RESEARCH ARTICLE

# Comparative Evaluation of Nearest Regularized Subspace and Machine Learning Techniques for Hyperspectral Image Classification

**Muhammad Osama**[1], **Sarwar Shah Khan**[2,*] **and Faheem Mahmood**[3]

[1] Department of Computer Science, Iqra National University, Swat Campus, Swat 19200, Pakistan
[2] Department of Computer Science, University of Engineering and Technology Mardan, Mardan 23200, Pakistan
[3] Department of Computer Science, City University of Science & Information Technology, Peshawar 25000, Pakistan

## Abstract

**Hyperspectral imaging (HSI) has become a powerful remote sensing and material analysis tool because it can capture detailed spectral information in hundreds of adjacent bands. Nevertheless, the high dimensionality and redundancy in HSI data make precise and efficient classification challenging. This paper presents an extensive comparative study of both traditional and state-of-the-art Machine Learning algorithms for HSI classification. Classical classifiers like Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) are compared with state-of-the-art methods like Collaborative and Sparse Representation-based approaches, Convolutional Recurrent Neural Networks (CRNN), Classification and Regression Trees (CRT), and Local Fisher Discriminant Analysis with Class Mean Modeling (LFDA-CMM). The specific emphasis is given to the Nearest Regularized Subspace (NRS) classifier family, which utilizes different distance measures—i.e., Spectral Angle Mapper, Manhattan, Euclidean, Cosine, and Chi-square distances—to achieve improved classification accuracy. Experimental results on two benchmark datasets, Indian Pines and University of Pavia, show that the proposed NRS-MD method consistently achieves better performance in accuracy, Kappa coefficient, and computational complexity. These results emphasize the capability of subspace models with regularization to meet the challenges of hyperspectral image classification and present valuable information for choosing appropriate methods in practical applications.**

**Keywords**: classification, hyperspectral image, indian pines, nearest regularized subspace, university of pavia.

## 1 Introduction

Hyperspectral imaging (HSI) is a fast advancing technology in the area of remote sensing. It creates a complex image cube with many spectral dimensions. Using narrow and adjacent spectral bands along the electromagnetic spectrum, HSI provides accurate spectral fingerprints that reveal

**Academic Editor:**
Feng Ding

**Submitted:** 25 July 2025
**Accepted:** 05 September 2025
**Published:** 05 November 2025

**Vol.** 2, **No.** 4, 2025.
10.62762/TIS.2025.224024

***Corresponding author:**
Sarwar Shah Khan
sskhan0092@gmail.com

**Citation**
Osama, M., Khan, S. S., & Mahmood, F. (2025). Comparative Evaluation of Nearest Regularized Subspace and Machine Learning Techniques for Hyperspectral Image Classification. *ICCK Transactions on Intelligent Systematics*, 2(4), 213–223.

© 2025 ICCK (Institute of Central Computation and Knowledge)

the material composition of objects. This distinctive capability makes HSI particularly valuable in military and security applications, where the ability to discriminate materials and detect camouflaged targets from aerial or satellite platforms addresses critical operational requirements [1]. As hyperspectral remote sensing technology advances swiftly, it makes it possible to quickly and easily collect hundreds of short spectral wavelengths for every single image pixel. This capability is made possible by airborne sensors like HyMAP, AVIRIS, Hyperion, and HYDICE, which collect detailed spectral data from above the Earth's surface. This rich spectral reflectance characteristic enables the precise discrimination of different materials with high accuracy. This feature proves immensely valuable across a range of areas. Food safety inspection and control have effectively used HSI [2], ecological science [3], ecological management, precision agriculture [4], machine vision, remote sensing, mineralogy and material Science, medical imaging, forensics and archaeology, environmental monitoring [5] and military applications, all of which have diverse and demanding requirements used in HSI.

There have been many approaches to HSI supervised classification established, such as Nearest Neighbor, K-Nearest Neighbor (KNN), and neural networks. They were used for a variety of purposes, such as satellite remote sensing imaging [6]. Typical classifiers include maximum likelihood estimation in logistic regression; Support Vector Machines (SVM) and Extreme Learning Machines (ELM) rely solely on spectral information.

However, these classifiers are inadequate because of the high correlation and redundancy among the spectral bands. Additionally, they ignore the important spatial diversity that is present in hyperspectral data, causing their poor performance. A simple method to improve classification accuracy is to develop a classifier that effectively employs both spatial and spectral information. This is achieved by gaining spatial information earlier, most commonly done through computing entropy. This combined approach seeks to fill the gap provided by traditional classifiers, ensuring better performance through consideration of the joint effect of both spectral and spatial factors [7], including low-rank representation, morphological attribute profiles, and fuzzy-based operations. A broad range of techniques for integrating such spatial preprocessing with spectral data has been systematically documented, spanning filter-based, transform-based, and classification-oriented paradigms [8]. The intersection of deep learning and hyperspectral image analysis has opened new avenues for spectral interpretation, spanning applications from close-range material characterization to large-scale airborne and satellite-based scene understanding across multiple disciplines [9].

In this research, a detailed comparison has been performed on hyperspectral image classification methods, including both conventional Machine Learning techniques and newer modern ones. Under the classical approaches, SVM and KNN were considered because of their popularity and underlying significance in pattern recognition applications. Apart from these, higher-level classification schemes were investigated, such as Collaborative Representation-based Classification (CRC) and Sparse Representation-based Classification (SRC), both of which exploit linear representation models to improve discriminative ability.

In addition, the work employs Convolutional Recurrent Neural Networks (CRNN), which merge the spatial feature extraction properties of convolutional layers with the temporal or sequential learning power of recurrent layers. The Classification and Regression Tree (CRT) algorithm is also in view for its interpretability and decision boundary complexity. Furthermore, the Local Fisher Discriminant Analysis with Class Mean Modeling (LFDA-CMM) was also tested for its efficiency in maintaining dimension reduction while keeping local discriminative information in hyperspectral data. Most significantly, the study explores a family of methods grounded on the nearest regularized subspace (NRS) classification concept. The methods seek to enhance the accuracy of classification through the inclusion of regularization in subspace modeling, enabling improved generalization and robustness.

Various distance measures were employed within this framework, resulting in several distinct versions:

- NRS-SAM: utilizing Spectral Angle Mapper as the distance metric,
- NRS-OPD: based on Optimal Pixel Distribution,
- NRS-MD: applying Mahalanobis Distance to account for variance and correlation in the data,
- NRS-ED: relying on the standard Euclidean Distance,
- NRS-CD: using Cosine Distance to measure

angular similarity, and

- NRS-X2: using the Chi-square Distance, ideal for histograms or frequency-based features.

By examining this wide range of methods, the research seeks to illuminate the relative merits and drawbacks of each methodology in coping with the high dimensionality and heterogeneity of hyperspectral imagery data, ultimately informing the choice of suitable models for practical applications.

## 2 Related Work

The variety of studies sheds light on the world of HSI in remote sensing. Researchers have been working hard to understand the complexities of hyperspectral data, focusing on challenges like making the data simpler and reducing unwanted signals. They've looked at different ways to sort images into categories, from traditional methods to newer machine learning methods, all to make the sorting more accurate. Additionally, they've explored ways to combine information about where things are with how they look, creating new methods that help understand these special images better. The Summary of literature review is shown in Table 1.

Aswathy et al. [10] introduce a de-noising technique based on sparsity theory. It effectively reduces noise in each band of the HSI by using low-pass sparse banded filter matrices, also known as AB filters. The effectiveness of this suggested method is confirmed through a thorough evaluation that includes both subjective and objective assessments. Visual interpretations involve subjective assessments of the de-noised images by human observers.

Zhu et al. [11] propose a spectral-spatial dependent global learning (SSDGL) framework for HSI classification under insufficient and imbalanced training samples. Unlike patch-based methods that suffer from redundant computation on overlapping regions, SSDGL adopts a patch-free global learning strategy that processes the entire image simultaneously. A global convolutional long short-term memory (GCL) module is incorporated to capture long short-term spectral dependencies, while a global joint attention mechanism (GJAM) reweights spectral and spatial features to highlight discriminative regions. To address class imbalance, a hierarchically balanced (H-B) sampling strategy and a weighted softmax loss are introduced to ensure adequate representation of minority classes during training. Cao et al. [12] introduced a semi-supervised feature learning technique that skillfully integrates spectral and spatial data is proposed and tailored for the classification of disjoint HSIs. The effectiveness of this suggested method is supported by empirical validation, which was carried out using three different HSI datasets.

Arslan et al. [13] used maximum likelihood, Mahalanobis distance, spectral angle mapper, and support vector machines—four commonly used classification techniques—to classify both dimensionally reduced and original (non-reduced) images. Assessing potential differences in classification accuracy between these two sets of images was the goal. The experiments carried out in this study show that dimensionality reduction techniques based on second-order statistics successfully reduced the complexity of the hyperspectral dataset.

Mou et al. [14] propose the use of deep recurrent neural networks (RNNs) for hyperspectral image classification, treating each pixel spectrum as a sequential input across spectral bands. Two network architectures are investigated: a standard RNN and a gated recurrent unit (GRU)-based network, both designed to capture long short-term spectral dependencies among contiguous bands. By modeling the spectral signature as a temporal sequence, the proposed approach effectively exploits the inherent inter-band correlations without requiring handcrafted features, achieving competitive classification accuracy on benchmark HSI datasets.

Ma et al. [16] propose a local manifold learning-based $k$-nearest-neighbor (LML-$k$NN) approach for hyperspectral image classification. Motivated by the observation that high-dimensional HSI data often lies on a low-dimensional nonlinear manifold, the method integrates local manifold learning techniques—such as locally linear embedding (LLE) and local tangent space alignment (LTSA)—with the $k$NN classifier to exploit the intrinsic geometric structure of the data. Rather than applying dimensionality reduction globally, the approach constructs local neighborhood representations that preserve the underlying manifold geometry, thereby improving inter-class separability and classification accuracy without the need for explicit feature engineering.

## 3 Methodology

### 3.1 Support Vector Machine (SVM)

SVM is a robust supervised learning model employed for classification, especially efficient

**Table 1.** Summary of literature review.

| Ref | Year | Method | Advantages | Disadvantages |
|---|---|---|---|---|
| Aswathy et al. [10] | **2015** | Sparsity-based de-noising using AB filters (low-pass sparse banded filter matrices) | Effectively reduces noise in each band; preserves signal quality | Computational overhead may increase with large datasets |
| Zhu et al. [11] | **2021** | Spectral-spatial dependency framework with global learning | Captures richer spectral–spatial relationships; flexible classification paradigms (pre-, integrated, post-, hybrid) | Complexity in defining neighborhood dependencies; may be dataset-sensitive |
| Cao et al. [12] | **2023** | Semi-supervised feature learning integrating spectral + spatial data for disjoint HSI classification | Effective with limited labeled data; exploits spectral–spatial structure | Semi-supervised approach may require careful tuning; sensitive to noisy labels |
| Arslan et al. [13] | **2017** | Classification using ML, Mahalanobis, SAM, and SVM on reduced vs. original images | Evaluates impact of dimensionality reduction; shows simpler models still effective | Performance depends on chosen reduction technique; second-order stats only |
| Ma et al. [16] | **2010** | Local manifold learning (LLE, LTSA) integrated with $k$NN classifier | Captures intrinsic manifold structure without global dimensionality reduction | Focused mainly on computational efficiency; limited to certain NRS scenarios |
| Mou et al. [14] | **2017** | Deep recurrent neural networks (RNN and GRU) treating pixel spectra as sequential data to capture long short-term spectral dependencies across bands | Exploits inter-band correlations without handcrafted features; effective spectral sequence modeling | Computationally demanding; requires structured training dictionary |

in high-dimensional domains like hyperspectral image (HSI) data. The fundamental concept of SVM is to find the best hyperplane that separates data points of various classes with maximum effectiveness. The decision function is described as follows:

$$f(\mathbf{x}) = \text{sign}(\mathbf{w}^T\mathbf{x} + b) \tag{1}$$

where $f(\mathbf{x})$ is the output label predicted by the model (either $+1$ or $-1$ in binary classification). sign$(\cdot)$ is a function that returns: $+1$ if the input is positive or $-1$ if the input is negative. $\mathbf{w}'\mathbf{x}$ is the dot product of the weight vector $\mathbf{w}$ and the input vector $\mathbf{x}$; this determines the position of $\mathbf{x}$ relative to the decision boundary. While the $b$ is the bias (intercept) term, it shifts the decision boundary.

$$\min_{\omega,b} \frac{1}{2}\|\omega\|^2 \quad \text{subject to} \quad y_i(\omega^T x + b) \geq 1 \tag{2}$$

In an optimization problem is $y_i(\omega^T x + b) \geq 1$ for all $i$. Ensures that all data points $(x_i, y_i)$ are correctly classified with a margin of at least 1. $y_i$ the true label of the $i^{th}$ data point, either +1 or -1. $\omega^T x + b$ is a Model's raw output for input $x_i$. SVM can be extended using kernel functions for non-linear classification by representing data in an extended-dimensional space [15].

### 3.2 K-Nearest Neighbors (KNN)

KNN is a non-parametric, instance-based learning method. It categorizes new input samples by calculating the distance to the k-nearest samples in the training set and assigning the class most frequent among them.

$$\hat{y} = \text{argmax}_c \sum_{i \in N_k(x)} \delta(y_i = c) \tag{3}$$

In equation (3), $\hat{y}$ is the predicted class label for a new input data point $x$. The notation $N_k(x)$ refers to the set of the $k$ nearest neighbors of $x$ in the training data, typically determined using a distance metric like Euclidean distance. The function $\delta(y_i = c)$ is an indicator function that returns 1 if the class label $y_i$ of the $i^{th}$ neighbor equals the class $c$, and 0 otherwise. The summation $\sum_{i \in N_k(x)} \delta(y_i = c)$ effectively counts how many neighbors belong to each class $c$. Finally, the argmax operator selects the class $c$ with the highest count, meaning the class most common among the $k$ nearest neighbors. Thus, KNN classifies a new data point by majority voting among its closest training examples. KNN uses different distance metrics, like Euclidean, Manhattan, or cosine distance are commonly used in HSI.

### 3.3 Collaborative Representation-based Classification (CRC)

CRC categorizes a test sample as a linear combination of all training samples with the use of -norm regularization. Unlike sparse models, CRC focuses on collaborative contributions from the whole dictionary [17, 18]. Mathematical representation of the model:

$$x = A\alpha + e \tag{4}$$

$$\hat{y} = \text{argmin}_c \|x - A_c\alpha_c\|_2 \tag{5}$$

where $A$ contains all training data. $\alpha$ tells us how to combine those samples to reconstruct the input. We narrow down $A$ to samples from just one class, so we can compare which class gives the best approximation of $x_e$.

### 3.4 Sparse Representation-based Classification (SRC)

SRC assumes that a test sample can be accurately reconstructed by only a few training samples and uses $l_1$-norm for sparse coding. Sparse coding mathematical representation is as follows:

$$\hat{\alpha} = \text{argmin}_\alpha \|\alpha\|_1 \quad \text{subject to} \quad x = A\alpha \tag{6}$$

$$\hat{y} = \text{argmin}_c \|x - A_c\alpha_c\|_2 \tag{7}$$

SRC is particularly effective when classes exhibit sparsity in the high-dimensional feature space [19].

### 3.5 Convolutional Recurrent Neural Network (CRNN)

CRNN combines convolutional layers (CNNs) to retrieve spatial attributes and recurrent layers to model spectral sequences in hyperspectral data. This hybrid network captures both spatial and spectral dependencies. CRNN is effective for pixel-wise classification in HSIs due to its deep representation capability [20].

### 3.6 Classification and Regression Tree (CRT)

CRT is a decision tree-based algorithm that recursively partitions the data space based on features that reduce classification error the most. It is simple, interpretable, and suitable for small-scale HSI tasks.

$$G = 1 - \sum_{i=1}^{n} p_i^2 \tag{8}$$

Equation (8) is commonly known as the Gini Index. It is a measure used in decision tree algorithms, such as Classification and Regression Trees, to evaluate how "pure" a node is. $G$ presents the Gini Impurity value of a node. $n$ is the number of classes or categories in the dataset. $p_i$ denotes the proportion of samples that belong to class $i$ at a particular node. The summation $\sum_{i=1}^{n} p_i^2$ calculates the sum of the squared probabilities of each class. Subtracting this sum from 1 gives the Gini Impurity. CRT constructs a tree in which nodes are decision rules and leaves are class labels [21].

### 3.7 Local Fisher Discriminant Analysis with Class Mean Modeling (LFDA-CMM)

LFDA-CMM is an improvement over standard Fisher Discriminant Analysis (FDA) that adds local structure and class mean modeling. It performs dimensionality reduction while retaining global and local class structures.

$$J(W) = \frac{W^T S_b W}{W^T S_w W} \tag{9}$$

The equation (9) is Fisher's Linear Discriminant criterion, often used in Linear Discriminant Analysis (LDA) for dimensionality reduction and classification. It aims to find the optimal projection vector $W$ that maximizes class separability. It's three main components:

$W^T$: This is the transpose of the weight (projection) vector $W$. It transforms the column vector $W$ into a row vector so that it can take part in a quadratic form operation.

$S_W$: This matrix calculates the scatter (or spread) of points within each class. It reflects the variability within individual classes and is calculated by summing the covariance matrix for each class. The smaller the within-class scatter value, the more compact and well-separated the classes are internally.

$W$ (projection vector): It is the vector along which the data will be projected. The desire is to determine the best $W$ that will have maximum class separation upon projection.

LFDA-CMM builds upon semisupervised local discriminant analysis frameworks that simultaneously preserve global class separability and local geometric structure in high-dimensional HSI data [22], extending them with class mean modeling to further stabilize the projection under limited training samples.

### 3.8 NRS Classifier

Let's suppose the three-dimensional hyperspectral image is stated, which can be considered as the collection of training pixels $X = \{x_i\}_{i=1}^{n} \in R^d$ (d is the dimension of spectral channels). Assumes a set of pixels is the dataset $D \in R^{d \times n}$ which contains $n$ training pixels of all classes and class labels $\omega_i \in \{1, 2, ..., C\}$ here the total number of classes is denoted by $C$, and for the class $i^{th}$ total available pixel samples are denoted as $n_1$ such as $\sum_{i=1}^{c} n_i = n$. The computation of the testing pixel $y$ to be classified and computed through a linear combination of the presented labeled training pixels of per-class [23, 24]. The pixel-wise classifier NRS seeks an approximation of the target sample $y$ for every class,

$$y_1 = X_i \alpha_i \tag{10}$$

where $X_i$ is the dimensionality size of total training pixels $d \times n_1$, and the weighting coefficients $\alpha_i$ is indicated by a vector of size $n_1 \times 1$. Firstly, for every class in computing $y_i$ via Eq. (10), then the class label of $y$ is found by the class with the minimum residual, which significantly contributes to classifying the misclassified pixels. That is,

$$\text{Class}(y) = \text{argmin}_{i=1,...,C}(r_i) \tag{11}$$

where $r_1 = \|y_1 - y\|_2^2$ is obtaining the residual between the approximation and the corresponding pixels. Nevertheless, the per-class weighting vector, $\alpha_i$, are unknown. In NRS, computing the $\alpha_i$ according to

$$\alpha_i = \text{argmin}_{\alpha} \|y - x_i \alpha\|_2^2 + \gamma \|r_{iy} - \alpha\|_2^2 \tag{12}$$

where $\gamma$ is a global regularization parameter that balances the minimization among the residual and regularization terms, $r_{iy}$ is a biasing Tikhonov matrix specific to every class 1 and pixel to be classified in their corresponding classes. Gemma $(r_{iy})$ is a diagonal matrix with the Euclidean norm to determine the dissimilarity among $y_i$ and $x_{i,n}$ in the main diagonal.

$$r_{iy} = \begin{bmatrix} \|y - x_{i,1}\|_2 & 0 & \cdots & 0 \\ 0 & \|y - x_{i,2}\|_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \|y - x_{i,n}\|_2 \end{bmatrix} \tag{13}$$

where $X_1, X_2, \ldots, X_n$ are column values of the $x_i$ for $i^{th}$ class. Then, the $\alpha_i$ weight vector will be computed for each testing pixel in a closed form.

$$\alpha = (X_1^T X_1 + \gamma^2 r_{iy}^T r_{iy})^{-1} X_1^T y \tag{14}$$

### 3.9 Nearest Regularized Subspace using Spectral Angle Mapper (NRS-SAM)

This method employs Spectral Angle Mapper (SAM) to estimate the angle between training and test spectra. Classification is performed by classifying the test sample into the class with the smallest angle in a regularized subspace [6]. The spectral angle mathematical equation is as follows.

$$\theta = \cos^{-1}\left(\frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}\right) \tag{15}$$

SAM is insensitive to variations in illumination and is widely utilized in spectral matching.

### 3.10 Nearest Regularized Subspace with Optimal Pixel Distribution (NRS-OPD)

NRS-OPD chooses the best-distributed representative pixels on class manifolds. Regularization is used to regularize classification, particularly when the classes intersect or have changing densities. This strategy enhances the classification reliability by concentrating on spatially representative and diverse pixels [6].

### 3.11 Nearest Regularized Subspace using Manhattan Distance (NRS-MD)

Nearest Regularized Subspace (NRS) is a classification method where every class corresponds to a subspace of its training samples. The test sample is projected on every class subspace, and the Manhattan Distance (L1-norm) between the test sample and its projection is calculated. In NRS-MD, Manhattan Distance (L1-norm) is utilized instead of the standard Euclidean distance (L2-norm), and regularization is used for improving numerical stability and preventing overfitting, particularly in high-dimensional space [6].

$$\mathrm{D_k} = \|\mathbf{x} - \hat{\mathbf{x}}_\mathrm{k}\|_1 = \sum_{i=1}^{\mathrm{d}} |\mathbf{x}_i - \hat{\mathbf{x}}_{k,i}| \tag{16}$$

### 3.12 Nearest Regularized Subspace using Chi-Square Distance (NRS-CS)

NRS-CS uses the chi-square distance to measure test and training sample dissimilarity across spectral bands.

It is useful when handling count-like data or non-linear spectral differences [25].

$$\mathrm{D}_{\chi^2}(\mathbf{x}, \boldsymbol{\mu}) = \sum_{i=1}^{n} \frac{(x_i - \mu_i)^2}{\mu_i} \tag{17}$$

NRS-CS is appropriate for applications where there is a need for sensitivity to relative differences, but not in absolute spectral values.

### 3.13 Nearest Regularized Subspace using Cosine Distance (NRS-Cosine)

NRS-Cosine employs cosine similarity to calculate the angle between the test vector and the class mean vector in a high-dimensional spectral space.

$$\mathrm{D}_{\cos}(\mathbf{x}, \boldsymbol{\mu}) = 1 - \frac{\mathbf{x} \cdot \boldsymbol{\mu}}{\|\mathbf{x}\| \|\boldsymbol{\mu}\|} \tag{18}$$

It is scale-invariant and performs well when the magnitude of the spectra changes, but the shape is comparable [25].

## 4 Experiments and Analysis

In this section, the performance of different popular hyperspectral image (HSI) classifiers is compared on a number of datasets.

### 4.1 Hyperspectral Datasets

In this paper, we employed two different datasets, like Indian Pines dataset and the University of Pavia dataset.

#### *4.1.1 Indian Pines*

The Indian pines dataset, which has been extensively used in applications for hyperspectral classification. The Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) gathered this dataset over Indiana's northwestern region. It offers a spatial resolution of 20 meters and a spectral range of 0.4 to 2.5 micrometers. The dataset has a spatial dimension of 145 × 145 pixels, and it contains 10,249 labeled pixels in total. The dataset has 200 spectral bands after water absorption channels are eliminated [6]. The reference map here is divided into 8 classes. This image is highly significant because it has mixed pixels, and all classes possess an unbalanced number of label pixels available. The challenge of class imbalance in Indian Pines has motivated spectral-spatial classification approaches that incorporate active learning and probabilistic inference to improve minority-class recognition [26].

Figure 1 shows the three-band false-colour scene and the reference map image, as commonly visualized in spectral-spatial classification studies employing this benchmark [27].

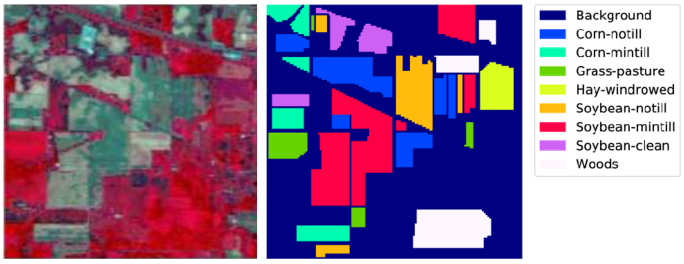



**Figure 1.** (a) Three-band false-colour scene of Indian Pines dataset. (b) reference map with 8 classes.

#### *4.1.2 University of Pavia*

The Pavia University dataset is a well-known hyperspectral image dataset that was captured over the city of Pavia, Italy, by a hyperspectral sensor mounted on an airborne. The image was gathered by the ROSIS (Reflective Optics System Imaging Spectrometer) sensor and contains high-resolution spectral data. The dataset includes a ground truth map that labels different land-cover classes such as buildings, trees, roads, and shadows. It contains 103 spectral bands. The image size is 610 x 340 pixels [25]. A three-band false-colour image and ground-truth map are shown in Figure 2.

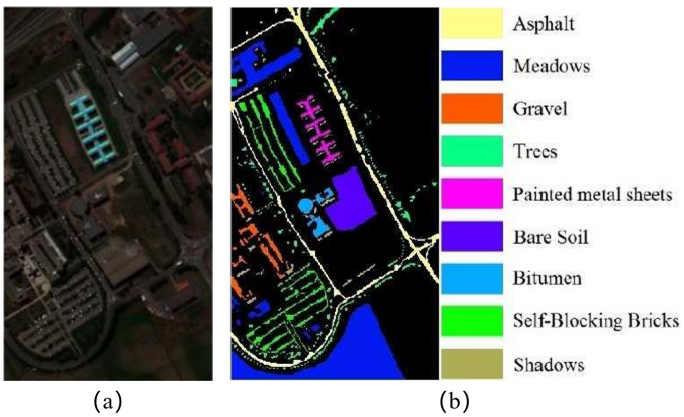



**Figure 2.** (a) Three-band false-colour scene of the University of Pavia dataset. (b) reference map with 9 classes.

### 4.2 Performance Assessment Measures

#### *4.2.1 Overall Accuracy (OA)*

Accuracy measures how effectively a classification model correctly identifies the objects within a hyperspectral image. It is determined as the ratio of

properly identified pixels to the number of pixels in the image.

#### 4.2.2 *Average Accuracy (AA)*

Average accuracy is obtained by calculating the mean accuracy across all individual classes. For every class, it is calculated by dividing the number of pixels correctly classified in a class by the number of pixels in the class. This makes every class contribute proportionally, irrespective of size.

#### 4.2.3 *Kappa Coefficient*

The Kappa coefficient measures the concordance between the classified prediction and the truth, taking into consideration the fact that it can happen by chance as well. The larger the Kappa value, the more robust the level of agreement over random chance.

#### 4.2.4 *Time Complexity*

Time complexity is the amount of time an algorithm takes to run. It usually varies with the size of the input data and is used to judge whether an algorithm can be used in real-time or big-scale applications. As the input size grows, the time complexity tends to rise as well.

### 4.3 Results and Analysis

**Table 2.** Classification results on the Indian Pines data set are given by the accuracy together with the OA, AA, Kappa coefficient, and computation time (in seconds).

| Methods | OA | AA | Kappa | Time |
|---|---|---|---|---|
| SVM | 81.04 | 82.73 | 80.44 | 29.72 |
| KNN | 75.81 | 76.91 | 73.01 | 29.88 |
| CRC | 80.29 | 80.91 | 80.60 | 32.01 |
| SRC | 87.62 | 88.90 | 85.44 | 36.43 |
| CRNN | 83.91 | 85.37 | 83.09 | 51.03 |
| CRT | 81.44 | 85.05 | 77.590 | 91.23 |
| LFDA-CMM | 84.00 | 88.91 | 81.97 | **26.28** |
| NRS-SAM | 86.09 | 87.20 | 84.91 | 29.82 |
| NRS-OPD | 86.32 | 88.10 | 85.22 | 33.13 |
| NRS-MD | **88.86** | **89.66** | **86.38** | 29.68 |
| NRS-ED | 85.71 | 88.77 | 83.56 | 31.94 |
| NRS-CS | 79.23 | 88.03 | 80.22 | 31.39 |
| NRS-X2 | 88.55 | 87.92 | 81.54 | 27.20 |

This section evaluates the performance and efficiency of different machine learning models on two benchmark datasets. For Indian Pines, 10% of labeled samples per class were randomly selected for training and the remaining 90% were used for testing. For University of Pavia, the same 10% per-class random split was applied. All experiments were repeated five times with different random seeds, and the reported metrics represent the mean values across runs. The regularization parameter $\gamma$ in NRS-based methods was set to 0.01 following [6]. The classifiers are compared using four primary metrics: overall accuracy, average accuracy, Kappa coefficient, and computation time. The models under comparison include SVM, KNN, CRC, SRC, CRNN, CRT, LFDA-CMM, and some Nearest Regularized Subspace (NRS) variants like NRS-SAM, NRS-OPD, NRS-MD, and NRS-ED.

**Table 3.** The classification results on the University of Pavia dataset include the accuracy with the OA, AA, Kappa coefficient, and the time of computation (in seconds).

| Methods | OA | AA | Kappa | Time |
|---|---|---|---|---|
| SVM | 83.72 | 84.54 | 86.32 | 23.33 |
| KNN | 82.84 | 83.46 | 84.54 | **22.33** |
| CRC | 85.77 | 86.67 | 85.33 | 23.33 |
| SRC | 82.33 | 85.76 | 84.44 | 23.49 |
| CRNN | 81.83 | 84.33 | 84.65 | 34.30 |
| CRT | 86.43 | 89.57 | 86.43 | 43.55 |
| LFDA-CMM | 87.54 | 88.22 | 89.00 | 23.23 |
| NRS-SAM | 86.43 | 88.76 | 83.76 | 52.43 |
| NRS-OPD | 86.22 | 90.43 | 89.32 | 43.65 |
| NRS-MD | **89.34** | **90.89** | **89.65** | 23.33 |
| NRS-ED | 88.98 | 89.33 | 89.44 | 24.54 |
| NRS-CS | 74.44 | 83.43 | 86.54 | 24.31 |
| NRS-$X^2$ | 89.33 | 85.76 | 86.33 | 25.46 |

Table 2 shows the classification results on the Indian Pines dataset. For highlighting performance, the best three algorithms are color-marked: the light blue marks the best, the light orange marks the second-best, and the light green marks the third-best. Interestingly enough, the NRS-MD algorithm yielded an overall accuracy of 88.86%, an average accuracy of 89.66%, and a Kappa coefficient of 86.38%. A visual comparison of various machine learning methods is presented in Figures 3 and 4. The top-performing results in each case are marked with a green circle for easy identification.

Table 3 shows the classification results on the University of Pavia dataset. Among the techniques, NRS-MD is distinguished by a Overall Accuracy (OA) of 89.34%, an average accuracy of 90.89%, and a Kappa Coefficient of 89.65%. Computational time is also a key factor in evaluating hyperspectral image classification methods. As shown in Table 2 (Indian Pines), LFDA-CMM achieves the lowest computation time of 26.28 seconds. On the University of Pavia dataset (Table 3), KNN records the fastest time at

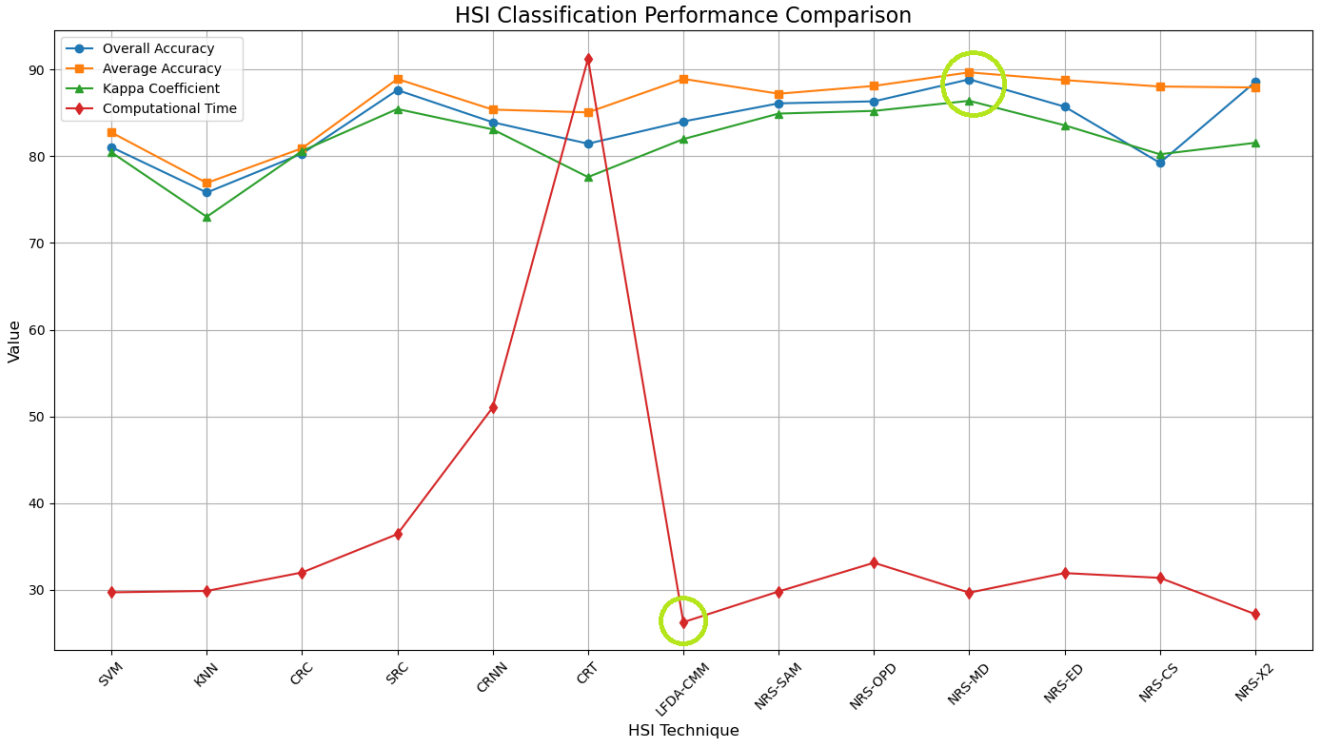



**Figure 3.** HSI classification performance comparison on different machine learning techniques.

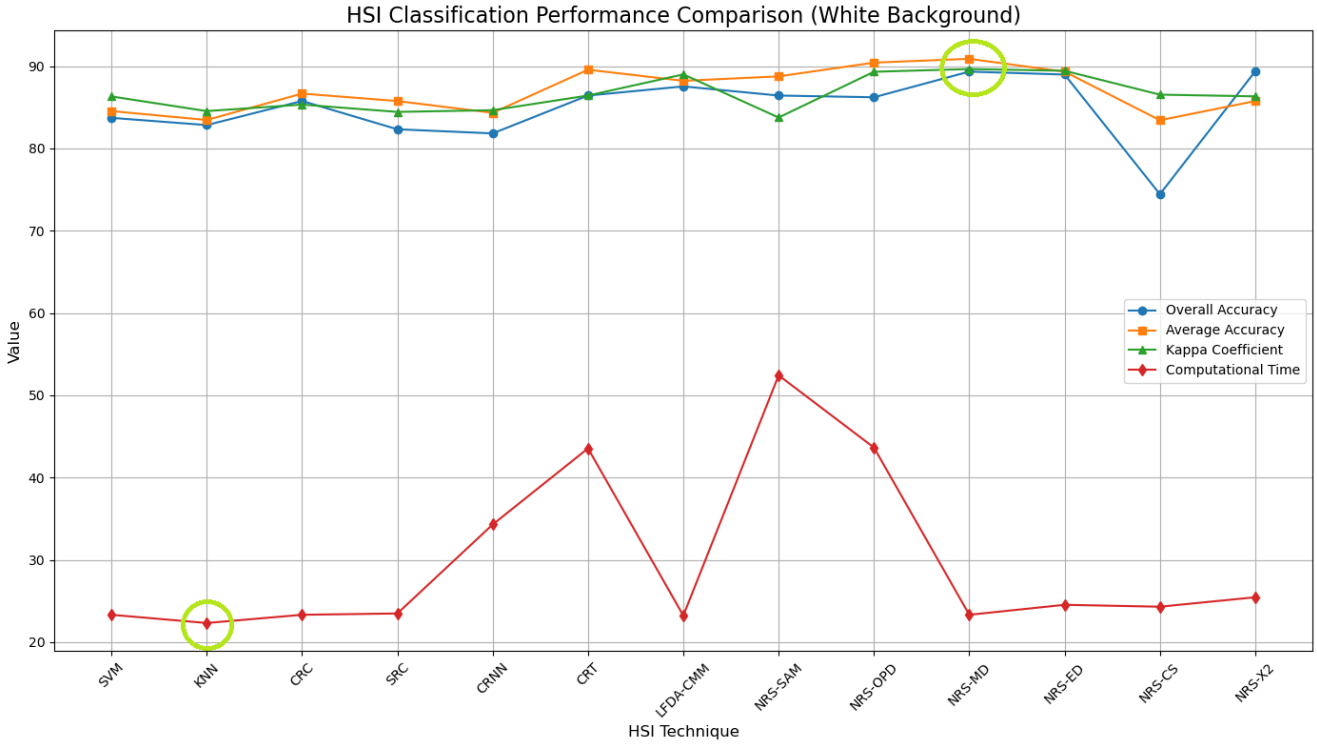



**Figure 4.** Comparison of HSI classification performance of different machine learning techniques.

22.33 seconds, while the proposed NRS-MD method also demonstrates competitive efficiency with 23.33 seconds. In Tables 2 and 3 , the colors blue, orange, and green indicate the highest, second-highest, and third-highest results, respectively.

## 5 Conclusion

This work presents a detailed comparative evaluation of several classification strategies for hyperspectral image data, encompassing both established Machine Learning techniques and more recent, advanced

models. Although conventional approaches such as SVM and KNN are a good starting point, they tend to fail with the inherent spectral redundancy and also lack the spatial sensitivity required for optimal results. On the other hand, modern approaches such as SRC, CRC, CRNN, and LFDA-CMM offer significant improvements by incorporating spectral-spatial information and advanced representation mechanisms. Among all the evaluated models, the Nearest Regularized Subspace (NRS) classifiers-particularly the variant utilizing Mahalanobis Distance (NRS-MD)-demonstrated the most balanced and effective performance across multiple metrics. NRS-MD attained maximum classification accuracy and Kappa coefficient on both Indian Pines and University of Pavia datasets while maintaining competitive computational efficiency, trailing only the fastest methods by a small margin. This suggests that integrating regularization with adaptive distance measures is a promising direction for HSI classification tasks. Overall, the findings support the conclusion that no single model fits all scenarios, but regularized subspace-based approaches stand out for their adaptability, robustness, and high accuracy. Future research may focus on hybrid models that further integrate spectral-spatial information and reduce the computational overhead while maintaining classification fidelity.

## Data Availability Statement

Data will be made available on request.

## Funding

This work was supported without any funding.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Ethical Approval and Consent to Participate

Not applicable.

## References

[1] Shimoni, M., Haelterman, R., & Perneel, C. (2019). Hyperspectral imaging for military and security applications: Combining myriad processing and sensing techniques. *IEEE Geoscience and Remote Sensing Magazine, 7*(2), 101-117. [CrossRef]

[2] Feng, Y. Z., & Sun, D. W. (2012). Application of hyperspectral imaging in food safety inspection and control: A review. *Critical Reviews in Food Science and Nutrition, 52*(11), 1039–1058. [CrossRef]

[3] Sun, G., Jiao, Z., Zhang, A., Li, F., Fu, H., & Li, Z. (2021). Hyperspectral image-based vegetation index (HSVI): A new vegetation index for urban ecological research. *International Journal of Applied Earth Observation and Geoinformation, 103*, 102529. [CrossRef]

[4] Ram, B. G., Oduor, P., Igathinathane, C., Howatt, K., & Sun, X. (2024). A systematic review of hyperspectral imaging in precision agriculture: Analysis of its current state and future prospects. *Computers and Electronics in Agriculture, 222*, 109037. [CrossRef]

[5] Rajabi, R., Zehtabian, A., Singh, K. D., Tabatabaeenejad, A., Ghamisi, P., & Homayouni, S. (2024). Hyperspectral imaging in environmental monitoring and analysis. *Frontiers in Environmental Science, 11*, 1353447. [CrossRef]

[6] Khan, S. S., Ran, Q., Khan, M., & Zhang, M. (2020). Hyperspectral image classification using nearest regularized subspace with Manhattan distance. *Journal of Applied Remote Sensing, 14*(3), 032604-032604. [CrossRef]

[7] He, L., Li, J., Liu, C., & Li, S. (2017). Recent advances on spectral–spatial hyperspectral image classification: An overview and new guidelines. *IEEE Transactions on Geoscience and Remote Sensing, 56*(3), 1579-1597. [CrossRef]

[8] Plaza, A., Benediktsson, J. A., Boardman, J. W., Brazile, J., Bruzzone, L., Camps-Valls, G., ... & Trianni, G. (2009). Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment, 113*, S110–S122. [CrossRef]

[9] Signoroni, A., Savardi, M., Baronio, A., & Benini, S. (2019). Deep learning meets hyperspectral image analysis: A multidisciplinary review. *Journal of Imaging, 5*(5), 52. [CrossRef]

[10] Aswathy, C., Sowmya, V., & Soman, K. P. (2015). Hyperspectral image denoising using low pass sparse banded filter matrix for improved sparsity based classification. *Procedia Computer Science, 58*, 26–33. [CrossRef]

[11] Zhu, Q., Deng, W., Zheng, Z., Zhong, Y., Guan, Q., Lin, W., ... & Li, D. (2021). A spectral-spatial-dependent global learning framework for insufficient and imbalanced hyperspectral image classification. *IEEE Transactions on Cybernetics, 52*(11), 11709-11723. [CrossRef]

[12] Cao, X., Li, C., Feng, J., & Jiao, L. (2023). Semi-supervised feature learning for disjoint hyperspectral imagery classification. *Neurocomputing, 526*, 9–18. [CrossRef]

[13] Arslan, O., Akyürek, Ö., & Kaya, Ş. (2017). A comparative analysis of classification methods for hyperspectral images generated with conventional dimension reduction methods. *Turkish Journal of*

*Electrical Engineering and Computer Sciences, 25*(1), 58–72. [CrossRef]

[14] Mou, L., Ghamisi, P., & Zhu, X. X. (2017). Deep recurrent neural networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing, 55*(7), 3639–3655. [CrossRef]

[15] Mercier, G., & Lennon, M. (2003). Support vector machines for hyperspectral image classification with spectral-based kernels. In *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings* (Vol. 1, pp. 288–290). [CrossRef]

[16] Ma, L., Crawford, M. M., & Tian, J. (2010). Local manifold learning-based $k$-nearest-neighbor for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing, 48*(11), 4099–4109. [CrossRef]

[17] Li, W., & Du, Q. (2014). Joint within-class collaborative representation for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*(6), 2200–2208. [CrossRef]

[18] Li, W., Du, Q., Zhang, F., & Hu, W. (2014). Collaborative-representation-based nearest neighbor classifier for hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters, 12*(2), 389–393. [CrossRef]

[19] Li, W., Du, Q., Zhang, F., & Hu, W. (2016). Hyperspectral image classification by fusing collaborative and sparse representations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*(9), 4178–4187. [CrossRef]

[20] Wu, H., & Prasad, S. (2017). Convolutional recurrent neural networks for hyperspectral data classification. *Remote Sensing, 9*(3), 298. [CrossRef]

[21] Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (2017). *Classification and regression trees*. Chapman and Hall/CRC.

[22] Liao, W., Pizurica, A., Scheunders, P., Philips, W., & Pi, Y. (2012). Semisupervised local discriminant analysis for feature extraction in hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing, 51*(1), 184–198. [CrossRef]

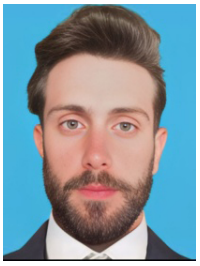
[23] Li, W., Tramel, E. W., Prasad, S., & Fowler, J. E. (2013). Nearest regularized subspace for hyperspectral classification. *IEEE Transactions on Geoscience and Remote Sensing, 52*(1), 477–489. [CrossRef]

[24] Li, W., & Du, Q. (2013). Unsupervised nearest regularized subspace for anomaly detection in hyperspectral imagery. In *2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS* (pp. 1055–1058). [CrossRef]

[25] Khan, S. S., Khan, M., Haider, S., & Damaševičius, R. (2022). Hyperspectral image classification using NRS with different distance measurement techniques. *Multimedia Tools and Applications, 81*(17), 24869-24885. [CrossRef]

[26] Li, J., Bioucas-Dias, J. M., & Plaza, A. (2012). Spectral–spatial classification of hyperspectral data using loopy belief propagation and active learning. *IEEE Transactions on Geoscience and Remote Sensing, 51*(2), 844–856. [CrossRef]

[27] Wei, Y., Zhou, Y., & Li, H. (2017). Spectral-spatial response for hyperspectral image classification. *Remote Sensing, 9*(3), 203. [CrossRef]



**Muhammad Osama** is a Master's candidate in Computer Science at Iqra National University, Swat Campus, Pakistan. He completed his undergraduate studies in Computer Science at Shaheed Benazir Bhutto University, Sheringal. His primary research focuses on Image Processing, Machine Learning, and Deep Learning. (Email: osama@gmail.com)



**Sarwar Shah Khan** is a Lecturer in the Department of Computer Science at the University of Engineering and Technology. He completed his Ph.D. in December 2024, specializing in the field of medical image processing. With seven years of experience in teaching and research, Dr. Sarwar has published over 42 research articles in reputable journals and conferences. His research interests extend to machine learning applications in image processing, and he has developed frameworks that integrate artificial intelligence and natural language processing for various applications. Currently, he is actively pursuing research in two major areas: Data Science (including text processing, natural language processing, and machine learning) and Digital Image Processing. His contributions have significantly enriched these domains. (Email: sskhan0092@gmail.com)



**Faheem Mahmood** is a Lecturer in Computer Science with over seven years of academic teaching experience across various universities in Pakistan. He holds a Master's degree in Computer Science from FAST—National University of Computer and Emerging Sciences, Islamabad. Throughout his career, he has delivered courses ranging from programming fundamentals to advanced artificial intelligence and conducted workshops on machine learning. His research interests focus on applying machine learning to astrophysics, engineering multi-agent software systems powered by large language models, and AI applications in interdisciplinary fields. Currently pursuing doctoral studies, he aspires to bridge theoretical AI advancements with practical applications. (Email: faheemmahmood1@gmail.com)