

Peer Reviewed Review openaccess

An outlook: machine learning in hyperspectral image classification and dimensionality reduction techniques

Tatireddy Subba Reddy^a and Jonnadula Harikiran^{b,*}^aResearch Scholar, School of Computer Science and Engineering, VIT-AP, University, Amaravathi, Andhra Pradesh-522237, India^bAssociate Professor, School of Computer Science and Engineering, VIT-AP University Amaravathi, Andhra Pradesh-522237, India**Contact**Tatireddy Subba Reddy: tatireddysubba12@gmail.comJonnadula Harikiran: jonnadulaharikiran@gmail.com <https://orcid.org/0000-0002-6534-5400>

Hyperspectral imaging is used in a wide range of applications. When used in remote sensing, satellites and aircraft are employed to collect the images, which are used in agriculture, environmental monitoring, urban planning and defence. The exact classification of ground features in the images is a significant research issue and is currently receiving greater attention. Moreover, these images have a large spectral dimensionality, which adds computational complexity and affects classification precision. To handle these issues, dimensionality reduction is an essential step that improves the performance of classifiers. In the classification process, several strategies have produced good classification results. Of these, machine learning techniques are the most powerful approaches. As a result, this paper reviews three different types of hyperspectral image machine learning classification methods: cluster analysis, supervised and semi-supervised classification. Moreover, this paper shows the effectiveness of all these techniques for hyperspectral image classification and dimensionality reduction. Furthermore, this review will assist as a reference for future research to improve the classification and dimensionality reduction approaches.

Keywords: remote sensing, hyperspectral images, classification, machine learning, dimensionality reduction

Introduction

Remote sensing is the interpretation and collection of data about an object, event or place without coming into direct contact with it. For remote sensing of Earth and its natural resources, the most commonly used platforms are satellites and aircraft.¹ In the visible region of the electromagnetic spectrum, aerial photography was the first form of remote sensing. However, the capture of

data in other wavelengths such as microwave, infrared and thermal infrared are now possible due to technology advancements.

Hyperspectral remote sensing involves the collection of information over a large number of wavelength bands.² The analysis and interpretation of hyperspectral images are concerned with spectra captured from a specific scene

CorrespondenceJonnadula Harikiran (jonnadulaharikiran@gmail.com)**Received:** 27 July 2021**Revised:** 26 October 2021**Accepted:** 17 November 2021**Publication:** 7 January 2022**doi:** 10.1255/jsi.2022.a1**ISSN:** 2040-4565**Citation**T.S. Reddy and J. Harikiran, "An outlook: machine learning in hyperspectral image classification and dimensionality reduction techniques", *J. Spectral Imaging* **11**, a1 (2022). <https://doi.org/10.1255/jsi.2022.a1>

© 2022 The Authors

This licence permits you to use, share, copy and redistribute the paper in any medium or any format provided that a full citation to the original paper in this journal is given and the use is not for commercial purposes.



by a satellite or airborne sensor at a long, medium or short distance. Hyperspectral remote sensing images with high spectral and spatial resolution have gained popularity in recent years.^{3,4} These images have a wide range of applications in the medical, mining, military and environmental areas due to their high resolving power. The acquisition of hyperspectral images is dependent on imaging spectrometers, which were developed in the 1980s. Hyperspectral Imaging (HSI) is utilised to capture electromagnetic radiation in the mid-infrared, near infrared, visible and ultra-violet ranges.⁵ In the chosen wavelength, each pixel can acquire a full spectrum in emission or reflection because the imaging spectrometer can record multiple very thin and continuous bands. As a result, hyperspectral images have a large number of bands, high spectral resolution and contain a lot of information.

Classification, dimensionality reduction, transformation, noise reduction and image correction are some of the more common approaches for processing hyperspectral remote sensing images. Unlike normal photos, hyperspectral images are rich in spectral information, which can reflect the chemical composition and physical structure of the item. This information is used to make the classification process easier.⁶⁻⁸ In the research field of HSI, classification is the most active part.

However, due to the similarity of the mixed and spectral pixels and the high-dimensional nature of hyperspectral data, hyperspectral image classification faces several challenges. Some of the most significant issues are listed below.

- Missing labelled samples: collecting hyperspectral image data is very easy in practical applications, but obtaining image-like label information is exceedingly challenging. As a result, the categorisation of hyperspectral images is sometimes complicated by a shortage of labelled samples.
- Hyperspectral image data has high dimensionality: hyperspectral images are created by combining hundreds of bands of spectral reflectance values gathered by spaceborne or airborne imaging spectrometers.
- Image quality: the interference of noise and background variables during the acquisition of hyperspectral images has a significant impact on the quality of the data. The classification accuracy of hyperspectral images is directly influenced by the image quality.
- The spectral information from HSI is affected by various factors like the surrounding environment, distribution of ground features, composition, sensors and atmospheric conditions. Moreover, the spatial

dimension of hyperspectral images is changed by spectral information. As a result, each pixel's ground feature is not a single feature.

Earlier in the study of the classification of hyperspectral images, people frequently concentrated on spectral information-based classification and created several classification algorithms such as polynomial logistic regression, neural networks, Random Forest (RF) and Support Vector Machine (SVM).^{9,10} Moreover, feature selection and extraction, which are dimension reduction techniques, were also proposed. Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA) and Independent Component Analysis (ICA) are examples of dimensionality reduction techniques. This article attempts to provide an overview of the frequently used machine learning-based HSI classification techniques. Also, it provides a review of dimensionality reduction approaches. Various extensive evaluations published in the previous five years have explained these strategies.

The main contribution of this work is as follows:

- We provide the most comprehensive and the most up-to-date overview of machine learning methods in HSI classification.
- We provide a complete review of dimensionality reduction techniques.
- We analysed the performance of the classification and dimensionality reduction approaches based on their accuracy.

Background of HSI classification

In the classification process, individual items (patterns/images, regions/pixels, objects) are classified based on their closeness to the group's description. The classification of image pixels inside a hyperspectral image into numerous categories is known as HSI classification.¹¹⁻¹³

The main aim of image classification is to determine the spectrum for each pixel in a hyperspectral image to identify materials, detect processes or locate objects. The technique of classifying pixels into different classes based on pixel values is known as image classification. The identification of distinct pixels in the hyperspectral image sequence is required for HSI classification. The spectral information is contained in this image series. Each spectral sequence of HSIs has a distinct electromagnetic spectrum, which allows each image to be recognised from other hyperspectral images. The classification of hyperspectral images is not an easy process because there are many things to consider.

- There are many spectral bands, but there are not a lot of training samples. This situation is called the curse of dimensionality or Hughes's Phenomenon.
- Differentiating a particular land cover such as wetlands, urban land-use, agriculture, forest and other vegetation mappings is difficult due to the large number of land cover classes.
- The data classes have a non-linear distribution.

To solve these issues, the right classifier should be chosen. Multi-class data, non-linear data and large amounts of data must be supported by the classifier. To get high accuracy, significant characteristics from the collected features must be picked and given to the classifiers; this procedure is called feature selection.¹⁴ Spectral and spatial classification are the two major categories of hyperspectral classification. In spectral classification, the pixels' reflectance values at different wavelengths are considered. Required spectral properties, such as variance, standard deviation, lowest reflectance, maximum reflectance and mean reflectance can be calculated and used for classification using these reflectances and wavelengths. The spatial arrangement of pixels and their contextual values, as well as textural characteristics and qualities, are detected in spatial classification. HSI classification is important in a variety of applications, including forest applications, urban planning, environmental damage assessment, agriculture monitoring, land-use/land-cover mapping and growth regulation etc. Determining an appropriate classification system, selecting training samples, image pre-processing and feature extraction, selecting adequate classification methodologies, post-classification processing and accuracy assessment are some of the important phases in image classification.

Images can be classified as multiple classifiers, contextual, knowledge-based, per-field, sub-pixel and per-pixel based on pixel information. Classification of images can be categorised as supervised, semi-supervised and cluster analysis based on the use of training samples. The identification of natural structures or groups is known as cluster analysis or unsupervised classification. The process of classifying (i.e. assigning unclassified pixels to one of the multiple informative classes) using samples of known identity is known as supervised classification. During the training process, semi-supervised methods blend a small number of labelled data with a large number of unlabelled data. Several authors have proposed various methodologies for the classification of hyperspectral images, which are briefly outlined in the following sections.

Supervised machine learning methods for classification

The model is trained via supervised learning using labelled training data, which is made up of a set of inputs and their matching outputs of class labels. The model iteratively updates its parameters throughout the training phase to accurately predict the intended outputs. The model is tested against the new input/test data in the testing step to ensure that it can correctly predict the labels. To ensure that the model can correctly predict the label, it is tested against the new test/input data in the testing stage. The model can predict the labels of new input data if it has been appropriately trained. Feature extraction, training and labelling processes are followed in the supervised technique.

Convolutional neural network

Paoletti *et al.*¹⁵ showed that using Deep Learning (DL) to train a conventional Artificial Neural Network (ANN) considerably improved the ANN's HSI classification efficiency. A new Bayesian approach based on integrated Deep Convolutional Neural Network (D-CNN) was implemented by Haut *et al.*¹⁶ who found that this hybridisation improved the effectiveness of the classifier. In addition, the deep recurrent neural network (DRNN) has been demonstrated to be more effective for HSI classification than the traditional ANN classifier, which is based on a single guided filter.¹⁷ A Grey Level Co-occurrence Matrix (GLCM) textural features and integrated CNN have recently been proposed by Zhao *et al.*¹⁸ with a limited training sample. For HSI classification, other ANN DL techniques, like deep belief network, have recently been employed with considerable benefits over traditional approaches (e.g., Reference 19).

Fast computing, optimisation, regularisation, loss function and activation function are all examples of recent breakthroughs in CNNs. Gu *et al.*,²⁰ who detailed these developments, also pointed out CNN's flaws, stating that the selection of appropriate hyper-parameters (e.g., kernel sizes of convolutional filters, learning rate) and computational efficiency are still difficult issues, particularly for large data sets. The deployment of a new CNN for HSI classification was premised on the computational restrictions of CNN algorithms to high-dimensional data stored in multidimensional data cubes. DL models are employed in a variety of fields, including medical diagnostics and detection of image-based cancer, which is a well-known CNN topic. More progress necessitates large and multi-data cubes.²¹ DL approaches try to acquire

discriminative and representative features from data in a hierarchical manner. For HSI classification, due to their ability to extract deep features from hyperspectral images, DL-based algorithms have shown promise for HSI classification. DL is the favoured tool for addressing the problem for classification currently, which can also be used discriminatively.²²

Support vector machine

Compared with other classification approaches like logistic regression-based techniques, neural networks and RFs, SVM is a strong classifier. Since its introduction, it has proved to be extremely effective in predicting urban land-use change, remote sensing image categorisation and tidal analysis.

To perform HSI classification, the combined technique of deep neural network with SVM has been performed by Okwuashi and Ndehedehe.²³ Wang *et al.*²⁴ used a combination strategy of genetic algorithm and modified ant colony algorithm to optimise the parameters of SVM. In the same way, Sukawattanavijit *et al.*²⁵ employed a genetic algorithm to optimise parameters and select features.

Because of the low data needs and excellent precision of classification, SVM is a widely applied technique in health assessment.^{26,27} It is especially effective in binary classification situations.²⁸ Zhang *et al.*,²⁹ for layer Alzheimer's disease (AD), employ coefficients of discrete wavelet-like test provision vector machine as well as functions of neural network classifications. Schizophrenia (SCZ) was detected with related learning algorithms for data analysis in regression and classification. This algorithm clearly shows promising results when used with fMRI in creating an information and set of features from multiple brain regions, providing for the classification of SCZ patients and healthy controls, with a large potential for translational influence.³⁰

SVM classifiers have the potential to make a significant contribution to medicine by allowing for accurate and speedy medical diagnostics. Because the detection of illness is the most important duty in health care, multiple lives could be saved if a disease is diagnosed early.³¹ Asuntha *et al.* were able to recognise cancer cells by assessing the features of the cancer cells.³¹ Vadali *et al.*³² classified cancer as malignant or benign. The invention was put to the test on real CT liver imaging. An SVM was created and evaluated to identify the tumour using both known and unknown data. To use a multi-class SVM classifier for the detection and prediction of lung cancer, Alam *et al.*³³ proposed an optimal method. In addition,

Qiao *et al.*³⁴ offer an SVM classifier based on a genetic approach for cataract diagnosis and feature weighting. For the automated detection of both coronary artery and regular disease issues, on the other hand, Dolatabadi *et al.*³⁵ proposed a method for the automatic identification of normal and Coronary Artery Disease states using the binary classifier (SVM).

Recurrent neural network (RNN)

In multi-temporal satellite pictures, land-cover categorisation was performed by Ienco *et al.*³⁶ using Long Short-Term Memory (LSTM) and Recurrent Neural Network (RNN). Landsat 8 images were classified by the RNN framework, which is based on a patch that was proposed by Sharma *et al.*³⁷ incorporating both spectral and spatial data in a local window. Moreover, single image-based RNN techniques are also implemented in HSI. Mou *et al.*³⁸ introduced an HSI classification technique based on the parametric rectified hyperbolic tangent function. In this approach, for the RNN input layer, each pixel in the image is assumed to be one sequential feature. Wu and Prasad³⁹ utilised the Convolutional RNN (CRNN) technique and examined the integration of RNN and CNN layers in the spectral feature domain. Moreover, in spectral bands, the patch-level local invariant data are extracted by CNN. It offers RNN layers with spatial-contextual characteristics. Instead of using the spectral vector as one piece of sequential data from all bands, Shi and Pun⁴⁰ suggested the spatial neighbours create the sequential data in the RNN model. In this approach, initially, the 3DCNN on the local image patch was employed to obtain the local spatial-spectral characteristics. Afterward, based on the eight-directional construction, the sequences were created. Even though RNN-based DL models have made a substantial contribution to HSI processing efforts, there are still certain crucial issues that must be solved. If only a small training set is available, the result is likely to be overfitted. It is generally difficult and time-consuming to get enough labelled training data for HSI classification.

Semi-supervised approaches for classification

In traditional classification systems, a classifier is generated with ground-truth data (labelled) at the learning/training stage. As is well known, there are limited labelled samples available in remote sensing applications, because obtaining labelled data from sites being observed is

costly and time consuming. Furthermore, the quantity of labelled samples has a strong influence on the performance of the classifier model. To overcome these limited training samples, Semi-Supervised Learning (SSL) is a new strategy that combines both unlabelled and labelled data at the same time. In SSL the unlabelled data is used for spectral signature perception.

SSL approaches are designed to operate with small numbers of labelled samples since labelled data in hyperspectral images are often not available or only in small quantities. Therefore, SSL is gaining researchers' interest.⁴¹ Self-training,⁴² Active Learning (AL),⁴³ graph-based learning⁴⁴ and transductive learning⁴⁵ are some of the semi-supervised learning methodologies which are used to increase learning performance. The primary goal of all investigations is to increase the training sample size. For the volumetric input of the image cube, He *et al.*⁴⁶ use a 3D bilateral filter with Generative Adversarial Networks (GANs). To extract the spatial and spectral features, the 3D bilateral filter was used by the authors, and then the extracted features were fed into a GAN. To produce a realistic sample, both unlabelled and labelled samples were used to train the GAN. To train the classifier, these realistic samples were used. During testing, it was discovered that while the 3D bilateral filter had an effect, it was not significant. Currently, the GAN classifier has received a lot of attention, but it is a complicated structure with a lot of network parameters that can be difficult to configure. To train spectral and spatial feature extractors based on stacked multi-loss auto-encoders, unlabelled datasets were used by Kemker and Kanan.⁴⁷ Furthermore, labelled data are also used to train the classifier. For classification, Kemker *et al.*⁴⁸ enhanced the approach by substituting the SVM classifier with a Semi-Supervised MultiLayer Perceptron (SS-MLP). Even though the datasets differ in spatial resolution, the number of wavelengths captured and spectral bands they found that using datasets from different sensors increased performance.

Recently, a few works that examined the combination of AL and DL for HSI classification have been published.⁴⁹⁻⁵¹ Particularly, Ahamad *et al.*⁴⁹ proposed an AL-based technique named Generalised Fuzziness Extreme Learning Machine Auto-Encoder (GFELM-AE). An approach to integrate a weighted incremental dictionary learning criterion with RBM was presented by Liu *et al.*⁵⁰ Haut *et al.*⁵¹ proposed a method to combine BCNN with six AL criteria (mutual information, breaking ties, maximum EP, random acquisition etc.). Moreover, the combination of these methods attained superior performance.

To create a semi-supervised HSI classification approach, Cui *et al.*⁵² analysed Rolling Guidance Filter (RGF) and Extended Label Propagation (ELP) for graph-based learning. To forecast pixel labels, this approach used a graph-based label propagation algorithm. Moreover, it used superpixels to correct some mislabelled pixels.

The transductive methodology necessitates fewer resources because it is primarily concerned with lowering inference error for a given set of unlabelled data, rather than attempting to enhance the overall quality of the acquired hypothesis. Hence, Huang *et al.*⁴⁵ proposed a strategy named transductive learning to overcome the limited labelling training sample and high dimensionality difficulties by extending non-local graph theory to the classification label space. Furthermore, Appice *et al.*⁵³ used the S2CoTraC algorithm to implement a transductive learning technique based on a co-training schema.

Cluster analysis for classification

Unlike supervised techniques, there has been limited study of HSI cluster analysis (i.e., clustering), which still has serious issues because of the high-dimensional and complex data observation. For cluster analysis HSI classification, a robust manifold matrix factorisation was presented by Zhang *et al.*⁵⁴ They used unified low-rank matrix factorisation to address the high feature dimensionality of the hyperspectral image. It performs data clustering and dimensionality reduction. In matrix factorisation, the reconstruction loss is measured using the 2,1-norm. It helps to decrease the mistakes caused by potential noisy observations. Moreover, to find the local optimum solution, the Augmented Lagrangian Method (ALM) was implemented as well as a new out-of-sample extension trick to enable the technique to handle large-scale hyperspectral remote sensing data. To classify hyperspectral images, the divide and conquer technique was presented. Also, they used the k-means clustering approach to separate the data produced after PCA into multiple clusters. Afterward, multi-class SVM is used to train these clusters separately. A Bayesian labelling method was suggested by Ghanbari *et al.*⁵⁵ that combined the findings of the Gaussian Mixture Model (GMM) with spatial-contextual data collected using Markov Random Fields (MRF). In addition, to increase MRF performance, a fuzzy segmentation-based technique is designed and implemented into the spatial energy.

Bo *et al.*⁵⁶ extend the classic KNN approach in a spectral–spatial collaborative manner. The set-to-point distance and weighted KNN framework are effectively integrated into this approach. According to Wang *et al.*,⁵⁷ cluster analysis algorithms are generally classified into four groups:

- 1) Centroid-based clustering approaches like fuzzy c-means and k-mean are used to decrease the distance within the cluster sample. These techniques are very sensitive to initialisation and noise, so they cannot produce a reliable result.
- 2) Density-based approaches using the fast search technique to create the clustering and obtain the density peak algorithm. Density-based methods are not suitable for HSI because finding the density peak in the sparse feature space is difficult.
- 3) Biological clustering techniques: for clustering HSI classification, these include an adaptive multi-objective differential evolution-based automatic fuzzy clustering approach and artificial immune networks. The biological techniques may not always completely match the properties of hyperspectral images. Therefore, their outcomes are not always good enough.
- 4) Graph-based approaches, like spectral clustering, perform well in HSI cluster analysis, but the affinity matrix and eigenvalue decomposition take too long for most of them. Because of the uniform data distribution induced by the substantial spectrum diversity, the existing HSI cluster analysis algorithm's accuracy is not satisfactory when compared to supervised procedures. Therefore, Xie *et al.*⁵⁸ and Zhao *et al.*⁵⁹ proposed techniques called adaptive density and fast spectral clustering for unsupervised hyperspectral remote sensing image clustering.

Moreover, Lin *et al.*⁶⁰ suggested an unsupervised model termed Multiple-Layer Feature-Matching Generative Adversarial Networks (MARTA GANs) to learn from unlabelled input. It has a discriminative model D as well as a generative model G. They considered D to be a feature extractor and employed a fusion layer to combine global and mid-level features to match the complicated aspects of remote sensing data. A Wasserstein GAN with Gradient Penalty (WGAN-GP) based unsupervised approach was presented by Wei *et al.*⁶¹ To produce more expanded high-resolution remote sensing images, a multi-feature layer is introduced by them. To cluster hybrid polarimetric data into seven clusters, unsupervised CNN is used by Chatterjee *et al.*⁶² They used an entropy-based loss function and an adaptive learning rate optimisation algorithm (Adam) for training.

Dimensionality reduction techniques

To decrease the dimensionality of the feature space, dimensionality reduction algorithms have been suggested. They aim to reduce computational complexity and alleviate statistical complexity by discarding redundant features, which can potentially increase the performance of the classifier. Hence, the purpose of dimensionality reduction is to minimise the complexity of incoming data while preserving part of the data's essential information. Traditional band selection methods or a feature extraction (projection) techniques can be used to reduce dimensionality in the HSI framework. Band selection methods choose the original spectral band's subset that should contain the image's most essential features. Band selection approaches can be classified as unsupervised and supervised depending on the availability of labelled samples. Unsupervised band selection is employed when no labelled samples are available. The goal of unsupervised band selection approaches is to select the most informative and distinct bands possible. Projection methods use the original band's non-linear or linear combinations to convert the data into a low-dimensional environment.

The most prominent dimensionality reduction approach is PCA. To attain the covariance matrix, it subtracts the population means from each sample. From these, it derives a transformation matrix by maximising sample scatter. To prevent overfitting and reduce noise, PCA is used in the pre-processing of other advanced dimensionality reduction algorithms.⁶³ LDA⁶⁴ and discriminant common vectors⁶⁵ are advanced versions of PCA. When samples are spread in a manifold structure, the PCA technique is unsuccessful in exposing the local structure of data because it is dependent on linear measurement.

For HSI classification, Ramamurthy *et al.*⁶⁶ employed a CNN-based classifier with an auto-encoder-based dimensionality reduction technique. Singular Value Decomposition (SVD) followed by the combination of QR decomposition and Inter-Band Block Correlation Coefficient (IBBC) was presented by Reshma *et al.*⁶⁷ to improve the spatial features. For dimensionality reduction, Li *et al.*⁶⁸ developed another spatial–spectral-based neighbour graph. For classification, the author used a Composite Kernel-based Support Vector Machine (SVM-CK). Charmisha *et al.*⁶⁹ and Hang and Liu⁷⁰ suggested dynamic mode decomposition and a Local Graph Discriminant Embedding (LGDE) based approach for dimensionality reduction. They extend the linear LGDE model into its kernel counterpart to

adequately reflect the non-linear feature of a hyperspectral image. Paul and Chaki⁷¹ proposed a method based on pooling. For efficient classification, Hidalgo *et al.*⁷² present the Self-Organising Map (SOM) which is an unsupervised dimensionality reduction technique. Within the group of input patterns, it identifies self-organising correlations automatically.

For the supervised technique, Gowtham⁷³ suggested a PCA-based approach. Sawant and Manoharan⁷⁴ and Prabukumar *et al.*⁷⁵ recommended a discrete cosine transform and weighted entropy-based algorithm. For hyperspectral images, an Expectation-Maximisation (EM) algorithm-based clustering with band fusion is offered by Prabukumar and Shrutika,⁷⁶ and a Band Correlation Clustering (BCC)-based feature extraction methodology is described by Ghorbanian and Mohammadzadeh.⁷⁷

By mapping high-dimensional data into lower-dimensional data, graph learning approaches have gained increasing interest from researchers as a useful tool for dimensionality reduction.⁷⁸ Many graph learning algorithms and variants have been presented to discover the fundamental geometric structure of high-dimensional data based on this concept.⁷⁹ For example, Spatial-Spectral Multiple Manifold Discriminant Analysis (SSMMDA)⁸⁰ and Isometric Feature Mapping (ISOMAP).⁸¹ For delivering extremely non-linear manifolds, ISOMAP strives to preserve geodesic distances of all similarity pairs, and it approximates the geodesic distance between two points by measuring the shortest path between these points.⁸² The multi-scale segmentation-based method of fusing classification findings has recently been extended to the field of dimensionality reduction, with promising results.⁸³

Multiple homogenous zones are common on a real surface. The differences between distinct homogenous regions are often overlooked when reducing dimensionality based on the optimal global projection direction. Because these multiple homogeneous zones are often of varying sizes, it might be difficult to accurately represent these differences for a single segmentation scale. To tackle these issues, Jiang *et al.*⁸³ segmented the hyperspectral image into superpixels of several scales. At different scales, the low-dimensional illustrations of HSI are created by applying PCA on each superpixel. By the use of the Majority Voting (MV) decision fusion approach, the final classification output is generated by fusing each scale's classification outcomes. This technique is known as Multi-scale segmentation-based superpixel-wise PCA (MSuperPCA). Moreover, the technique that depends on single-scale segmentation is known as Superpixel-wise PCA (SuperPCA). Based on this, Zhang *et al.*⁸⁴ developed

a SuperKPCA approach for dimensionality reduction and achieved better results.

Wavelet transforms for image and spectra compression

A modified dictionary learning process, a novel ordering system and the use of Differential Pulse Code Modulation (DPCM) have been proposed by Ertem *et al.*⁸⁵ The coefficients of sparse can be expressed more effectively than typical SSASR due to improvements in dictionary learning. To reconstruct the hyperspectral image, the superpixel map is kept or stored for a long time (side information). When utilising the new ordering method, it is no longer necessary to store the superpixel map as side information. Furthermore, using DPCM reduces the sparse coefficients' magnitudes.

Discrete Wavelet Transform (DWT)-based compression algorithms have been shown to provide greater compression performance than DCT-based approaches, despite DCT's low computational complexity.⁸⁶

A lossy HSI compression scheme depending on the regression of wavelet coefficients of 3D was described by Zikiou *et al.*⁸⁷ The hyperspectral images are weakly represented using the 3D wavelet transform. The wavelet details are then subjected to an SVR (SVM regression), which generates weights and vector supports that represent texture features for the wavelet. 3D-bicubic interpolation with factor 2 increases the size of DWT's high-frequency components even further.

Nagendran and Vasuki⁸⁸ proposed employing hybrid transform to compress hyperspectral images. The image is first broken into 1D, then clustered and tiled. In order to obtain Integer Karhunen-Loeve Transform (IKLT) bands in the spectral dimension, every cluster is subjected to the IKLT, which is then applied to the entire image. To decorrelate the spatial data in spatial dimension applying IKLT bands, the integer wavelet transform (IDWT) is then used. Hybrid transform is the result of combining IDWT with IKLT. Then the coefficients of decorrelated wavelet are used to the Spatial-orientation Tree Wavelet (STW), Wavelet Difference Reduction (WDR) and Adaptively Scanned Wavelet Difference Reduction (ASWDR) algorithms.

Hernández-Cabronero *et al.*⁸⁹ aimed to determine the most beneficial trade-off of compression-complexity inside the state-of-the-art in HSI compression. For performance assessment, the authors combined the

most competitive spectrum decorrelation techniques with the low-complexity compressors.

Evaluation of results

Following the classification of hyperspectral images, we must assess the quality of the classification findings before evaluating the classifier's performance. The accuracy of classification is measured to graphically indicate the capability of the suggested approach, which is followed by a relative examination of existing methods and further enhancements depending on their weaknesses. Moreover, the performance of the techniques is evaluated on some hyperspectral datasets. The most commonly used datasets are Salinas, Pavia University, Kennedy Space Center and Indian Pines, which are described below.

The Salinas dataset: the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor was used to collect these data with 3.7 m spatial resolution. The dataset was captured over the area of Salinas Valley, USA. After 20 water absorption bands have been eliminated, the image has 512×217 pixels and 204 bands. This dataset mainly includes wine fields, barren soils and vegetable fields. There are 16 other classes as well.

The Pavia University dataset: the Reflective Optics System Imaging Spectrometer (ROSIS-03) optical sensor was used to collect these data. The dataset contains

103 spectral bands with a 0.43–0.86 μm spectral range and the size of the image is 610×340 pixels. It has 1.3 m spatial resolution.

Indian Pines dataset: this dataset was obtained by AVIRIS. Indian Pines covers 220 spectral bands and is 145×145 pixels in size. It covers the 0.4–2.45 μm wavelength range. Moreover, each pixel has a spatial resolution of 20 m.

Kennedy Space Center dataset: it includes 512×614 -pixel sizes and was obtained from Florida in 1996 by AVIRIS. It has 13 ground-truth classes and 18 mpx^{-1} spatial resolution. There are 176 bands left once the noise bands are removed.

Comparison of the approaches

The comparison of supervised, unsupervised and semi-supervised approaches is given in Table 1. From this, it can be observed that most of the recent hyperspectral image classification algorithms are based on supervised and semi-supervised approaches. Unsupervised machine learning-based classification techniques are utilised rarely. Moreover, compared with unsupervised approaches, the accuracy of semi-supervised and supervised machine learning algorithms is higher. However, the computation time of unsupervised algorithms is much less. Compared with supervised approaches, they provide the results in a very short time.

Table 1. Comparison of supervised, semi-supervised and unsupervised approaches based on accuracy.

Year	Reference	Types of classification	Techniques	Dataset	Accuracy
2018	Paoletti <i>et al.</i> ¹⁵	Supervised	Improved 3-D deep CNN model	Indian Pines, University of Pavia	98.6 %
2018	Haut <i>et al.</i> ¹⁶	Supervised	D-CNN	Indian Pines, Salinas Valley, Kennedy Space Center	99 %
2018	Guo <i>et al.</i> ¹⁷	Supervised	RNN	Indian Pines and Kennedy Space Center	93.80 %
2019	Zhao <i>et al.</i> ¹⁸	Supervised	CNN	University of Pavia, Houston, Zhangye	98.4 %
2019	Li <i>et al.</i> ¹⁹	Supervised	Deep belief network (DBN)	Indian Pines and Pavia University	97.26 %
2018	Li <i>et al.</i> ²¹	Supervised	AdaBoost-weighted composite kernel extreme learning machines (WCKELM)	Indian Pines and Pavia University	96.46 %

2020	Okwuashi and Ndehedehe ²³	Supervised	Deep SVM	Indian Pines and Pavia University	98.86 %
2019	Wang <i>et al.</i> ²⁴	Supervised	SVM-modified binary-coded Ant colony algorithm	Heart, Credit, Forest, Ionosphere and Spectf	91.24 %
2020	Sukawattanavijit <i>et al.</i> ²⁵	Supervised	SVM-genetic algorithm (GA)	Thaichote (THEOS), RADARSAT-2 (RS2), LANDSAT-8, RS2 data	95 %
2017	Ienco <i>et al.</i> ³⁶	Supervised	LSTM	THAU, Reunion Island	75.15 %
2018	Sharma <i>et al.</i> ³⁷	Supervised	Patch-based RNN	Indian Pines, Pavia University, Houston	97.21 %
2017	Mou <i>et al.</i> ³⁸	Supervised	RNN	Indian Pines, Pavia University, Houston	88.63 %
2017	Wu and Prasad ³⁹	Supervised	CRNN	University of Houston, Indian Pines	89 %
2018	Shi and Pun ⁴⁰	Supervised	Multi-scale hierarchical RNN	Pavia University, Pavia Center scene and Salinas	96.32 %
2020	Zhang ⁴¹	Semi-supervised	Locally linear embedding	Indian Pines, Botswana (BOT), Kennedy Space Centre (KSC)	99.65 %
2018	Li <i>et al.</i> ⁴²	Semi-supervised	Self-training with iterative region merging using semantics (ST-IRGS)	University of Pavia, Salinas, Kennedy Space Center	91.7 %
2020	Cao <i>et al.</i> ⁴³	Semi-supervised	Active learning-based CNN	Indian Pines, Pavia University, Pavia Center	97.45 %
2021	He <i>et al.</i> ⁴⁴	Semi-supervised	Constrained label propagation with particle competition and cooperation (CLPPCC)	Indian Pines, Pavia University, Salinas	98.57 %
2021	Huang <i>et al.</i> ⁴⁵	Semi-supervised	Transductive learning method	The University of Houston, Urban, JasperR, PaviaU Indian P and Salinas-A	99.13 %
2017	He <i>et al.</i> ⁴⁶	Semi-supervised	3DBF-GAN	Indian Pines, University of Pavia and Salinas	92.30 %
2017	Kemker and Kanan ⁴⁷	Semi-supervised	MICA and SCAE	Indian Pines, Salinas Valley and Pavia University	99.2 %
2018	Kemeker <i>et al.</i> ⁴⁸	Semi-supervised	SS-MLP	Indian Pines and Pavia University	91.32 %
2020	Ahmad <i>et al.</i> ⁴⁹	Semi-supervised	GFELM-AE	Salinas	99.89 %
2017	Liu <i>et al.</i> ⁵⁰	Semi-supervised	Active deep learning	PaviaC, PaviaU and Botswana	96.25 %
2018	Haut <i>et al.</i> ⁵¹	Semi-supervised	AL-based B-CNN	Indian Pines, Salinas Valley and Kennedy Space Center	99 %

2018	Cui <i>et al.</i> ⁵²	Semi-supervised	ELP-RGF	Kennedy Space Center, University of Pavia and Indian Pines	85.21 %
2017	Appice <i>et al.</i> ⁵³	Semi-supervised	S ² CoTraC	Indian Pines, Pavia University and Salinas Valley	95 %
2019	Zhang <i>et al.</i> ⁵⁴	Cluster analysis	Robust manifold matrix factorisation	Washington DC dataset, Salinas-A	91.86 %
2018	Ghanbari <i>et al.</i> ⁵⁵	Cluster analysis	GMMMRF	Indian Pines, Pavia University and Salinas	94.5 %
2018	Bo <i>et al.</i> ⁵⁶	Cluster analysis	Spectral-Spatial KNN	AVIRIS Indian Pines and University of Pavia	93.8 %
2017	Wang <i>et al.</i> ⁵⁷	Cluster analysis	Fast spectral clustering with anchor graph (FSCAG)	Indian Pines, Salinas, Pavia Center	81.55 %
2018	Xie <i>et al.</i> ⁵⁸	Cluster analysis	K-means fast density peak-based clustering (K-FDPC)	UCI Iris, Jasper Ridge HSI data set and Salinas-A	83.5 %
2019	Zhao <i>et al.</i> ⁵⁹	Cluster analysis	FSC	Salinas and Salinas-A, Indian Pines, Urban, Jasper Ridge, Samson, Kennedy Space Center and Pavia University	The accuracy of the approach was not computed
2017	Lin <i>et al.</i> ⁶⁰	Cluster analysis	MARTA GAN	UC Merced Land Use and Brazilian coffee scenes	94.86 %
2020	Wei <i>et al.</i> ⁶¹	Cluster analysis	WGAN-GP	NWPU-RESISC45, AID and UC Merced Land Use	92.45 %
2020	Chatterjee <i>et al.</i> ⁶²	Cluster analysis	Deep CNN	RISAT-I	86.08 %

From the analysis of supervised and semi-supervised approaches, it is observed that the classification results of semi-supervised approaches are superior to the supervised approaches. From the semi-supervised approaches, the auto-encoder and GAN-based techniques achieve high classification accuracy. Therefore, semi-supervised approaches are better than supervised and unsupervised approaches.

We analysed different dimensionality reduction approaches based on the evaluation metric accuracy and this is given in Table 2.

From Table 2 it can be seen that the minimum noise fraction, auto-encoder, SSNG and PCA techniques provide the best results, helping to provide better classification

accuracy. Moreover, they preserve the relevant information to enhance the semantic interpretation of hyperspectral images.

Conclusion

Hyperspectral image recognition and classification are crucial functions of hyperspectral image processing. This study reviews current advances in hyperspectral image classification utilising various supervised, unsupervised and semi-supervised machine learning algorithms. Moreover, a brief review of various dimensionality reduction techniques is also discussed. In addition,

Table 2. Comparison of dimensionality reduction techniques.

Year	Author	Approach	Dataset	Accuracy
2020	Uddin <i>et al.</i> ⁶³	Comparison of non-linear and linear variants of PCA (kernel entropy component analysis, minimum noise fraction, spectrally segmented PCA, folded-PCA, segmented-PCA and kernel-PCA)	Indian Pines and Washington DC Mall	Minimum noise fraction = 98.93%
2020	Jayaprakash <i>et al.</i> ⁶⁴	Randomised independent component analysis (RFICA) and randomised linear discriminant analysis (RFLDA)	Salinas Scene and Pavia University	RFLDA = 95.39%
2018	Fordellone <i>et al.</i> ⁶⁵	Partial least squares discriminant analysis (PLS-DA)	Real dataset	Accuracy not computed
2020	Ramamurthy <i>et al.</i> ⁶⁶	Autoencoder	Pavia University dataset	98%
2018	Reshma <i>et al.</i> ⁶⁷	Singular value decomposition (SVD)	Salinas-A and Pavia University	97.3%
2019	Li <i>et al.</i> ⁶⁸	Spatial-spectral neighbour graph (SSNG)	Indian Pines, Pavia University and Salinas	98.90%
2018	Charmisha <i>et al.</i> ⁶⁹	Dynamic mode decomposition (DMD)	Salinas-A and Indian Pines	90.56%
2018	Hang and Liu ⁷⁰	LGDE	Pavia University Scene, Indian Pines and Kennedy Space Center	92.06%
2019	Paul and Chaki ⁷¹	Pooling	Kennedy Space Centre and Botswana	95%
2020	Hidalgo <i>et al.</i> ⁷²	SOM	Pavia University, Salinas Valley, Indian Pines	81.9%
2021	Gowtham ⁷³	PCA	Pavia University and Indian Pines	98.2%
2020	Sawant and Manoharan ⁷⁴	3D DCT	Salinas, Pavia University and Indian Pines	92.83%
2018	Prabukumar <i>et al.</i> ⁷⁵	3-D DCT	Salinas, Pavia University and Indian Pines	94.62%
2018	Prabukumar and Shrutika ⁷⁶	EM	Botswana, Salinas, Pavia University and Indian Pines	84.92%
2018	Ghorbanian and Mohammadzadeh ⁷⁷	Band correlation clustering (BCC)	Washington DC, Houston, and Pavia University	88.16%
2017	Xia <i>et al.</i> ⁷⁸	KPCA	Indian Pines AVIRIS, University of Pavia ROSIS	74%

2017	Feng <i>et al.</i> ⁷⁹	Graph-based discriminant analysis with spectral similarity (denoted as GDA-SS)	Salinas, University of Pavia	93 %
2019	Shi <i>et al.</i> ⁸⁰	SSMMDA	PaviaU, Heihe, Washington DC Mall	95.65 %
2019	Orts Gómez <i>et al.</i> ⁸¹	ISOMAP	Indian Pines, Salinas-A, Pavia	95.76 %
2017	Li <i>et al.</i> ⁸²	ISOMAP	Indian Pines, Salinas-A, Pavia	91 %
2018	Jiang <i>et al.</i> ⁸³	SuperPCA	Indian Pines, University of Pavia, Salinas Scene	94.96 %
2019	Zhang <i>et al.</i> ⁸⁴	SuperKPCA	Indian Pines, Pavia University, Salinas	89.90 %

the performance of these approaches (accuracy) is also analysed. From the analysis, it can be seen that the semi-supervised approaches, such as auto-encoder and GAN-based techniques, provide better classification results and non-linear variants of PCA (minimum noise fraction) based dimensionality reduction technique help to improve the classification accuracy.

References

1. A. Diakite, G. Jiangsheng and F. Xiaping, "Hyperspectral image classification using 3D 2D CNN", *IET Image Proc.* **15(5)**, 1083–1092 (2021). <https://doi.org/10.1049/ipr2.12087>
2. U.B. Gewali, S.T. Monteiro and E. Saber, "Machine learning based hyperspectral image analysis: a survey", *arXiv preprint arXiv:1802.08701* (2018). <https://arxiv.org/abs/1802.08701>
3. S.K. Roy, J.M. Haut, M.E. Paoletti, S.R. Dubey and A. Plaza, "Generative adversarial minority over-sampling for spectral-spatial hyperspectral image classification", *IEEE Trans. Geosci. Remote Sens.* **60**, 5500615 (2021). <https://doi.org/10.1109/TGRS.2021.3052048>
4. A. Paul, S. Bhoomik and N. Chaki, "SSNET: an improved deep hybrid network for hyperspectral image classification", *Neural Comput. Appl.* **33(5)**, 1575–1585 (2021). <https://doi.org/10.1007/s00521-020-05069-1>
5. L. Pitak, K. Laloon, S. Wongpichet, P. Sirisomboon and J. Posom, "Machine learning-based prediction of selected parameters of commercial biomass pellets using line scan near infrared-hyperspectral image", *Processes* **9(2)**, 316 (2021). <https://doi.org/10.3390/pr9020316>
6. H. Su, Y. Yu, Q. Du and P. Du, "Ensemble learning for hyperspectral image classification using tangent collaborative representation" *IEEE Trans. Geosci. Remote Sens.* **58(6)**, 3778–3790 (2020). <https://doi.org/10.1109/TGRS.2019.2957135>
7. Z. Meng, E. Merkurjev, A. Koniges and A.L. Bertozzi, "Hyperspectral image classification using graph clustering methods", *Image Process. On Line* **7**, 218–245 (2017). <https://doi.org/10.5201/ipol.2017.204>
8. A. Mohan and M. Venkatesan, "HybridCNN based hyperspectral image classification using multiscale spatio-spectral features", *Infrared Phys. Technol.* **108**, 103326 (2020). <https://doi.org/10.1016/j.infrared.2020.103326>
9. J. Su, D. Yi, C. Liu, L. Guo and W.H. Chen, "Dimension reduction aided hyperspectral image classification with a small-sized training dataset: experimental comparisons", *Sensors* **17(12)**, 2726 (2017). <https://doi.org/10.3390/s17122726>
10. F. Luo, L. Zhang, B. Du and L. Zhang, "Dimensionality reduction with enhanced hybrid-graph discriminant learning for hyperspectral image classification", *IEEE Trans. Geosci. Remote Sens.* **58(8)**, 5336–5353 (2020). <https://doi.org/10.1109/TGRS.2020.2963848>
11. B. Alhayani and H. Ilhan, "Hyperspectral image classification using dimensionality reduction techniques", *Int. J. Innov. Res. Elec. Electron.*

- Instrum. Control Eng.* **5**, 71–74 (2017). <https://doi.org/10.17148/IJREEICE.2017.5414>
12. H.C. Li, W.Y. Wang, L. Pan, W. Li, Q. Du and R. Tao, “Robust capsule network based on maximum correntropy criterion for hyperspectral image classification”, *IEEE J. Select. Top. Appl. Earth Observ. Remote Sens.* **13**, 738–751 (2020). <https://doi.org/10.1109/JSTARS.2020.2968930>
 13. R. Vaddi and P. Manoharan, “Hyperspectral image classification using CNN with spectral and spatial features integration”, *Infrared Phys. Technol.* **107**, 103296 (2020). <https://doi.org/10.1016/j.infrared.2020.103296>
 14. M.E. Paoletti, J.M. Haut, R. Fernandez-Beltran, J. Plaza, A.J. Plaza and F. Pla, “Deep pyramidal residual networks for spectral–spatial hyperspectral image classification”, *IEEE Trans. Geosci. Remote Sens.* **57(2)**, 740–754 (2018). <https://doi.org/10.1109/TGRS.2018.2860125>
 15. M.E. Paoletti, J.M. Haut, J. Plaza and A. Plaza, “A new deep convolutional neural network for fast hyperspectral image classification”, *ISPRS J. Photogram. Remote Sens.* **145**, 120–147 (2018). <https://doi.org/10.1016/j.isprsjprs.2017.11.021>
 16. J.M. Haut, M.E. Paoletti, J. Plaza, J. Li and A. Plaza, “Active learning with convolutional neural networks for hyperspectral image classification using a new Bayesian approach”, *IEEE Trans. Geosci. Remote Sens.* **56(11)**, 6440–6461 (2018). <https://doi.org/10.1109/TGRS.2018.2838665>
 17. Y. Guo, S. Han, H. Cao, Y. Zhang and Q. Wang, “Guided filter based deep recurrent neural networks for hyperspectral image classification”, *Procedia Comput. Sci.* **129**, 219–223 (2018). <https://doi.org/10.1016/j.procs.2018.03.048>
 18. W. Zhao, S. Li, A. Li, B. Zhang and Y. Li, “Hyperspectral images classification with convolutional neural network and textural feature using limited training samples”, *Remote Sens. Lett.* **10(5)**, 449–458 (2019). <https://doi.org/10.1080/2150704X.2019.1569274>
 19. C. Li, Y. Wang, X. Zhang, H. Gao, Y. Yang and J. Wang, “Deep belief network for spectral–spatial classification of hyperspectral remote sensor data”, *Sensors* **19(1)**, 204 (2019). <https://doi.org/10.3390/s19010204>
 20. J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai and T. Chen, “Recent advances in convolutional neural networks”, *Pattern Recogn.* **77**, 354–377 (2018). <https://doi.org/10.1016/j.patcog.2017.10.013>
 21. L. Li, C. Wang, W. Li and J. Chen, “Hyperspectral image classification by AdaBoost weighted composite kernel extreme learning machines”, *Neurocomputing* **275**, 1725–1733 (2018). <https://doi.org/10.1016/j.neucom.2017.09.004>
 22. B. Pan, Z. Shi and X. Xu, “R-VCANet: A new deep-learning-based hyperspectral image classification method”, *IEEE J. Select. Topics Appl. Earth Obs. Remote Sens.* **10(5)**, 1975–1986 (2017). <https://doi.org/10.1109/JSTARS.2017.2655516>
 23. O. Okwuashi and C.E. Ndehedehe, “Deep support vector machine for hyperspectral image classification”, *Pattern Recogn.* **103**, 107298 (2020). <https://doi.org/10.1016/j.patcog.2020.107298>
 24. M. Wang, Y. Wan, Z. Ye and X. Lai, “Remote sensing image classification based on the optimal support vector machine and modified binary coded ant colony optimization algorithm”, *Inform. Sci.* **402**, 50–68 (2017). <https://doi.org/10.1016/j.ins.2017.03.027>
 25. C. Sukawattanavijit, J. Chen and H. Zhang, “GA-SVM algorithm for improving land-cover classification using SAR and optical remote sensing data”, *IEEE Geosci. Remote Sens. Lett.* **14(3)**, 284–288 (2017). <https://doi.org/10.1109/LGRS.2016.2628406>
 26. Y. Cheng, H. Zhu, K. Hu, J. Wu, X. Shao and Y. Wang, “Reliability prediction of machinery with multiple degradation characteristics using double-Wiener process and Monte Carlo algorithm”, *Mech. Syst. Signal Process.* **134**, 106–333 (2019). <https://doi.org/10.1016/j.ymssp.2019.106333>
 27. W. Fu, J. Tan, X. Zhang, T. Chen and K. Wang, “Blind parameter identification of MAR model and mutation hybrid GWO-SCA optimized SVM for fault diagnosis of rotating machinery”, *Complexity* **2019**, 1–17 (2019). <https://doi.org/10.1155/2019/3264969>
 28. Z. Li, R. Outbib, S. Giurgea, D. Hissel, S. Jemei, A. Giraud and S. Rosini, “Online implementation of SVM based fault diagnosis strategy for PEMFC systems”, *Appl. Energy* **164**, 284–293 (2016). <https://doi.org/10.1016/j.apenergy.2015.11.060>
 29. J. Zhang, Y. Gao, Y. Gao, B.C. Munsell and D. Shen, “Detecting anatomical landmarks for fast Alzheimer’s disease diagnosis”, *IEEE Trans. Med. Imaging* **35(12)**, 2524–2533 (2016). <https://doi.org/10.1109/TMI.2016.2582386>

30. E. Veronese, U. Castellani, D. Peruzzo, M. Bellani and P. Brambilla, "Machine learning approaches: from theory to application in schizophrenia", *Comput. Math Meth. Med.* **2013**, 867924 (2013). <https://doi.org/10.1155/2013/867924>
31. A. Asuntha, A. Brindha, S. Indirani and A. Srinivasan, "Lung cancer detection using SVM algorithm and optimization techniques", *J. Chem. Pharm. Sci.* **9(4)**, 3198–3203 (2016). [https://www.jchps.com/issues/Volume%209_Issue%204/jchps%209\(4\)%20286%200450716%203198-3203.pdf](https://www.jchps.com/issues/Volume%209_Issue%204/jchps%209(4)%20286%200450716%203198-3203.pdf)
32. S. Vadali, G.V.S.R. Deekshitulu and J.V.R. Murthy, "Analysis of liver cancer using data mining SVM algorithm in MATLAB", in *Soft Computing for Problem Solving*. Springer, Singapore, pp. 163–175 (2019). https://doi.org/10.1007/978-981-13-1592-3_12
33. J. Alam, S. Alam and A. Hossan, "Multi-stage lung cancer detection and prediction using multi-class SVM classifier", in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*. IEEE, pp. 1–4 (2018). <https://doi.org/10.1109/IC4ME2.2018.8465593>
34. Z. Qiao, Q. Zhang, Y. Dong and J.J. Yang, "Application of SVM based on genetic algorithm in classification of cataract fundus images", in *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*. IEEE, pp. 1–5 (2017). <https://doi.org/10.1109/IST.2017.8261541>
35. A.D. Dolatabadi, S.E.Z. Khadem and B.M. Asl, "Automated diagnosis of coronary artery disease (CAD) patients using optimized SVM", *Comput. Meth. Progr. Biomed.* **138**, 117–126 (2017). <https://doi.org/10.1016/j.cmpb.2016.10.011>
36. D. Ienco, R. Gaetano, C. Dupaquier and P. Maurel, "Land cover classification via multitemporal spatial data by deep recurrent neural networks", *IEEE Geosci. Remote Sens. Lett.* **14(10)**, 1685–1689 (2017). <https://doi.org/10.1109/LGRS.2017.2728698>
37. A. Sharma, X. Liu and X. Yang, "Land cover classification from multi-temporal, multi-spectral remotely sensed imagery using patch-based recurrent neural networks", *Neural Networks* **105**, 346–355 (2018). <https://doi.org/10.1016/j.neunet.2018.05.019>
38. L. Mou, P. Ghamisi and X.X. Zhu, "Deep recurrent neural networks for hyperspectral image classification", *IEEE Trans. Geosci. Remote Sens.* **55(7)**, 3639–3655 (2017). <https://doi.org/10.1109/TGRS.2016.2636241>
39. H. Wu and S. Prasad, "Convolutional recurrent neural networks for hyperspectral data classification", *Remote Sens.* **9(3)**, 298 (2017). <https://doi.org/10.3390/rs9030298>
40. C. Shi and C.M. Pun, "Multi-scale hierarchical recurrent neural networks for hyperspectral image classification", *Neurocomputing* **294**, 82–93 (2018). <https://doi.org/10.1016/j.neucom.2018.03.012>
41. Z. Zhang, "Semi-supervised hyperspectral image classification algorithm based on graph embedding and discriminative spatial information", *Microproc. Microsyst.* **75**, 103070 (2020). <https://doi.org/10.1016/j.micpro.2020.103070>
42. F. Li, D.A. Clausi, L. Xu and A. Wong, "ST-IRGS: a region-based self-training algorithm applied to hyperspectral image classification and segmentation", *IEEE Trans. Geosci. Remote Sens.* **56(1)**, 3–16 (2018). <https://doi.org/10.1109/TGRS.2017.2713123>
43. X. Cao, J. Yao, Z. Xu and D. Meng, "Hyperspectral image classification with convolutional neural network and active learning", *IEEE Trans. Geosci. Remote Sens.* **58(7)**, 4604–4616 (2020). <https://doi.org/10.1109/TGRS.2020.2964627>
44. Z. He, K. Xia, T. Li, B. Zu, Z. Yin and J. Zhang, "A constrained graph-based semi-supervised algorithm combined with particle cooperation and competition for hyperspectral image classification", *Remote Sens.* **13(2)**, 193 (2021). <https://doi.org/10.3390/rs13020193>
45. B. Huang, L. Ge, G. Chen, M. Radenkovic, X. Wang, J. Duan and Z. Pan, "Nonlocal graph theory based transductive learning for hyperspectral image classification", *Pattern Recognition* 107967 (2021). <https://doi.org/10.1016/j.patcog.2021.107967>
46. Z. He, H. Liu, Y. Wang and J. Hu, "Generative adversarial networks-based semi-supervised learning for hyperspectral image classification", *Remote Sens.* **9(10)**, 1042 (2017). <https://doi.org/10.3390/rs9101042>
47. R. Kemker and C. Kanan, "Self-taught feature learning for hyperspectral image classification", *IEEE Trans. Geosci. Remote Sens.* **55(5)**, 2693–2705 (2017). <https://doi.org/10.1109/TGRS.2017.2651639>
48. R. Kemker, R. Luu and C. Kanan, "Low-shot learning for the semantic segmentation of remote sensing imagery", *IEEE Trans. Geosci. Remote Sens.* **56(10)**,

- 6214–6223 (2018). <https://doi.org/10.1109/TGRS.2018.2833808>
49. M. Ahmad, S. Shabbir, D. Oliva, M. Mazzara and S. Distefano, “Spatial-prior generalized fuzziness extreme learning machine autoencoder-based active learning for hyperspectral image classification”, *Optik* **206**, 163712 (2020). <https://doi.org/10.1016/j.ijleo.2019.163712>
50. P. Liu, H. Zhang and K.B. Eom, “Active deep learning for classification of hyperspectral images”, *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.* **10(2)**, 712–724 (2016). <https://doi.org/10.1109/JSTARS.2016.2598859>
51. J.M. Haut, M.E. Paoletti, J. Plaza, J. Li and A. Plaza, “Active learning with convolutional neural networks for hyperspectral image classification using a new Bayesian approach”, *IEEE Trans. Geosci. Remote Sens.* **56(11)**, 6440–6461 (2018). <https://doi.org/10.1109/TGRS.2018.2838665>
52. B. Cui, X. Xie, S. Hao, J. Cui and Y. Lu, “Semi-supervised classification of hyperspectral images based on extended label propagation and rolling guidance filtering”, *Remote Sens.* **10(4)**, 515 (2018). <https://doi.org/10.3390/rs10040515>
53. A. Appice, P. Guccione and D. Malerba, “A novel spectral-spatial co-training algorithm for the transductive classification of hyperspectral imagery data”, *Pattern Recognition* **63**, 229–245 (2017). <https://doi.org/10.1016/j.patcog.2016.10.010>
54. L. Zhang, L. Zhang, B. Du, J. You and D. Tao, “Hyperspectral image unsupervised classification by robust manifold matrix factorization”, *Inform. Sci.* **485**, 154–169 (2019). <https://doi.org/10.1016/j.ins.2019.02.008>
55. H. Ghanbari, S. Homayouni, A. Safari and P. Ghamisi, “Gaussian mixture model and Markov random fields for hyperspectral image classification”, *Eur. J. Remote Sens.* **51(1)**, 889–900 (2018). <https://doi.org/10.1080/22797254.2018.1503565>
56. C. Bo, H. Lu and D. Wang, “Spectral-spatial K-Nearest Neighbor approach for hyperspectral image classification”, *Multimed. Tools App.* **77(9)**, 10419–10436 (2018). <https://doi.org/10.1007/s11042-017-4403-9>
57. R. Wang, F. Nie and W. Yu, “Fast spectral clustering with anchor graph for large hyperspectral images”, *IEEE Geosci. Remote Sens. Lett.* **14(11)**, 2003–2007 (2017). <https://doi.org/10.1109/LGRS.2017.2746625>
58. H. Xie, A. Zhao, S. Huang, J. Han, S. Liu, X. Xu and X. Tong, “Unsupervised hyperspectral remote sensing image clustering based on adaptive density”, *IEEE Geosci. Remote Sens. Lett.* **15(4)**, 632–636 (2018). <https://doi.org/10.1109/LGRS.2017.2786732>
59. Y. Zhao, Y. Yuan and Q. Wang, “Fast spectral clustering for unsupervised hyperspectral image classification”, *Remote Sens.* **11(4)**, 399 (2019). <https://doi.org/10.3390/rs11040399>
60. D. Lin, K. Fu, Y. Wang, G. Xu and X. Sun, “MARTA GANs: Unsupervised representation learning for remote sensing image classification”, *IEEE Geosci. Remote Sens. Lett.* **14(11)**, 2092–2096 (2017). <https://doi.org/10.1109/LGRS.2017.2752750>
61. Y. Wei, X. Luo, L. Hu, Y. Peng and J. Feng, “An improved unsupervised representation learning generative adversarial network for remote sensing image scene classification”, *Remote Sens. Lett.* **11(6)**, 598–607 (2020). <https://doi.org/10.1080/2150704X.2020.1746854>
62. A. Chatterjee, J. Saha, J. Mukherjee, S. Aikat and A. Misra, “Unsupervised land cover classification of hybrid and dual-polarized images using deep convolutional neural network”, *IEEE Geosci. Remote Sens. Lett.* **18(6)**, 969–973 (2020). <https://doi.org/10.1109/LGRS.2020.2993095>
63. M.P. Uddin, M.A. Mamun and M.A. Hossain, “PCA-based feature reduction for hyperspectral remote sensing image classification”, *IETE Tech. Rev.* 1–21 (2020). <https://doi.org/10.1080/02564602.2020.1740615>
64. C. Jayaprakash, B.B. Damodaran, S. Viswanathan and K.P. Soman, “Randomized independent component analysis and linear discriminant analysis dimensionality reduction methods for hyperspectral image classification”, *J. Appl. Remote Sens.* **14(3)**, 036507 (2020). <https://doi.org/10.1117/1.JRS.14.036507>
65. M. Fordellone, A. Bellincontro and F. Mencarelli, “Partial least squares discriminant analysis: a dimensionality reduction method to classify hyperspectral data”, *ArXiv preprint arXiv: 1806.09347* (2018). <https://arxiv.org/abs/1806.09347>
66. M. Ramamurthy, Y.H. Robinson, S. Vimal and A. Suresh, “Auto encoder based dimensionality reduction and classification using convolutional neural networks for hyperspectral images”, *Microproc. Microsyst.* **79**, 103280 (2020). <https://doi.org/10.1016/j.micpro.2020.103280>

67. R. Reshma, V. Sowmya and K.P. Soman, "Effect of Legendre–Fenchel denoising and SVD-based dimensionality reduction algorithm on hyperspectral image classification", *Neural Comput. Appl.* **29(8)**, 301–310 (2018). <https://doi.org/10.1007/s00521-017-3145-y>
68. D. Li, X. Wang and Y. Cheng, "Spatial-spectral neighbour graph for dimensionality reduction of hyperspectral image classification", *Int. J. Remote Sens.* **40(11)**, 4361–4383 (2019). <https://doi.org/10.1080/01431161.2018.1562587>
69. K.S. Charmisha, V. Sowmya and K.P. Soman, "Dimensionality reduction by dynamic mode decomposition for hyperspectral image classification using deep learning and kernel methods", in *Advances in Signal Processing and Intelligent Recognition Systems. SIRS 2018*, Ed by S. Thampi, O. Marques, S. Krishnan, K.C. Li, D. Ciuonzo and M. Kolekar. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-5758-9_22
70. R. Hang and Q. Liu, "Dimensionality reduction of hyperspectral image using spatial regularized local graph discriminant embedding", *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.* **11(9)**, 3262–3271 (2018). <https://doi.org/10.1109/JSTARS.2018.2847042>
71. A. Paul and N. Chaki, "Dimensionality reduction of hyperspectral images using pooling", *Pattern Recogn. Image Anal.* **29(1)**, 72–78 (2019). <https://doi.org/10.1134/S1054661819010085>
72. D.R. Hidalgo, B.B. Cortés and E.C. Bravo, "Dimensionality reduction of hyperspectral images of vegetation and crops based on self-organized maps", *Inform. Process. Agric.* **8(2)**, 310–327 (2021). <https://doi.org/10.1016/j.inpa.2020.07.002>
73. B. Gowtham, "Hyperspectral image analysis using principal component analysis and Siamese network", *Turk. J. Comput. Math. Educ.* **12(7)**, 1191–1198 (2021).
74. S.S. Sawant and P. Manoharan, "Unsupervised band selection based on weighted information entropy and 3D discrete cosine transform for hyperspectral image classification", *Int. J. Remote Sens.* **41(10)**, 3948–3969 (2020). <https://doi.org/10.1080/01431161.2019.1711242>
75. M. Prabukumar, S. Sawant, S. Samiappan and L. Agilandeewari, "Three-dimensional discrete cosine transform-based feature extraction for hyperspectral image classification", *J. Appl. Remote Sens.* **12(4)**, 046010 (2018). <https://doi.org/10.1117/1.JRS.12.046010>
76. M. Prabukumar and S. Shrutika, "Band clustering using expectation–maximization algorithm and weighted average fusion-based feature extraction for hyperspectral image classification", *J. Appl. Remote Sens.* **12(4)**, 046015 (2018). <https://doi.org/10.1117/1.JRS.12.046015>
77. A. Ghorbanian and A. Mohammadzadeh, "An unsupervised feature extraction method based on band correlation clustering for hyperspectral image classification using limited training samples", *Remote Sens. Lett.* **9(10)**, 982–991 (2018). <https://doi.org/10.1080/2150704X.2018.1500723>
78. J. Jiang, J. Ma, C. Chen, Z. Wang, Z. Cai and L. Wang, "SuperPCA: A superpixelwise PCA approach for unsupervised feature extraction of hyperspectral imagery", *IEEE Trans. Geosci. Remote Sens.* **56(8)**, 4581–4593 (2018). <https://doi.org/10.1109/TGRS.2018.2828029>
79. L. Zhang, H. Su and J. Shen, "Hyperspectral dimensionality reduction based on multiscale superpixelwise kernel principal component analysis", *Remote Sens.* **11(10)**, 1219 (2019). <https://doi.org/10.3390/rs11101219>
80. G.Y. Shi, H. Huang, J.M. Liu, Z.Y. Li and L.H. Wang, "Spatial-spectral multiple manifold discriminant analysis for dimensionality reduction of hyperspectral imagery", *Remote Sens.* **11**, 2414 (2019). <https://doi.org/10.3390/rs11202414>
81. F.J. Orts Gómez, G. Ortega López, E. Filatovas, O. Kurasova and G.E.M. Garzón, "Hyperspectral image classification using Isomap with SMACOF", *Informatica* **30(2)**, 349–365 (2019). <https://doi.org/10.15388/Informatica.2019.209>
82. W. Li, L. Zhang, L. Zhang and B. Du, "GPU parallel implementation of isometric mapping for hyperspectral classification", *IEEE Geosci. Remote Sens. Lett.* **14(9)**, 1532–1536 (2017). <https://doi.org/10.1109/LGRS.2017.2720778>
83. J. Jiang, J. Ma, C. Chen, Z. Wang, Z. Cai and L. Wang, "SuperPCA: A superpixelwise PCA approach for unsupervised feature extraction of hyperspectral imagery", *IEEE Trans. Geosci. Remote Sens.* **56(8)**, 4581–4593 (2018). <https://doi.org/10.1109/TGRS.2018.2828029>
84. L. Zhang, H. Su and J. Shen, "Hyperspectral dimensionality reduction based on multiscale superpixelwise kernel principal component analysis", *Remote*

- Sens.* **11(10)**, 1219 (2019). <https://doi.org/10.3390/rs11101219>
85. A. Ertem, A.C. Karaca, O. Urhan and M.K. Güllü, "Superpixel based compression of hyperspectral image with modified dictionary and sparse representation", *Int. J. Remote Sens.* **41(16)**, 6307–6324 (2020). <https://doi.org/10.1080/01431161.2020.1737338>
86. W. Fu, S. Li, L. Fang and J.A. Benediktsson, "Adaptive spectral-spatial compression of hyperspectral image with sparse representation", *IEEE Trans. Geosci. Remote Sens.* **55(2)**, 671–681 (2017). <https://doi.org/10.1109/TGRS.2016.2613848>
87. N. Zikiou, M. Lahdir and D. Helbert, "Support vector regression-based 3D-wavelet texture learning for hyperspectral image compression", *The Visual Computer* **36(7)**, 1473–1490 (2020). <https://doi.org/10.1007/s00371-019-01753-z>
88. R. Nagendran and A. Vasuki, "Hyperspectral image compression using hybrid transform with different wavelet-based transform coding", *Int. J. Wavelets Multires. Inform. Process.* **18(1)**, 1941008 (2020). <https://doi.org/10.1142/S021969131941008X>
89. M. Hernández-Cabronero, J. Portell, I. Blanes and J. Serra-Sagristà, "High-performance lossless compression of hyperspectral remote sensing scenes based on spectral decorrelation", *Remote Sens.* **12(18)**, 2955 (2020). <https://doi.org/10.3390/rs12182955>