RF			HybridSN CNN			Modified HybridSN CNN		
	PA	UA	f score	PA	UA	f score	PA	UA	f score	PA	UA	f score
Dense urban fabric	71.69	75.29	73.45	84.80	55.98	67.44	**93.69**	80.31	86.49	91.60	**84.17**	**87.73**
Mineral extraction sites	96.15	83.33	89.29	95.56	71.67	81.90	**100.00**	**93.33**	**96.55**	96.23	85.00	90.27
Non irrigated arable land	90.16	80.74	85.19	88.52	75.82	81.68	95.24	**86.07**	**90.42**	**95.43**	85.66	90.28
Fruit trees	80.36	63.38	**70.87**	**100.00**	21.13	34.88	90.91	56.34	69.57	68.42	**73.24**	70.75
Olive groves	91.38	91.67	91.53	85.25	87.07	86.15	**91.86**	**94.85**	**93.33**	90.02	94.45	92.18
Broad leaved forest	88.61	34.83	50.00	86.30	31.34	45.99	90.97	**65.17**	**75.94**	**94.55**	51.74	66.88
Coniferous forest	87.50	38.89	53.85	**91.50**	50.22	64.85	89.52	66.44	76.28	80.78	**73.78**	**77.12**
Mixed forest	62.25	55.03	58.42	66.33	53.68	59.34	**84.18**	86.01	**85.08**	77.89	**88.70**	82.95
Dense sclerophyllous vegetation	72.41	87.35	79.18	73.11	87.20	79.54	85.03	89.05	86.99	**87.57**	**90.42**	**88.98**
Sparce sclerophyllous vegetation	82.32	82.32	82.32	78.10	85.26	81.52	85.34	89.73	87.48	**90.22**	**87.79**	**88.99**
Sparsely vegetated areas	82.81	58.40	68.50	83.70	52.34	64.41	**91.81**	86.50	89.08	90.93	88.43	89.66
Rocks and sand	97.12	84.70	90.49	92.29	84.70	88.33	**99.75**	89.73	**94.47**	96.34	***90.18***	93.16
Water	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Coastal water	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.51	100.00	99.75
OA	81.61			80.86			89.29			**89.69**		
κ	77.82			76.85			87.21			**87.74**		
Training time (s)	**5.47**			16.21			7.61			8.42		
Classification time (s)	18.75			**7.16**			11.64			12.27		

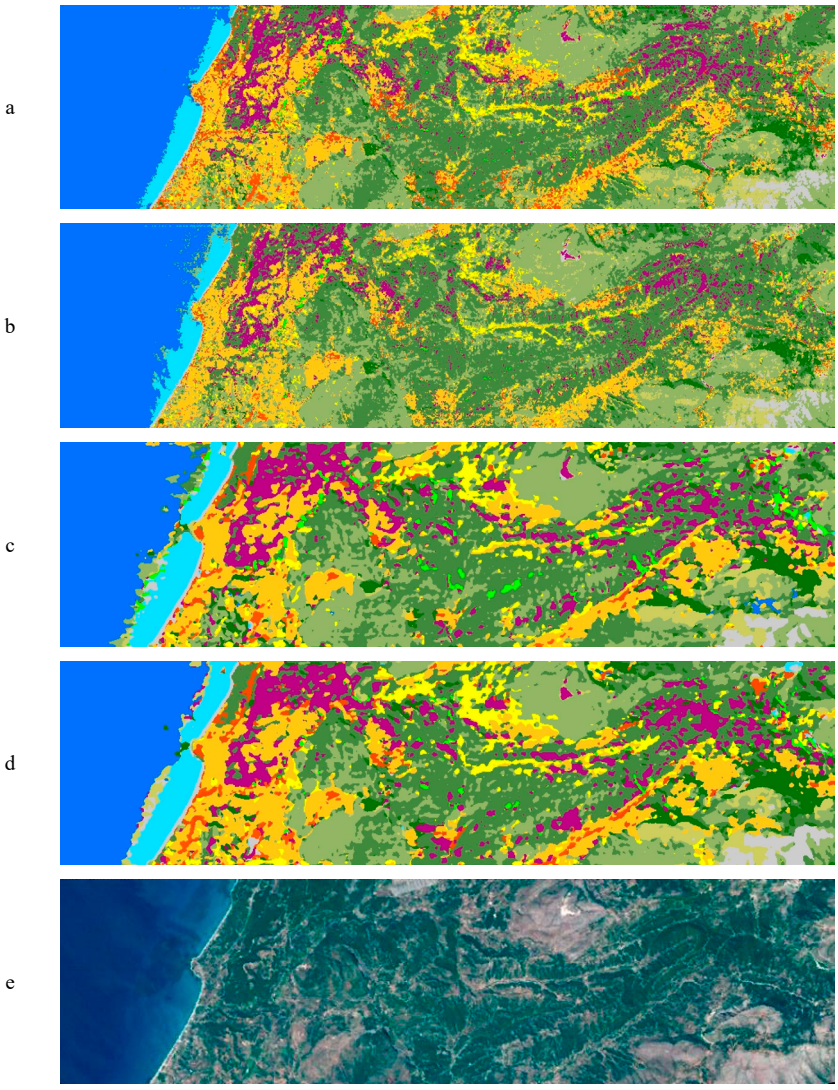

Figure 3. Classification maps for the HyRANK Loukia dataset: (a) SVM, (b) RF, (c) HybridSN CNN, (d) Modified HybridSN CNN, and (e) true color composition

Table 7. The performance analysis of the classification algorithms on Houston 2013 HSI dataset.

Class Name	SVM			RF			HybridSN CNN			Modified HybridSN CNN		
	PA	UA	f score	PA	UA	f score	PA	UA	f score	PA	UA	f score
Healthy grass	96.12	98.06	97.08	97.95	96.36	97.15	96.03	**99.76**	97.86	**98.48**	99.51	**98.99**
Stressed grass	98.76	97.02	97.88	95.97	98.32	97.13	**100.00**	98.09	99.04	**100.00**	**98.55**	99.27
Synthetic grass	100.00	**99.58**	**99.79**	100.00	98.88	99.44	100.00	97.90	98.94	99.72	98.60	99.16
Trees	**98.60**	98.68	98.64	98.76	98.33	98.55	97.25	99.38	98.31	98.36	**100.00**	**99.17**
Soil	99.83	99.23	99.53	98.22	99.06	98.64	**99.91**	**100.00**	**99.96**	98.73	**100.00**	99.36
Water	100.00	97.38	98.67	100.00	**98.03**	**99.01**	100.00	94.10	96.96	100.00	96.39	98.16
Residential	93.22	94.20	93.71	92.31	90.36	91.32	**99.75**	90.44	94.87	97.82	**97.82**	**97.82**
Commercial	94.48	91.22	92.82	95.46	89.66	92.47	93.53	95.98	94.74	98.81	95.57	97.16
Road	90.17	93.14	91.63	83.71	88.13	85.86	97.03	98.07	**97.55**	99.03	94.78	96.86
Highway	93.38	92.36	92.86	89.76	91.65	90.70	**97.55**	99.38	**98.45**	96.32	99.92	98.09
Railway	91.44	93.26	92.34	85.24	90.99	88.02	**99.93**	97.16	**98.52**	98.09	98.37	98.23
Parking lot 1	91.74	92.38	92.06	87.55	89.11	88.32	**99.07**	**99.84**	**99.46**	97.94	99.77	98.84
Parking lot 2	78.60	72.68	75.52	88.24	63.05	73.54	94.07	**97.20**	95.61	**98.34**	93.52	**95.87**
Tennis court	93.08	99.56	96.21	91.04	99.56	95.11	**97.45**	**100.00**	**98.71**	**97.45**	**100.00**	**98.71**
Running track	**100.00**	97.91	98.94	98.05	98.05	98.05	95.35	**100.00**	97.62	95.86	**100.00**	**97.89**
OA	94.58			92.72			97.83			**98.29**		
κ	94.14			92.12			97.66			**98.15**		
Training time (s)	**7.12**			19.59			8.77			9.25		
Classification time (s)	43.76			**20.12**			33.87			36.75		

The evaluation of classification performance using the Houston 2013 dataset is summarized in Table 7. Notably, the Modified HybridSN CNN algorithm achieved the highest OA at 98.29%, contrasting with the RF algorithm's lowest OA of 92.72%. When analyzing class-based accuracy metrics, it is observed that both CNN models generally exhibit higher accuracy values and outperform other algorithms, particularly in the Parking lot 2 class. A comparison of CNN algorithms reveals that their class accuracies are similar.

The classification maps of the Houston 2013 data are given in Figure 4. Overall, the algorithms extracted the shapes of human-made objects successfully, taking the advantage of the superior spatial and spectral resolution of the dataset. However, a salt-and-pepper effect was noticeable in classification maps of the SVM and the RF, stemming from pixel-wise classification. Misclassified pixels were observed, particularly by SVM and RF, in areas such as stadium corners and regions affected by cloud shadows, resulting in significant pixel misclassification. Notably, the shaded area has proven difficult for all algorithms, likely due to insufficient training data for this region. This indicates that the lighting conditions of the area being classified have a direct impact on classification performance. Moreover, it is evident that the Modified HybridSN CNN algorithm performs better at building classification in the shaded area than the HybridSN CNN algorithm.

Table 8 represents the performance analysis results for the Salinas Scene dataset. Based on the calculated OA and κ values, the Modified HybridSN CNN algorithm outperformed SVM and RF, achieving an OA of 99.85% and a κ value of 99.83. Overall, the Modified HybridSN CNN algorithm exhibits slightly higher class-based performance metrics compared to other algorithms. Notably, in classes such as Grapes_untrained and Vinyard_untrained, CNN outperformed other algorithms. Upon reviewing processing times, CNN stands out for its shortest processing time, in terms of both training and classifying the whole scene.

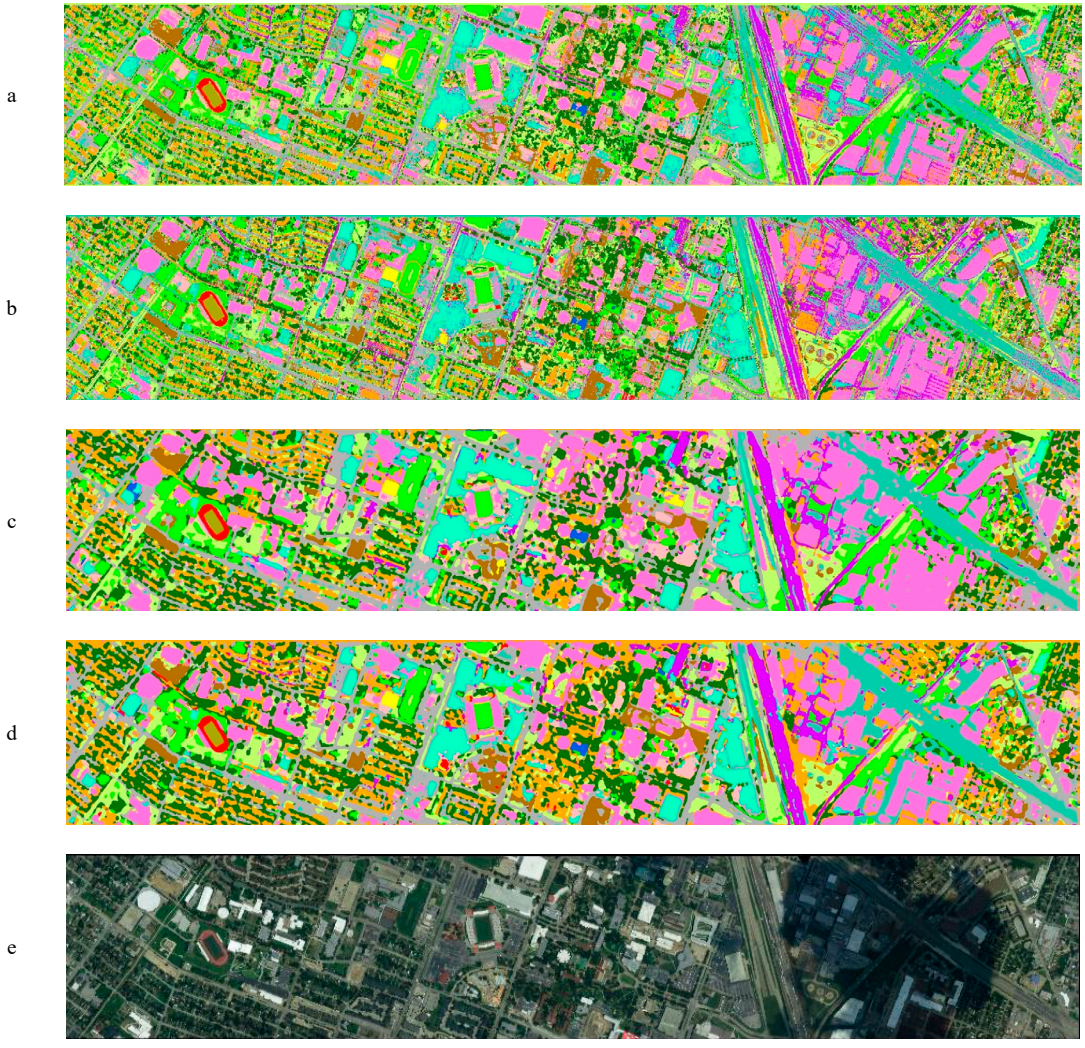

Figure 4. Classification maps for the Houston 2013 dataset: (a) SVM, (b) RF, (c) HybridSN CNN, (d) Modified HybridSN CNN, and (e) true color composition

Table 8. The performance analysis of the classification algorithms on Salinas Scene HSI dataset.

Class Name	SVM			RF			HybridSN CNN			Modified HybridSN CNN		
	PA	UA	f score	PA	UA	f score	PA	UA	f score	PA	UA	f score
Brocoli_green_weeds_1	100.00	100.00	100.00	100.00	99.83	99.92	100.00	100.00	100.00	100.00	100.00	100.00
Brocoli_green_weeds_2	100.00	99.94	99.97	99.79	99.88	99.84	100.00	**100.00**	**100.00**	100.00	**100.00**	**100.00**
Fallow	98.88	99.49	99.19	96.67	99.66	98.15	100.00	100.00	**100.00**	99.94	100.00	99.97
Fallow_rough_plow	98.89	99.68	99.29	98.89	99.68	99.29	99.35	97.93	98.64	**99.44**	**99.76**	**99.60**

(*Continued*)

Class Name	SVM			RF			HybridSN CNN			Modified HybridSN CNN		
	PA	UA	f score	PA	UA	f score	PA	UA	f score	PA	UA	f score
Fallow_smooth	98.67	98.88	98.78	99.62	98.55	99.08	98.89	100.00	99.44	99.88	99.71	99.79
Stubble	100.00	99.92	99.96	100.00	99.78	99.89	100.00	100.00	100.00	100.00	100.00	**100.00**
Celery	99.78	99.97	99.88	100.00	99.50	99.75	100.00	99.84	99.92	100.00	100.00	**100.00**
Grapes_untrained	83.02	91.41	87.01	81.03	91.01	85.73	99.88	98.37	99.12	99.89	99.52	**99.70**
Soil_vinyard_develop	99.34	99.96	99.65	99.13	99.87	99.50	100.00	**100.00**	**100.00**	99.98	100.00	99.99
Corn_senesced_green_weeds	98.41	96.54	97.47	97.23	95.22	96.22	99.63	99.53	99.58	**100.00**	99.86	99.93
Lettuce_romaine_4wk	96.94	98.75	97.84	94.88	96.46	95.67	98.26	100.00	99.12	**99.69**	**100.00**	99.84
Lettuce_romaine_5wk	99.48	99.77	99.63	98.30	**100.00**	99.14	**100.00**	99.83	99.91	**100.00**	**100.00**	**100.00**
Lettuce_romaine_6wk	98.67	98.91	98.79	98.44	99.15	98.79	**100.00**	99.88	99.94	**100.00**	**100.00**	**100.00**
Lettuce_romaine_7wk	96.90	97.30	97.10	98.62	96.26	97.43	99.90	100.00	99.95	**100.00**	**100.00**	**100.00**
Vinyard_untrained	84.71	70.89	77.19	83.29	67.07	74.31	97.70	99.82	98.74	**99.27**	**99.83**	**99.55**
Vinyard_vertical_trellis	99.33	99.69	99.51	99.08	99.08	99.08	99.75	99.51	99.63	**100.00**	**100.00**	**100.00**
OA	93.88			93.06			99.52			**99.85**		
κ	93.18			92.25			99.47			**99.83**		
Training time (s)	61.45			56.56			26.84			**25.38**		
Classification time (s)	16.07			10.27			6.76			**6.85**		

Figure 5 illustrates the classification maps generated for the Salinas Scene dataset. Despite algorithms achieving high OA, instances of misclassifications were observed, particularly in the Grapes_untrained and Vinyard_untrained classes. Notably, the classification map of both CNN algorithms contains less salt-and-pepper effects compared to other classification maps. Analysis based on ground truth photos by Gualtieri et al. (1999) suggests that confusion, particularly in mixed classes, is expected due to similar texture and color characteristics.

A statistical analysis comparing the classification algorithms using McNemar's test was performed. The test results are presented in Table 9. Note that the χ^2 table value is 3.841 at 95% level of confidence. Upon analysis of the results, it is evident that there are statistically significant variances among all classification accuracies, except for the comparison between HybridSN CNN and Modified HybridSN CNN on the Loukia dataset from HyRANK. These findings suggest that employing the CNN algorithm for HSI classification generally yields higher accuracy. However, it's noteworthy that the choice of activation function and optimizer within the CNN algorithm significantly impacts model performance.

Table 9. McNemar's test results for HyRANK Loukia, Houston 2013, and Salinas Scene datasets.

Model 1 - Model 2	HyRANK Loukia		Houston 2013		Salinas Scene	
	χ^2	Significant?	χ^2	Significant?	χ^2	Significant?
SVM – RF	7.604	Yes	94.494	Yes	81.999	Yes
SVM – HybridSN CNN	406.961	Yes	284.944	Yes	2463.601	Yes
SVM – Modified HybridSN CNN	442.350	Yes	410.614	Yes	2815.944	Yes
RF – HybridSN CNN	478.365	Yes	573.005	Yes	2902.725	Yes
RF – Modified HybridSN CNN	513.757	Yes	694.404	Yes	3212.820	Yes
HybridSN CNN – Modified HybridSN CNN	2.413	No	16.420	Yes	98.462	Yes

7. Discussion

This study compares the performance of four supervised ML classification algorithms for HSI classification. Three benchmark datasets, namely HyRANK Loukia, Houston 2013, and Salinas Scene, were utilized, each possessing varying levels of spectral and spatial resolution. Notably, HyRANK Loukia and Houston 2013 datasets are infrequently utilized benchmarks in the literature. In the data preparation stage, PCA was employed to reduce the spectral dimensionality of the HSI datasets, retaining the first 15 principal components for each dataset. The classification algorithms were evaluated using common performance evaluation metrics.

Upon reviewing the classification results of the datasets, it is evident that the classification performance for the HyRANK Loukia dataset was comparatively lower than that of the other datasets. This discrepancy could potentially be attributed to factors such as an insufficient number of collected ground truth data and the presence of mixed pixels, resulting in similar spectral characteristics between classes. Additionally, the lower accuracy observed in vegetation classes may be attributed to the low density of vegetation and the relatively low spatial resolution of the Hyperion sensor. Notably, the UAs for these classes were notably low in the classification results obtained using SVM and RF algorithms. Similarly, Christovam et al. (2019) and Si et al. (2021) reported low accuracy values were calculated for the same classes in their studies. In the case of the Houston 2013 dataset, a significant number of pixels located within cloud shadows were misclassified. Some studies utilizing this dataset as a benchmark also observed misclassifications in shadow areas, while In a few cases, the shadowy areas were removed. It is crucial to retain these effects to impartially test classification algorithms and understand their performance underpotential real-life scenarios. Visual interpretation of the classification maps generated for the Salinas Scene dataset revealed numerous misclassified pixels, despite high accuracy values being obtained. Given that the classes predominantly represent vegetation species in the image, the low accuracy may stem from similar spectral characteristics and low vegetation density. Therefore, assessing classification performance should involve not only quantitative metrics but also visual interpretation of the results. Upon reviewing the classification accuracies across datasets, it was clear that both CNN algorithms consistently outperformed SVM and RF. Notably, while SVM and RF classification maps exhibited a salt-and-pepper effect, CNN tended to eliminate details such as roads and small objects visible on the true color images of the HSIs. Statistical analysis using McNemar's test revealed significant differences in pairwise comparisons of classification maps generated by the algorithms. Moreover, upon scrutinizing the processing times of the algorithms, it becomes challenging to definitively determine the superiority of one algorithm over another, as the training and classification times vary across different datasets. However, when focusing solely on classification success, the CNN's exceptional performance renders it an appealing option for tasks demanding precise classification. The classification outcomes derived from

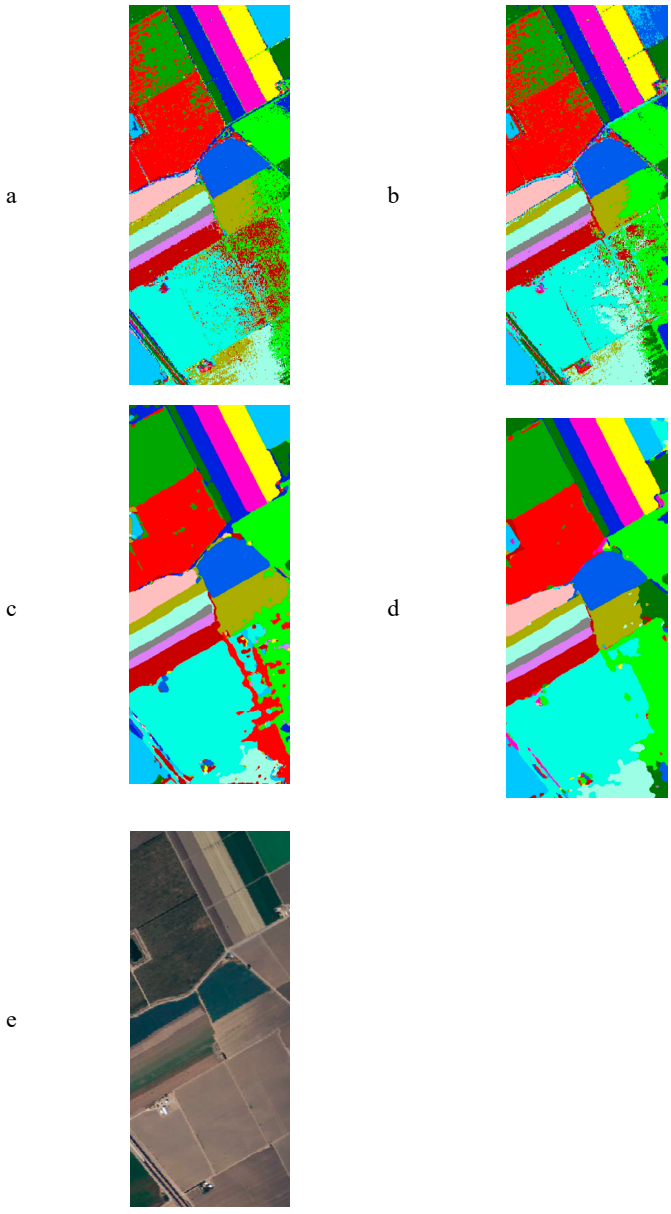

Figure 5. Classification maps for the Salinas Scene dataset: (a) SVM, (b) RF, (c) HybridSN CNN, (d) Modified HybridSN CNN, and (e) true color composition

the Modified HybridSN CNN and HybridSN CNN algorithms underscored the significance highlighted in literature regarding the selection of activation functions and optimizers.

8. Conclusion

This study examined the comparative analysis of four supervised ML algorithms for HSI classification utilizing three benchmark datasets. The investigation provided valuable insights into the strengths and limitations of SVM, RF, and CNN in handling varying spectral and spatial characteristics of HSI. It was observed that the performance of classification algorithms varied across datasets, with factors such as insufficient ground truth data, mixed pixels, and spectral similarity between classes affecting classification performances. While SVM and RF performed well performance in certain cases, particularly in datasets with less complex spectral characteristics, CNN models consistently outperformed them, especially in datasets with complex spectral signatures and spatial features. The capacity of CNN to extract features and capture complex spatial patterns made it a compelling choice for tasks requiring high precision classification. Moreover, the HybridSN CNN architecture underwent modifications with the expectation that it would influence the classification performance of activation functions and optimizers in the CNN algorithm. In line with literature recommendations, the Modified HybridSN CNN algorithm, integrating the Adamax optimizer and Mish activation function, exhibited a slight enhancement in classification performance. Furthermore, the study emphasized the significance of not only relying on quantitative metrics but also incorporating visual interpretation of classification results to gain a comprehensive understanding of algorithm performance. Overall, the findings contribute to advancing the field of HSI classification by providing valuable insights into the comparative performance of ML algorithms across various datasets. These insights can inform the selection of appropriate algorithms based on specific application requirements, ultimately enhancing the accuracy and efficiency of HSI analysis in various domains.

Acknowledgments

The study constitutes the master's thesis of the first author. The authors thank the ISPRS Commission III, WG III/4 for providing the HyRANK dataset and to the National Center for Airborne Laser Mapping, the Hyperspectral Image Analysis Laboratory at the University of Houston, and the IEEE GRSS Image Analysis and Data Fusion Technical Committee for providing the Houston dataset used in the experiments.

Funding

This study was supported by Afyon Kocatepe University Scientific Research Projects Coordination Unit within the 20.FEN.BİL.12 project.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., & Isard, M. (2016). Tensorflow: A system for large-scale machine learning. *12th symposium on operating systems design and implementation*. https://doi.org/10.48550/arXiv.1605.08695

Abdikan, S., Sekertekin, A., Narin, O. G., Delen, A., & Balik Sanli, F. (2023). A comparative analysis of SLR, MLR, ANN, XGBoost and CNN for crop height estimation of sunflower using Sentinel-1 and Sentinel-2. *Advances in space research*, *71*(7), 3045-3059. https://doi.org/10.1016/j.asr.2022.11.046

Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., & Sousa, J. J. (2017). Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sensing*, *9*(11), 1110. https://doi.org/10.3390/rs9111110

Agarwal, M., Rajak, A., & Shrivastava, A. K. (2021). Assessment of optimizers impact on image recognition with convolutional neural network to adversarial datasets. *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/1998/1/012008

Akar, O., & Tunc Gormus, E. (2021). Land use/land cover mapping from airborne hyperspectral images with machine learning algorithms and contextual information. *Geocarto International*, 1-28. https://doi.org/10.1080/10106049.2021.1945149

Akın, A. T., & Cömert, Ç. (2023). The development of an augmented reality audio application for visually impaired persons. *Multimedia Tools and Applications*, *82*(11), 17493-17512. https://doi.org/10.1007/s11042-022-14134-x

Ardouin, J.-P., Lévesque, J., & Rea, T. A. (2007). *A demonstration of hyperspectral image exploitation for military applications*. 10th International Conference on Information Fusion. https://doi.org/10.1109/ICIF.2007.4408184

Ariff, N. A. M., & Ismail, A. R. (2023). *Study of Adam and Adamax Optimizers on AlexNet Architecture for Voice Biometric Authentication System*. 17th International Conference on Ubiquitous Information Management and Communication (IMCOM). https://doi.org/10.1109/IMCOM56909.2023.10035592

Basantia, N., Nollet, L. M., & Kamruzzaman, M. (2018). *Hyperspectral Imaging Analysis and Applications for Food Quality*. CRC Press. https://doi.org/10.1201/9781315209203

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24-31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Bera, S., & Shrivastava, V. K. (2020). Analysis of Various Optimizers on Deep Convolutional Neural Network Model in the Application of Hyperspectral Remote Sensing Image Classification. *International Journal of remote sensing*, *41*(7), 2664-2683. https://doi.org/10.1080/01431161.2019.1694725

Bhosle, K., & Musande, V. (2020). Evaluation of CNN model by comparing with convolutional autoencoder and deep neural network for crop classification on hyperspectral imagery. *Geocarto International*, 1-15. https://doi.org/10.1080/10106049.2020.1740950

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). *A training algorithm for optimal margin classifiers*. Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory. https://doi.org/10.1145/130385.130401

Breiman, L. (2001). Random Forests. *Machine learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press. https://doi.org/10.1201/9781315139470

Chan, J. C. W., & Paelinckx, D. (2008). Evaluation of Random Forest and Adaboost Tree-based Ensemble Classification and Spectral Band Selection for Ecotope Mapping Using Airborne Hyperspectral Imagery. *Remote Sensing of Environment*, *112*(6), 2999-3011. https://doi.org/10.1016/j.rse.2008.02.011

Chen, S., Jin, M., & Ding, J. (2020). Hyperspectral remote sensing image classification based on dense residual three-dimensional convolutional neural network. *Multimedia Tools and Applications*, 1-24. https://doi.org/10.1007/s11042-020-09480-7

Chen, Y., Jiang, H., Li, C., Jia, X., & Ghamisi, P. (2016). Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(10), 6232-6251. https://doi.org/10.1109/TGRS.2016.2584107

Cheng, G., Yan, B., Shi, P., Li, K., Yao, X., Guo, L., & Han, J. (2022). Prototype-CNN for Few-Shot Object Detection in Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-10. https://doi.org/10.1109/TGRS.2021.3078507

Chollet, F. (2015). *Keras*. https://github.com/fchollet/keras

Christovam, L. E., Pessoa, G. G., Shimabukuro, M. H., & Galo, M. L. B. T. (2019, 10–14 June 2019). *Land Use and Land Cover Classification Using Hyperspectral Imagery: Evaluating the Performance of Spectral Angle Mapper, Support Vector Machine and Random Forest*. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, Enschede, The Netherlands. https://doi.org/10.5194/isprs-archives-XLII-2-W13-1841-2019

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine learning*, *20*(3), 273-297. https://doi.org/10.1007/BF00994018

Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2022). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, *503*, 92-108. https://doi.org/10.1016/j.neucom.2022.06.111

Erturk, A., Iordache, M. D., & Plaza, A. (2015). Sparse unmixing-based change detection for multitemporal hyperspectral images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(2), 708-719. https://doi.org/10.1109/JSTARS.2015.2477431

Fırat, H., Asker, M. E., Bayındır, M. İ., & Hanbay, D. (2022). Hybrid 3D/2D Complete Inception Module and Convolutional Neural Network for Hyperspectral Remote Sensing Image Classification. *Neural Processing Letters*, 1-44. https://doi.org/10.1007/s11063-022-10929-z

Foody, G. M. (2004). Thematic map comparison. *Photogrammetric Engineering & Remote Sensing*, *70*(5), 627-633. https://doi.org/10.14358/PERS.70.5.627

Ghanbari, H., Mahdianpari, M., Homayouni, S., & Mohammadimanesh, F. (2021). A meta-analysis of convolutional neural networks for remote sensing applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 3602-3613. https://doi.org/10.1109/JSTARS.2021.3065569

Gualtieri, J., Chettri, S. R., Cromp, R., & Johnson, L. (1999). *Support Vector Machine Classifiers as Applied to AVIRIS Data*. Proceedings Eighth JPL Airborne Geoscience Workshop, Pasadena.

Hang, R., Li, Z., Ghamisi, P., Hong, D., Xia, G., & Liu, Q. (2020). Classification of Hyperspectral and LiDAR Data Using Coupled CNNs. *arXiv preprint arXiv:2002.01144v1*, *1*, 1-12. https://doi.org/10.48550/arXiv.2002.01144

Hao, W., Yizhou, W., Yaqin, L., & Zhili, S. (2020). *The role of activation function in CNN*. 2nd International Conference on Information Technology and Computer Application (ITCA). https://doi.org/https://doi.org/10.1109/ITCA52113.2020.00096

Heiden, U., Heldens, W., Roessner, S., Segl, K., Esch, T., & Mueller, A. (2012). Urban structure type characterization using hyperspectral remote sensing and height information. *Landscape and urban Planning*, *105*(4), 361-375. https://doi.org/10.1016/j.landurbplan.2012.01.001

Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). *A practical guide to support vector classification*. Taipei, Taiwan.

Karantzalos, K., Karakizi, C., Kandylakis, Z., & Antoniou, G. (2018). *HyRANK Hyperspectral Satellite Dataset I (Version v001)*. https://doi.org/10.5281/zenodo.1222202

Kavzoglu, T., & Colkesen, I. (2009). A Kernel Functions Analysis for Support Vector Machines for Land Cover Classification. *International Journal of Applied Earth Observation and Geoinformation*, *11*(5), 352-359. https://doi.org/10.1016/j.jag.2009.06.002

Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980*, *1*, 1-15. https://doi.org/10.48550/arXiv.1412.6980

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *ImageNet classification with deep convolutional neural networks*. Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, Lake Tahoe, Nevada. https://doi.org/10.1145/3065386

Kulkarni, V. Y., & Sinha, P. K. (2012). Pruning of random forest classifiers: A survey and future directions. International Conference on Data Science & Engineering (ICDSE). https://doi.org/10.1109/ICDSE.2012.6282329

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, *521*(7553), 436-444. https://doi.org/10.1038/nature14539

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, *86*(11), 2278-2324. https://doi.org/10.1109/5.726791

Li, Y., Zhang, H., Xue, X., Jiang, Y., & Shen, Q. (2018). Deep Learning for Remote Sensing Image Classification: A Survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(6), e1264. https://doi.org/10.1002/widm.1264

Loggenberg, K., Strever, A., Greyling, B., & Poona, N. (2018). Modelling water stress in a Shiraz vineyard using hyperspectral imaging and machine learning. *Remote Sensing*, *10*(2), 202. https://doi.org/10.3390/rs10020202

Lu, G., & Fei, B. (2014). Medical hyperspectral imaging: a review. *Journal of biomedical optics*, *19*(1), 010901. https://doi.org/10.1117/1.jbo.19.1.010901

Luo, Y., Zou, J., Yao, C., Zhao, X., Li, T., & Bai, G. (2018). *HSI-CNN: A Novel Convolution Neural Network for Hyperspectral Image.* 2018 International Conference on Audio, Language and Image Processing (ICALIP), Beijing. https://doi.org/10.1109/ICALIP.2018.8455251

Melgani, F., & Bruzzone, L. (2004). Classification of Hyperspectral Remote Sensing Images with Support Vector Machines. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(8), 1778-1790. https://doi.org/10.1109/TGRS.2004.831865

Meng, Z., Zhao, F., Liang, M., & Xie, W. (2021). Deep Residual Involution Network for Hyperspectral Image Classification. *Remote Sensing*, *13*(16), 3055. https://doi.org/10.3390/rs13163055

Misra, D. (2019). Mish: A self regularized non-monotonic activation function. *ArXiv preprint arXiv:1908.08681*. https://doi.org/10.48550/arXiv.1908.08681

Mountrakis, G., Im, J., & Ogole, C. (2011). Support Vector Machines in Remote Sensing: A Review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(3), 247-259. https://doi.org/10.1016/j.isprsjprs.2010.11.001

Nair, V., & Hinton, G. E. (2010). *Rectified Linear Units Improve Restricted Boltzmann Machines.* The 27th International Conference on Machine Learning (ICML 2010), Haifa, Israel.

Pal, M. (2005). Random Forest Classifier for Remote Sensing Classification. *International Journal of remote sensing*, *26*(1), 217-222. https://doi.org/10.1080/01431160412331269698

Pal, M., & Mather, P. (2005). Support Vector Machines for Classification in Remote Sensing. *International Journal of remote sensing*, *26*(5), 1007-1011. https://doi.org/10.1080/01431160512331314083

Park, B., & Lu, R. (2015). *Hyperspectral imaging technology in food and agriculture*. Springer. https://doi.org/10.1007/978-1-4939-2836-1

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *The Journal of machine Learning research*, *12*, 2825-2830. https://doi.org/10.48550/arXiv.1201.0490

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*, 93-104. https://doi.org/10.1016/j.isprsjprs.2011.11.002

Roy, S. K., Krishna, G., Dubey, S. R., & Chaudhuri, B. B. (2019). HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification. *IEEE Geoscience and Remote Sensing Letters*, *17*(2), 277-281. https://doi.org/10.1109/LGRS.2019.2918719

Sahin, E. K., Colkesen, I., & Kavzoglu, T. (2020). A comparative assessment of canonical correlation forest, random forest, rotation forest and logistic regression methods for landslide susceptibility mapping. *Geocarto International*, *35*(4), 341-363. https://doi.org/10.1080/10106049.2018.1516248

Seyrek, E. C., & Uysal, M. (2024). A comparative analysis of various activation functions and optimizers in a convolutional neural network for hyperspectral image classification. *Multimedia Tools and Applications*, *83*(18), 53785-53816. https://doi.org/10.1007/s11042-023-17546-5

Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., & Homayouni, S. (2020). Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 6308-6325. https://doi.org/10.1109/JSTARS.2020.3026724

Si, Y., Gong, D., Guo, Y., Zhu, X., Huang, Q., Evans, J., He, S., & Sun, Y. (2021). An Advanced Spectral–Spatial Classification Framework for Hyperspectral Imagery Based on DeepLab v3+. *Applied Sciences*, *11*(12), 5703. https://doi.org/10.3390/app11125703

Stuart, M. B., McGonigle, A. J., & Willmott, J. R. (2019). Hyperspectral imaging in environmental monitoring: a review of recent developments and technological advances in compact field deployable systems. *Sensors*, *19*(14), 3071. https://doi.org/10.3390/s19143071

Teke, M., Deveci, H. S., Haliloğlu, O., Gürbüz, S. Z., & Sakarya, U. (2013). *A short survey of hyperspectral remote sensing applications in agriculture.* 6th International Conference on Recent Advances in Space Technologies (RAST). https://doi.org/10.1109/RAST.2013.6581194

Ustuner, M. (2024). Randomized Principal Component Analysis for Hyperspectral Image Classification. *ArXiv preprint arXiv:2403.09117*. https://doi.org/10.48550/arXiv.2403.09117

Van der Meer, F. D., Van der Werff, H. M., Van Ruitenbeek, F. J., Hecker, C. A., Bakker, W. H., Noomen, M. F., Van Der Meijde, M., Carranza, E. J. M., De Smeth, J. B., & Woldai, T. (2012). Multi-and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation*, *14*(1), 112-128. https://doi.org/10.1016/j.jag.2011.08.002

Vani, S., & Rao, T. M. (2019). An experimental approach towards the performance assessment of various optimizers on convolutional neural network. 3rd international conference on trends in electronics and informatics (ICOEI). https://doi.org/10.1109/ICOEI.2019.8862686

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer - Verlag. https://doi.org/10.1007/978-1-4757-3264-1

Wang, Y., Li, Y., Song, Y., & Rong, X. (2020). The influence of the activation function in a convolution neural network model of facial expression recognition. *Applied Sciences*, *10*(5), 1897. https://doi.org/10.3390/app10051897

Waske, B., Benediktsson, J. A., Árnason, K., & Sveinsson, J. R. (2009). Mapping of Hyperspectral AVIRIS Data Using Machine-Learning Algorithms. *Canadian Journal of Remote Sensing*, *35*(sup1), S106-S116. https://doi.org/10.5589/m09-018

Xia, J., Yokoya, N., & Iwasaki, A. (2016). Hyperspectral image classification with canonical correlation forests. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(1), 421-431. https://doi.org/10.1109/TGRS.2016.2607755

Zhang, Z. H., Yang, Z., Sun, Y., Wu, Y. F., & Xing, Y. D. (2019). *Lenet-5 Convolution Neural Network with Mish Activation Function and Fixed Memory Step Gradient Descent Method*. 16th International Computer Conference on Wavelet Active Media Technology and Information Processing. https://doi.org/10.1109/ICCWAMTIP47768.2019.9067661

Zhong, Y., Hu, X., Luo, C., Wang, X., Zhao, J., & Zhang, L. (2020). WHU-Hi: UAV-borne hyperspectral with high spatial resolution (H2) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF. *Remote Sensing of Environment*, *250*, 112012. https://doi.org/10.1016/j.rse.2020.112012