Koc. J. Sci. Eng., 8(2): (2025) 121-132 https://doi.org/10.34088/kojose.1658929

Kocaeli University

Kocaeli Journal of Science and Engineering

http://dergipark.org.tr/kojose

Refining Sparse Coding Dictionaries Using High Dimensional Model Representation for Hyperspectral Imagery

Süha TUNA [1, *]

[1] *Computational Science and Engineering, Informatics Institute, Istanbul Technical University, Istanbul, 34469, Türkiye,* ***ORCID:*** *0000-0002-9492-6896*

Article Info

Research paper

Received : *March 16, 2025*
Accepted : *June 23, 2025*

Keywords

Classification
Dictionary Refinement
High Dimensional Model Representation
Hyperspectral Imagery
Machine Learning
Optimization
Sparse Coding

Abstract

Hyperspectral imaging, known for its rich spectral and spatial details, finds applications in remote sensing, biomedical engineering, and quality control. Nonetheless, the high dimensionality and large data volume of hyperspectral images create substantial challenges in efficient processing and classification. Sparse coding-based techniques are widely employed to tackle these challenges. Therefore, an appropriate dictionary should be constructed to improve the efficacy of the sparse coding-based methods. This study introduces a dictionary refinement method that enhances sparse coding-based classification by exploiting an efficient denoising and decorrelation technique named High Dimensional Model Representation. This technique decomposes the 3-D hyperspectral data into manageable components, effectively reducing noise and correlations. Then, a refined dictionary is acquired by using random spectral signals of the hyperspectral data under consideration. The sparse coding-based classifier adopting the refined dictionary is exploited to improve the classification accuracy. Experimental results on widely used HS datasets show that the proposed method significantly boosts classification accuracy. This method leverages the benefits of denoising and decorrelation of the High Dimensional Model Representation method to generate a refined dictionary and provide a robust and efficient solution for hyperspectral image classification.

1. Introduction

Hyperspectral (HS) images capture electromagnetic energy across hundreds of narrow spectral bands, forming a 3-D tensor structure because of their rich spectral and spatial information [1]. These images have diverse applications, including food quality assessment [2], artwork authentication [3], drug forgery detection [4], and medical imaging [5]. In remote sensing, HS imagery is widely used for land cover classification, mineral exploration, environmental monitoring, and precision agriculture [6]. However, the high dimensionality and strong spectral correlation of HS data result in excessive computational complexity and storage demands, necessitating efficient feature extraction techniques [7].

The classification of endmembers is a crucial task in HS imagery. Researchers have developed several algorithms of various types to address this problem. The approaches can be grouped into three main categories. These are Machine Learning techniques, deep learning approaches and hybrid methods, respectively. Traditional machine learning techniques for HS image classification primarily focus on spectral and spatial information, with spectral-based classification treating each pixel as an independent entity. Among the widely used classifiers, Support Vector Machines (SVM) have gained prominence due to their effectiveness in high-dimensional feature spaces, benefiting from kernel functions such as radial basis function (RBF) and polynomial kernels [8]. Another commonly employed technique is K-Nearest Neighbors (KNN), which assigns labels based on the majority class of neighboring pixels, making it a simple yet effective non-parametric approach [9]. Additionally, decision trees and random forests provide hierarchical decision structures that efficiently classify HS data based on spectral attributes [10]. To enhance classification accuracy, spatial-spectral approaches have been developed, integrating spatial information with spectral features. These traditional techniques, while effective, often struggle with high-dimensional data and computational inefficiencies, paving the way for deep learning-based approaches that leverage

* Corresponding Author: suhatuna@itu.edu.tr

both spectral and spatial information more effectively. Deep learning methods transform HS image classification by automatically extracting hierarchical spectral and spatial features. Convolutional Neural Networks (CNNs) are among the most prominent approaches, where 1-D CNNs focus on spectral information, 2-D CNNs capture spatial patterns, and 3-D CNNs extract spectral-spatial features simultaneously [11]. Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, leverage sequential spectral dependencies to improve classification accuracy [12]. More recently, transformer-based models have demonstrated remarkable success in HS classification by employing self-attention mechanisms to model long-range dependencies across spectral-spatial domains. These methods have significantly improved classification performance; however, they often require large amounts of labeled data and substantial computational resources. CNN-RNN models combine spatial and spectral feature extraction, whereas attention-based CNNs refine classification by focusing on the most relevant spectral-spatial features [13].

Sparse representation-based methods have gained attention owing to their efficiency in feature selection and classification. These approaches model HS pixels as sparse linear combinations of dictionary elements (atoms) of representative spectral signatures. By enforcing sparsity constraints, these methods improve robustness against noise and enhance classification performance, even with limited training samples. Sparse Representation-based Classification (SRC) and dictionary learning techniques have been widely explored and have demonstrated competitive results in HS image analysis [14]. Recent advancements include structured sparsity models and deep sparse representation approaches, which further refine feature selection and improve computational efficiency [15].

In recent years, deep learning methods such as Convolutional Neural Networks (CNNs) have demonstrated exceptional performance in a wide range of image classification tasks. Nevertheless, their success is heavily dependent on the availability of large-scale, labelled datasets, which are often impractical to obtain in remote sensing due to the substantial cost and complexity associated with manual annotation. This constraint has led to increasing interest in alternative techniques that maintain high performance even with limited training data. SRC has gained significant attention in this regard, as it exhibits strong resilience in data-scarce environments. Beyond its data efficiency, SRC is also computationally advantageous, particularly when combined with fast greedy algorithms such as Orthogonal Matching Pursuit (OMP). Furthermore, unlike deep neural networks, which are often criticized for their lack of interpretability, SRC provides a transparent classification mechanism. The resulting sparse coefficient vector explicitly indicates which training samples are used to represent each test input, thereby offering meaningful insight into the basis of the classification decision. These characteristics make SRC especially well-suited for remote sensing applications, where data availability, computational efficiency, and interpretability are critical considerations.

Extensive research has been conducted on extracting discriminative features from HS images, particularly through tensor decomposition methods. These extracted features play a crucial role in enhancing classification accuracy and enabling data compression [16]. Classical tensor decomposition approaches, such as Tucker Decomposition and CANDECOMP/PARAFAC, have been widely adopted for HS data analysis [17]. However, more advanced techniques, including High Dimensional Model Representation (HDMR) and its variants, have demonstrated superior performance in capturing the intrinsic structure of HS data and improving classification tasks [18-21].

In this paper, we propose an approach to enhance SRC by integrating HDMR for efficient feature extraction. Initially, HDMR decomposition is applied to the 3-D HS tensor to mitigate noise and reduce spectral redundancy, facilitating efficient dictionary elements. The extracted features are then utilized to improve classification performance while reducing computational complexity. The proposed methodology aims to provide a robust framework for HS image analysis, particularly in applications in which efficient data processing and classification are critical. A central motivation of this study arises from the recognition that the performance of SRC is highly sensitive to the quality of the underlying dictionary. In HS imagery, raw spectral signatures are often contaminated by noise and exhibit strong inter-band correlations, largely due to sensor imperfections and atmospheric interference. Directly employing such unprocessed data as dictionary atoms can significantly decrease classification performance. This challenge underscores the necessity for a principled dictionary refinement strategy that explicitly mitigates both noise and spectral redundancy. In this work, we propose a systematic framework for constructing a more robust and discriminative dictionary, with the goal of enhancing SRC accuracy in HS image analysis.

A distinguishing feature of our approach is its holistic modelling of HS data through the application of High Dimensional Model Representation (HDMR) to the entire 3-D hyperspectral cube, rather than operating on individual pixels or local patches. This enables a global decomposition that systematically captures both spatial and spectral interactions, offering a more comprehensive

representation of the data structure. Moreover, HDMR provides a precise mathematical framework for hierarchical data expansion, where higher-order terms encapsulate noise and intricate variable dependencies. By truncating these higher-order residual components, we achieve a theoretically-grounded denoising and decorrelation process that surpasses conventional heuristic filtering methods. The central contribution of this work lies in the development of an HDMR-based dictionary refinement framework that enhances the quality of the representation used in sparse classification. Notably, we demonstrate that incorporating this refined dictionary into a standard SRC classifier significantly improves classification performance without requiring any modification to the classifier itself.

2. Related Work

Sparse representation-based classification (SRC) has emerged as a powerful and widely explored approach for HS image analysis due to its ability to capture the underlying structure of high-dimensional spectral data. Various methods have extended the classical SRC framework to enhance classification performance and robustness. For instance, a novel semi-supervised joint dictionary learning (S2JDL) algorithm was proposed to integrate both labeled and unlabeled data within a unified objective function that jointly minimizes reconstruction and classification errors. By leveraging the reconstruction term from unlabeled data and a discrimination-aware term from labeled samples, this method significantly improves classification accuracy across multiple benchmark HS imagery datasets [22].

Robustness to noise has also been a major concern in sparse representation for HSI, as real-world HS data are often contaminated with mixed noise types rather than ideal Gaussian noise. To address this, a unified model has been developed that integrates Gaussian and sparse noise priors with sparse representation frameworks. The resulting classification schemes, including joint SRC and superpixel-level joint SRC, demonstrate superior performance on both synthetic and real data, highlighting the importance of realistic noise modeling in practical applications [23].

Denoising remains a fundamental preprocessing step in HS image analysis, given the high sensitivity of spectral channels to noise due to limited photon capture at high spectral resolutions. An improved method combining singular value decomposition (SVD) with an adaptive block algorithm has been introduced for accurate noise estimation. Furthermore, the integration of low-rank theory with sparse representation has shown substantial improvements in denoising performance, surpassing traditional algorithms by over 3.0 dB on the Indian Pines dataset, and offering reliable noise intensity estimates [24].

Lastly, recent advancements in robust HS image denoising have emphasized the preservation of rare pixel types, which are often compromised in conventional low-rank approximation frameworks. The RhyDe algorithm exemplifies this direction by explicitly modeling low-rank structure, exploiting self-similarity, and employing collaborative sparsity to retain infrequent spectral signatures. This method has demonstrated enhanced denoising and rare pixel detection capabilities in both semi-realistic and real hyperspectral scenarios [25].

Collectively, these studies underscore the ongoing evolution of SRC-based methods, incorporating semi-supervised learning, robust noise modeling, spatial-spectral integration, and advanced denoising strategies to address the unique challenges posed by HS imagery.

The key advantage of the proposed HDMR-based framework lies in its inherent ability to model complex non-linear relationships within HS data. Traditional dimensionality reduction techniques such as Principal Component Analysis (PCA) are constrained by their linear nature and may fail to capture the intricate, non-linear interactions between spatial and spectral dimensions that are often present in HS imagery. In contrast, HDMR provides a flexible and general function decomposition methodology capable of representing these non-linear interdependencies. Additionally, while PCA may offer incidental denoising by retaining components with the largest variance, it does not explicitly distinguish between signal and noise. HDMR, by design, separates the data into structured, low-order components representing a meaningful signal and a high-order residual term that encapsulates noise and complex variable interactions, thereby offering a more principled approach to signal refinement. Furthermore, HDMR expansion terms possess clear interpretability where each component corresponds to specific contributions, such as the mean effect or variable interactions, unlike the abstract principal components derived from PCA. This interpretability facilitates a deeper understanding of the HS data's underlying structure, enhancing both analysis and decision-making processes.

3. Materials and Methods

3.1. Sparse Representations of HS Signals

HS data involves high correlation among its spectral signals because of the sensor limitations and environmental issues. These signals also contain noise. The correlation and noise complicate the determination of the true class of

the signals. Thus, decorrelation and denoising of the spectral signals become an important task in HS imagery to establish high classification accuracies.

Sparse representation of a signal enables to describe a given signal in terms of a sparse linear combination of a predefined signal dictionary [14,26]. The signals of the dictionary are called atoms. This approach works efficiently for the data having similar or highly correlated features. Since HS data has high correlation along with its spectral dimension, the spectral signals can be expressed by an appropriate linear combination of the atoms. The aim of the sparse representation is to exhibit the target signal as a linear combination of a few number of dictionary atoms. Therefore, the noise and other artefacts can be removed accordingly. On the other hand, employing well-structured dictionaries is also an option to improve the sparse coding capability for the spectral signals. If the target signal is depicted as $\boldsymbol{x}$ and the corresponding $B \times n$ dictionary (B is the number of bands and n is the number of atoms) is stated as $\boldsymbol{D}$, the sparse coding of $\boldsymbol{x}$ is given as

$$\boldsymbol{x} = \boldsymbol{D}\boldsymbol{s} \tag{1}$$

where $\boldsymbol{s}$ is a sparse vector whose non-zero entries indicate the contributions of the corresponding atoms from $\boldsymbol{D}$. So, the mathematical definition of the sparse representation for the signal $\boldsymbol{x}$ and given $\boldsymbol{D}$ and $\boldsymbol{s}$ is as follows

$$\min_{\boldsymbol{s}} \|\boldsymbol{s}\|_0 \quad s.t. \quad \boldsymbol{D}\boldsymbol{s} = \boldsymbol{x} \tag{2}$$

Since the L_0 norm is not a true norm in the mathematical sense and either not convex, this problem cannot be solved using continuous optimization techniques. Besides, the optimal solution of the corresponding problem cannot be determined in a polynomial-time scale, which concludes that the problem in Eq. (2) is NP-hard. Therefore, the problem in Eq. (2) can be solved by adopting a suitable greedy algorithm. In this concern, instead of sweeping all possible atom subsets, one can search the tree of possibilities while pruning many unlikely states. By following this approach, many possible optimality tests may be skipped and computation time can be reduced drastically [27].

In this work, an efficient greedy algorithm named the Orthogonal Matching Pursuit (OMP) is employed for exploiting the best support from the last step of the algorithm [28]. As the initial step, the OMP finds the atom which best matches $\boldsymbol{D}\boldsymbol{s}$ to $\boldsymbol{x}$ by sweeping through all n atoms (columns) of $\boldsymbol{D}$. Once this atom is found, the corresponding column will be added to an empty set called the support set, S. Given the previously chosen atoms, the next one to best fit $\boldsymbol{D}\boldsymbol{s}$ to $\boldsymbol{x}$ is again obtained by sweeping $n-1$ possibilities. Consequently, the support set S grows at each iteration. The algorithm stops when the error $\|\boldsymbol{D}\boldsymbol{s} - \boldsymbol{x}\|_2$ is small enough or a prefixed number of atoms is achieved. The pseudo-code for the OMP algorithm is provided in Algorithm 1.

Algorithm 1. OMP Algorithm

Input: $\boldsymbol{D}, \boldsymbol{s}, \boldsymbol{x}, k$
Output: $\hat{\boldsymbol{s}}$

for $i = 1,2,\ldots,k$
  Compute correlations: $\boldsymbol{c} = \boldsymbol{D}^{\boldsymbol{T}}\boldsymbol{r}$
  Select index: $j^* = \underset{j}{\operatorname{argmax}}|c_j|$
  Update support set: $S = S \cup \{j^*\}$
  Solve LS problem: $\hat{\boldsymbol{s}} = \underset{s}{\operatorname{argmin}}\|\boldsymbol{D}(:,S)\boldsymbol{s} - \boldsymbol{x}\|_2$
  Update residual: $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{D}(:,S)\hat{\boldsymbol{s}}$
end

In Alg. 1, k stands for the sparsity level and the notation $\boldsymbol{D}(:,S)$ refers to the columns of the matrix $\boldsymbol{D}$ that are indexed by the elements in the set S.

3.2. High Dimensional Model Representation

High Dimensional Model Representation (HDMR) is a statistical technique that enables the decomposition of a given function, model or data into lower-dimensional components [18,20,29,30]. This methodology provides a mathematical foundation for analyzing complex systems in which the output is influenced by multiple variables or dimensions.

HDMR is particularly advantageous when dealing with high-dimensional datasets, as it facilitates the examination and modeling of the relationships between input variables and outputs while reducing overall computational complexity [18,20]. Consequently, HDMR is widely employed in applications such as data compression [21], decorrelation [31], denoising and feature extraction [18].

As previously discussed, HS images inherently exhibit a high-dimensional structure, comprising two spatial dimensions and one spectral dimension. Suppose the HS cube is denoted as H, where its spatial and spectral dimensions are represented by the variables x_1, x_2 and x_3, respectively. The HDMR expansion of the trivariate function $H(x_1, x_2, x_3)$ is expressed as follows:

$$\begin{aligned} H(x_1,x_2,x_3) = {} & h_0 + h_1(x_1) + h_2(x_2) + h_3(x_3) \\ & + h_{12}(x_1,x_2) + h_{13}(x_1,x_3) + h_{23}(x_2,x_3) \\ & + h_{123}(x_1,x_2,x_3) \end{aligned} \tag{3}$$

In Eq. (3), the expansion constitutes a finite (2^3 terms) and additive decomposition [29,30]. Within this formulation, h_0 corresponds to the constant HDMR term, whereas $h_i(x_i)$'s and $h_{ij}(x_i, x_j)$'s represent the univariate (one-dimensional) and bivariate (two-dimensional) HDMR components, respectively. The final term, h_{123}, is referred to as the residual term, which encapsulates noise and correlations among the variables. Each HDMR term is explicitly formulated to characterize different properties of the data: the constant term h_0 signifies the overall average effect, the univariate terms capture the individual contributions of each variable, and the bivariate terms describe the interactions between specific variable pairs [25,26].

All HDMR terms obey the vanishing under integration condition [29,30]. More precisely, if η and μ are two HDMR terms that depend on a shared variable x, then:

$$\int_a^b \eta(\dots, x, \dots)\mu(\dots, x, \dots)\, dx = 0. \tag{4}$$

As stated in Eq. (4), the HDMR terms exhibit mutual orthogonality, implying that each term can be uniquely determined [29,30]. By applying the conditions outlined in Eq. (4), the constant HDMR term, h_0, is explicitly computed as follows

$$h_0 = \frac{1}{(b_1 - a_1)(b_2 - a_2)(b_3 - a_3)} \times \int_{a_1}^{b_1}\int_{a_2}^{b_2}\int_{a_3}^{b_3} H(x_1, x_2, x_3)dx_1dx_2dx_3 \tag{5}$$

Similarly, the univariate HDMR terms are determined explicitly as

$$h_1(x_1) = \frac{1}{(b_2 - a_2)(b_3 - a_3)} \times \int_{a_2}^{b_2}\int_{a_3}^{b_3} H(x_1, x_2, x_3)dx_2dx_3 - h_0$$

$$h_2(x_2) = \frac{1}{(b_1 - a_1)(b_3 - a_3)} \times \int_{a_1}^{b_1}\int_{a_3}^{b_3} H(x_1, x_2, x_3)dx_1dx_3 - h_0$$

$$h_3(x_3) = \frac{1}{(b_1 - a_1)(b_2 - a_2)} \times \int_{a_1}^{b_1}\int_{a_2}^{b_2} H(x_1, x_2, x_3)dx_1dx_2 - h_0 \tag{6}$$

Using the same methodology, the bivariate HDMR terms are evaluated uniquely as follows

$$h_{12}(x_1, x_2) = \frac{1}{(b_3 - a_3)}\int_{a_3}^{b_3} H(x_1, x_2, x_3)dx_3 - h_1(x_1) - h_2(x_2) - h_0$$

$$h_{13}(x_1, x_3) = \frac{1}{(b_2 - a_2)}\int_{a_2}^{b_2} H(x_1, x_2, x_3)dx_2 - h_1(x_1) - h_3(x_3) - h_0$$

$$h_{23}(x_2, x_3) = \frac{1}{(b_1 - a_1)}\int_{a_1}^{b_1} H(x_1, x_2, x_3)dx_1 - h_2(x_2) - h_3(x_3) - h_0 \tag{7}$$

After the HDMR terms have been uniquely determined, as described in Eq. (5), (6), and (7), they can be truncated at any chosen level, as indicated in Eq. (3). This leads to the derivation of the following HDMR approximants, which offer a denoised and decorrelated representation of $H(x_1, x_2, x_3)$ where

$$\begin{aligned} H_0 &= h_0 \\ H_1 &= H_0 + h_1(x_1) + h_2(x_2) + h_3(x_3) \\ H_2 &= H_1 + h_{12}(x_1, x_2) + h_{13}(x_1, x_3) + h_{23}(x_2, x_3) \end{aligned} \tag{8}$$

In Eq. (8), H_0, H_1, and H_2 correspond to the zeroth, first, and second-order HDMR approximants, respectively [29,30]. By utilizing these approximants, an HS dataset that is effectively denoised and decorrelated can be obtained to a specified degree. Using the approximated HS data, a refined dictionary consisting of denoised and decorrelated atoms can be constructed. This dictionary is subsequently employed to facilitate more efficient sparse coding-based analyses.

3.3. The Proposed Method

In this study, the dictionary $\boldsymbol{D}$ is constructed by directly extracting random training samples from the HS image under consideration. However, enhanced robustness can be achieved by refining this dictionary through an appropriate decorrelation strategy. To this end, HDMR is initially applied to the corresponding HS dataset, and subsequent classification of a test sample is performed based on the resulting sparse vector $\boldsymbol{s}$.

To address the reconstruction problem, we begin by estimating the sparse vector $\boldsymbol{s}$ for a given spectral sample $\boldsymbol{x}$. With a dictionary $\boldsymbol{D}$ composed of training samples, the sparse representation $\boldsymbol{s}$, satisfying the equation $\boldsymbol{Ds} = \boldsymbol{x}$, is obtained by solving the following optimization problem

$$\boldsymbol{s}^* = \text{argmin}\|\boldsymbol{s}\|_0 \quad s.t. \quad \boldsymbol{Ds} = \boldsymbol{x} \tag{9}$$

To account for the inherent approximation errors present in empirical data, the strict equality constraint in Eq. (4) is relaxed into an inequality constraint.

$$\boldsymbol{s}^* = \text{argmin}\|\boldsymbol{s}\|_0 \quad s.t. \quad \|\boldsymbol{Ds} - \boldsymbol{x}\|_2 \leq \sigma \tag{10}$$

In this context, σ denotes the error tolerance. Alternatively, the problem in Eq. (9) can be formulated as one that minimizes the approximation error while enforcing a predefined level of sparsity.

$$\boldsymbol{s}^* = \text{argmin}\|\boldsymbol{Ds} - \boldsymbol{x}\|_2 \quad s.t. \quad \|\boldsymbol{s}\|_0 \leq K_0 \tag{11}$$

In Eq. (11), K_0 denotes the predetermined upper bound on the sparsity level.

The optimization problems described above are NP-hard; however, they can be efficiently approximated using greedy pursuit algorithms such as Orthogonal Matching Pursuit (OMP) [32] and Subspace Pursuit (SP) [33]. Both algorithms are employed to determine the support of the sparse vector that approximately satisfies the problem in Eq. (9), albeit through different atom selection strategies. Specifically, the OMP algorithm incrementally adds one index to the support set until either K_0 atoms are selected or the approximation error falls below a predefined threshold. In contrast, the SP algorithm maintains a fixed support set of size K_0 and, during each iteration, it expands this set by considering K_0 new candidates, subsequently removing K_0 less significant indices from the combined set of $2K_0$ candidates. This backtracking mechanism enables SP to effectively isolate the K_0 most significant atoms. Both algorithms exhibit a computational complexity on the order of $O(BNK_0)$.

The class label of $\boldsymbol{x}$ can be predicted directly from the properties of the recovered sparse vector $\boldsymbol{s}$ [34]. We define the m-th residual, which quantifies the error between the test sample and its reconstruction using the training samples corresponding to the m-th class, as follows

$$r^m(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{D^m s^m}\|_2, \quad m = 1, \ldots, M \tag{12}$$

where M denotes the total number of distinct classes present in the HS dataset. Here, $\boldsymbol{s}^{\text{m}}$ denotes the subset of the recovered sparse coefficients that corresponds specifically to the training samples of the m-th class. Consequently, the class label of the spectral signal $\boldsymbol{x}$ is assigned based on the minimal residual error as

$$Class(\boldsymbol{x}) = \underset{m=1,\ldots,M}{\text{argmin}}\, r^m(\boldsymbol{x}) \tag{13}$$

Thus, the class label of $\boldsymbol{x}$ is determined and can be directly compared with the corresponding ground truth image for evaluation.

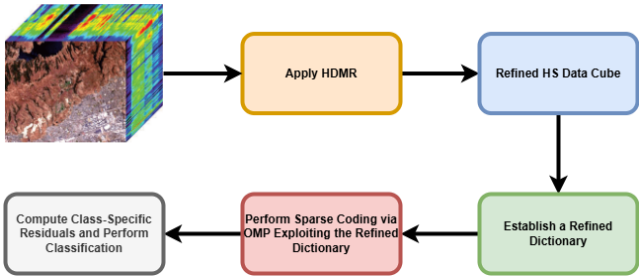

Figure 1. Flowchart of the proposed HDMR-enhanced sparse coding framework for HS image classification.

Figure 1 illustrates the overall flowchart of the proposed method, which integrates HDMR with SRC.

4. Results and Discussion

4.1. Dataset Specifications

To highlight the effectiveness and robustness of the proposed method, three publicly available HS datasets were chosen: Botswana, Indian Pines, and Pavia University. These datasets vary in spatial resolution and spectral band count. Additionally, each dataset was captured using a distinct HS sensor. This diversity provides an opportunity to evaluate the performance of the proposed method across different configurations and conditions. The characteristics of the datasets utilized in this study are summarized in Table 1.

Table 1. Specifications of the HS datasets

Dataset	**Sensor**	**Size** (R×C×B)	**Bit depth**
Botswana	HYPERION	1476 × 256 × 145	16 bit
Indian Pines	AVIRIS	145 × 145 × 200	16 bit
Pavia Uni.	ROSIS	610 × 340 × 103	16 bit

Each dataset has its unique ground truth images created by the experts. According to the ground truths Botswana dataset contains 14 distinct classes. The number of distinct classes of Indian Pines and Pavia Uni. datasets are 16 and 9, respectively. The corresponding ground truth images are displayed in Figure 2.

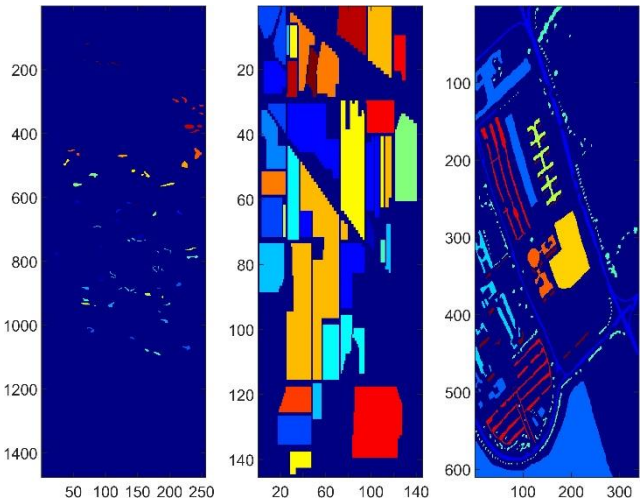

Figure 2. Ground truth images of the HS datasets. Botswana (left), Indian Pines (middle) and Pavia Uni. (right)

4.2. Experimental Setup

In this subsection, we provide a comprehensive description of the experiments conducted to ensure the reproducibility of the results. The experiments are performed using three HS datasets. To investigate the impact of the sparsity level, values ranging from 1 to 10, with increments of 1, are selected. The core principle underlying sparse representation is the reconstruction of a signal using the minimal number of representative atoms, typically favouring solutions that lie within a low-integer sparsity range. This principle is considered by constraining the number of atoms, commonly in the range of 1 to 10, which aligns directly with the objective of achieving useful sparse representations. However, as both our experimental findings and prior studies in the literature suggest, classification accuracy does not necessarily improve monotonically with increasing sparsity. Extremely sparse representations may result in underfitting, failing to capture sufficient discriminative information, whereas overly dense representations risk overfitting by incorporating atoms from irrelevant or noisy classes. To better understand this trade-off, we systematically examine classifier performance across a range of sparsity levels. To emphasize the contribution of HDMR to the classification process, we present the overall accuracy (OA) and Cohen's Kappa (Kappa) metrics for both scenarios: with and without HDMR. To assess the effectiveness of the training ratio, the dictionary atoms for each dataset are constructed using 10%, 20%, and 30% of the spectral signals from each relevant class, following the removal of background pixels. It is noteworthy that each experimental condition was replicated 10 times to verify the stability and accuracy of the results.

4.3. Comparative Evaluations

To demonstrate the effectiveness of HDMR in enhancing the overall method, we provide a comprehensive performance analysis along with a comparative evaluation of the proposed HS classification approach. To this end, multiple classification experiments are conducted for each dataset. The classification results obtained using sparse coding are compared with those incorporating HDMR, allowing for a direct assessment of the impact of HDMR on the overall process. The performance of the proposed algorithm is evaluated using OA and Kappa metrics. Additionally, classification results are reported for varying sparsity levels to examine the effect of incorporating a greater number of dictionary atoms.

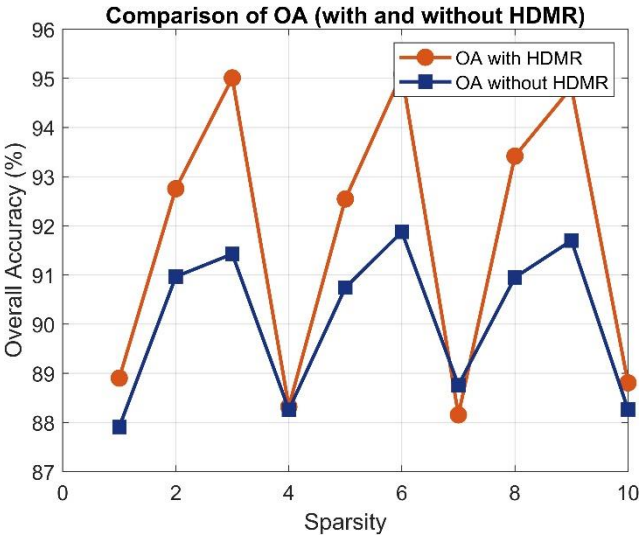

Figure 3. Comparison of the overall accuracy with respect to the sparsity level with and without HDMR for Botswana dataset.

The OA and Kappa values obtained for the Botswana dataset are presented in Figure 3 and Figure 4, respectively. It is evident that incorporating HDMR into the classification process enhances accuracy. Specifically, the OA achieved by the sparse coding-based classifier is slightly below 88% when utilizing a single atom, whereas the inclusion of HDMR increases this value to 89%. For a sparsity level of 2, the standard classifier attains an accuracy of 91%, while HDMR further improves it to approximately 93%. When three atoms are selected from the dictionary, the OA of the standard classifier reaches 91.5%, whereas HDMR leads to a significant increase, achieving 95%.

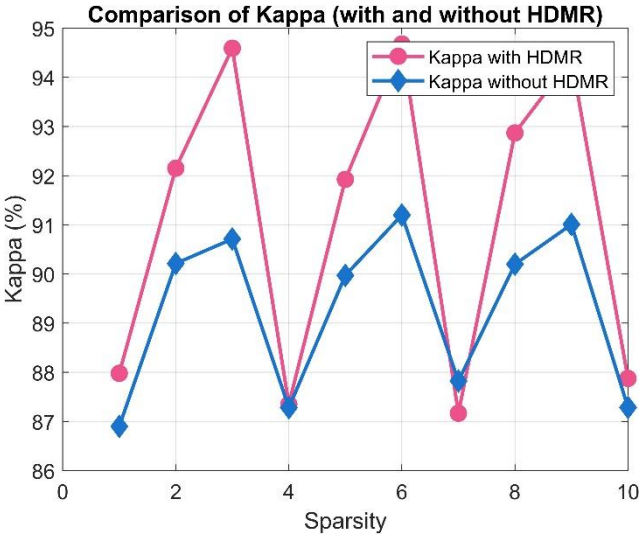

Figure 4. Comparison of the Cohen's Kappa with respect to the sparsity level with and without HDMR for Botswana dataset.

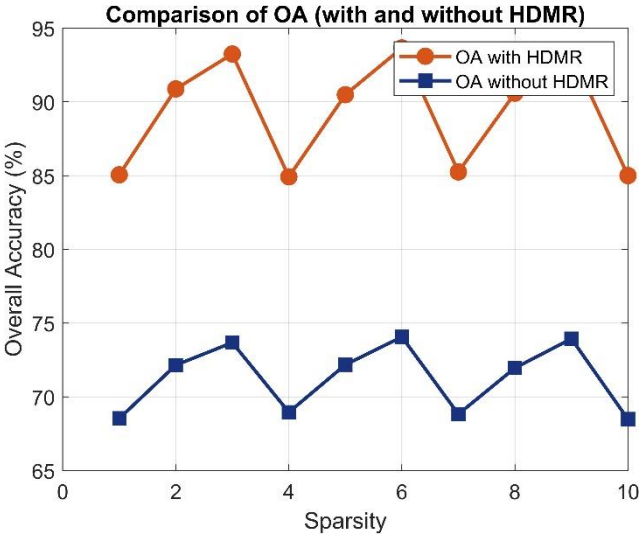

Figure 5. Comparison of the overall accuracy with respect to the sparsity level with and without HDMR for Indian Pines dataset.

The fluctuations observed in OA and Kappa metrics across different sparsity levels in sparse representation-based classification of the corresponding experiments can be attributed to the stability of sparse coding solutions and the effectiveness of dictionary selection. However, when the sparsity level is too low, the model fails to capture sufficient discriminatory information, resulting in underfitting, whereas excessive sparsity leads to inter-class interference, thereby diminishing classification performance. These fluctuations can be explained using the Restricted Isometry Property (RIP), wherein an increasing RIP constant at suboptimal sparsity levels induces instability in signal reconstruction, leading to classification errors [35]. While HDMR enhances feature selection and improves classification performance, fluctuations persist due to variations in dictionary stability across different sparsity values. Since Kappa quantifies classification agreement beyond chance, its variations are inherently linked to changes in OA.

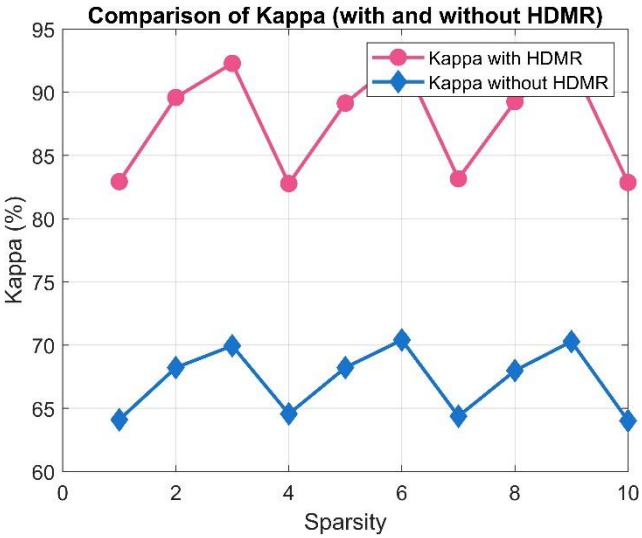

Figure 6. Comparison of the Cohen's Kappa with respect to the sparsity level with and without HDMR for Indian Pines dataset.

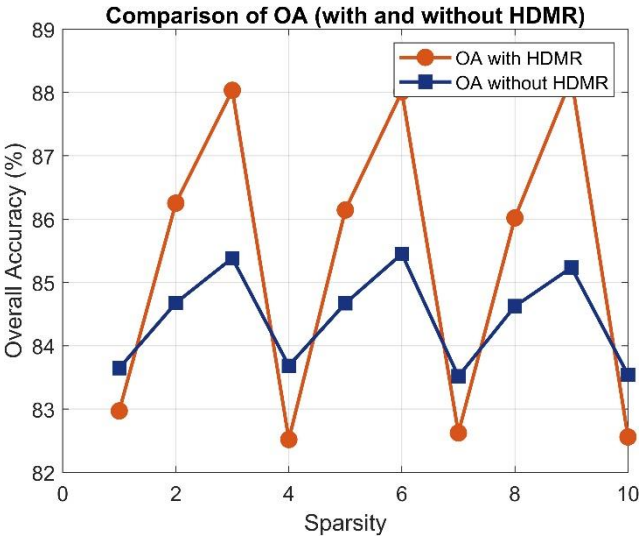

Figure 7. Comparison of the overall accuracy with respect to the sparsity level with and without HDMR for Pavia Uni. dataset.

Similar seesaw-like patterns can be observed for the Indian Pines and Pavia University HS datasets, as illustrated in Figures 5, 6, 7 and 8, respectively. In Figure 5 and Figure 6, the incorporation of HDMR leads to a significant improvement in both OA and Kappa for the Indian Pines dataset. Specifically, for a sparsity level of 1, the conventional classification approach achieves an OA of 69%, whereas the inclusion of HDMR increases it to 85%. When two atoms are selected from the corresponding dictionary, the OA reaches 72% without HDMR, while

HDMR enhances it to 91%. For a sparsity level of 3, the OA values are 74% and 94% for the cases without and with HDMR, respectively.

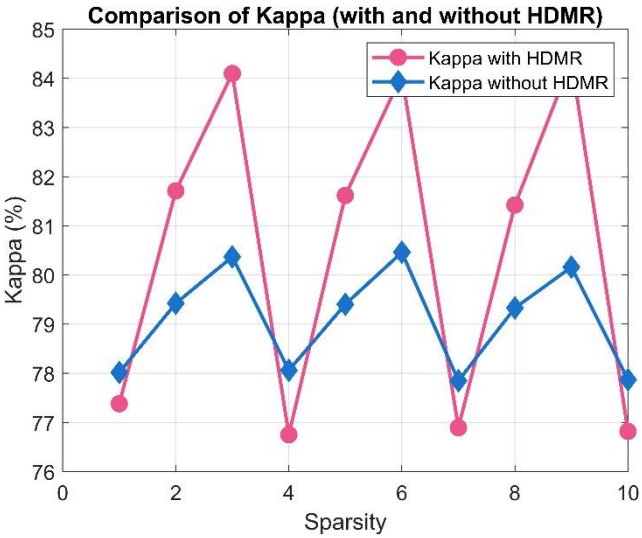

Figure 8. Comparison of the Cohen's Kappa with respect to the sparsity level with and without HDMR for Pavia Uni. dataset.

As depicted in Figure 7, HDMR effectively enhances OA for the Pavia University HS dataset. Specifically, at a sparsity level of 2, the OA improves from 84% to 86.3%, while at a sparsity level of 3, it increases from 85.4% to 88%.

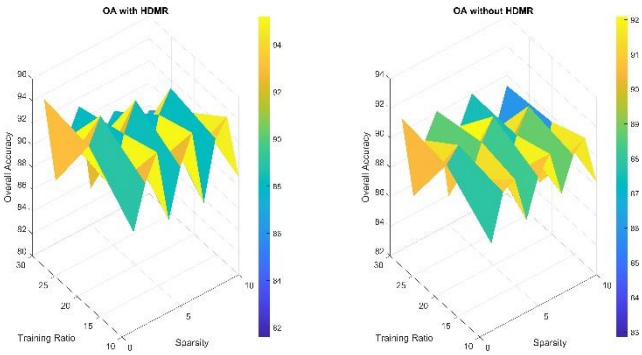

Figure 9. Impact of the training ratio and sparsity level on overall accuracy for Botswana dataset. With HDMR (left) and without HDMR (right)

To demonstrate the impact of the training ratio on the classification performance, multiple experiments are conducted. In this context, OA values are reported for both with and without HDMR cases across different sparsity levels and training ratios. The corresponding results for the Botswana, Indian Pines, and Pavia University datasets are presented in Figures 9, 10, and 11, respectively.

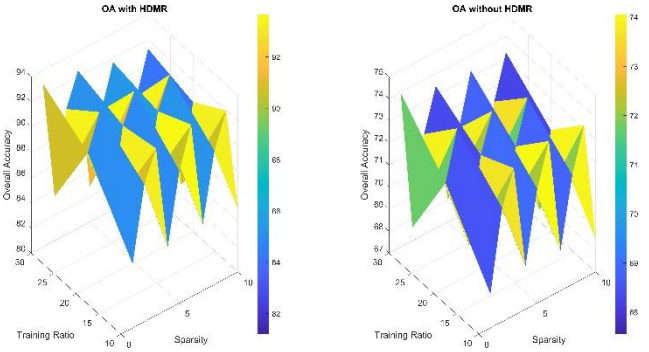

Figure 10. Impact of the training ratio and sparsity level on overall accuracy for Indian Pines dataset. With HDMR (left) and without HDMR (right)

Figures 9, 10 and 11 illustrate the variation in OA across different training ratios and sparsity levels for sparse representation-based classification of HS images, both with and without HDMR. The results reveal that incorporating HDMR generally enhances classification accuracy by improving feature selection, leading to better signal representation and reducing redundant information. The fluctuations observed in both cases can be attributed to the interaction between sparsity and training ratio. At lower training ratios, the available training samples are insufficient to construct a robust dictionary, leading to increased classification instability. As sparsity increases, the classifier may fail to properly represent test samples due to excessive regularization, causing a decline in accuracy.

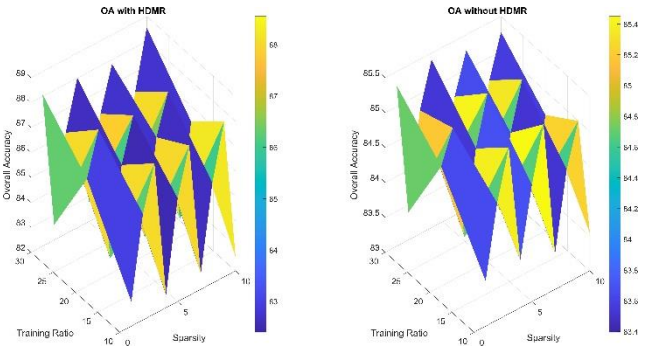

Figure 11. Impact of the training ratio and sparsity level on overall accuracy for Pavia Uni. dataset. With HDMR (left) and without HDMR (right)

Conversely, very low sparsity may result in overfitting, particularly when the training ratio is small. The improvement with HDMR is particularly noticeable in higher training ratios, where the feature extraction process becomes more effective in reducing the curse of dimensionality while preserving discriminative information. However, fluctuations persist, suggesting that an optimal trade-off between training ratio and sparsity is necessary to ensure stable and high-performance classification.

5. Conclusion

This study introduces an efficient dictionary refinement approach tailored for sparse coding-based classification in HS imagery. To achieve this, a data decomposition technique known as High Dimensional Model Representation is applied to the given HS dataset. HDMR serves as an effective tool for denoising and decorrelating HS data. Consequently, instead of directly selecting normalized spectral signals as dictionary atoms, the proposed method constructs the dictionary using refined (i.e., denoised and decorrelated) spectral signals.

Experimental evaluations conducted on HS datasets with varying characteristics demonstrate that integrating HDMR into the dictionary construction process leads to a significant improvement in classification accuracy.

Another key finding of this study relates to the selection of sparsity levels. The results indicate that the proposed method performs optimally when a moderate number of dictionary atoms is selected. Specifically, choosing too few atoms results in underfitting, whereas incorporating an excessively large number of atoms leads to overfitting, both of which negatively impact classification performance. These insights underscore the importance of adaptive sparsity tuning and robust dictionary construction in achieving reliable and accurate HS image classification.

Future work will focus on several key directions to further advance this research. First, the development of adaptive mechanisms to automatically determine the optimal sparsity level and HDMR truncation order will be explored. Second, the integration of the proposed HDMR-based refinement framework with advanced dictionary learning algorithms, such as K-SVD, will be investigated to replace the current reliance on random sampling and enhance dictionary quality. Finally, the generalizability of this framework will be assessed by applying it to alternative classification models and diverse data types, thereby broadening its applicability across different domains.

Declaration of Ethical Standards

The author of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

Conflict of Interest

The author declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Bhargava, A., Sachdeva, A., Sharma, K., Alsharif, M. H., Uthansakul, P., & Uthansakul, M. 2024. Hyperspectral imaging and its applications: A review. Heliyon, **10**(12).

[2] Sun, D. W., Pu, H., & Yu, J. 2024. Applications of hyperspectral imaging technology in the food industry. Nature Reviews Electrical Engineering, **1**(4), pp. 251-263.

[3] Polak, A., Kelman, T., Murray, P., Marshall, S., Stothard, D. J., Eastaugh, N., & Eastaugh, F. 2017. Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication. Journal of Cultural Heritage, **26**, pp. 1-11.

[4] Wilczyński, S., Koprowski, R., Marmion, M., Duda, P., & Błońska-Fajfrowska, B. 2016. The use of hyperspectral imaging in the VNIR (400–1000 nm) and SWIR range (1000–2500 nm) for detecting counterfeit drugs with identical API composition. Talanta, **160**, pp. 1-8.

[5] Karim, S., Qadir, A., Farooq, U., Shakir, M., & Laghari, A. A. 2023. Hyperspectral imaging: a review and trends towards medical imaging. Current Medical Imaging Reviews, **19**(5), pp. 417-427.

[6] Qian, S. E. 2022. Overview of hyperspectral imaging remote sensing from satellites. Advances in Hyperspectral Image Processing Techniques, pp. 41-66.

[7] Rasti, B., Hong, D., Hang, R., Ghamisi, P., Kang, X., Chanussot, J., & Benediktsson, J. A. 2020. Feature extraction for hyperspectral imagery: The evolution from shallow to deep: Overview and toolbox. IEEE Geoscience and Remote Sensing Magazine, **8**(4), pp. 60-88.

[8] Tarabalka, Y., Fauvel, M., Chanussot, J., & Benediktsson, J. A. 2010. SVM-and MRF-based method for accurate classification of hyperspectral images. IEEE Geoscience and Remote Sensing Letters, **7**(4), pp. 736-740.

[9] Huang, K., Li, S., Kang, X., & Fang, L. 2016. Spectral–spatial hyperspectral image classification based on KNN. Sensing and Imaging, **17**, pp. 1-13.

[10] Zhang, Y., Cao, G., Li, X., & Wang, B. 2018. Cascaded random forest for hyperspectral image classification. IEEE journal of selected topics in applied earth observations and remote sensing, **11**(4), pp. 1082-1094.

[11] Bera, S., Shrivastava, V. K., & Satapathy, S. C. 2022. Advances in Hyperspectral Image Classification Based on Convolutional Neural Networks: A Review. CMES-Computer Modeling in Engineering & Sciences, **133**(2).

[12] Khan, A., Vibhute, A. D., Mali, S., & Patil, C. H. 2022. A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications. Ecological Informatics, **69**, 101678.

[13] Jaiswal, G., Sharma, A., & Yadav, S. K. 2021. Critical insights into modern hyperspectral image applications through deep learning. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, **11**(6), e1426.

[14] Peng, J., Sun, W., Li, H. C., Li, W., Meng, X., Ge, C., & Du, Q. 2021. Low-rank and sparse representation for hyperspectral image processing: A review. IEEE Geoscience and Remote Sensing Magazine, **10**(1), pp. 10-43.

[15] Bodrito, T., Zouaoui, A., Chanussot, J., & Mairal, J. 2021. A trainable spectral-spatial sparse coding model for hyperspectral image restoration. Advances in Neural Information Processing Systems, **34**, pp. 5430-5442.

[16] Wang, M., Hong, D., Han, Z., Li, J., Yao, J., Gao, L., Zhang, B. & Chanussot, J. 2023. Tensor decompositions for hyperspectral data processing in remote sensing: A comprehensive review. IEEE Geoscience and Remote Sensing Magazine, **11**(1), pp. 26-72.

[17] Yang, L., Zhou, J., Jing, J., Wei, L., Li, Y., He, X., ... & Nie, B. 2022. Compression of hyperspectral images based on Tucker decomposition and CP decomposition. Journal of the Optical Society of America A, **39**(10), pp. 1815-1822.

[18] Tuna, S., Korkmaz Özay, E., Tunga, B., Gürvit, E., & Tunga, M. A. 2022. An efficient feature extraction approach for hyperspectral images using Wavelet High Dimensional Model Representation. International Journal of Remote Sensing, **43**(19-24), pp. 6899-6920.

[19] Şen, M. E., & Tuna, S. 2025. A new feature extraction scheme based on support optimization in Enhanced Multivariance Products Representation for Hyperspectral Imagery. Journal of the Franklin Institute, **362**(2), 107464.

[20] Taşkın, G., Kaya, H., & Bruzzone, L. 2017. Feature selection based on high dimensional model representation for hyperspectral images. IEEE Transactions on Image Processing, **26**(6), pp. 2918-2928.

[21] Tuna, S., Töreyin, B. U., Demiralp, M., Ren, J., Zhao, H., & Marshall, S. 2020. Iterative enhanced multivariance products representation for effective compression of hyperspectral images. IEEE Transactions on Geoscience and Remote Sensing, **59**(11), pp. 9569-9584.

[22] Xue, Zhaohui, et al. 2017. Discriminative sparse representation for hyperspectral image classification: A semi-supervised perspective. Remote Sensing **9**(4), pp. 386.

[23] Huang, S., Zhang, H., & Pižurica, A. 2017. A robust sparse representation model for hyperspectral image classification. Sensors, **17**(9), 2087.

[24] Tao, W., Liu, N., Chen, Y., Su, J., Xiao, H., & Li, X. 2023. Research on Denoising Methods for Hyperspectral Images Based on Low-Rank Theory and Sparse Representation. IEEE International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD), pp. 1-5

[25] Zhuang, L., Gao, L., Zhang, B., Fu, X., & Bioucas-Dias, J. M. 2020. Hyperspectral image denoising and anomaly detection based on low-rank and sparse representations. IEEE Transactions on Geoscience and Remote Sensing, 60, pp.1-17.

[26] Ülkü, İ., & Töreyin, B. U. 2015. Sparse representations for online-learning-based hyperspectral image compression. Applied optics, **54**(29), pp. 8625-8631.

[27] Zhang, Z., Xu, Y., Yang, J., Li, X., & Zhang, D. 2015. A survey of sparse representation: algorithms and applications. IEEE Access, **3**, pp. 490-530.

[28] Li, C., Ma, Y., Mei, X., Liu, C., & Ma, J. 2016. Hyperspectral image classification with robust sparse representation. IEEE Geoscience and Remote Sensing Letters, **13**(5), pp. 641-645.

[29] Rabitz, H., & Aliş, Ö. F. 1999. General foundations of high-dimensional model representations. Journal of Mathematical Chemistry, **25**(2), pp. 197-233.

[30] Tuna, S., Tunga, B., Baykara, N. A., & Demiralp, M. 2009. Fluctuation free matrix representation based univariate integration in hybrid high dimensional model representation (HHDMR) over plain and factorized HDMR. WSEAS Transactions on Mathematics, **8**(5), pp. 225-230.

[31] Arar, M. E., & Sedef, H. 2023. An efficient lung sound classification technique based on MFCC and

HDMR. Signal, Image and Video Processing, **17**(8), pp. 4385-4394.

[32] Wang, J., Kwon, S., & Shim, B. 2012. Generalized orthogonal matching pursuit. IEEE Transactions on signal processing, **60**(12), pp. 6202-6216.

[33] Dai, W., & Milenkovic, O. 2009. Subspace pursuit for compressive sensing signal reconstruction. IEEE transactions on Information Theory, **55**(5), pp. 2230-2249.

[34] Chen, Y., Nasrabadi, N. M., & Tran, T. D. 2011. Hyperspectral image classification using dictionary-based sparse representation. IEEE transactions on geoscience and remote sensing, **49**(10), pp. 3973-3985.

[35] Davenport, M. A., & Wakin, M. B. 2010. Analysis of orthogonal matching pursuit using the restricted isometry property. IEEE Transactions on Information Theory, **56**(9), pp. 4395-4401.