REPUBLIC OF TÜRKİYE

YILDIZ TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

HYPERSPECTRAL IMAGE CLASSIFICATION USING AN ACTIVE DEEP LEARNING FRAMEWORK INCLUDING EDGE PRESERVING FILTERS

Zainab Dheyaa Mohammed AL-SAMMARRAIE

MASTER OF SCIENCE THESIS

Department of Computer Engineering

Computer Engineering Program

Supervisor

Assoc. Prof. Dr. Ali Can KARACA

February, 2024

REPUBLIC OF TÜRKİYE

YILDIZ TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

HYPERSPECTRAL IMAGE CLASSIFICATION USING AN ACTIVE DEEP LEARNING FRAMEWORK INCLUDING EDGE PRESERVING FILTERS

A thesis submitted by Zainab Dheyaa Mohammed AL-SAMMARRAIE in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE is approved by the committee on 06.02.2024 in Department of Computer Engineering, Computer Engineering Program.

Assoc. Prof. Dr. Ali Can KARACA
Yıldız Technical University
Supervisor

Approved By the Examining Committee

Assoc. Prof. Dr. Ali Can KARACA, Supervisor
Yildiz Technical University _____

Assoc. Prof. Dr. Hamza Osman İLHAN, Member
Yildiz Technical University _____

Asst. Prof. Dr. Coşku ÖKSÜZ, Member
Kastamonu University _____

I hereby declare that I have obtained the required legal permissions during data collection and exploitation procedures, that I have made the in-text citations and cited the references properly, that I haven't falsified and/or fabricated research data and results of the study and that I have abided by the principles of the scientific research and ethics during my Thesis Study under the title of "Hyperspectral Image Classification Using An Active Deep Learning Framework Including Edge Preserving Filters" supervised by my supervisor, Assoc. Prof. Dr. Ali Can KARACA. In the case of a discovery of false statement, I am to acknowledge any legal consequence.

Zainab Dheyaa Mohammed AL-SAMMARRAIE

Signature

Dedicated to soul my parents

and my sister and best friends

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my supervisor, Assoc. Prof. Dr. Ali Can Karaca for his excellent guidance, patience, encouragement, and always supporting me in difficult moments. He helped me throughout this study; this study would not have been possible without his great support and guidance. I would also like to thank my family very much for their support and encouragement.

With best wishes.

Zainab Dheyaa Mohammed AL-SAMMARRAIE

TABLE OF CONTENTS

TABLE OF CONTENTS v
LIST OF SYMBOLS vii
LIST OF ABBREVIATIONS viii
LIST OF FIGURES ix
LIST OF TABLES x
ABSTRACT xi
1 INTRODUCTION 1
1.1 Fundamental Information 1
1.2 Literature Review 7
1.2.1 The Research about Active Learning 7
1.2.2 The Research about Dimensional Reduction 9
1.2.3 The Research about Edge-Preserving Filters 10
1.3 Motivation 11
1.4 Outlines of the Thesis 13
2 HYPERSPECTRAL IMAGE CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORKS AND ACTIVE LEARNING 14
2.1 CNN+AL+MRF Framework 17
2.2 Improved CNN+AL+MRF Framework 19
2.2.1 Dimension Reduction Techniques 22
2.2.2 Average Band Technique (AVG) 22
2.2.3 Principal Component Analysis (PCA) 25
2.2.4 Data Augmentation 27
2.2.5 Edge-Preserving Filters (EPF) 27
2.3 Algorithm Details 37
3 DATASETS 41
3.1. Indian Pines 41
3.2 Pavia University 42
3.3 Salinas Dataset 44
4 EXPERIMENTS 46
4.1 Experimental Setup 46
4.2 Criteria for Evaluating Performance 47
4.1.1 Overall Accuracy (OA) 47
4.1.2 Average Accuracy (AA) 47
4.1.3 Computation Time (s) 47

4.3 Effect of Edge Preserving Filtering on Datasets 48
4.4 Performance Evaluation of The Proposed Framework 49
4.4.1 Results of Indian Pines (IP) 50
4.4.2. Results of Pavia University (PU) 56
4.4.3 Results of Salinas 61
4.5 Overall Comparison 68
5 CONCLUSION **70**
REFERENCES **72**
PUBLICATION FROM THE THESIS **79**

LIST OF SYMBOLS

DA	Data Augmentation
DimR	Dimension Reduction
R	Iteration Number
L	Labeled Pixels
D	Number of Actively Chosen Pixels in Each Round
PSB (zi)	The Best Class Membership Probability of Sample zi
RB	The Desired Reduced Number of Bands
PSB (zi)	The Second-Best Class Membership Probability of Sample zi
B	The Total Number of Bands
U	Unlabeled Pixels

LIST OF ABBREVIATIONS

AIF	Averaging-based Image Fusion
AL	Active Learning
AVG	Averaging Band Process
BNN	Bayesian Neural Network
CIE	Commission on Illumination
CNN	Convolutional Neural Networks
DA	Data Augmentation
Deep AL	Deep Active Learning
DL	Deep Learning
DR	Dimensionality Reduction
ELM	Extreme Learning Machines
EM	Electromagnetic
EMR	Electromagnetic Radiation
EPF	Edge-Preserving Filtering
FE	Feature Extraction
GT	Ground Truth
HDR	High Dynamic Range
HIC	Hyperspectral Image Classification
HS	Hyperspectral Image
IEPF	Iterative Edge-Preserving Filtering
MNF	Minimum Noise Fraction
MRF	Markov Random Fields
MSI	Multi Spectral Imaging
NIR	Near-Infrared
PCA	Principal Component Analysis
RGB	Red, Green, Blue
RS	Remote Sensing
SP	Stabilizing Prediction
3D	Three Dimensions

LIST OF FIGURES

Figure 1.1 Representation of Hyperspectral Image Cube **3**
Figure 1.2 The Pixel Spectral of Various of Objects **4**
Figure 1.3 Some Applications of HIS **5**
Figure 1.4 The Pool-Based AL Cycle **8**
Figure 2.1 Active Deep Learning Framework **18**
Figure 2.2 CNN Architecture **2**
Figure 2.3 An Example of Markov Random Field **19**
Figure 2.4 Enhancing CNN+AL+MRF Active Deep Learning **22**
Figure 2.5 Illustrate Example of AVG Process **23**
Figure 2.6 Principal Component Analysis **25**
Figure 2.7 Result of WLS on three datasets with and without DR **29**
Figure 2.8 Result of WMF on three datasets with and without DR **31**
Figure 2.9 Result of FGS on three datasets with and without DR **33**
Figure 2.10 Result of BF on three datasets with and without DR **34**
Figure 2.11 Result of SD on three datasets with and without DR **36**
Figure 3.1 Original Image (A) [98] and GT of IP dataset (B) **41**
Figure 3.2 Original Image (A) [98] and GT of IP dataset (B) **43**
Figure 3.3 Original Image (A)[98] and GT of SA (B) **44**
Figure 4.1 Classification maps for IP data set for using 2% training samples of the methods without DR and EPF **55**
Figure 4.2 Classification maps for IP data set for using 2% training samples of the methods with AVG and EPF **55**
Figure 4.3 Classification maps for IP dataset using 2% training samples of the methods with PCA and EPF **56**
Figure 4.4 Classification maps for PU dataset using 1% training samples of the methods without dimension reduction **59**
Figure 4.5 Classification maps for PU data set for using 1% training samples of the methods with AVG and EPF **60**
Figure 4.6 Classification maps for PU dataset using 1% training samples of the methods with PCA and EPF **61**
Figure 4.7 Classification maps for SA dataset using 2% training samples of the methods without dimension reduction **66**
Figure 4.8 Classification maps for SA dataset using 2% training samples of the methods with AVG and EPF **67**
Figure 4.9 Classification maps for SA dataset using 2% training samples of the methods with PCA and EPF **68**

LIST OF TABLES

Table 3.1 The dataset of hyperspectral in the experiments **41**

Table 3.2 Number of test and train samples (2%) in IP dataset **42**

Table 3.3 Number of test and train samples (1%) in PU dataset **43**

Table 3.4 Number of test and train samples (2%) in SA dataset **45**

Table 4.1 Hyperparameters about Active Learning during different experiments **49**

Table 4.2 Hyperparameters about different EPF methods **49**

Table 4.3 The accuracy percentages (%) and computational times for IP methods of classification without DR and EPF **51**

Table 4.4 The accuracy percentages (%) and computational times for IP methods of classification with AVG and EPF **52**

Table 4.5 The accuracy percentages (%) and computational times for IP methods of classification with PCA and EPF **53**

Table 4.6 The accuracy percentages (%) and computational times for PU methods of classification without DR and EPF **57**

Table 4.7 The accuracy percentages (%) and computational times for PU methods of classification with AVG and EPF **57**

Table 4.8 The accuracy percentages (%) and computational times methods of classification of PU with PCA and EPF **58**

Table 4.9 The accuracy percentages (%) and computational times methods of classification of SA without DR and with EPF **62**

Table 4.10 The accuracy percentages (%) and computational times methods of classification of SA with AVG and EPF **63**

Table 4.11 The accuracy percentages (%) and computational times methods of classification of SA with PCA and EPF **64**

Table 4.12 Summarized the Best Results (%) for Used Techniques **69**

ABSTRACT

Hyperspectral Image Classification Using An Active Deep Learning Framework Including Edge Preserving Filters

Zainab Dheyaa Mohammed AL-SAMMARRAIE

Department of Computer Engineering

Computer Engineering Program

Master of Science Thesis

Supervisor: Assoc. Prof. Dr. Ali Can KARACA

Hyperspectral images with high spectral resolution allow the identification and differentiation of land cover materials. It is undeniable that hyperspectral image classification is critical for applications including military operations, environmental monitoring, geological research, and agriculture. However, pixel labeling is a tedious and resource-intensive procedure and thus calls for an enhanced approach that achieves good results even with a small sample size. Since hyperspectral images are high-dimensional, traditional algorithms have difficulty recognizing them efficiently, due to their computational complexity. In response, cutting-edge technologies such as DL, AL, and ML have great potential to manage and extract actionable insights from hyperspectral data.

Classifiers based on AL have become a particularly effective strategy for hyperspectral images, especially in a very small set of training samples. Iterative sampling techniques entail regularly feeding the classifier model with new labeled pixels corresponding to those that will increase the model's performance compared to random sampling. Thus, active learning algorithms have become more popular in many remote sensing applications. Using a set of unlabeled samples, AL iteratively selects the most informative samples. It is also important to select the

classifier in AL. To increase accuracy as the number of samples decreases, CNN-based classification techniques use this methodology by selectively training on the most relevant data.

The objective of this thesis is to enhance the existing CNN-AL-MRF framework in the literature by integrating dimensionality reduction methods with edge-preserving filters, coupled with strategically implemented AL iterations. The comprehensive analysis incorporates five different edge-preserving filters: weighted least squares (WLS), joint histogram weighted mean filter (joint WMF), fast global smoothing filter (FGS), binary filter (BF), and static/dynamic (S/D) with techniques averaging (AVG) and principal component analysis (PCA). The synergy of dimensionality reduction methods with edge-preserving smoothing filters enhances the analysis of hyperspectral images. This improvement allows for the accurate extraction and classification of features by preserving crucial edges while mitigating noise and unnecessary details. This method not only improves the classification performance but also maintains the image quality. Examining contextual information with and without Markov random fields is a key element of this project. We tested the proposed framework on three well-known datasets to see how well it balances resource usage and provides a cost-effective way to properly classify HSI.

In conclusion, our thesis provides a powerful hyperspectral classification framework for limited sample numbers by applying dimensionality reduction techniques and edge preserving filters.

Keywords: Hyperspectral imaging, dimension reduction, edge preserving filters, convolutional neural network, active learning.

ÖZET

Hiperspektral Görüntülerin Kenar Koruyucu Filtreler İçeren Bir Aktif Derin Öğrenme Çerçevesiyle Sınıflandırılması

Zainab Dheyaa Mohammed AL-SAMMARRAIE

Bilgisayar Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Danışman: Doç. Dr. Ali Can KARACA

Yüksek spektral çözünürlüğe sahip hiperspektral görüntüler, arazi örtüsü malzemelerinin tanımlanmasına ve ayırt edilmesine olanak sağlar. Hiperspektral görüntü sınıflandırmanın askeri operasyonlar, çevresel izleme, jeolojik araştırmalar ve tarım gibi uygulamalar için kritik olduğu inkar edilemez. Ancak piksel etiketleme sıkıcı ve kaynak yoğun bir prosedürdür ve bu nedenle küçük bir örnek boyutuyla bile iyi sonuçlar elde eden gelişmiş bir yaklaşım gerektirir. Hiperspektral görüntüler yüksek boyutlu olduğundan, geleneksel algoritmalar, hesaplama karmaşıklıkları nedeniyle bunları verimli bir şekilde tanımakta zorluk çeker. Buna karşılık, derin öğrenme, aktif öğrenme ve makine öğrenimi gibi en son teknolojiler, hiperspektral verilerden eyleme geçirilebilir içgörüleri yönetme ve çıkarma konusunda büyük bir potansiyele sahiptir. Aktif öğrenmeye (AL) dayalı sınıflandırıcılar, özellikle çok küçük bir eğitim örnekleri kümesinde hiperspektral görüntüler için özellikle etkili bir strateji haline gelmiştir. Yinelemeli örnekleme teknikleri, sınıflandırıcı modelinin, rastgele örneklemeyle karşılaştırıldığında modelin performansını artıracak olanlara karşılık gelen yeni etiketli piksellerle düzenli olarak beslenmesini gerektirir. Böylece aktif öğrenme algoritmaları birçok uzaktan algılama uygulamasında daha popüler hale gelmiştir. Aktif öğrenme, bir dizi etiketlenmemiş örnek kullanarak yinelemeli olarak en bilgilendirici örnekleri

seçer. Aktif öğrenmede sınıflandırıcının seçimi de önemlidir. Örnek sayısı azaldıkça doğruluğu artırmak için CNN tabanlı sınıflandırma teknikleri, en ilgili veriler üzerinde seçici olarak eğitim vererek bu metodolojiyi kullanır.

Bu tezin amacı, stratejik olarak uygulanan aktif öğrenme yinelemeleriyle birlikte, boyut azaltma yöntemlerini kenar koruyucu filtrelerle entegre ederek literatürdeki mevcut CNN-AL-MRF çerçevesini geliştirmektir. Kapsamlı analiz beş farklı kenar koruyucu filtre içerir: Ağırlıklı En Küçük Kareler (WLS), Ortak Histogram Ağırlıklı Ortanca Filtre (Joint WMF), Hızlı Global Görüntü Yumuşatma (FGS), Bilateral Filtre (BF), and Static/dynamic (SD) ortalama (AVG) ve temel bileşen analizi (PCA) teknikleri ile. Kenar koruyucu yumuşatma filtreleri ile boyutsallık azaltma yöntemlerinin sinerjisi, hiperspektral görüntülerin analizini geliştirir. Bu iyileştirme, gürültüyü ve gereksiz ayrıntıları azaltırken önemli kenarları koruyarak özelliklerin doğru şekilde çıkarılmasına ve sınıflandırılmasına olanak tanır. Bu yöntem sınıflandırma performansını artırmanın yanı sıra görüntü kalitesini de korur. Bağlamsal bilgilerin Markov rastgele alanları ile ve Markov olmadan incelenmesi bu projenin önemli bir unsurudur. Önerilen çerçeveyi, kaynak kullanımını ne kadar iyi dengelediğini ve HSI'yi uygun şekilde sınıflandırmak için uygun maliyetli bir yol sağladığını görmek için iyi bilinen üç veri kümesi üzerinde test ettik.

Sonuç olarak tezimiz, boyut azaltma teknikleri ve kenar koruyucu filtreler uygulayarak sınırlı sayıda örnek için güçlü bir hiperspektral sınıflandırma çerçevesi sunmaktadır.

Anahtar Kelimeler: Hiperspektral görüntüleme, boyut azaltımı, kenar koruyan filtreler, evrişimsel sinir ağları, aktif öğrenme.

1
INTRODUCTION

This section provides fundamental information about motivation of the thesis, literature review in this field, and outline of the thesis.

1.1 Fundamental Information

The study of the Earth and its land cover is one of the most important studies because of its importance on our lives and the lives of all creatures. There are many factors that affect and contribute to some of the changes that occur because of climate change or human activities and experiments, and many of these factors are important, so it is necessary to monitor the changes that occur. Land cover classification is of great importance to know what is happening, so it is necessary to investigate and monitor all types of events. This is done either through field reconnaissance, which requires great effort and high cost, and is not a good option, so remote sensing is the ideal solution for this task [1].

Remote sensing refers to the process of obtaining information about objects or areas that have no physical contact. Using remote sensing, information such as images or other types of data can be collected from remote or potentially dangerous places without going there. Remote sensing is widely used in sensor technology carried by aircraft or launched into outer space to identify and classify objects on Earth, providing valuable information about those areas that may be inaccessible, which is important such as scientific, environmental, and industrial fields.

Historically, this method has mapped land cover using multispectral images according to the needed degrees of resolution. Despite their effectiveness, Multispectral images does not provide spectral information as detailed as hyperspectral images (HSIs), which has caused a paradigm change in how we monitor the Earth's surface [2].

In terms of accurate pixel identification and object detection, HSI devices surpass conventional imaging technologies, thanks to their unsurpassed capabilities. By moving away from multispectral imaging toward hyperspectral imaging, which

offers substantially better spectral resolution, the precision of remote sensing applications is improved, enabling precise categorization in species recognition.

Hyperspectral imaging has become one of the basic techniques in remote sensing, as each pixel has a distinct spectral fingerprint, making it possible to detect terrestrial elements that are not visible to the naked eye [3].

One of the fundamental techniques in remote sensing is that hyperspectral imaging, where each pixel has a distinct spectral signature, makes it possible to detect terrestrial elements that are not visible to the naked eye [3]. In the past 10 years, hyperspectral image analysis has become one of the most advanced and fastest developing remote sensing techniques. Advances in hyperspectral imaging technology have allowed hyperspectral sensors to collect large amounts of data across a wide band of electromagnetic frequencies [3].

Therefore, this technology is a means of combining the spectrum of ground objects identified by their unique physical composition with a spatial image that reflects the shape, texture, and layout of ground objects to achieve accurate detection, recognition, and analysis of their features.

The resulting of this combining is the hyperspectral images contain an abundance of spectral information that reflects the unique physical of properties earth features and rich spatial information for it. The hyperspectral image is often represented as a three-dimensional cube with three axes: x and y, representing the width and height of the image containing spatial information, and the z-axis indicating the number of spectral bands [4].

The cube structure is especially valuable if the images contain the largest number of bands. The example in Figure 1.1 shows the x and y axes (spatial) and the z axis (spectral bands).

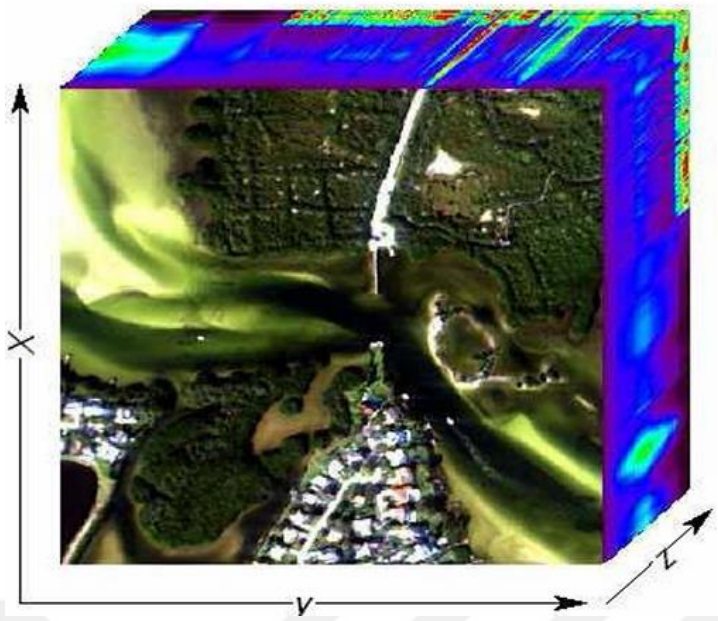


Figure 1.1 Representation of Hyperspectral Image Cube [4]

Hyperspectral imaging examines the spectrum of light reflected by an object at a variety of wavelengths and represents a major technological convergence by capturing data across many narrow, adjacent bands within the electromagnetic spectrum, unlike traditional or multispectral imaging [5].

Due to their wide band of spectral data acquired across different wavelengths, they provide highly accurate information, allowing to facilitate accurate identification and classification of features such as soil, water, urban areas, and forests, figure 1.2 shows a model of the spectra of an object for a defense sector company [6].

Notice the variation and differences in wavelengths that help in distinguishing different types of materials in the same classification. For example, distinguishing two different types of trees based on the spectral changes [6]. Thus, hyperspectral images have proven superior to more complex classifications [7].

HS imaging sensors are widely used for monitoring earth's surface through RS due to their high spectral resolution, these sensors record more than hundreds of spectral bands [7]. Defense research, urban inquiry, crop analysis, geological mapping, mineral exploitation, and military surveillance are just a few of the many fields that find use for hyperspectral images (HSI).

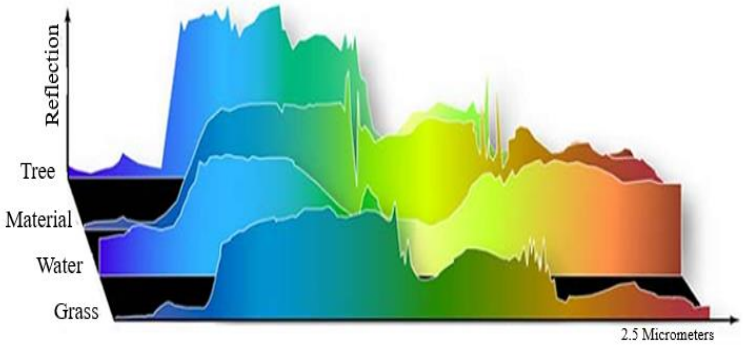


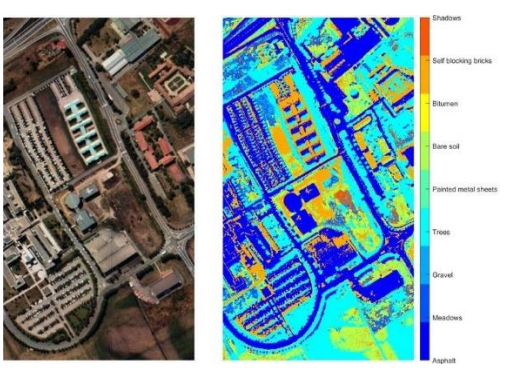
Figure 1.2 The Pixel Spectral of Various of Objects [6]

Figure 1.2 shows show some of applications of HSI, for example, urban inquiry provides a detailed description of the urban environment and classifies the types of land cover within the areas specified for classification, such as buildings, tree, roads, etc.

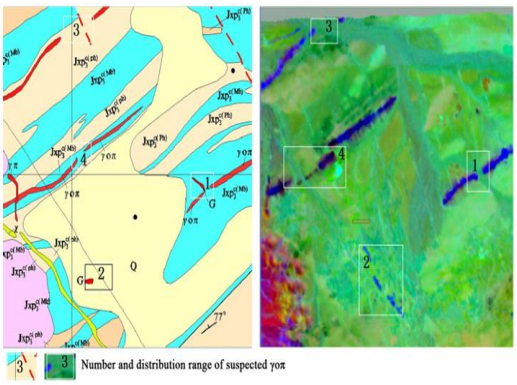
Its aim is to evaluate and monitor the health of the land cover and distribution, monitor air and water quality, and much more information about land cover. Which contributes to reducing damage and overcoming the consequences of potential disasters, and many things, thus taking advantage of the ability of hyperspectral sensors to analyze the complexities of urban landscapes on a fine scale.

Geological mapping describes the Earth's surface and subsurface geology, providing detailed information about the composition of geological materials, resource exploration and geological risk assessment.

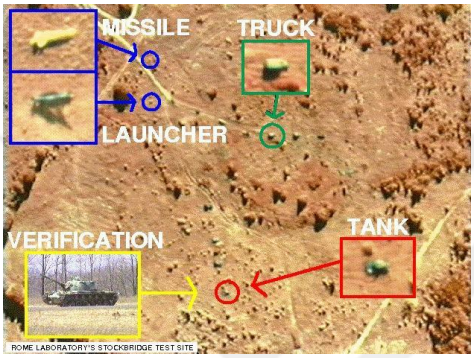
They enhance the ability to distinguish between geological features and various other materials. Defense research that uses the HSI includes various military and defense fields, for example, classifying and identifying military-classified objects such as targets deployed on land, sea, and air. The HSI helps in discovering the camouflage patterns used in these fields to conceal equipment and thus leads to enhanced security and sustainability for defense facilities.



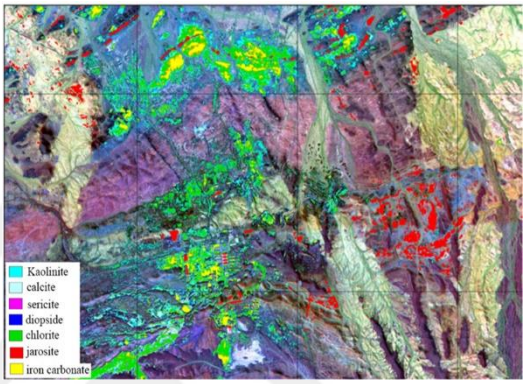
Urban Inquiry (A) [8]



Geological Mapping (B) [7]



Defense Research (C) [9]



Mineral Exploitation (D) [7]

Figure 1.3 Some Applications of HIS

HSI is widely used in the discovery of important minerals and resources due to its ability to identify, characterize and distribute minerals that depended on available information about geological structures, result of its unique signatures in decision making, resource assessment and environmental management in the mining industry.

Therefore, HSI is more suitable for classification applications, and have become more popular recently [5]. But on the other hand, these images contain a large number of spatial bands , which create highly correlated or adjacent bands that often lead to a large amount of repetition and high effort in training the model, so the need arose to use advanced computers to address this problem, but it is not an ideal solution, which led to the evolution and development of algorithms that address these problems, and the best of them are machine learning techniques, that have the ability to deal with this data and train it better using different algorithms. All of which serve one purpose, which is appropriate computational capacity and optimal

performance with high quality for the model trained for classification and spectral image analysis.

There are many machine learning techniques, such as deep learning and active learning, they specialize in specific challenges and tasks to enhance learning capabilities and predictive decision-making, thus achieving high accuracy in classifying this data. Although there have been several attempts to integrate DL, AL, and HSI classification, spatial information should be considered to lessen the impact of this issue [10].

The methods of these techniques have emerged and increased in popularity due to their effective ability in hyperspectral image classification (HIC). The most prominent of these methods is convolutional neural networks (CNN) and for training spectral data, which information collected from variety wavelengths [11]. CNNs are used to classify HSI based on spectral data, with the advancement of Earth observation techniques, but many problems appear in this process.

The difficult task in HSI remote sensing [12] is classification (HIC), which entails using a predefined discriminant function and a certain set of spectral and spatial image features by giving each pixel vector a label. The nature of data annotation is time-consuming and expensive and therefore requires specialized knowledge that may not be easily accessible, which is the main reason for the difficulty of HIC. To is used by extracting important spatio-spectral features to resolve the defects leverage unlabeled relevant data to improve classification performance [13], a CNN and high cost of manual labeling [14]. Image recognition and segmentation processes widely use CNNs, which consider the spatial correlation between pixels. By sparingly including unlabeled training samples, this technique improves generalizability and resilience to overlay with poorly labeled data [15] by integrating spatial and spectral information from the primary hyperspectral image.

There are many methods that have achieved good accuracy results, but they still face some obstacles, such as the curse of dimensions and the lack of clarity of this type of image, which prompted researchers to continue studying this problem to solve it. In turn, through this thesis, we seek to solve some of the problems facing HIC by using dimensionality reduction techniques and edge-preserving filters as

pre-processing techniques before the training process in order to enhance the model's ability and efficiency.

1.2 Literature Review

Researchers are now actively studying spectral imaging for remote sensing of the objects and the features on the ground. In the field of remote sensing, this technology has made things easier and more convenient by serving as a substitute for big aperture, high spatial resolution satellite imaging systems. Let's first review the definition of the electromagnetic (EM) spectrum to properly comprehend this technology. The phrase "EM spectrum" refers to the entire range of electromagnetic radiation [16].

Waves or photons can be used to depict electromagnetic energy. According to the wave theory, light moves in a direct path and the power it carries vibrate in a wave-like manner unless it is affected by an external object. When spectral imaging was first applied, it was known as MSI and just a few chosen bands from the visible and infrared portions of the electromagnetic spectrum were used [17].

Two of the most well-known MSI systems are the land remote sensing satellite (Landsat) [18] in addition to ASTER, the advanced spaceborne thermal emission and reflection radiometer [18]. Hundreds of contiguous spectral bands are used in the more recent version, HSI, to distinguish between different natural and man-made materials.

Unfortunately, the cost of the equipment makes it difficult to collect these images, which restricts the availability of big data sets for learning-based categorization techniques. The problem of inadequate training data has been the subject of numerous investigations.

1.2.1 The Research about Active Learning

Active Learning (AL) is a ML method that intelligently selects the most useful data points, and queries about them. It starts with an initial set of labeled images for training, after which the model predicts labels for unlabeled data.

AL algorithms focus on uncertainties in the model, and present to one of the oracle annotators for labeling the uncertain samples. and then the labeled data is used to

relearn. This process is repeated, with the algorithm continually selecting the most valuable samples to label. The strategy behind active learning is effective for enhancing model accuracy even when working with a limited amount of labeled data. For all of this, AL has become an important branch of ML, which has prompted researchers to conduct a lot of research into it, and re-review some of it.

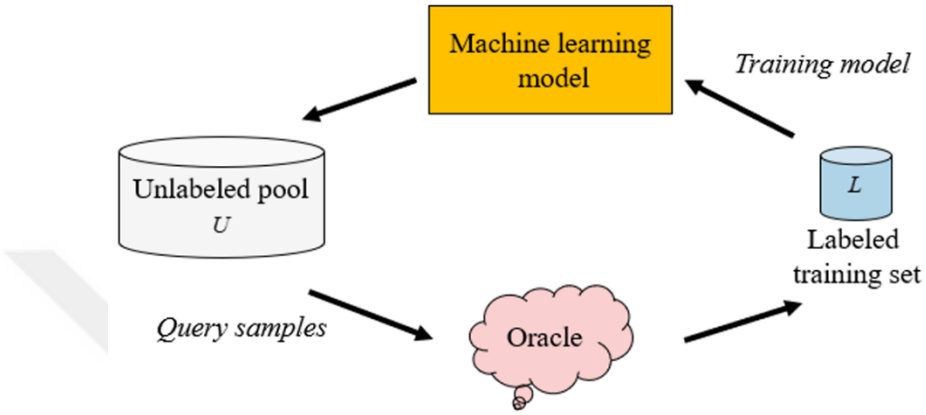


Figure 1.4 The Bool-Based AL Cycle [19]

A technique that leverages unlabeled relevant data and uses convolutional neural network (CNN) to improve classification performance by extracting important spatio-spectral features has been presented to solve the drawbacks and cost of manual labeling [20]. AL significantly reduces the required number of classifier samples,[20] and AL assumes that not all training samples for the classifier are equally important. F. Shi et al. 2020 presents a framework that combines AL and DL for hyperspectral agricultural image classification. It uses three AL algorithms (lowest confidence, margin sampling, and entropy) to select informative samples for manual annotation. Additionally, introduce “image pooling” to make the most of the augmented images [21].

An active deep learning method is proposed for simple supervised polarimetric synthetic aperture radar (PolSAR) image classification. This method integrates AL and fine-tuned CNNs to reduce the need for annotation and data augmentation due to label sparsity. Additionally, it effectively utilizes MRF to ensure label smoothness [22]. It emphasizes additional fine-tuning for robustness and accelerated convergence, expanding the PolSAR image database and exploring deep generative models for PolSAR classification on parallel computing architectures is recommended [22].

A Bayesian convolutional neural network (B-CNN) [23] guided by AL was used to classify hyperspectral images. [23] It combines spectral, spatial, and spectral-spatial features to achieve a good and robust classification by combining both AL with CNN. Using BNN networks, it gave good results, suggesting post-processing methods such as conditional random field.

Active deep learning [24] is proposed for hyperspectral image classification for remote sensing applications. It uses weighted incremental dictionary learning (WI-DL), which effectively selects samples to maximize representativeness and uncertainty. This method achieves greater accuracy with fewer training examples and outperforms competing methods at the time [24].

An improved active deep learning approach is introduced for the semi-supervised categorization of hyperspectral images (HSI) [25]. Integrating the multi-graph randomization technique and the anchor graph algorithm achieves the automation of fine-grained labeling for unlabeled data, effectively solving the problem of insufficiently labeled samples [25]. Adding pseudo-labeled samples enhances model fine-tuning and generalizability without requiring substantial manual labeling [25].

Cao et al. [26] have developed a new method for classifying hyperspectral images using a combination of DL and AL. The method involves selecting the most informative pixels from a pool of candidate labels, fine-tuning the CNN iteratively, and introducing a MRF to enforce class label smoothness. The method outperforms classical and DL-based algorithms with fewer labeled samples but has drawbacks like computing complexity and algorithm sensitivity.

It can be concluded from these studies that there is a common problem which is the lack of sufficient labeled data, and therefore some models need to be improved by adding some techniques before or after the training process.

1.2.2 The Research about Dimensional Reduction

PCA helps in band selection and noise reduction, while MNF helps in information extraction and noise redistribution but degrades spatial information [27]. By retrieving edge information, the triple filter improves classification accuracy by 1.5% compared to other models cited in this study, which demonstrates the value

of edge information being better for ML models without increasing their complexity [27].

It combines the features of global and local data, reduces the band numbers through averaging based on the fusion, and then uses edge-preserving filtering to divide the image into detailed layers. These techniques, which use dimensionality reduction and AIF techniques, work with edge-preserving filters. This approach is based on the principle of experiment to specify parameters [28].

The research of SSARIN [29], a specialized network for classifying invariant hyperspectral images, consists of band selection, local spatial feature optimization and light weighting. It also proposes the use of PCA in data pre-processing to reduce spectral dimensions, was conducted on five datasets of hyperspectral images, but although [29]. The SSARIN network extracts rotation-invariant spatial spectral features and has increased computational time, leading to recommendations to create lighter rotation-invariant features to improve performance and computational efficiency.

SuperPCA [30] is a new method that enables feature extraction from hyperspectral images and captures these diverse spectral features within homogeneous regions using super pixel segmentation that contains spatial information. It outperforms traditional methods, so it is considered a powerful method in the direction of noise, so it represents an effective modulation in hyperspectral image analysis [31].

1.2.3 The Research about Edge-Preserving Filters (EPF)

To solve the problems of methods not performing well on a variety of image contents and lack of widely used datasets for objective comparison, Zhu et al. [32] propose an edge-preserving image smoothing criterion. The benchmark consists of an image dataset with basic methods that may produce competitive results for a variety of image content as well as ground truth results for image smoothing. Trained deep networks provide superior qualitative and quantitative results and operate faster than most state-of-the-art smoothing methods [32].

An edge-preserving adaptive approach for hyperspectral image classification inside MRFs [33]. It improves categorization within several classes and edge preservation by using numerous adaptive maps. These maps improve label dependence by

finding the best balance between region uniformity and feature preservation. This makes classification more accurate while keeping fine details in HIS [32].

Applies a bilateral filter with a multispectral domain (MNF) [27] to HSI, merging the information of spatial and spectral. Using extreme learning machines (ELM) and Principal Component Analysis (PCA), this method dramatically increases classification accuracy over benchmark datasets. The MNF-based approach performs better in HSI classification than single spatial domain BF methods and efficiently maximizes spatial features as compared to other filtering techniques [33]. It [34] enhances the accuracy of classification by applying EPF to the maps, and thus the filtering process achieves a local improvement of the probabilities. This is unlike some methods that improve classification maps globally.

One of the advantages of it is the focus on spectral information and its dominance over spatial information, as well as its computational ability. It [34] is computationally efficient and, therefore, will be very useful in real applications, indicated that in the future, further improvement can be achieved by adapting the filtering size to spatial information and thus better filtering and improved classification accuracy.

IEPF repeatedly feeds the smooth probability maps and combines them with the processed image cube, which enhances the classification [35]. By repeatedly feeding the smooth probability maps and combining them with the processed image cube, so IEPF outperforms EPF and thus enhances the classification accuracy, and this is what this study indicated. Through the results of the experiments, it showed significant improvements compared to EPF.

1.3 Motivation

As it is seen from the literature, the active learning process is an important technique for hyperspectral imaging. CNN+AL+MRF method is considered very popular research and has broad prospects in research and development of hyperspectral image classification. It has presented a new approach from previous research by combining deep and AL to enhance hyperspectral images, and this is what made it benefit from both methods. So, it has got the discriminating ability to distinguish it from DL. AL reduces labeling efforts and costs by strategically selecting the most

useful pixels for labeling, while CNNs enhance the capability of the model by learning complex spatial and spectral features in image classification, and finally MRF provides a spatial regularization step that creates improved labels between pixels contiguous, this model gave high results in spectral image classification and contributed effectively to hyperspectral image classification research.

However, its limitations are that it deals with raw spectral data, the raw data in HSI relies on the original that unprocessed, obtained from the hyperspectral sensor, which gives rise to several problems, including:

- CNN+AL uses full dimensions which increases training time. In addition, the Hughes phenomenon/curse of dimensionality leads to increased variability and redundancy of data.
- Risks of overfitting, as the model learns on noise and significant discounts to the training set for using raw data before the training process.
- Spectral variability is an important problem in hyperspectral imaging. As intra-class variability increases, a greater number of data is needed to be added in AL process. Therefore, there is a need for an efficient feature extraction that decreases the intra-class variances within the data.

As a result of the above, the model will suffer from increased computational complexity, making it difficult to process and analyze data efficiently, requiring high training times and resources such as increased memory. Finally, our motivations in this thesis are explained below:

As mentioned earlier, the CNN+AL method trains on full-dimensional data, which increases significant computational efforts, so we suggest using dimensionality reduction. Data processing creates low-dimensional data and extracts useful features in training the model. It contributes to maintaining the most accurate bands, which requires little storage space and memory. Preprocessing techniques have a major role in improving the analysis and classification of different spectral data and understanding their features more fully.

In order to solve these problems, we use edge-preserving filtering technology that effectively reduces noise in the data, improves the visual quality of the image and enhances the performance of the subsequent stage, providing a compromise

between processing noise and unwanted changes while ensuring that important image features are visible.

Learning generally begins with a small number of data sets that are randomly selected from the selected data points. The initial step aims to create a valuable training set to train the model. The performance of the model depends largely on the quality of the initial training set, so pre-processing techniques are very important in the quality of the model.

The raw spectral data is processed using dimensionality reduction and edge-preserving filter techniques before the training process for the purposes of producing high-accuracy, low-dimensional data, free of noise and unwanted or not useful elements. This is reflected in improving the quality of training and analysis and enabling the model to perform highly efficient classification.

1.4 Outlines of the Thesis

This thesis is organized into five chapters. In the first chapter, we presented the thesis, then the basic concepts and literature, then the motivations for this thesis. The second chapter talks about hyperspectral image classification using CNN and AL and method CNN+AL and how it was improved using dimension reduction techniques and edge-preserving filters. The third chapter is about the data that was used, which is three data (Indian Pines, Pavia University, and Salinas) each of which contains different categories and bands, and our proposed techniques were applied to each of them. The fourth chapter reviewed the results of the dimension reduction techniques and the five filter techniques that were applied to three collected data, illustrated with tables and pictures that prove that the framework that we improved gave better results than the original. Chapter Five is a summary of what has been achieved in our research regarding classification of hyperspectral images and improving their performance.

2
HYPERSPECTRAL IMAGE CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORKS AND ACTIVE LEARNING

Machine learning is a subfield of artificial intelligence (AI) that focuses on the automatic discovery of concealed patterns in data without human intervention. Training data, also known as sample data, is an essential prerequisite for developing a machine-learning model. After being trained using the training data, the model will have the capability to produce prognosis or decisions using new, previously unseen data, referred to as testing data. The fundamental concept underlying machine learning is that computers are trained to autonomously acquire knowledge and make judgments by leveraging previously observed data (training data) rather than relying on explicit programming to do a specific task. There are two main categories of machine learning algorithms: supervised methods and unsupervised methods [36]. During unsupervised procedures, the model does not receive labels during training. However, supervised methods involve including data points along with their corresponding labels. Unsupervised learning techniques used with unlabeled data can assist in finding new data clusters [37], enabling new categorizations when labeling but their applicability is more constraint. The supervised technique performs training on the labeled dataset in each cycle of AL learning before conducting semi-supervised training across all datasets. AL can utilize to lower the cost of sample annotation while maintaining the robust learning capabilities of DL.

Supervised statistical algorithms require labeled training data to estimate parameters accurately. Acquiring labeled data can be costly and time-consuming. However, the significant number of dimensions in hyperspectral data poses a challenge in developing classifiers with only a limited number of labeled data points. [38], whereas labeled data is optional. While labeled data can be expensive and time-consuming to acquire, the extremely high dimensionality makes it challenging to create classifiers with just a few labeled data points at hyperspectral data can choose one or more samples at random from the unlabeled pool U in the initial state, submit this sample to the Oracle query label to create the labeled dataset

L, and then train the model using supervised learning on L. Machine learning has two subfields the AL and DL.

DL is the model needs many data to optimize a lot of parameters if it wants to learn how to extract high-quality features. An innovative method to produce features linked to deep structures is provided by the DL-based approach. Another name for DL is representation learning. It realizes the automatic extraction of data features and has its roots in the research of artificial neural networks. Due to its complicated structure, DL has great learning capabilities, but this also means that it needs a lot of labeled samples to do the necessary training. DL models inspired from the organization of the human brain to create appropriate models, due to these models can possess the ability to learn and make decisions that somewhat resemble human intellect.

Deep learning-based strategy, which currently rules the HSI classification community, achieves higher HSI performance when compared to non-deep learning-based methods. Deep network training requires many training samples, which can be costly. Additionally, appropriately labeling a sizable volume of HSI data by humans remains very challenging. For HSI classification approaches, the non-deep learning-based approach has been used for decades. While AL has the potential to significantly cut labeling costs, DL offers great learning power in the context of high-dimensional data processing and automatic feature extraction.

AL attempts to annotate the fewest number of samples while maximizing a model's performance increase. AL (also called “query learning" or sometimes “optimal experimental design” in the statistics literature) is a subfield of machine learning and, more generally artificial intelligence. The key hypothesis is that the learning algorithm will perform better with less training if it is given the freedom to select the facts from which it learns—to be "curious," if you will. AL is an iterative process that uses a set of unlabeled samples to identify the most informative ones.

Classification method that relies on CNNs and combined it with an AL strategy to get higher accuracy with fewer data. The method can train a deep network faster and with fewer samples compared to classic supervised learning methods. Researchers originally discovered and utilized the advantage of using unlabeled data to accelerate the classification process in the context of remote sensing data.

AL focuses on the communication between the classifier and the user. The pixels with the most uncertain classification results are those that the model returns to the user. The AL methodology has proven effective when applied to massive datasets requiring precise example selection. Stabilizing Prediction (SP) leverages the stable predictive capabilities inherent in a well-trained model. The pre-divided stop set serves to eliminate additional labeling costs, which aligns with the principles of AL as it doesn't necessitate specific labeling information.

While originally proposed for AL, SP is also applicable to DeepAL. DeepAL approaches are currently the subject of research directions that are largely concerned with task-independent models, training method optimization, and AL selection strategy enhancement.

AL strategically selects the most informative data points to label, allowing CNNs to learn more effectively from limited labeled samples. This synergy enhances the classification accuracy of hyperspectral images while minimizing the annotation effort. The combined approach helps CNNs make more informed decisions about which data to learn from, thus maximizing the model's performance with minimal labeled data [39].

In numerous methods for processing remote sensing images, AL algorithms are gaining popularity, and this algorithm's iterative sampling techniques involve feeding a classifier model new labeled pixels that correspond to the ones that will increase the model's performance on a regular basis. Active learning (AL) combined with CNNs in HSI is a burgeoning area aiming to optimize the training process for these models [40].

Hyperspectral imaging (HSI) enables precise object recognition by capturing and analyzing broad spectral information [2]. However, managing such large data effectively poses challenges and the “curse of dimensionality” becomes a concern when dealing with data volumes in the band of several hundreds, even with relatively small spatial dimensions and multiple spectral bands, the curse arises from the inconsistency between feature dimensions and available data samples, necessitating the use of efficient feature extraction and dimensionality reduction techniques.

For feature extraction (FE) and DR, researchers use average band and PCA as a dimensionality reduction approach in hyperspectral imaging (HSI). These techniques are used with edge-preserving filters as a pre-processing step to preserve spatial information, which is important in improving HSI classification. Not only does it enable greater complexity in the training phase, but it also provides exceptional classification accuracy while maintaining its sharpness and stability.

During the process of smoothing and trimming, edge-preserving filters preserve the information on the edge. It not only enhances the complexity during training but also supports great classification accuracy, retaining sharpness and stability. During the process of smoothing and trimming, edge-preserving filters preserve the information on the edge.

2.1 CNN+AL+MRF Framework

The CNN + AL + MRF [26] method is a noteworthy and influential contribution to hyperspectral image classification. The popularity of this method is confirmed by a highly cited of papers on hyperspectral image classification. Its innovative combination of deep and AL and the nature of its feeding into convolutional neural network training make it a high-impact approach that is widely recognized within the research community.

It involves training a CNN by using labeled data and refining it by selecting and merging information pixels effectively and repeatedly. Moreover, the incorporation of MRFs significantly enhances the classification results.

Active learning has gained paramount importance in image processing, particularly in the domain of hyperspectral image (HSI) classification, witnessing notable advancements.

In general, Figure 2 illustrate a visual representation of the active deep learning framework, highlighting the mechanism of the iterative process of selecting samples and training the model [26]. With a limited number of labeled pixels, the convolutional neural network (CNN) is trained, and the more informative pixels are selected among them to be labeled. The previous steps are repeated to label them and combine them with the training set.

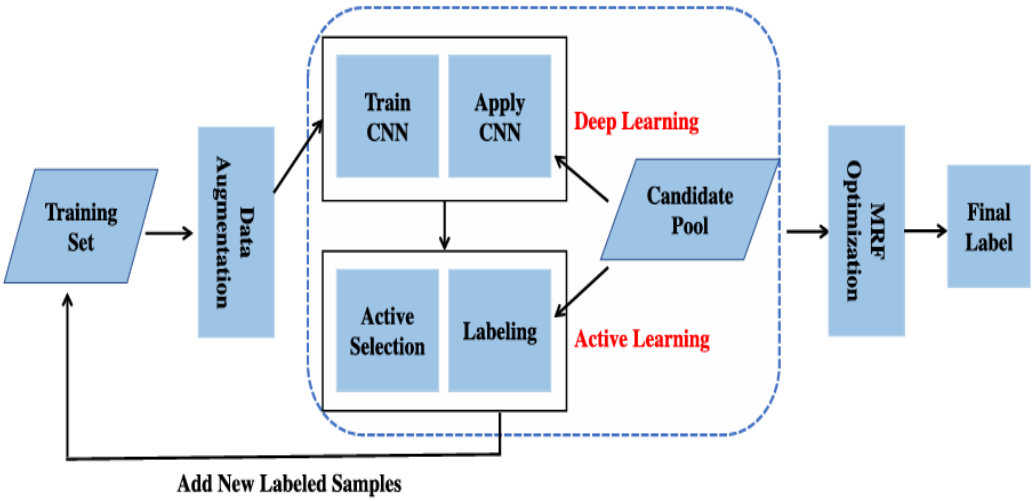


Figure 2.1 Active Deep Learning Framework [26]

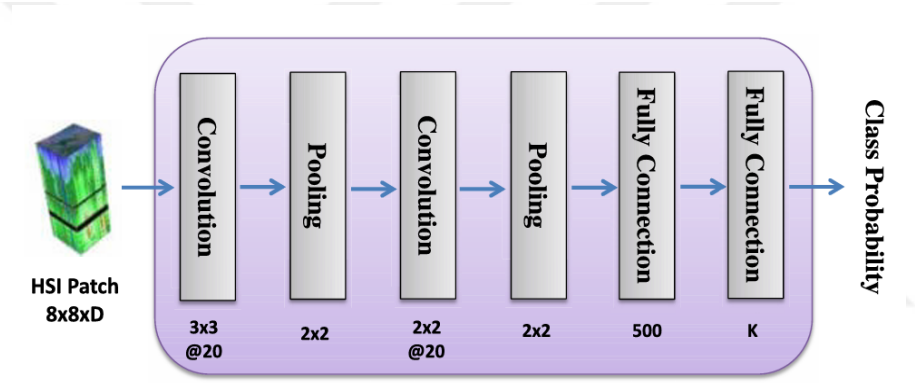


Figure 2.2 CNN Architecture [26]

CNN architecture in [26] contains a single input layer with two pairs of convolution and max pooling, two fully connected layers, and a SoftMax output layer. The experimental parameters that will be fed to the training CNN are prepared, such as the batch size and learning rate, and then fed to the training CNN. The input patch size is 8 × 8 × D, where D represents the number of HSI spectral bands. The first convolutional layer receives the input patch, which contains 20 filters of size 3×3, while the second convolutional layer contains 20 filters of size 2×2. between each convolutional layer, there is a max pooling layer with a kernel size of 2 x 2 and a pitch of 2 pixels [26]. At the beginning of each round (except the first round), set the CNN parameters to the values obtained in the final iteration of the previous round. These parameters serve as a starting point for training the current round and are subsequently adjusted based on the current data set, and then a MRF [40]

enforces spatial consistency between adjacent pixels to mitigate classification noise, ensuring that adjacent pixels potentially belonging to the same category have similar or diverse labels [40].

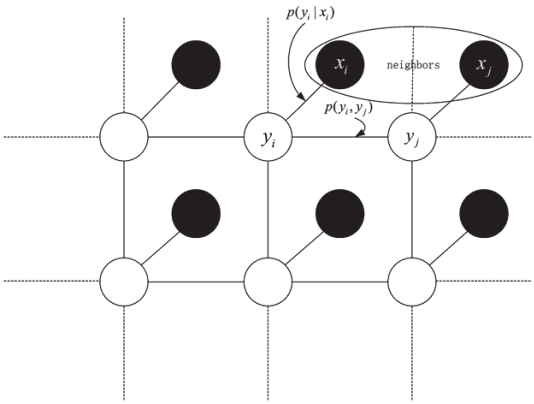


Figure 2.3 An Example of Markov Random Field [40]

To represent the spatial correlations between the resulting pixels, MRF is used. It represents the data as a graph with nodes and potentials, as shown in figure 2.3. It saves energy using energy function and smoothness requirements and captures spatial coherence using heuristics by taking spatial dependencies along with spectral information.

Therefore, it is applied as a post-processing step that utilizes contextual information to achieve greater representation, consistent with the land cover classes, the result is an improved classifier map with comprehensive quality, combining AL and DL and exploiting the benefits of both approaches led to improved model performance, opening broad horizons for conducting much research within this effective approach to hyperspectral image classification.

2.2 Improved CNN+AL+MRF Framework

CNN+AL+MRF [26] faces obstacles because it uses raw data and does not pre-process the pixels before the training process. Therefore, the approach suffers from several disadvantages. These include the curse of dimensionality, which leads to computational complexity that may hinder algorithm performance. Additionally, visualizing high-dimensional data without reducing dimensions is complex, and there is difficulty in selecting features. Furthermore, the model looks at all features,

making it impossible to distinguish useful ones. Moreover, the approach does not utilize any filtering technique to preserve spatial details and improve extraction [41]. If the features are noisy or redundant, this may increase the risk of misclassification, generalization, visualization difficulties, and increased noise. Additionally, not using any DR and filtering technique [42] can further exacerbate these issues.

The lack of techniques that act as pre-processing steps before training has a major role in consuming many resources and great computational accuracy, which negatively affects image analysis and processing tasks. We therefore seek to improve the performance of this model and remove these shortcomings by using dimensionality reduction and edge-preserving filter techniques as pre-processing before the model training process.

We previously touched on some of the obstacles facing CNN+AL+MRF [26], and therefore we seek to improve the performance of this model and remove these disadvantages by using dimension reduction techniques and edge-preserving filter techniques as pre-processing before the model training process.

On three standard HSI datasets, the proposed framework exhibits superior classification performance compared to existing methods. Using AL techniques to classify hyperspectral datasets may overcome the challenges posed by multispectral training and potential errors. The use of dimensionality reduction techniques is the dominant strategic approach to address this problem [43–46]. The DR methods used here reduce the dimensionality of the input data by projecting it onto low-dimensional subspaces, and this leads to the removal of redundant information within spectral bands [47]. We applied a DR strategy popular in the field of remote sensing image classification before acquiring the spatial data of the selected pixel. Five edge-preserving smoothing filtering techniques, along with dimensionality reduction techniques, were applied to the two main parts of the processing pipeline: averaging band (AVG) and PCA.

These techniques are used for filtering. Smoothing, also known as edge preservation, is an artistic method. Image processing technology that reduces noise or texture while maintaining the clarity of sharp edges. These are the five filters used in our proposed methodology:

- Weighted least square (WLS)
- Joint-Histogram Weighted Median Filter (JointWMF)
- Fast global image smoother (FGS)
- Bilateral Filter (BF)
- Static/Dynamic (SD)

Tests were conducted to evaluate the effect of combining DR and edge preservation filters on enhancing resolution with five filters that use smoothing techniques for image processing, with the specific goal of reducing noise while preserving edge integrity. The primary goal of this approach is to achieve a harmonious balance between reducing spectral dimensions and preserving edge by reducing noise for example, through the use of advanced image processing techniques, thus obtaining excellent classification performance, and contributing to subsequent studies for the purpose of obtaining optimal and problem-free classification of spectral images. In this study conducted a to evaluate the effectiveness of with and without Markov Random Fields (MRF) and compared processing times to evaluate computational efficiency. The evaluation is done through the values of Overall Accuracy (OA), Average Accuracy (AA), for PCA, and AVG or without it in conjunction of the five filtering methods. In addition, we considered execution times to evaluate the impact of DR and edge-preserving filters on classification efficiency. In addition to employing a DA strategy which involves translating, flipping, rotating, and cropping, to create more samples in order to mitigate the problem of inadequate labeled samples. CNN requires retraining for more than once according of mechanism is designed which specific at each hyper data.

The method incorporates the Active Learning (AL) strategy, which selects the most informative and uncertain samples by querying for labeling into the graph-based CNN model.

This results in the development of a novel supervised HIS classification method when merge the DR and EPF techniques. Edge-preserving filters are essential for enhancing the overall quality of images by removing undesired artifacts and maintaining the sharpness and clarity of object borders, textures, and structures as a pre-processing technique that have a crucial role in improving the quality of image data, leading to more precise and dependable outcomes in later classification tasks.

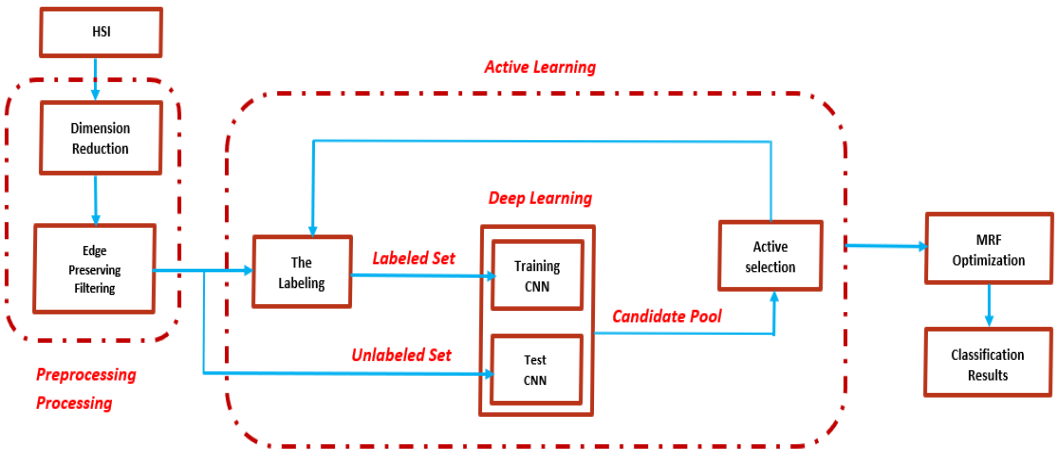


Figure 2.4. Enhancing CNN+AL+MRF Active Deep Learning

2.2.1 Dimension Reduction Techniques

Dimension reduction strategies are essential for managing the inherent high dimensionality of hyperspectral data. Dimensionality reduction is commonly conducted in the steps of pre-processing to minimize the interconnection that connects among bands and the spectral dimension of the input HSI. Executing dimensionality reduction and DA are two categories of ways to decrease the influence of these congestions. DR technique is used to lower the spectral band of the input hyperspectral image (HSI). The reduced HSI is then coupled with our underlying feature vector to generate a new set of spatial features. The process of producing this final spatial feature vector is described, as well as its use in a classification challenge. These strategies improve model performance by prioritizing relevant dimensions of the data while removing fewer important dimensions. The combination of a small sample size and many bands offers advantages such as simplicity, low memory usage, short training times, prevention of the curse of dimensionality problems, and reduced complexity of HS data. The increase in the number of dimensions, known as the curse of dimensionality, poses challenges in analysing hyperspectral data. consequently, DR is performed as a pre-processing stage previous classification in our proposal.

2.2.2 Average Band Technique (AVG)

It includes the averaging technique (AVG), which is technique used for pre-processing hyperspectral data, which is done by combining adjacent spectral bands

and calculating their average values [47]. This process effectively reduces the overall number of spectral bands in the dataset and creates a single band that represents the result of the average spectral response across the entire spectrum and thus simplifies its structure [47]. This simplification makes the subsequent analysis or classification generally more computationally efficient [48]. The experiment is carried out to demonstrate that the strategy based on averaging can eliminate image noise and the corresponding common band is that obtained by averaging these bands. Image noise or noise in individual bands can be effectively reduced by averaging adjacent bands, providing a smoother representation of the overall spectral information [48]. The rationale for this is that hyperspectral images have different bands and appear with varying degrees of noise, with only a few bands appearing with very high levels of noise. A good way to deal with noisy data in this case is to average the nearby bands [49], and consequently aid analysis and interpretation in certain genres by emphasizing common image features.

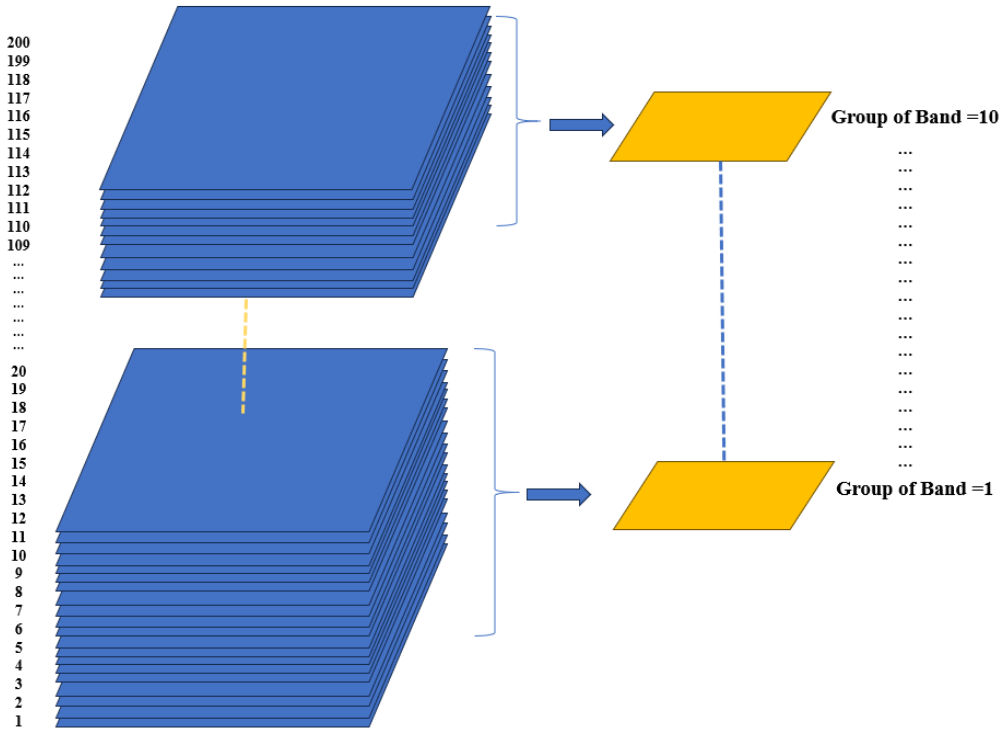


Figure 2.5 Illustrate Example of AVG Process

AVG is particularly useful in hyperspectral imaging, where datasets often contain many bands, some of which may be noisy or less informative. By averaging adjacent bands, AVG smoothed out differences and reduces the impact of noise, making the data set more manageable for analysis. Computational efficiency

improves with the help of dimensionality reduction, because of reducing the computational burden when processing data. However, while AVG simplifies the data set, there is a trade-off in that it may result in the loss of fine spectral detail or subtle differences that may sometimes be important for certain applications. Therefore, the effectiveness of AVG depends largely on the characteristics of the data set and the specific goals of the analysis or classification tasks. But in general, AVG is a valuable pre-processing step to enhance data clarity and manageability, especially when the focus of denoising and simplification is critical for the final analysis.

AVG method effectively reduces the influence of noisy or less informative bands, which are often present in hyperspectral data [50], once the data set is reduced, the SP can be split to a series of consecutive groups corresponding to the number of dimensions provided. For example, if there are 200 bands and the dimension is reduced to 10, the data is condensed into 10 dimensions, the bands will be divided into ten groups, each group comprising a subset of the original bands as in figure 2.5. For example, the first group might contain bands 1 to 20, the second group bands 21 to 40, and so on. The spectral values of these 10 bands are averaged to create one representative band for each group that captures the collective spectral characteristics of that group.

Let the hyperspectral image as HSI and HSIn the n denotes of nth band of image.

The dimension reduction of average band [50] is applied as a:

$$\mathrm{HSI}_{i}^{R} = \frac{B}{RB} \times \sum_{n=1+(i-1)\frac{RB}{B}}^{i(\frac{RB}{B})} HSI_n \qquad i = 1,2,..,B \tag{2.1}$$

In equation (2.1) [50], HSIR and RB represent the reduced HSI and total number of bands in HIS, B the number of reduced bands. The representative band obtained from this equation, which calculates the average spectral values in each group, represents a condensed summary of the information contained within each subgroup [50]. This helps mitigate the effect of noise of the individual bands by summing these bands, thus reducing computational efficiency, which reduces the time and cost of processing the data entering the model for training.

2.2.3 Principal Component Analysis (PCA)

PCA is a widely used DR technique, first developed by Karl Pearson in 1901 [51] and since then it has been widely used in many academic disciplines. The aim of PCA technique is to determine the most accurate representation of the data by exploring a lower-dimensional subspace consisting of unrelated linear combinations of the original variables, known as principal components. PCA assigns the input data to the dimensions that show the greatest amount of variance, thus retaining the largest variances in the data.

Principal Component Analysis was used on a 3D dataset to examine and reduce the dimensionality of the data into a certain number of principal components and is considered our second DR methods. PCA is a usually applied technique that seeks to capture the most important differences in a data set.

Figure 2.6 shows how it works in reducing the dimensions of the data. In the first steps, it collects the observed data that is linearly related (if it is not linearly related, it will fail to capture the variance with the smallest possible number of components) and centering them by subtracting the mean and dividing it by the standard deviation for each variable.

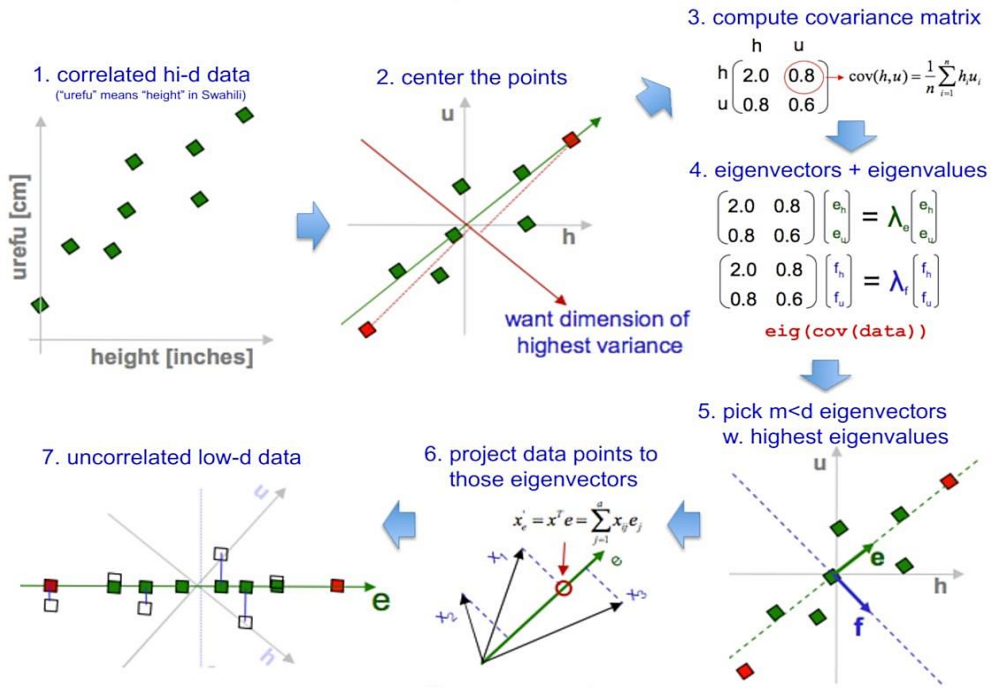


Figure 2.6 Principal Component Analysis [52]

Next step is to calculate the covariance matrix for the standardized data set from which to find out how the variables change in relation to each other. Then calculate the eigenvalue and eigenvector: where the eigenvectors represent the principal components and the eigenvalues represent the amount of variance captured by each principal component, and then sort them in descending order. Then a projection phase involves projecting the original data onto the low-dimensional space formed by the selected principal components.

PCA effectively reduces the number of dimensions in the data by transforming the original spectral bands into a new set of variables (principal components) that are not related to each other. It is a type of condensing spectral information into a smaller set of principal components while preserving the most important information [52].

PCA and convolution layers are applied to enhance the ability to extract the most important features and increase classification accuracy by highlighting and emphasizing spectral bands that contain important information and reducing the effects of less relevant dimensions. It provides insight into the most influential features within data set. So PCA helps visualize high-dimensional data by projecting it onto lower-dimensional spaces, usually 2D or 3D. This visualization allows a better understanding of the data structure and relationships between different spectral bands [52]. The proposal was made based on the core ranking principles of PCA and the significant influence of convolution, as stated in reference [51]. Conclude that PCA is a mechanics applied to lower the dimensions of a data set by transforming it from a higher dimensional space to a lower dimensional space. Due to a high dimensionality of HSI, PCA can be used to reduce the quantity of these spectral bands. Adjacent bands of hyperspectral images have comparable spectral properties and thus convey similar information about an object. PCA uses mathematical and statistical features to evaluate correlations between domains and thus reduce dimensionality.

Principal component analysis is mainly according to eigenvalue decomposition of covariance matrices. Overall, PCA serves as a powerful tool to reduce the complexity of hyperspectral data while preserving its essential properties, facilitating efficient analysis and interpretation.

2.2.4 Data Augmentation

Data augmentation is a widely applied approach that enhances the ability of deep neural networks to generalize. It has a crucial function in RS situations where there is a scarcity of accurate ground truth data and obtaining additional samples is expensive or unfeasible [53].

Manual annotation of image data in hyperspectral images is a prevalent issue due to its difficulty, high cost, and susceptibility to human bias. Both data augmentation and classifier fusion are viable strategies for addressing the limitations of small sample sizes [54]. The effectiveness of data augmentation in increasing training sets has been well-established. Created the data mixture model by combining the training samples and their corresponding labels [54].

Employ the data augmentation (DA) approach [55] which involves translating, flipping, rotating, and cropping, to produce more samples and mitigate the problem of limited labeled data.

2.2.5 Edge-Preserving Filters (EPF)

The filter technique provides optimal preservation of edges while filtering images. By using weighted averages of local pixel windows obtained through clustering, the filter technique reduces intensity shift abnormalities and effectively deals with halo artifacts commonly found in other filters [56]. Due to its customizable smoothing control, low computational demand, and proven effectiveness in many applications, it is a viable option for preserving edges, enhancing images, and performing numerous other imaging tasks [56].

Edge-preserving image filtering is the procedure for reducing the impact of unimportant features while maintaining the clarity of important edges. In recent decades, there has been a significant focus on edge-preserving filters, which are essential used to a wide band of electronic imaging and imaging tasks. These applications include denoising [57,58], image enhancing [59] flash/non-flash fusion [60], HDR tone mapping [61] dehazing [62], stereo matching [63,64], and shadow removal [65-67].

One can categorize EPF as a worldwide or as a regional, the majority of optimization issues with globally filtered involve both data and regulation terms [56]. In general, the data term refers to the discrepancy between the input and output

images, and the regulation term refers to the penalty for lack of smoothness in the gradient domain. The optimization issues often have a common data term but vary in terms of the regulation term

The deep learning-based filters demonstrate practical efficiency; nonetheless, they encounter challenges in adapting to varying degrees of smoothing. In addition, the deep learning-based filters necessitate extensive training data and time-intensive training procedures. Typically, edge-preserving filtering is regarded as a type of post-processing. However, we utilize edge-preserving filters as a sort of pre-processing [68]. The primary objective of this process is to diminish noise and enhance the visual appearance of an image while simultaneously maintaining the integrity of its intricate edges, details, and characteristics [69].

These filters are particularly advantageous when aiming to enhance image quality by eliminating noise or minimizing undesired artifacts while simultaneously retaining the sharpness and clarity of object borders, textures, and structures. One of the most renowned filters for preserving edges is:

1. *Weighted Least Square (WLS)*

Farbman et al. [70] developed a method that uses image guidance and optimization to create a filter that preserves edges and is suitable for gradually reducing image details and manipulating details at several scales [70]. Weighted Least Squares (WLS) filter is a recognized method for preserving edges while enhancing details and performing HDR tone mapping. It achieves this by computing weights from a guidance image, creating a weighted least square model, and solving a large linear system to obtain the filtering output [70]. Although highly effective in suppressing halo artifacts, the computational expense associated with WLS remains a significant drawback. Moreover, the existing deep learning-based adaptations of the WLS framework are predominantly supervised learning-based, which [71], prioritize efficiency over the preservation of image quality. These adaptations often boost productivity but might compromise the overall quality of the output.

2. *Joint-Histogram Weighted Median (JointWMF) Filter*

The image histogram illustrates the pixel distribution within certain regions, typically expressed as bins, displaying a normalized probability distribution. In the

context of local mode filtering [72], the creation of a relaxed histogram involves describing individual pixels using Gaussian distributions.

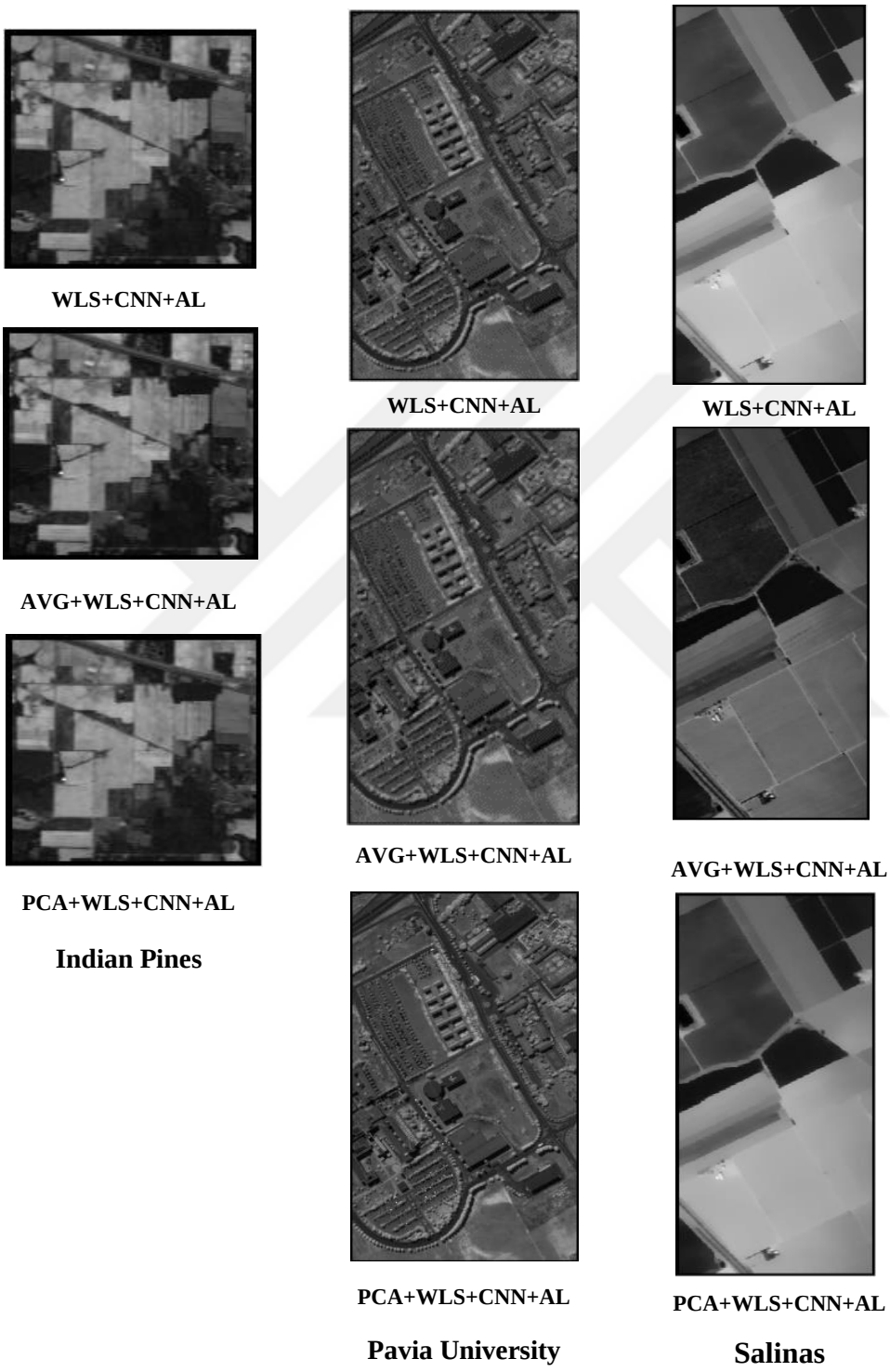


Figure 2.7 Result of WLS on three datasets with and without DR

This approach seeks to represent each pixel by using a Gaussian distribution within the histogram. This results in a more accurate depiction of the pixel values in the image.[73] A weighted median filter (WMF) is implemented using a joint histogram.

Construct the histogram using a weight based on the similarity measure between the reference pixel and its nearby pixels. The final solution is obtained by searching for the global mode on the histogram. Joint filtering involves calculating the weight using a signal that is different from the signal being filtered. It is possible to expand the histogram into a weighted filtering process. In order to minimize norms, the filtered value is used as a solution, enhancing its resilience to data outliers compared to norm minimization [73].

A local mode filtering technique was introduced by Weijer and Boomgaard [74] to specifically preserve edges and details in an image while effectively reducing noise. This method is based on histograms and operates in a nonlinear manner [74]. To calculate the histogram, apply a dataset which include of neighboring pixels of a reference pixel [74]. It then recursively searches for a local mode on the histogram, starting with the intensity of the pixel as the initial answer. The local mode filtering iteratively converges to the nearest local mode on the histogram based on the given pixel and its related histogram [75]. demonstrated the equivalence between local mode filtering and bi-lateral filtering [75]. The bi-lateral filtering approach, as proven in, offers an optimal solution by minimizing the norm [75].

Weighted median filtering is a filtering technique that applies a weighted median filter on both the input image and its corresponding histogram. It is beneficial for maintaining the integrity of image borders and structures. It is particularly useful when dealing with extraneous noise affecting a large portion of the data matrix, as it helps eliminate this noise effectively, ensuring a more widespread and accurate data representation [76].

3. Fast Global Image Smoother (FGS) Filter

Fast Global Smoother is a highly efficient edge smoothing method depend on the WLS formulation. The first technique used, but it’s worth noting that it is much fast about 30 times faster than the WLS-depend on filter in [77].

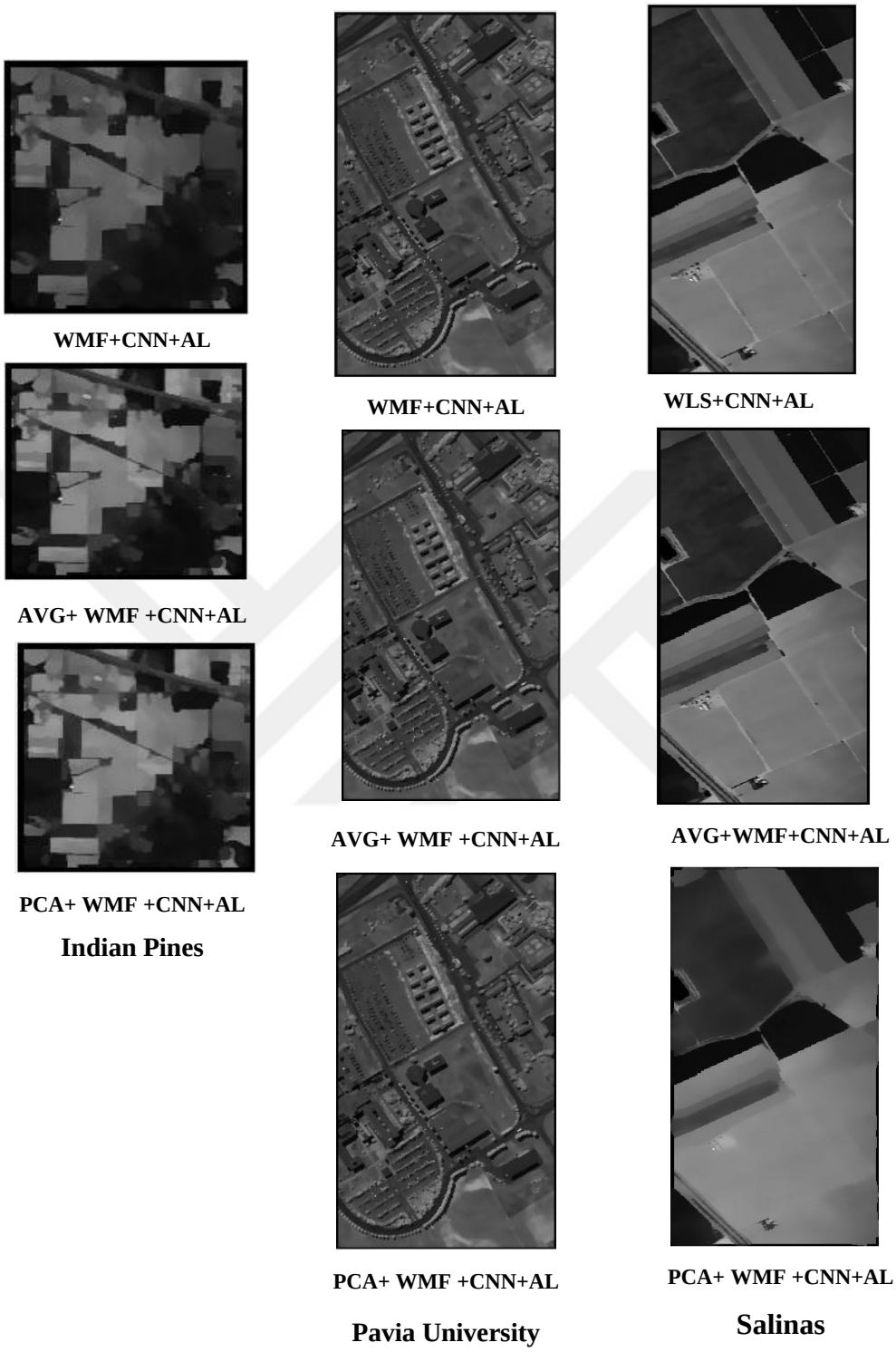


Figure 2.8 Result of WMF on three datasets with and without DR

FGS is a method that improves image smoothing by solving a linear problem using a non-homogeneous Laplacian matrix [78]. This approach uses a two-step procedure, which includes a subsequent corrective phase after the initial smoothing process, resulting in a more advanced form of image smoothing overall [78]. This

filter reduces noise and achieves a smoother look in images quickly and effectively. Unlike local techniques that only consider nearby pixels within a given window, FGS works on a global scale, including the entire image. Process [79] FGS adjusts the value of each pixel by considering the entire image by calculating a weighted average.

The FGS algorithm was specially developed to preserve sharp edges and complex details while efficiently reducing noise. It provides superior results in terms of quality and computational efficiency compared to some traditional techniques [79]. This world-class technology allows fast image processing while preserving basic image attributes. By using FGS instead of Gaussian filter, it is possible to prevent unwanted blur effect and preserve important features such as edges and outlines [79].

The filters chosen are concerned with focusing on both comprehensiveness and importance, so the FGS approach emphasizes computing economy while maintaining excellent edge preservation, considering the entire image rather than just local pixel neighborhoods.

By using weighted averages calculated across the entire image, FGS adjusts pixel values to achieve smoother images while preserving essential details and edges. This method is particularly effective in reducing noise without causing unwanted effects, ensuring fast and high-quality image processing [78].

4. *Bilateral Filter (BF)*

Bilateral filtering (BF) is well known for its effectiveness in retaining edges during image smoothing in the field of image processing [80]. However, when using BF to reduce noise in an image, a fundamental difficulty is the ability to distinguish between image noise and complex features, which often leads to the retention of both noise and edges during the noise reduction process. It is crucial to maintain a balance between noise reduction and edge preservation during the filtering process.

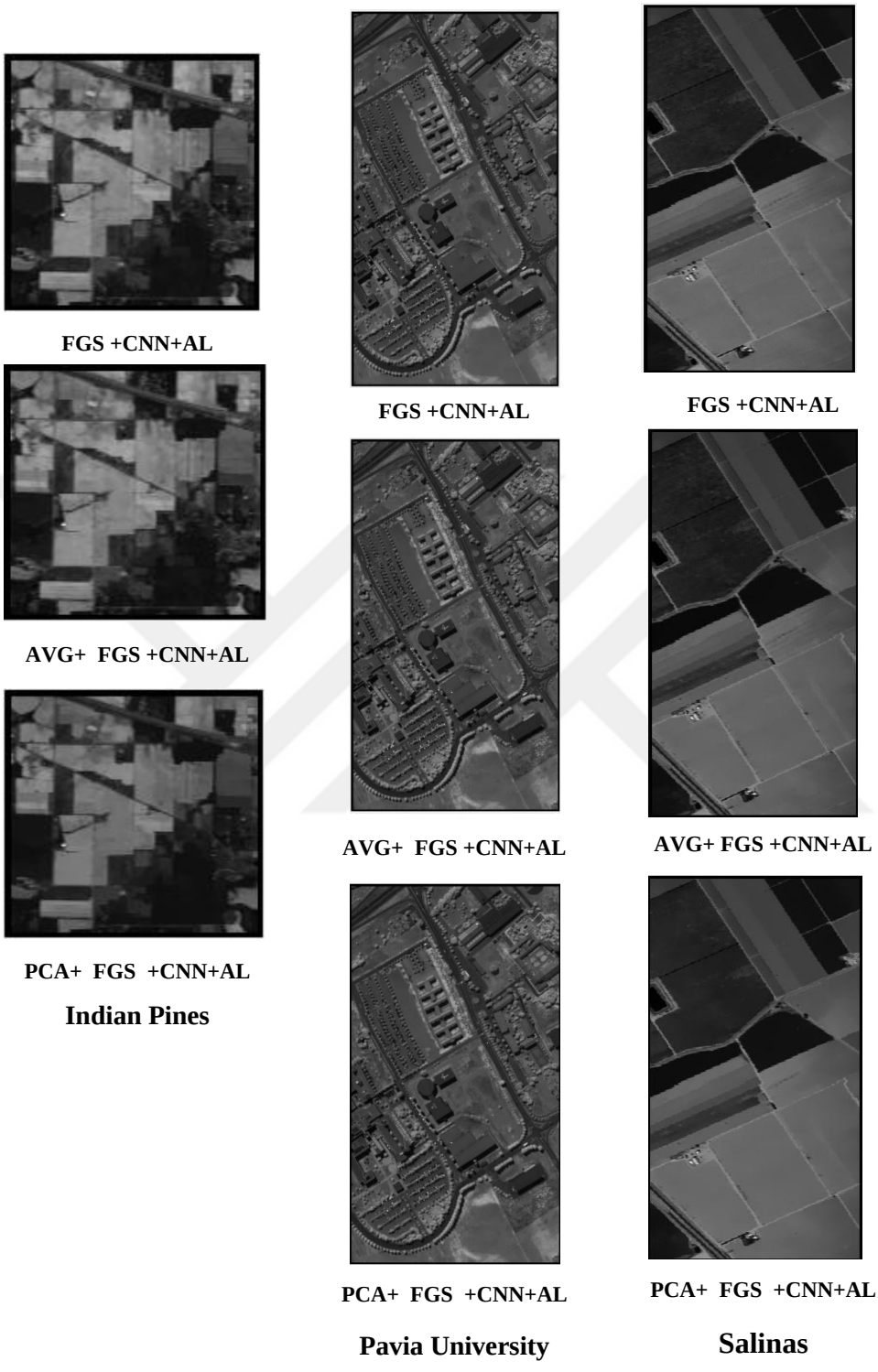


Figure 2.9 Result of FGS on three datasets with and without DR

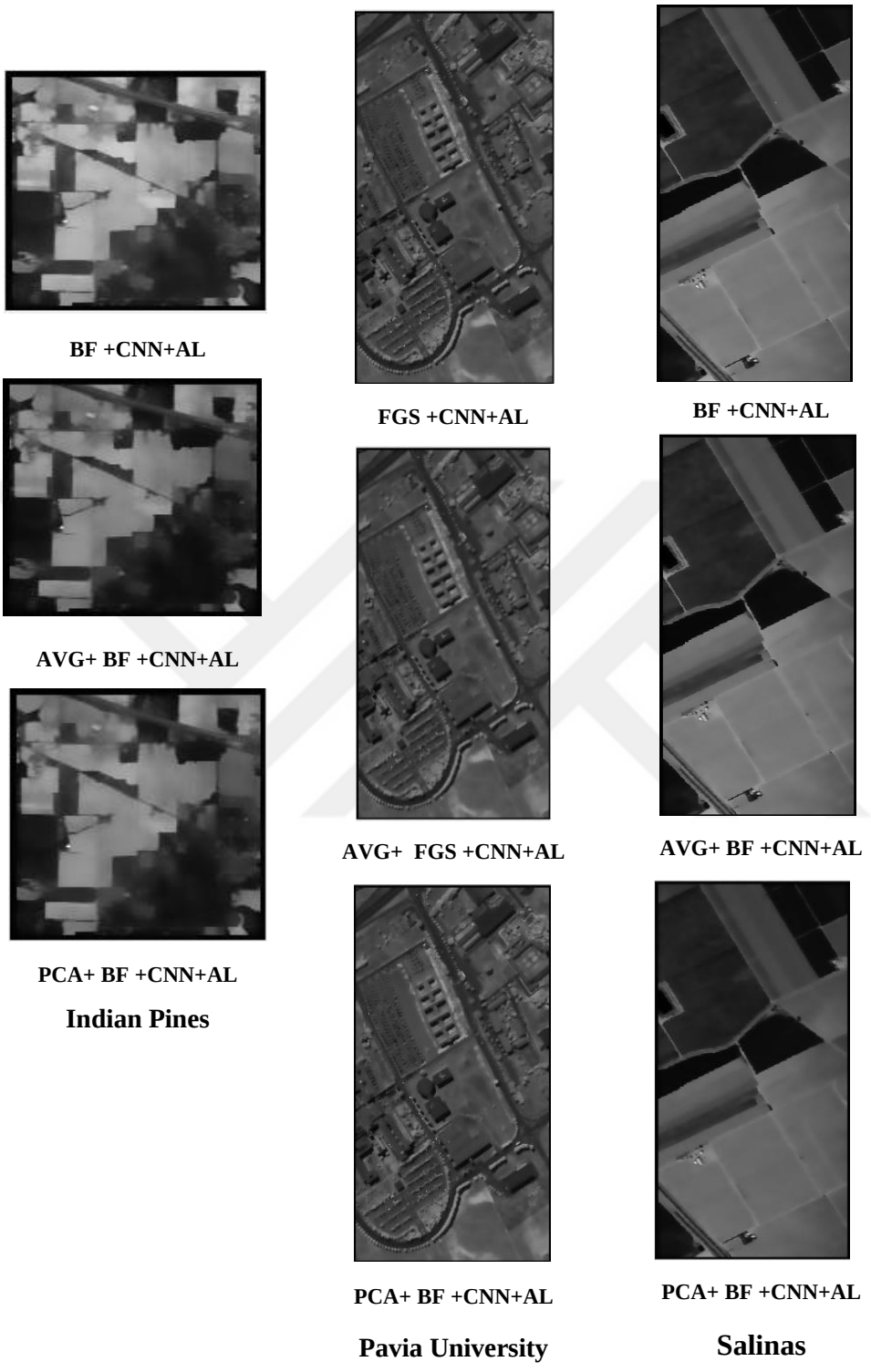


Figure 2.10 Result of BF on three datasets with and without DR

The fusion process is expedited and computationally uncomplicated because of its non-iterative nature and the availability of efficient implementation approaches [81]. The filter accomplishes the task of smoothing the image while preserving edges and minor details by considering both spatial and intensity similarities. Researchers and practitioners use local filters to smooth data by explicitly defining spatially changing kernels, such as the widely used bilateral filter (BF) [82-84]. A vector-based bilateral filtering method tailored for enhancing hyperspectral images.

This technique leverages decorrelation within hyperspectral data to simplify parameter determination, focusing solely on estimating noise variance per PCA band instead of the full covariance matrix. Application to a HYDICE (HS Digital Imagery Collection Experiment, which includes 210 spectral bands from 397 to 2496 nm with a ground resolution of two meters) [85] hyperspectral image demonstrates its efficacy by effectively removing noise while preserving edges, particularly in bands with poor signal-to-noise ratios. Comparative evaluations indicate superior performance over traditional scalar bilateral filtering on individual bands or other existing methods, both qualitatively and quantitatively [86,87].

5. Static/Dynamic (SD) Filtering

Hyperspectral imaging utilizes a combination of static and dynamic techniques to process hyperspectral data, [88] a process known as static/dynamic (SD) filtering. Unchanging or predetermined filters are universally applied to the dataset in the static component, regardless of the content of the data. On the contrary, it involves the use of adaptive or changeable filters that adjust themselves according to the content or characteristics of the hyperspectral data on the dynamic side. This hybrid strategy aims to use both pre-set filters and adaptive algorithms to improve the quality, level of detail, and reduce noise in hyperspectral images [88].

Adaptive parameter adjustments responsive to input changes along with fixed parameters are integrated into each other. So, image processing usually uses it for tasks such as feature extraction and noise reduction. Its purpose is to leverage pre-existing knowledge with the ability to adapt to different characteristics of the input [89]. SD filtering uses both a static guidance image and a dynamic target image to iteratively enhance the quality of the target image, adjusting its parameters dynamically during the filtering process [90].

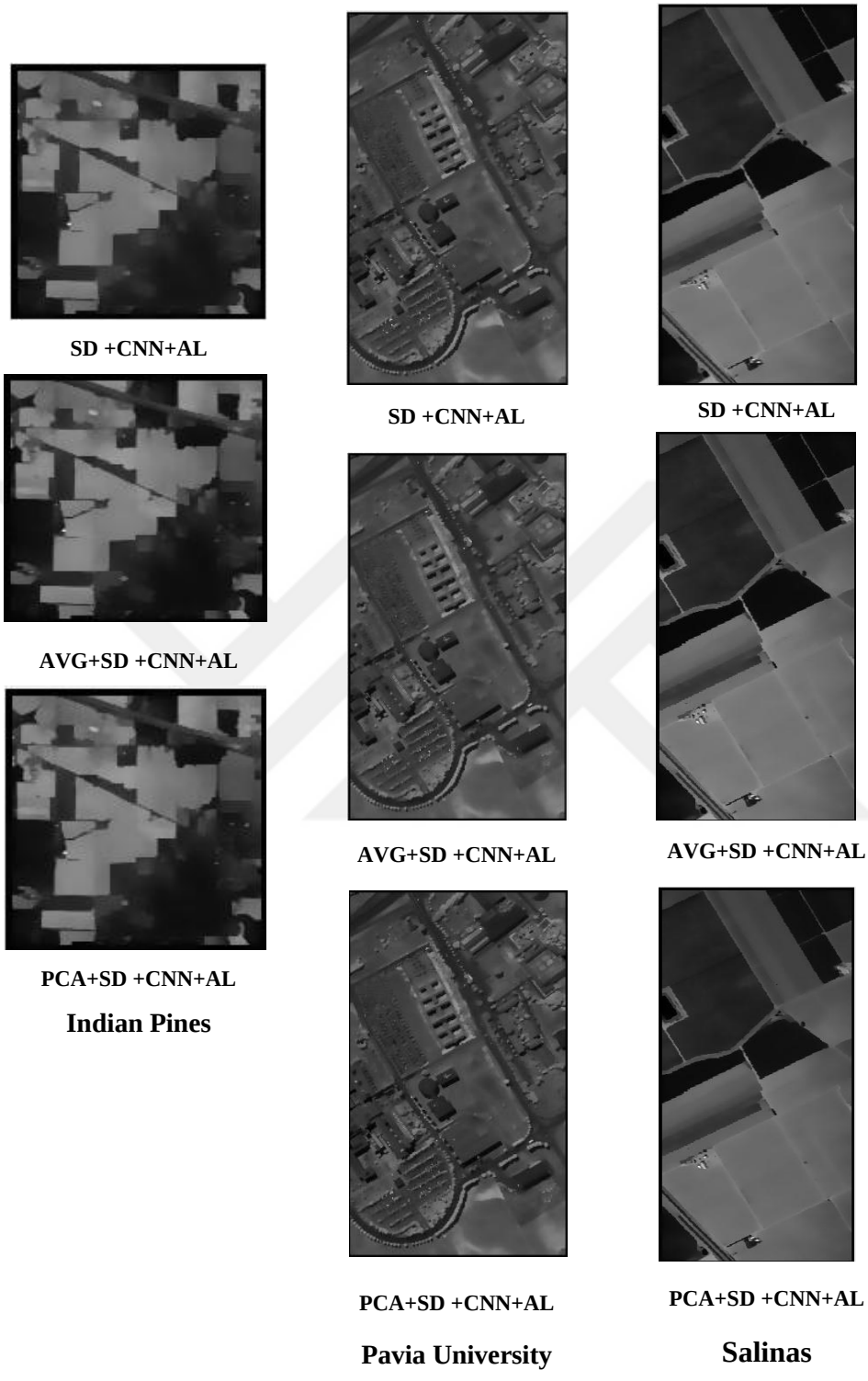


Figure 2.11 Result of SD on three datasets with and without DR

2.3 Algorithm Details

The active learning process begins after dimensionality reduction and use of edge-preserving filters. The combined use of these techniques gives synergy and enhances various aspects such as image analysis and machine learning tasks. Dimensionality reduction extracts essential features, reduces noise, and removes irrelevant data, which leads to increase the accuracy and effectiveness of the overall model. This joint application also contributes to enhancing image clarity by maintaining sharp edges and preserving important details, in addition, the joint approach improves computational efficiency by, simplifying operations and facilitating faster model training and, inference.

Data augmentation techniques are used, including translation, rotation, flipping, and cropping, to compensate for the restricted training data after dimension reduction and before filtering techniques, all of which are considered as pre-processing [91]. This model is repeated several times depending on the specific structure of each spectral feature, but it is similar in terms of the working mechanism. During the initial iteration of AL, we choose the training labels randomly. We used an eight-layer CNN architecture. After training on the labeled set, we test all pixels in the unlabeled set and actively select the most suspicious pixels using the best versus second best (BvSB) method [26], as in this define:

$$BvSB\ (z_i) = P_B\ (z_i) - P_{SB}\ (z_i) \tag{2.2}$$

where $P_B\ (z_i)$ denotes the best class membership probability of sample z_i, which can be computed as $P_B\ (z_i) = P_B\ (z_{i_k})$ where k = argmax $P(z_{i_l})$, and $P_{SB}\ (z_i)$ is the second-best class membership probability of sample zi, which can be computed in a similar way in equation (2.2) [26].

The BvSB metric is ideal for multi-class HSI classification, as it addresses performance issues due to small class probabilities for non-significant classes by measuring the probability difference between noisy classes [92].

During this stage, a low BvSB value of the sample indicates the presence of a noisy pixel. Once the pixels are labeled, they are included in the label set, and the iteration continues until the specified number of iterations is reached. The AL process

repeatedly includes the most ambiguous data until it achieves the specified number of iterations.

In the final stage, use MRF to correct misclassification in the spatial domain, following the approach described by Cao et al. (2020). Algorithm 1. Provides a comprehensive and detailed explanation of the technique, including step-by-step instructions.

Algorithm I. Algorithm of the Proposed Framework

Input: Number of training samples *D*, unlabeled pixel set *U*, maximum number of iteration *R*, number of actively chosen pixels in each round *B*, flag of use MRF, edge-preserving filter type and its parameters, dimensionality reduction method.

Initialization: set *r* to 0.

Main Loop:

While $r < R$ or terminating standards that were not met:

1. *Dimension Reduction*:

• Apply dimension reduction to *D*, reducing it to *DimR*.

2. *Data Augmentation*:

• Augment *D* to create a new training set *DA* using *DimR* and *D*.

3. *Edge-Preserving Filtering*:

• Apply an edge-preserving filter to the hyperspectral image data in *D* and *U*. This filter helps enhance image details, preserve sharp edges, and manage noise. You can choose filter parameters (e.g., bilateral filter, guided filter) to suit your needs.

4. *CNN Training/Fine-Tuning*:

• Train or fine-tune a Convolutional Neural Network (CNN) using the augmented training set *DA*.

5. *Active Pixel Selection*:

• Use the trained CNN to estimate class probabilities for pixels in the unlabeled candidate pool *U*.

• Actively choose the top *B* informative pixels from U based on class probabilities.

6. *Update Training Set*:

• Add patches corresponding to adding the chosen *B* pixels to the existing training set *D*, using it as the new training set for the following iteration.

7. *Increment Round*:

• Increment r by 1.

8. *MRF Utilization* (*Optional*):

• If Use MRF is false, go to Labeled.

End of Main Loop

To enhance classification efficiency through pre-processing, DR techniques such as AVG and PCA were employed. This approach is particularly advantageous when dealing with low sample numbers and diverse bands, offering simplicity, low memory usage, fast training times, prevention of the curse of dimensionality problem, and reduction in the complexity of hyperspectral data. Averaging band and PCA play pivotal roles in simplifying the data representation, making it more manageable for subsequent analysis. Averaging helps in reducing noise and smoothing out variations, therefore improving the quality of the dataset. To enhance the classification efficiency through pre-processing, DR techniques such as AVG and PCA were used.

On the other hand, PCA and AVG identify and retain the most important features and discard less relevant information. Integrating both dimensionality reduction techniques with edge-preserving filtering improves model performance by emphasizing the essential dimensions of the data while eliminating less important dimensions, refined, improved dataset, laying a strong foundation for subsequent classification.

The use of these pre-processing methods goes beyond enhancing model performance by addressing bottlenecks associated with hyperspectral data, ensuring efficient use of computational resources and alleviating issues related to limited samples. Thus, this approach not only improves classification accuracy, but also provides scientific and practical advantages in terms of resource efficiency and effectiveness.

Because the most important obstacle facing hyperspectral data is its dimensions, which constitutes a challenge in analysing it. Hence, dimensionality reduction is performed as a preliminary step before classification [93]. The AVG method for averaging sequential spectral bands is the first technique used to effectively reduce the dimensionality or influence of noisy or less distinct bands that are often present in hyperspectral data. Using this technique, the total amount of spectral bands is diminished through the process of averaging the band values.

The data set is first diminished to a lesser extent by dividing the spectral bands into an equal number of successive groups.

As for the second method, we used principal component analysis (PCA), which also reduces the dimensions of the data to a certain number of elementary components [94]. It is a commonly used method that works to capture the most important differences in the data and transforms these original spectral bands into a new set of variables that are not related to each other, replace the original spectral bands with the new set of spectral bands [94]. During all experiments, we used the number of reduced bands as 10 for PCA and AVG methods.

Then comes the role of the edge-preserving filter, which is an image processing technique whose purpose is to reduce noise and improve image clarity by preserving sharp edges and complex features. Simply put, it is a filter specifically created to improve image quality. Additionally, edge-preserving filtering methods improve differences between different classes while minimizing differences within the same class. This helps reduce spectral variation across all categories of the spectral image. It works with dimension reduction techniques as pre-processing.

Our approach combines DR techniques such as PCA and AVG with modern filters to preserve the edge and as much important information as possible while removing any extraneous elements. The process involves eliminating ineffective bands, enhancing the smoothness of these images without compromising its sharpness. Moreover, it preserves the complex features that are usually associated with boundaries between classes in the process of hyperspectral image classification, implement on a 3D dataset.

This approach is particularly suitable for scenarios with a limited amount of labeled data. More precisely, we use DR techniques such as PCA and AVG in conjunction with five distinct edge-preserving filter methods on a 3D dataset to achieve a satisfactory solution to hyperspectral image classification problems of which the dimensions are large. And the appearance of noise in the images. Furthermore, we enhance the accuracy of classification by incorporating MRF post-processing techniques in addition to the pre-treatment activities.

3
DATASETS

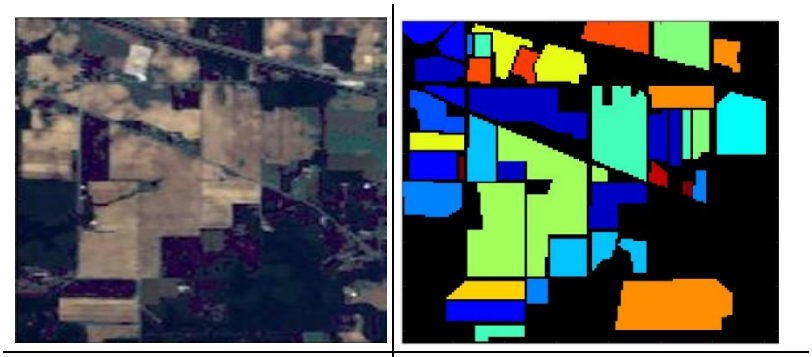
This study utilized three hyperspectral datasets obtained by remote sensing. Datasets are ubiquitous in images from hyperspectral literature. The datasets and their attributes are concisely outlined in Table 3.1.

Table 3.1 The dataset of hyperspectral in the experiments

#	HSI Dataset	Name of Sensor	No. Bands	No. Classes	Total of Samples
1	Indian Pines	AVIRIS	200	16	10249
2	Pavia University	ROSIS	103	9	42776
3	Salinas	AVIRIS	204	16	54129

3.1 Indian Pines

The Indian Pines dataset was captured by an Airborne Visible/Infrared Imaging Spectrometer Sensor (AVIRIS) above the Indian Pines test site located in northwest Indiana. The dataset consists of a grid measuring 145 by 145 pixels, with 220 spectral reflectance bands spanning from 0.4 to 2.5 micrometers.



Original RGB (A) [95] Ground truth (GT) (B)

Figure 3.1 Original Image (A) [95] and GT of IP dataset (B)

The Indian Pines habitat comprises an intricate combination of indigenous perennial vegetation, farmed regions, man-made structures, transportation

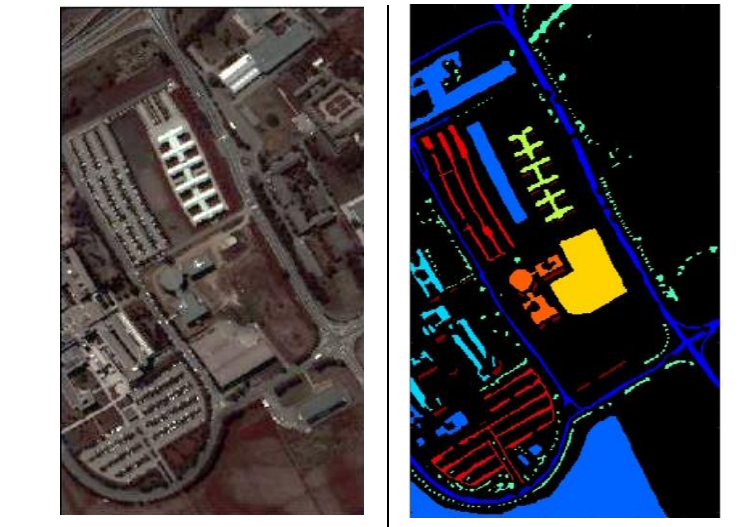
infrastructure, and residential zones. Due to the disparity in sample quantities, the categorization of 16 distinct classes presents a distinct challenge. In order to address this issue, we employ an innovative approach by randomly selecting 2% of the samples from each class for the purpose of training. Table 3.2 presents the quantity of training and test samples.

Table 3.2 Number of test and train samples (2%) in IP dataset

No	Classes	Train	Test	Total Samples
1	Alfalfa	1	45	46
2	Corn-notill	29	1399	1428
3	Corn-mintill	17	813	830
4	Corn	5	232	237
5	Grass-pasture	10	473	483
6	Grass-trees	15	715	730
7	Grass-pasture-mowed	1	27	28
8	Hay-windrowed	10	468	478
9	Oats	1	19	20
10	Soybean-notill	20	952	972
11	Soybean-mintill	50	2405	2455
12	Soybean-clean	12	581	593
13	Wheat	4	201	205
14	Woods	26	1239	1265
15	Buildings-Grasstrees	8	378	386
16	Stone-Steel-Towers	2	91	93
Total		211	10038	10249

3.2 Pavia University

The Pavia University dataset is a well utilized hyperspectral remote sensing dataset employed for classification tasks and image analysis. The data is acquired using an airborne hyperspectral sensor called the Imaging Spectrometer with Reflective Optics System Imaging Spectrometer (ROSIS) Sensor.



Original RGB [95] (A) | Ground truth (B)

Figure 3.2 Original Image (A) [95] and GT of PU dataset (B)

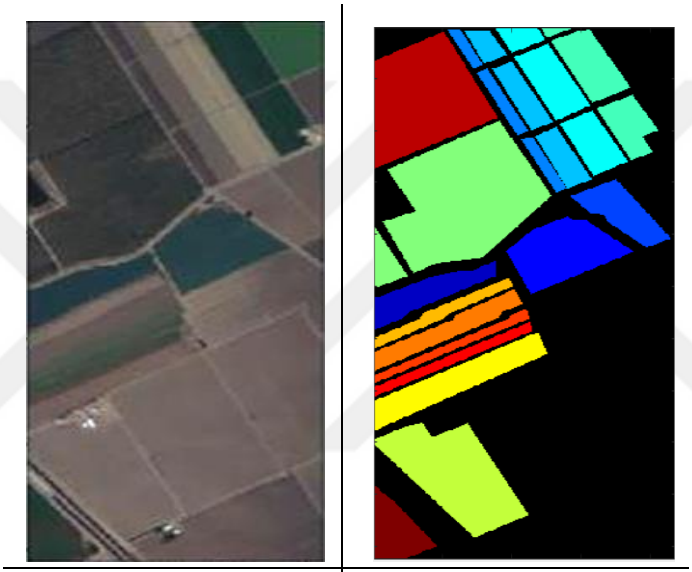
The dataset consists of 9 distinct classes, as indicated in Table 3.3. The dataset has dimensions of 610 pixels in width, 340 pixels in height, and 103 spectral bands. This dataset encompasses a collection of difficulties linked to the vast number of dimensions in hyperspectral data. In Table 3.3, choose 1% of the samples from every class.

Table 3.3 Number of test and train samples (1%) in PU dataset

No	Classes	Train	Test	Total Samples
1	Asphalt	67	6564	6631
2	Meadows	187	18462	18649
3	Gravel	21	2078	2099
4	Trees	31	3033	3064
5	Painted metal sheets	14	1331	1345
6	Bare Soil	51	4987	5029
7	Bitumen	14	1316	1330
8	Self-Blocking Bricks	37	3645	3682
9	Shadows	10	947	947
Total		432	42344	42776

3.3 Salinas Dataset

The 224-band AVIRIS sensor captured the image above the Salinas Valley in California, revealing a high level of detail with 3.7-meter pixels. The covered region contains a total of 512 lines, with 217 samples distributed evenly among them. We excluded the 20 water absorption bands in a similar manner to the Indian Pines scenario. Only at-sensor radiance data was available for this image. The area encompasses vineyards, agricultural plots, and dry terrain. The training and testing experiments are conducted on Salinas data, which is similar to the Indian data mentioned earlier.



Original RGB [95] (A) | Ground truth (B)

Figure 3.3 Original Image (A) [95] and GT of SA (B)

In the experiment, we select 2% of samples from each class, following the same scenario as the previous data set. Keeping the training rounds and size the same ensures that the mechanisms used are identical to those in the Indian data. Table 3.4 displays the aggregate quantity of training and test samples for each class. The Salinas dataset stands out from others because of its abundant spectral information and extensive dimensions.

Table 3.4 Number of test and train samples (2%) in SA dataset

No	Classes	Train	Test	Total Samples
1	Brocoli green weeds 1	41	1968	2009
2	Brocoli green weeds 2	38	3688	3726
3	Fallow	40	1936	1976
4	Fallow rough plow	28	1366	1394
5	Fallow smooth	54	2624	2678
6	Stubble	80	3879	3959
7	Celery	72	3507	3579
8	Grapes untrained	226	11045	11271
9	Soil vinyard develop	125	6078	6203
10	Corn senesced green weeds	66	3212	3278
11	Lettuce romaine 4wk	22	1046	1068
12	Lettuce romaine 5wk	39	188	1927
13	Lettuce romaine 6wk	19	897	916
14	Lettuce romaine 7wk	22	1048	1070
15	Vinyard untrained	146	7122	7268
16	Vinyard vertical trellis	37	1770	1807
Total		1055	53074	54129

4
EXPERIMENTS

The results were achieved using the updated CNN+AL+MRF method, which combines proposed methods such as dimensionality reduction (DR) and edge-preserving filter (EPF) for HIC. An evaluation was performed to determine the impact of these different strategies on the accuracy of classification and performance using specific methodologies in the three groups. Furthermore, this approach is evaluated by displaying and analyzing quantitative results, such as OA, AA, and computing time.

These results prove the usefulness of the framework approach in improving classification accuracy, computational efficiency, effectiveness, and adaptability in the field of HSI analysis.

4.1 Experimental Setup

The proposed approach relied on three popular spectral datasets – Indian Pines, Pavia University, and Salinas to analyze the effects of different phases. Dimensionality reduction methods such as AVG and PCA have been applied to simplify the domains and improve the performance of the classifier as pre-processing combined with the use of edge-preserving filters such as WLS, JointWMF, FGS, BF, and SD.

Several different configurations were tested in order to obtain or achieve classification accuracy that is good and higher than the previous one. Repeated operations and careful comparisons were performed in several scenarios to obtain results in detailed tables that demonstrate the efficiency of these techniques when applied to these three data.

This comprehensive investigation demonstrates the effectiveness and usability of proposed technique to develop CNN+AL+MRF method in the field of hyperspectral imaging. Moreover, the use of MRF has greatly enhanced the classification accuracy and thus led to the creation of a systematic method that includes multiple iterations and distinct experimental groups allocated, the experimental groups being precisely named.

When combined with PCA, CNN-AL can be used with or without MRF (W MRF or WO MRF) combined with an edge-preserving filter as a bilateral filter (BF). This combination is referred to as “PCA+BF+CNN-Al W/MRF”. This designation is a combination of abbreviations of procedures used in the relevant combination to indicate the experimental configuration.

4.2 Criteria for Evaluating Performance

For assessing the performance of feature selection methods on HSI datasets, in this study employed three distinct assessment criteria. These criteria were numerically compared using the following:

4.1.1 Overall Accuracy (OA)

Measures the overall classification performance as a percentage, this statistic assesses the model's efficiency to accurately classes the complete hyperspectral image. Consider it a comprehensive metric, symmetric to a global assessment reflecting the model's precision throughout the entire image. For instance, if the OA is 90%, it indicates that the model correctly categorized 90 out of 100 pixels.

4.1.2 Average Accuracy (AA)

Measuring the average level of categorical accuracy as a percentage, AA offers a broader and more comprehensive perspective compared to simple accuracy measures by assessing the average accuracy of specific groups within an image. It provides valuable information regarding the model's performance in each category, or the individual category shown in the image, the AA delves deeper. For example, if an image includes categories such as “vegetation,” “water,” and “buildings,” AA calculates the average accuracy for each of these categories.

4.1.3 Computation Time (s)

Provides a metric for calculating the time it takes to perform calculations for each technique, measured in seconds. This metric determines the time required for the model to analyze and classify the hyperspectral image. Evaluating the effectiveness of the model is critical, especially in applications where real-time analysis is required or situations where time is of the utmost importance.

The improved CNN+AL+MRF method effectively modulate classification accuracy in hyperspectral imaging by integrating dimensionality reduction techniques and edge-preserving filters. Evaluating the classification performance of this approach depends heavily on these integrated methods and the size of the data used.

The integration of these methods has contributed to achieving remarkable and enhanced classification results, as observed by the results of high accuracy and effective performance in hyperspectral image classification.

This method is performed on a computer with an NVIDIA GeForce RTX 3060 graphics card and 64GB of memory. Implement this method in a MATLAB environment using the MatConvNet package.

4.3 Effect of Edge Preserving Filtering on Datasets

The figures from Figure 4.1 to Figure 4.9 present an illustrated analyses of the implementation of five different filtering techniques: WLS filter, JointWMF filter, FGS filter, BF, and finally SD filter with dimensionality reduction techniques. This implementation is considered as a pre-processing step. Apply it to three different data sets and then apply DA to them so that they are prepared for the CNN structure to train the model with the available labeled data. After this initial training phase, AL methodology is applied to identify the most informative or uncertain samples from the unlabeled dataset and then using MRF it helps in improving the classification accuracy by incorporating contextual information into the classification process.

Moreover, the usefulness of this approach in improving classification accuracy and computational efficiency is demonstrated by displaying and analyzing quantitative results, such as OA, AA, and computing time in the following tables for implementing this approach.

Clearly, all techniques involve the filtration of band images and the integration of spatial data, which has an impact on the optical fidelity of hyperspectral images. The Edge preservation (EP) in homogenous zones is important for minimizing class variance and translating spectral signatures into more comparable patterns.

Specifically, SD and BF methods exhibit greater degrees of image smoothing when compared to other techniques in the overall context.

Note that hyperparameters about both active learning process and different edge preserving filters are shared in Table 4.1 and Table 4.2, respectively. More information can be found from the related references [71,76,78,80,88].

Table 4.1 Hyperparameters about Active Learning during different experiments

Type Data	Learning Rate	Epochs	Number Times Iteration	Number of samples
Indian Pines	0.002	300	5	15
Pavia University	0.001	300	4	5
Salinas	0.002	300	5	15

Table 4.2 Hyperparameters about different EPF methods

Filter Type	Parameter 1 & Its Value	Parameter 2 & Its Value	Parameter 3 & Its Value
WLS	Smoothness term, 1	Degree of control over the affinities, 1.2	-
JointWMF	Window radius, 10	Standard deviation of kernel, 50	Iteration, 1
FGS	Standard deviation,0.005	Lambda, 100	-
BF	Window size, 5	Standard deviation, 3	-
SD	Static guidance, 100	Dynamic guidance, 400	Step, 3

4.4 Performance Evaluation of The Proposed Framework

This section demonstrates the impact of the amount of training data on the accuracy of classification, specifically for each feature selection approach.

4.4.1 Results of Indian Pines (IP)

The new methodology includes randomly selecting 2% samples from each class for training purposes, so the total training samples are 211 out of a total of 10,249, and the rest of the samples are for testing, as shown in table 3.2, which specifies the specific number of trainings from the Indian Pines (IP) dataset and samples. Test for each category. Training consists of 300 epochs in each round. Under AL methods, 15 training samples are selected in each round and this is repeated in subsequent rounds to enlarge the training set, ensuring consistency over successive iterations. The iterative selection process ensures that the learning model robustly captures diverse patterns from the data set. Implemented as pre-processing by implementing dimensionality reduction techniques and edge-preserving filters together, this combination ensures a more manageable and interpretable data set, by providing a more concise and robust representation of the data.

In addition, the combined application facilitates accurate image segmentation by preserving boundaries and structures, making it particularly resource efficient and ensuring efficient computational resource management.

Overall, this fusion results in a focused, noise-free, and information-rich dataset or image, which ultimately enhances the results of analysis, classification, and processing.

Table 4.2 shows a detailed look at the classification accuracy and computation times of different methods that uses without DR along with different edge-preserving filters (WLS, JWMF, FGS, BF, and SD) and (with and without) the MRF. Furthermore, we incorporate the performance metrics of the original CNN+AL approach in the first row for the sake of reference. It is important to mention that this chart deliberately does not include any methods that reduce dimensions to focus solely on the impact of edge-preserving filters on classification results.

The results demonstrate a consistent pattern in which the incorporation of edge-preserving filters generally leads to higher overall accuracy in comparison to the original technique.

Nevertheless, a fascinating finding arises when the MRF is in operation, the WLS filter displays a decline in AA measurements in comparison to another filter's

techniques. In contrast, the FGS, BF, and SD filters consistently demonstrate AA scores exceeding 95% in both versions without and with MRF.

The SD+CNN+AL combination exhibits the best level of accuracy compared to the other configurations that were examined, making its performance particularly remarkable. The model earned an OA of 99.00% and an AA of 98.62%, highlighting the effectiveness of this combination in improving classification accuracy. These findings emphasize the need to use edge-preserving filters to improve the accuracy of classification in the field of hyperspectral imaging.

Table 4.3 The accuracy percentages (%) and computational times for IP methods of classification without DR and EPF

No	Classes	CNN+AL		WL +CNN+AL		JWMF+CNN+AL		FGS+CNN+AL		BF + CNN+AL		SD + CNN+AL	
		W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF
1.	**Alfalfa**	97.43	93.69	**98.34**	92.37	97.55	90.55	**97.67**	97.16	97.65	81.34	97.67	85.41
2.	**Corn-notill**	98.15	81.03	98.27	87.08	98.13	82.45	98.11	96.97	**98.35**	98.61	**98.47**	95.03
3.	**Corn-mintill**	99.56	93.83	**100**	71.62	**100**	67.70	**100**	100	**100**	92.73	100	96.33
4.	**Corn**	98.14	92.36	98.55	93.61	96.88	93.37	96.88	100	96.88	97.09	97.48	96.86
5.	**Grass-pasture**	99.04	97.93	98.90	96.70	99.31	97.79	99.31	99.17	**99.31**	99.72	**99.31**	99.45
6.	**Grasstrees**	**100**	98.05	**100**	99.16	**100**	97.89	96.88	100	**100**	97.22	**100**	99.79
7.	**Grass_pasture-mowed**	97.86	88.42	98.19	87.06	98.69	85.56	98.81	98.05	**98.72**	98.07	**98.81**	97.07
8.	**Hay-windrowed**	97.61	94.52	99.38	92.09	99.58	92.12	99.53	97.61	**99.57**	96.30	**99.58**	95.36
9.	**Oats**	96.47	92.06	98.65	90.88	**99.14**	85.88	98.80	96.58	**98.81**	96.08	98.80	95.38
10.	**Soybean-notill**	99.52	99.52	100	94.20	**100**	96.17	**100**	99.51	**100**	99.52	**100**	99.02
11.	**Soybean-mintill**	99.60	99.44	99.84	99.04	**100**	99.04	**100**	99.60	**100**	99.68	**100**	99.60
12.	**Soybean-clean**	98.65	79.45	**100**	70.03	**100**	74.05	**100**	99.16	**100**	96.91	**100**	97.73
13.	**Wheat**	81.72	80.65	85.87	**98.91**	87.64	**93.55**	91.11	82.22	90.00	88.89	89.89	91.01
14.	**Woods**	96.00	67.31	**98.08**	73.08	**98.00**	73.08	97.96	93.88	97.96	89.80	97.78	88.89
15.	**Buildings-Grass-Trees**	76.00	92.00	73.08	0.00	**100**	0.00	**100**	78.26	**100**	36.00	**100**	64.00
16.	**Stone-Steel-Towers**	89.47	91.99	60.00	20.00	**100**	36.84	**100**	100	**100**	100	**100**	89.47
Overall Accuracy (OA)		95.33	92.70	98.78	91.17	98.92	90.19	98.94	97.93	**98.95**	95.11	**99.00**	95.26
Average Accuracy (AA)		95.00	90.15	94.20	89.11	98.43	89.13	**98.64**	96.02	98.58	91.75	**98.61**	93.15
Computation Time (s)		252.5	166.6	196.2	194.6	195.3	189.1	192.7	189.2	202.6	200.0	210.9	202.9

The SD+CNN+AL combination exhibits the best level of accuracy compared to the other configurations that were examined, making its performance particularly remarkable. The model earned an OA of 99.00% and an exceptional AA of 98.62%, highlighting the effectiveness of this combination in improving classification accuracy. These findings emphasize the need to use edge-preserving filters to improve the accuracy of classification in the field of hyperspectral imaging.

Table 4.4 The accuracy percentages (%) and computational times for IP methods of classification with AVG and EPF

No	Classes	AVG + CNN+AL		AVG+WLS+ CNN +AL		AVG+JWMF+ CNN +AL		AVG +FGS+ CNN +AL		AVG + BF+ CNN +AL		AVG +SD+ CNN +AL	
		W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF
1.	**Alfalfa**	97.64	97.42	**97.64**	96.35	97.20	92.33	**97.71**	97.57	97.62	85.58	97.42	87.08
2.	**Corn-notill**	98.64	97.40	98.64	97.27	93.18	92.44	98.64	97.29	**98.64**	98.27	**98.64**	94.20
3.	**Corn-mintill**	**100**	100	**100**	99.56	**100**	97.3	**100**	100	**100**	93.01	**100**	95.61
4.	**Corn**	**98.97**	98.14	98.76	98.35	91.30	98.14	**99.18**	98.14	98.96	97.72	98.76	96.4
5.	**Grass-pasture**	99.32	99.04	99.31	99.31	99.72	99.31	**99.45**	99.03	**99.40**	99.72	99.31	99.45
6.	**Grasstrees**	**100**	100	**100**	100	**100**	99.79	**100**	100	**100**	97.06	**100**	99.79
7.	**Grass_pasture-mowed**	98.41	97.99	98.40	98.61	**98.72**	97.86	98.49	97.96	**98.61**	98.40	98.40	97.65
8.	**Hay-windrowed**	99.25	98.13	99.58	98.87	**99.62**	94.69	99.25	97.74	99.54	94.64	**99.75**	93.38
9.	**Oats**	98.83	96.33	**99.15**	96.11	98.31	95.28	98.82	96.64	98.65	95.79	**98.89**	95.44
10.	**Soybean-notill**	**100**	99.03	**100**	99.03	**100**	99.52	**100**	99.03	**100**	99.52	**100**	99.04
11.	**Soybean-mintill**	**100**	99.68	**100**	99.60	**100**	99.68	**100**	99.60	**100**	99.44	**100**	99.12
12.	**Soybean-clean**	**100**	98.65	100	97.55	**100**	99.46	**100**	98.65	**100**	97.30	**100**	98.64
13.	**Wheat**	88.17	81.74	89.01	83.52	**89.25**	73.12	88.17	81.72	88.17	87.10	**89.01**	86.81
14.	**Woods**	**98.11**	98.11	**98.08**	98.08	96.00	90.00	98.04	96.08	98.00	86.00	98.04	90.20
15.	**Buildings-Grasstrees**	75.00	83.33	73.08	76.92	**100**	66.67	72.00	76.00	73.08	38.46	**100**	61.54
16.	**Stone-Steel-Towers**	**100**	95.00	**100**	100	**100**	100	**100**	90.00	**100**	100	**100**	90.00
Overall Accuracy (OA)		98.09	96.25	**98.98**	98.13	98.20	95.93	98.92	97.97	98.97	95.25	**99.03**	94.90
Average Accuracy (AA)		96.25	94.21	96.98	96.20	**97.71**	93.48	96.86	95.34	96.92	91.75	**98.64**	92.78
Computation Time (s)		165.4	161.0	179.5	**114.5**	188.9	**152.5**	204.7	189.9	205.7	190.4	202.7	190.47

The implementation of dimensionality reduction techniques, in particular AVG and PCA, are well captured and evaluated in Tables 4.2 and 4.3 respectively. The tables provide a detailed representation of the significant impact that different methods have on the hyperspectral image classification process. The AVG and PCA approaches, along with the implementation of the five-filters approach, significantly influence the classification results.

The quantitative evaluation of each reduction technique with each of the five filters shown in the aforementioned tables shows the OA, AA, and computation time, showing the significant improvement in performance for all categories when using these methods combined together. The collective improvement is clearly, as all filters exceed the basic CNN+AL technique. The SD filter is undoubtedly the most outstanding performer among these three tables.

Table 4.5 The accuracy percentages (%) and computational times for IP methods of classification with PCA and EPF

No	Classes	PCA + CNN+AL		PCA+WLS+ CNN+AL		PCA+JWMF + CNN+AL		PCA+FGS + CNN+AL		PCA+ BF+ CNN+AL		PCA+SD + CNN+AL	
		W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF
1.	Alfalfa	97.43	93.36	97.64	96.49	**97.64**	92.05	97.63	97.35	97.63	85.40	**97.63**	89.81
2.	Corn-notill	98.15	83.50	98.15	98.03	**98.27**	91.46	98.15	97.65	**98.15**	98.02	98.15	94.31
3.	Corn-mintill	99.56	77.33	**100**	99.13	**100**	97.39	**100**	100	**100**	93.39	**100**	96.51
4.	Corn	**98.14**	93.35	96.89	98.34	96.90	98.35	96.90	98.14	96.90	97.52	**97.70**	97.08
5.	Grass-pasture	**99.04**	96.28	99.31	99.45	99.31	99.31	99.31	99.17	99.31	99.7	**99.31**	99.45
6.	Grasstrees	**100**	98.95	**100**	100	**100**	99.58	**100**	100	**100**	97.27	**100**	100
7.	Grass-pasture-mowed	97.86	84.53	**98.71**	98.29	98.72	97.87	98.71	97.96	98.72	98.40	**98.72**	97.54
8.	Hay-windrowed	97.61	91.41	99.54	98.58	99.54	94.93	99.54	97.49	**99.58**	94.78	**99.54**	92.47
9.	Oats	**96.47**	89.04	98.82	95.78	98.82	94.94	98.82	96.63	98.82	95.79	**98.82**	94.60
10.	Soybean-notill	99.52	94.23	**100**	99.04	**100**	99.49	**100**	100	**100**	99.52	**100**	99.04
11.	Soybean-mintill	99.60	98.81	**100**	99.52	**100**	99.52	**100**	99.68	**100**	99.44	**100**	98.96
12.	Soybean-clean	98.65	74.86	**100**	97.83	**100**	98.65	**100**	98.64	**100**	97.57	**100**	98.92
13.	Wheat	81.72	96.74	**89.13**	83.70	88.17	73.12	88.17	81.72	88.17	88.17	**90.04**	85.87
14.	Woods	96.00	74.51	**98.00**	98.00	98.04	96.08	**98.04**	98.04	98.00	86.00	97.96	87.76
15.	Buildings-Grasstrees	76.00	0.00	**100**	76.00	**100**	69.23	**100**	76.92	**100**	38.46	**100**	73.08
16.	Stone-Steel-Towers	89.47	26.32	**100**	95.00	**100**	**100**	**100**	95.00	**100**	100	**100**	90.00
Overall Accuracy (OA)		97.97	90.72	98.92	98.08	98.92	95.82	98.91	95.90	**98.92**	93.29	**99.02**	95.07
Average Accuracy (AA)		95.33	88.91	98.51	95.82	**98.46**	93.84	98.46	94.83	98.46	91.84	**98.62**	93.46
Computation Time (s)		**151.7**	**144.4**	179.4	177.5	193.3	192.7	190.7	189.9	199.9	198.5	203.9	200.0

Combining the AVG+CNN+AL and PCA+CNN+AL settings with the SD filter results in impressive OA values of 99.03% and 99.02% respectively. According to the results of this study, edge-preserving filters, and dimensionality reduction techniques (AVG and PCA) work better together to improve classification accuracy and make it more accurate.

This is further evidence of the success of the updated CNN+AL+MRF method, especially what was observed in the SD filter. The SD filter design stands out for its persistent dominance, making it the most effective strategy for gaining higher classification precision. Combinations such as AVG+CNN+AL and PCA+CNN+AL, combined with the SD filter, achieve impressive OA scores of 99.03% and 99.02%, respectively, highlighting their outstanding performance.

The numerical findings highlight the effectiveness of dimensionality reduction approaches combined with edge-preserving filters. They also underline the improved performance of the SD filter configuration in achieving higher classification precision. When assessing the accuracy percentages, it is crucial to consider the contextual relevance of different land cover types.

The variations in accuracy seen among distinct land cover categories, such as "Grasstrees" and "Soybeans," obtaining nearly 100% accuracy compared to categories like "Wheat" and "Buildings-Grass-Trees" with lesser accuracy, indicate the diverse performance of classification across different types of land cover.

The combination of CNN, AL, AVG, SD, and MRF yields the most effective approach, with an OA of 99.03% and an AA of 98.64%. This not only showcases superior outcomes but also highlights its computational efficiency. To summarize, the inclusion of dimensionality reduction techniques improves both the accuracy of classification and the speed of computation.

This highlights the effectiveness of combining AVG and PCA methods in the proposed CNN+AL+MRF approach for hyperspectral image classification. Tables 4.1, 4.2, and 4.3 provide a comprehensive analysis and detailed information on how these methods affect hyperspectral image classification. This provides insight into subtle differences in performance as metrics such as OA, AA, and computation time have large numerical values compared to the original method.

4.4.1.1 Classification of Map IP

Classification maps were created for this group to evaluate the effectiveness of the approach across the real spectrum data set, and as shown in the figures 4.1, 4.2, and 4.3.

They are the results of the classification that were carried out in three stages in CNN+AL original and AVG + EPF+CNN +AL and PCA +EPF+ CNN +AL, so visualize These maps the spatial distribution of IP classes within the image and how these categories are distributed.

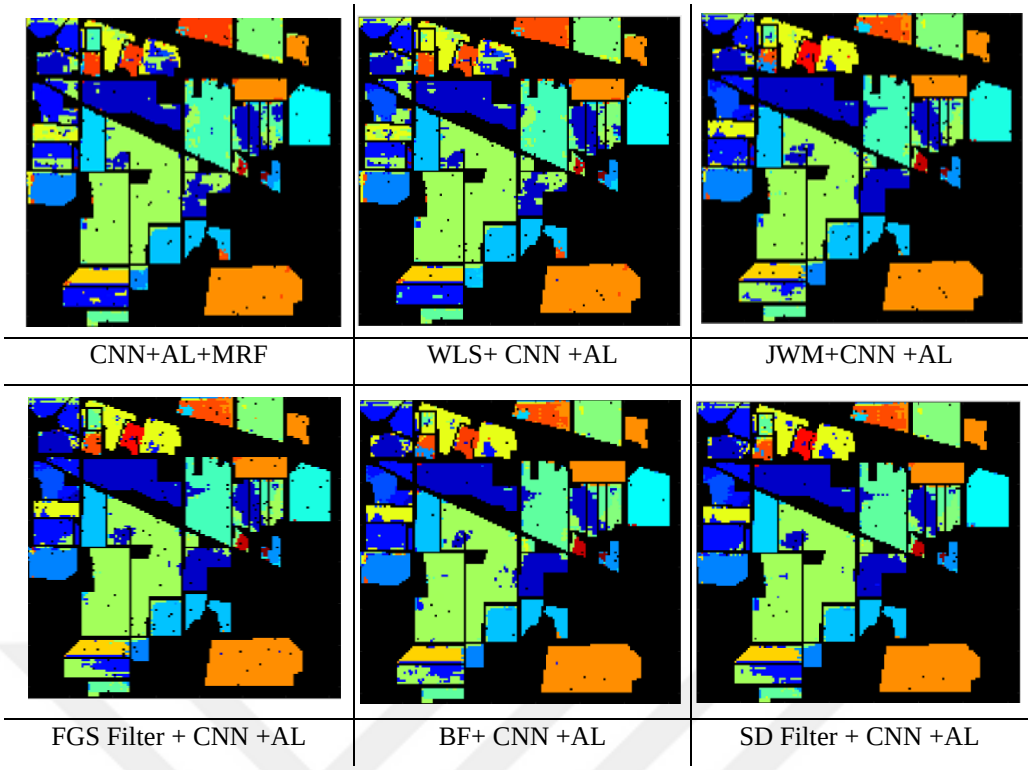


Figure 4.1 Classification maps for IP data set for using 2% training samples of the methods without DR and EPF

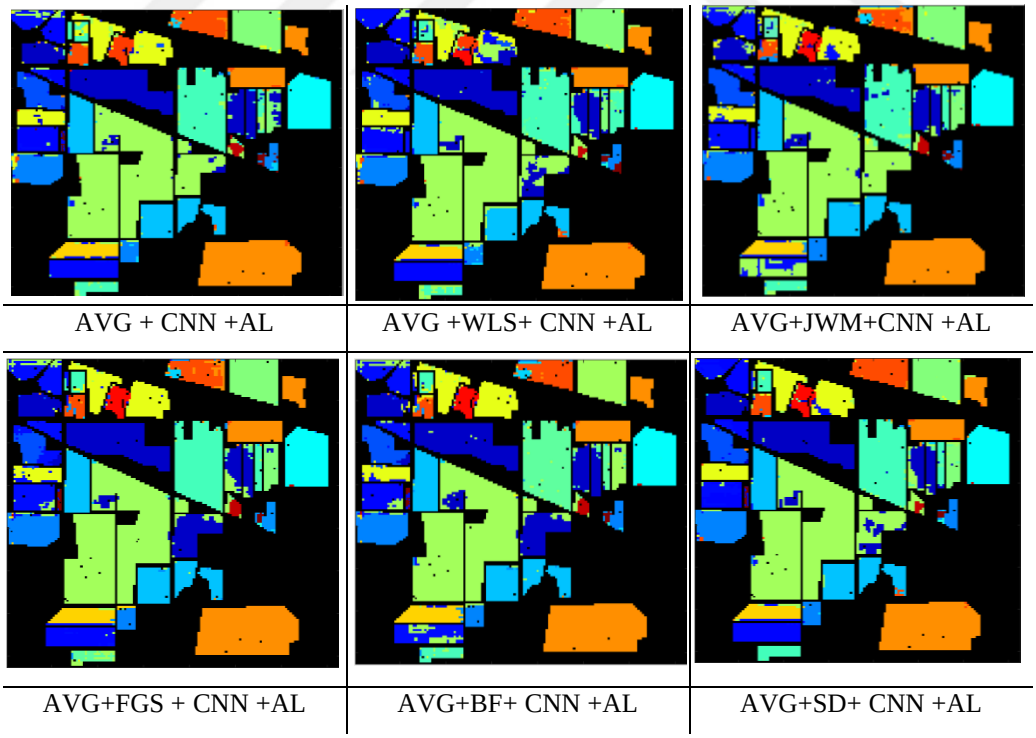


Figure 4.2 Classification maps for IP data set for using 2% training samples of the methods with AVG and EPF

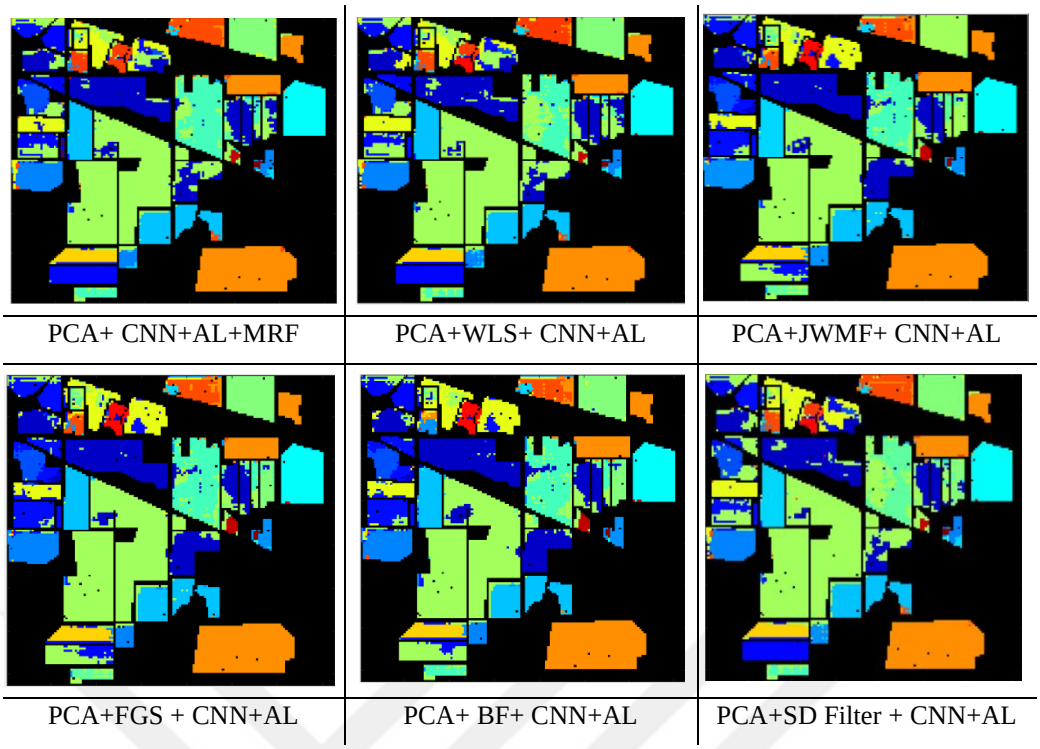


Figure 4.3 Classification maps for IP dataset using 2% training samples of the methods with PCA and EPF

4.4.2. Results of Pavia University (PU)

University of Pavia dataset was carefully evaluated by the final results of the improved CNN+AL+MRF method, as shown in Table 4.4, Table 4.5 and Table 4.6. This evaluation, illustrated by advanced calculations, provides a comprehensive analysis of the complex effect of edge-preserving filters combined with DR techniques in the field of hyperspectral image classification. CNN training sessions consist of 300 epochs, and these are performed four times in an iterative manner, where 5 samples are selected in each epoch. This procedure in the experimental design, a rigorous and random sampling strategy is used, where 1% of each category of samples is selected, as shown in Table 3.3, and the remainder of the samples are designated for tests. Continues until the repetition condition ends, and after that comes the turn of MRF as a final improvement. The function of MRF as a regular processing contribute to enhances accuracy levels, highlighting its effectiveness in adjusting spatial linkages and improving classification outcomes. It is considered a one of solution to image labeling problems this is due to its predictive power.

Table 4.6 The accuracy percentages (%) and computational times for PU methods of classification without DR and EPF

No	Classes	CNN+AL		WLS + CNN+AL		JWMF+CNN+ AL		FGS + CNN+AL		BF + CNN+AL		SD+ CNN+AL	
		W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF
1.	**Asphalt**	99.26	96.63	99.32	91.78	99.36	80.22	99.36	94.61	**99.39**	67.01	**99.92**	52.37
2.	**Meadows**	94.70	91.28	95.03	99.83	**95.17**	98.07	94.73	94.71	94.95	98.14	**95.12**	98.51
3.	**Gravel**	99.76	98.52	**99.96**	86.08	99.76	99.67	99.78	96.64	**99.78**	99.71	99.76	99.87
4.	**Trees**	97.22	98.01	**98.90**	85.95	**99.04**	89.95	97.92	97.29	97.22	84.31	98.01	73.94
5.	**Painted metal sheets**	99.93	99.93	**99.98**	99.03	99.93	99.85	99.93	98.21	99.94	99.93	**100**	99.33
6.	**Bare Soil**	**100**	96.95	**100**	81.58	**100**	94.92	**100**	91.96	**100**	96.70	**100**	96.16
7.	**Bitumen**	97.52	96.46	98.03	94.73	97.63	96.84	97.59	96.73	**98.13**	96.70	**98.03**	74.72
8.	**Self-Blocking Bricks**	39.56	40.24	42.20	37.88	**44.07**	40.15	**42.66**	40.13	41.99	26.11	40.62	23.98
9.	**Shadows**	98.84	99.26	**99.24**	98.31	**99.24**	97.67	98.85	97.36	**98.93**	96.73	98.83	65.20
Overall Accuracy (OA)		92.06	90.81	93.29	90.00	94.02	89.26	**94.22**	90.13	94.10	86.24	**94.99**	83.30
Average Accuracy (AA)		91.86	89.83	**92.52**	88.35	**92.69**	88.59	92.31	89.74	92.26	85.04	92.25	76.01
Computation Time (s)		**398.3**	**371.0**	420.0	410.2	498.4	436.2	419.33	400.9	581.7	522.8	573.1	514.97

Table 4.7 The accuracy percentages (%) and computational times for PU methods of classification with AVG and EPF

No	Classes	AVG + CNN +AL		AVG+WLS + CNN +AL		AVG +JWMF+ CNN +AL		AVG +FGS+ CNN +AL		AVG + BF + CNN +AL		AVG +SD + CNN +AL	
		W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF
1.	**Asphalt**	98.73	86.65	98.98	81.27	98.97	83.02	**99.30**	87.41	98.73	86.03	**99.27**	94.69
2.	**Meadows**	99.07	95.31	99.59	98.42	100	98.39	**99.99**	98.33	99.08	97.86	**99.98**	98.54
3.	**Gravel**	72.32	68.31	75.21	90.27	**80.87**	82.64	**81.12**	85.54	79.24	81.06	80.02	72.84
4.	**Trees**	93.24	68.93	94.12	89.06	**95.02**	89.78	94.73	89.19	**95.22**	91.31	94.60	93.45
5.	**Painted metal sheets**	99.85	98.88	**100**	97.83	99.99	97.46	**100**	96.79	**100**	98.29	99.99	99.78
6.	**Bare Soil**	99.12	78.57	99.98	91.94	99.98	92.56	99.12	94.71	**99.99**	96.71	**100**	87.83
7.	**Bitumen**	94.80	53.54	95.07	58.14	**97.21**	83.79	**95.99**	78.58	95.15	84.05	95.90	58.30
8.	**Self-Blocking Bricks**	92.00	66.65	93.08	48.11	**94.52**	70.84	93.21	70.75	**93.89**	81.07	93.03	88.57
9.	**Shadows**	62.16	65.22	63.02	97.57	64.96	96.61	66.99	98.31	**70.01**	97.78	**70.10**	97.67
Overall Accuracy (OA)		95.76	84.47	96.86	88.31	97.13	91.03	**97.99**	91.89	97.10	92.73	**98.00**	92.97
Average Accuracy (AA)		90.14	75.78	91.01	83.63	**92.39**	88.34	92.27	88.85	92.37	90.46	**92.54**	87.96
Computation Time (s)		**356.7**	**332.8**	411.2	395.2	410.8	411.0	402.2	395.9	458.8	458.8	536.6	509.26

The (SD) filter also demonstrates exceptional performance in all three tables, achieving high OA scores of 94.99 %, 98.00, and 97.00%, along with AA scores of 92.25%, 92.54%, and 91.38% in Table 4.4, Table 4.5, and Table 4.6, respectively. The consistent performance of the SD filter demonstrates its capacity to effectively maintain a balance between precision and accuracy when applied to various datasets.

Nevertheless, several categories in the University of Pavia dataset pose difficulties and exhibit relatively lower levels of individual accuracy, the difficulties caused by the large dimensions of the hyper data in the University of Pavia dataset

Table 4.8 The accuracy percentages (%) and computational times methods of classification of PU with PCA and EPF

No	Classes	PCA + CNN+AL		PCA+WLS + CNN+AL		PCA+JWMF+ CNN+AL		PCA+FGS + CNN+AL		PCA+ BF+ CNN+AL		PCA+SD+ CNN+AL	
		W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF
1.	**Asphalt**	**99.26**	85.35	96.32	79.50	97.24	81.27	98.73	84.01	**98.73**	80.31	98.26	80.32
2.	**Meadows**	97.58	96.51	**99.46**	98.90	99.39	98.42	99.07	97.93	99.07	98.27	**99.62**	99.31
3.	**Gravel**	**99.62**	33.35	53.46	48.44	72.71	**90.27**	72.38	84.98	72.33	71.98	80.88	73.61
4.	**Trees**	**97.88**	91.30	88.99	91.35	89.66	89.06	93.24	91.18	93.24	85.78	**94.78**	87.54
5.	**Painted metal sheets**	99.78	95.09	**100**	99.85	**100**	97.83	99.85	99.03	99.85	98.30	99.32	97.94
6.	**Bare Soil**	96.32	87.57	93.57	71.98	95.30	91.94	**99.12**	94.29	99.12	88.09	**99.69**	88.41
7.	**Bitumen**	**96.68**	76.02	89.15	47.35	92.81	58.14	94.80	63.36	94.80	60.39	**94.81**	69.44
8.	**Self-Blocking Bricks**	43.88	78.63	**93.95**	66.03	78.52	48.11	**92.08**	60.79	91.99	82.99	92.06	82.59
9.	**Shadows**	97.99	40.49	**98.84**	99.26	**98.62**	97.57	62.22	97.57	62.16	98.96	62.99	98.62
Overall Accuracy (OA)		93.24	86.80	94.48	85.42	95.57	88.31	**95.77**	89.98	95.76	90.10	**97.00**	90.69
Average Accuracy (AA)		**92.10**	76.03	90.41	78.07	91.58	83.63	90.17	85.90	90.14	85.01	**91.38**	86.42
Computation Time (s)		**350.7**	**319.9**	429.9	410.2	403.2	382.2	375.7	331.9	487.1	441.9	355.7	340.0

The highest point of the discoveries is evident in the CNN+AL+AVG+SD filter configuration, utilizing MRF technology, which emerges as the most skilled in providing superior classification accuracy.

The setup attains an amazing OA of 98.00% and AA of 92.54%, signifying a significant enhancement compared to the previous CNN+AL approach. It is important to emphasize that although the SD filter produces significant enhancements in accuracy, it also entails higher computing complexity in comparison to alternative approaches. To summarize, these findings offer a thorough comprehension of the complex interaction among edge-preserving filtering techniques, MRF technology, and classification precision in the field of hyperspectral image analysis. The results highlight the adaptability and effectiveness of these methods in overcoming the difficulties presented by the large size of hyperspectral datasets, especially in the complex landscape of the University of Pavia dataset.

4.4.2.1 Classification Map of PU

The figures below are the results of dividing the maps of the Pavia University dataset, where notice in the figures 4.4, 4.5, and 4.6 the effect of dimension reduction techniques with the five filters on the maps, which help in identifying the regions and then determining the contents of the regions such as pastures, gravel, trees, and others within the image.

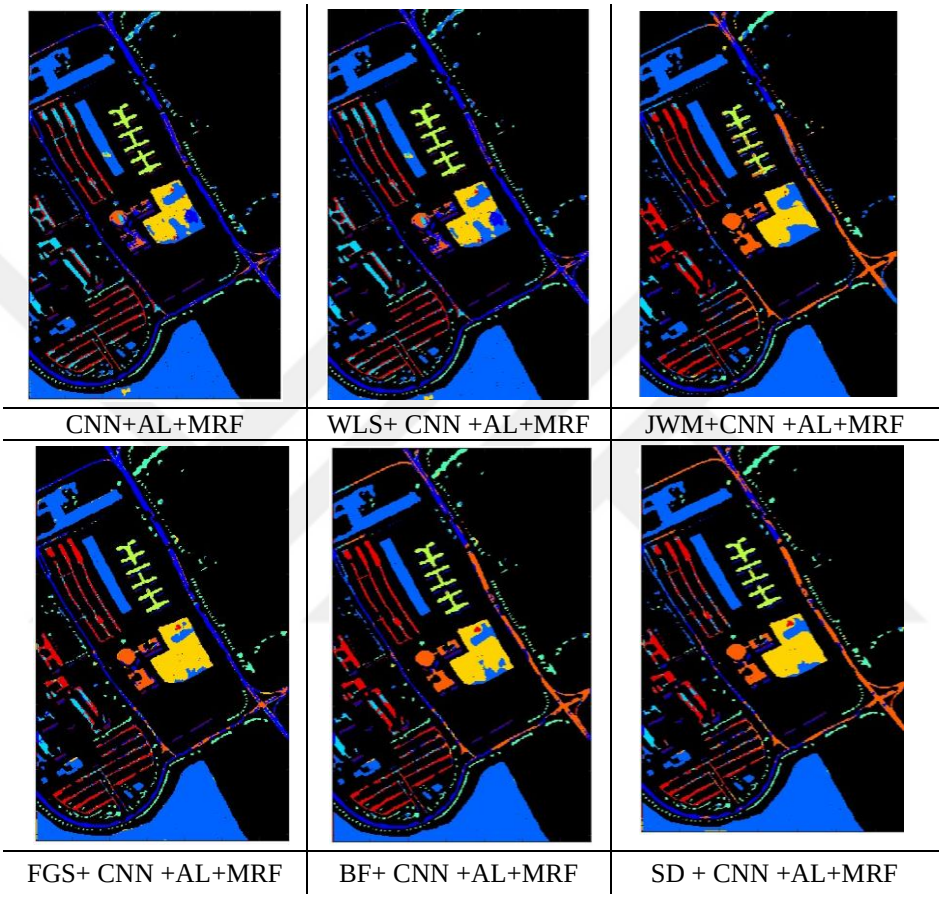


Figure 4.4 Classification maps for PU dataset using 1% training samples of the methods without dimension reduction

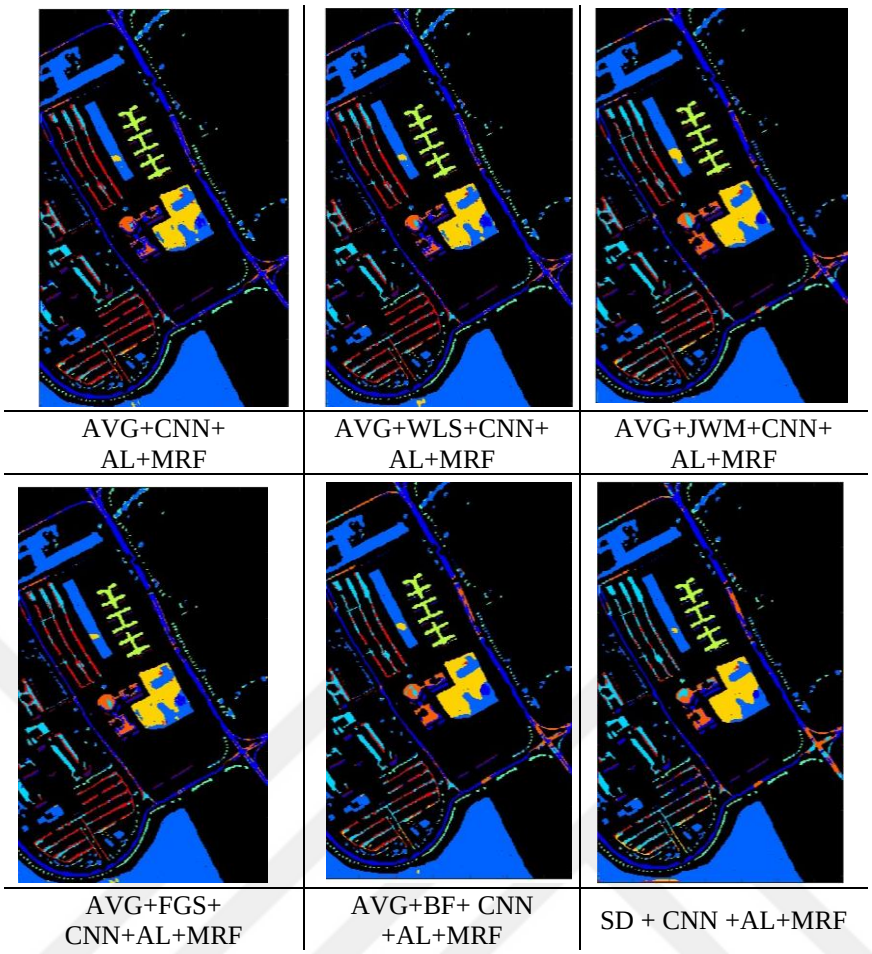


Figure 4.5 Classification maps for PU data set for using 1% training samples of the methods with AVG and EPF

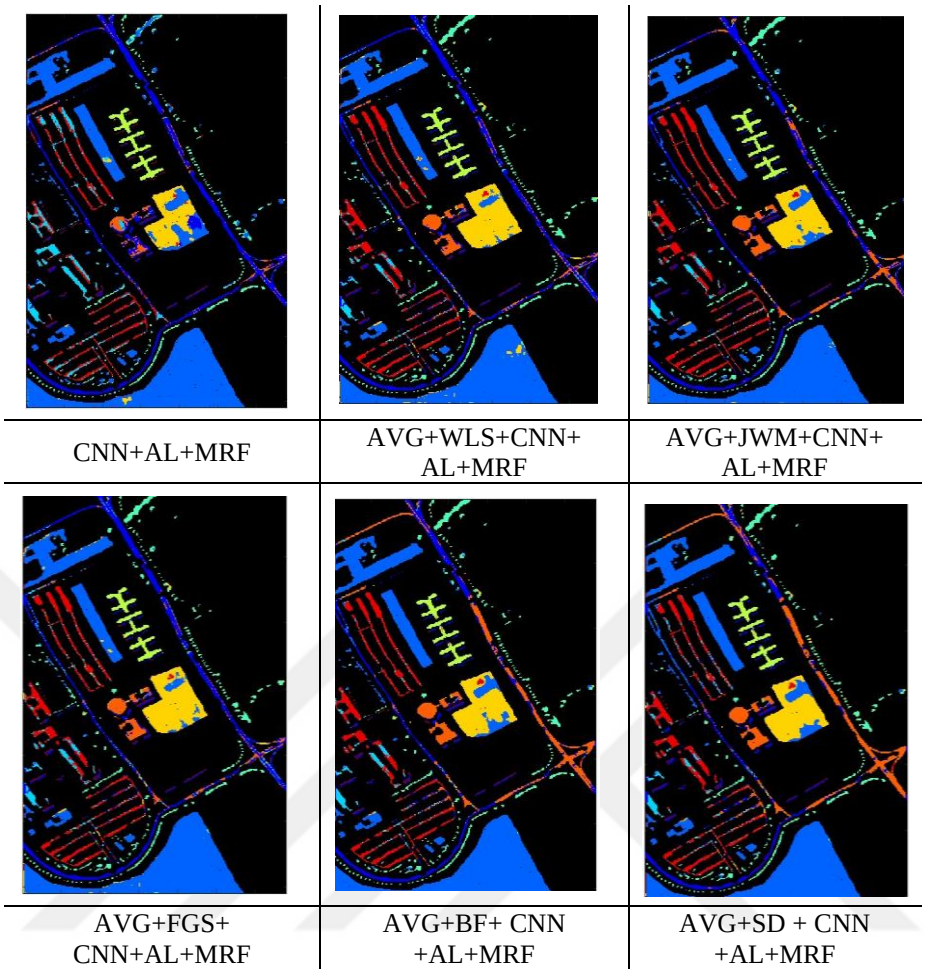


Figure 4.6 Classification maps for PU dataset using 1% training samples of the methods with PCA and EPF

4.4.3 Results of Salinas

The training and testing strategy applied on the Salinas dataset and the implementation mechanism of the CNN+AL model are very similar to or the same as the Indian dataset in terms of mechanism and number of parameters, ensuring consistency in the sampling methodology. Therefore, a unified process is used across these data sets in terms of choosing the training epoch in each epoch and repeating these rounds by adding 15 training samples. This is done by systematically taking 2% of the samples from each category, so the total training samples it is 1055 out of a total of 54129, and the rest of the samples are for testing. This may be the difference between Salinas and the Indian data set, which is the number of samples and their categories. By having training sessions and their

duration consistent with previous data sets, this contributes to creating a unified framework for evaluation, which enhances the ability to draw conclusions and make decisions based on model performance in a reliable and objective manner.

Table 4.9 The accuracy percentages (%) and computational times methods of classification of SA without DR and with EPF

No	Classes	CNN +AL		WLS+CNN +AL		JWMF+ CNN+AL		FGS + CNN+AL		BF+ CNN +AL		SD + CNN +AL	
		W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	WO MRF	W MRF
1.	**Brocoli green weeds 1**	99.33	98.82	**100**	98.48	**100**	100.0	**100**	99.95	**100**	98.83	**100**	100
2.	**Brocoli green weeds 2**	99.50	99.75	**100**	99.81	**100**	99.89	**100**	99.89	**100**	99.53	**100**	99.78
3.	**Fallow**	87.19	86.45	89.09	87.13	86.42	86.50	89.24	93.39	**89.82**	88.94	**90.11**	91.78
4.	**Fallow rough plow**	93.48	93.26	93.48	93.78	94.24	96.19	94.10	**94.65**	94.21	92.83	**94.42**	75.39
5.	**Fallow smooth**	**99.37**	99.14	99.96	99.81	99.96	99.85	99.96	99.43	99.96	99.69	**99.96**	99.81
6.	**Stubble**	**98.39**	97.99	**100**	99.92	**100**	99.92	**100**	99.95	**100**	99.92	**100**	99.87
7.	**Celery**	98.21	97.27	99.37	98.20	**99.32**	98.54	99.37	98.06	99.37	97.38	**99.37**	98.46
8.	**Grapes untrainedrrrr**	91.77	83.75	**92.13**	82.56	**94.04**	76.04	91.45	75.98	92.05	76.42	91.97	73.90
9.	**Soil vinyard develop**	98.19	99.67	99.79	98.92	**99.79**	**99.77**	99.79	99.75	99.79	99.77	99.77	99.77
10.	**Corn senesced green weeds**	96.47	95.20	96.69	95.01	96.66	94.11	96.67	95.08	**97.12**	94.27	**97.23**	94.60
11	**Lettuce romaine 4wk**	**98.22**	96.92	**100**	98.18	**100**	98.09	**100**	98.18	**100**	97.80	**100**	96.65
12.	**Lettuce romaine 5wk**	99.25	98.09	99.95	98.25	**100**	98.09	99.95	97.99	**99.95**	97.82	**100**	97.67
13.	**Lettuce romaine 6wk**	97.61	97.77	**97.97**	97.77	97.66	97.77	98.37	97.66	**98.37**	98.33	97.66	98.11
14.	**Lettuce romaine 7wk**	96.66	96.61	98.38	95.04	98.38	97.33	98.37	97.23	**98.37**	98.38	**98.38**	96.47
15.	**Vinyard untrained**	84.29	77.29	**87.86**	80.38	85.20	91.84	**88.18**	88.33	**88.18**	93.33	87.02	93.05
16.	**Vinyard vertical trellis**	98.02	96.32	**98.59**	95.81	**98.59**	96.15	98.47	96.21	**98.53**	98.02	98.47	96.78
Overall Accuracy (OA)		95.56	92.58	97.04	93.59	96.50	93.11	**97.02**	93.24	97.02	94.01	**97.16**	92.86
Average Accuracy (AA)		96.00	94.65	97.08	94.94	96.89	95.63	97.12	95.73	**97.23**	95.70	**97.16**	94.51
Computation Time (s)		**869.0**	**869.0**	1001	987.5	1000.8	971.1	959.2	935.2	1045	1000	1089.6	1023.5

Table 3.4 functions as a guide, clearly showing the complete allocation of training and test samples for each class in the Salinas dataset. It is crucial to mention that the Salinas dataset is remarkable due to its unique features, which include a considerable amount of spectral information and significantly greater dimensions compared to the other datasets being considered. The consistent use of experimental design and systematic sampling in various datasets allows for a fair and consistent evaluation of the CNN+AL+MRF methodology's performance and effectiveness in handling hyperspectral image classification, even though there may be subtle variations in the attributes of the datasets. The results obtained from the Salinas dataset, which are presented precisely in tables 4.7, 4.8, and 4.9, show consistent alignment with the patterns discovered in previous experiments. Here, once again,

DR techniques and edge-preserving filters prove their ability to significantly improve the effectiveness of the active learning strategy and give high results for the implemented model. As a routine procedure, the combination of both DR and EPF showed distinct improvements, and this is noticeable when reading the OA scores, which are at least 98%. These scores exceed the baseline set by the original CNN+AL approach by a large margin of at least 3%.

Table 4.10 The accuracy percentages (%) and computational times methods of classification of SA with AVG and EPF

NO	Classes	AVG + CNN +AL		AVG+WLS+ CNN + AL		AVG+JWMF+ CNN+ AL		AVG +FGS+ CNN+ AL		AVG + BF + CNN + AL		AVG +SD + CNN + AL	
		W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF
1.	Brocoli green weeds 1	**100**	100	**100.0**	98.20	**100**	97.76	**100**	99.75	**100**	99.95	**100**	99.75
2.	Brocoli green weeds 2	**99.88**	99.72	**100.0**	99.97	**100**	99.41	**100**	99.92	**100**	99.52	**100**	96.23
3.	Fallow	**100**	99.95	99.54	99.90	**99.64**	100	99.56	99.80	99.54	100	99.54	100
4.	Fallow rough plow	**99.86**	98.99	99.64	83.61	99.55	93.52	99.64	99.86	99.64	99.07	**99.64**	97.26
5.	Fallow smooth	**98.66**	97.46	98.65	100	98.55	99.96	98.65	99.25	98.65	98.99	**98.65**	99.59
6.	Stubble	**99.95**	98.99	**99.99**	98.65	99.30	98.66	99.21	98.66	99.19	98.53	99.19	98.45
7.	Celery	98.68	98.15	**99.78**	92.61	99.78	98.26	**99.98**	99.56	99.78	98.91	**99.78**	96.64
8.	Grapes untrained	94.95	88.12	**96.82**	68.11	95.93	84.88	**96.84**	91.48	96.33	85.01	95.81	37.17
9.	Soil vinyard develop	99.80	99.45	99.79	99.79	99.79	99.68	**99.80**	99.79	**99.79**	99.79	99.79	99.79
10.	Corn senesced green weeds	89.63	83.11	98.50	85.11	98.53	95.41	**98.72**	95.36	98.50	93.82	**98.53**	94.00
11	Lettuce romaine 4wk	94.26	89.44	98.11	2.17	**98.72**	90.57	**98.13**	94.73	98.11	95.39	98.11	65.82
12.	Lettuce romaine 5wk	95.86	90.92	**97.81**	97.65	97.71	96.78	97.71	97.24	**97.76**	94.44	97.71	94.99
13.	Lettuce romaine 6wk	**96.68**	93.47	96.61	98.14	96.61	97.59	96.60	96.38	**97.20**	97.59	96.60	97.26
14.	Lettuce romaine 7wk	97.28	95.50	97.38	97.85	97.38	90.63	**97.47**	97.66	97.38	96.25	**97.38**	92.96
15.	Vinyard untrained	**89.92**	78.00	87.48	96.46	88.39	97.90	88.36	84.51	88.40	93.03	**89.32**	97.10
16.	Vinyard vertical trellis	**99.06**	97.39	98.23	98.28	**98.28**	94.23	98.23	93.95	98.23	95.34	98.22	95.11
Overall Accuracy (OA)		97.06	94.28	**98.01**	88.34	98.01	95.82	**98.05**	96.53	98.01	95.39	98.00	91.29
Average Accuracy (AA)		**97.15**	94.29	98.02	88.53	98.01	95.95	**98.06**	96.74	**98.03**	96.60	98.02	91.38
Computation Time (s)		859.6	853.1	885.6	870.0	934.5	984.2	842.7	803.2	848.5	827.2	840.1	823.9

CNN+AL+SD wMRF combination stands out as the best performer in Table 10, with an exceptional OA score of 97.16% and a similar high AA score of 97.16%. The AVG and PCA algorithms show exceptional performance in the CNN+AL+FGS wMRF approach, obtaining OA scores of 98.05% and 99.37%, respectively.

It highlights the variability in achieving optimal performance based on the data set and on the diversity and adaptability of these strategies with respect to the unique characteristics of the data. Classification efficiency is greatly improved through the use of various procedures and strategies, especially through the use of filtering

methods that eliminate noise or roughness from the data while retaining important edge features.

Table 4.11 The accuracy percentages (%) and computational times methods of classification of SA with PCA and EPF

No.	Classes	PCA+CNN+AL		PCA+WLS+ CNN+AL		PCA+JWMF+ CNN+AL		PCA+FGS + CNN+AL		PCA+ BF+ CNN+AL		PCA+SD + CNN+AL	
		W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF	W MRF	WO MRF
1.	Brocoli green weeds 1	**100**	99.65	**100**	**100**	**100**	99.25	**100**	100	**100**	100	**100**	100
2.	Brocoli green weeds 2	**98.81**	99.87	**100**	99.97	**100**	99.46	**100**	100	**100**	99.89	**100**	99.78
3.	Fallow	**99.23**	100	**100**	**100**	**100**	100	**100**	99.95	**100**	100	**100**	100
4.	Fallow rough plow	98.16	98.16	**99.35**	80.61	**99.42**	92.41	99.35	99.35	99.35	99.06	**99.35**	96.39
5.	Fallow smooth	98.25	98.34	99.17	100	99.32	97.30	**99.40**	99.47	99.36	99.25	**99.32**	99.29
6.	Stubble	**98.24**	98.13	**100**	100	**100**	99.21	**100**	100	**100**	99.97	**100**	99.87
7.	Celery	98.97	98.51	**99.97**	99.78	99.97	99.44	**100**	99.75	**100**	100	99.97	99.83
8.	Grapes untrained	98.78	96.98	98.77	84.31	**99.99**	92.41	99.52	94.79	99.78	93.19	**99.88**	89.72
9.	Soil vinyard develop	99.32	98.84	**99.79**	99.79	**99.79**	99.79	99.81	99.77	99.79	99.77	**99.81**	99.77
10.	Corn senesced green weeds	**99.71**	98.44	99.79	99.88	99.66	99.26	99.69	99.17	99.66	99.14	**99.69**	99.23
11	Lettuce romaine 4wk	**100**	99.71	**100**	7.99	**100**	89.55	**100**	99.62	**100**	99.81	**100**	95.36
12.	Lettuce romaine 5wk	98.12	98.59	**99.95**	99.63	99.53	97.70	**99.63**	99.84	99.53	99.48	99.58	99.37
13.	Lettuce romaine 6wk	98.70	96.60	**100**	99.12	**99.89**	99.23	**100**	99.01	**100**	99.67	**100**	99.23
14.	Lettuce romaine 7wk	98.81	95.21	99.91	98.87	**100**	96.05	**99.91**	96.32	99.91	98.59	99.91	99.15
15.	Vinyard untrained	84.02	68.51	**97.91**	99.65	83.00	96.90	**93.75**	93.52	89.36	95.03	86.65	97.23
16.	Vinyard vertical trellis	98.10	97.66	98.89	99.50	98.89	99.56	**98.89**	97.89	**98.89**	99.05	**98.94**	98.55
Overall Accuracy (OA)		97.47	96.41	**99.33**	94.27	98.09	97.05	**98.92**	97.74	98.32	97.68	98.04	97.06
Average Accuracy (AA)		97.95	96.45	**99.59**	91.82	98.72	97.34	**99.37**	98.65	**99.10**	98.87	98.94	98.30
Computation Time (s)		854.0	849.2	968.0	952.0	882.6	825.9	864.7	**821.3**	**817.2**	799.8	933.6	903.4

These improvements contribute cumulatively to HIC in terms of performance and efficiency, which demonstrates the various methods that increase the accuracy of classification results and opens the way for planning using other techniques that contribute effectively to this field.

As noted above, although accuracy and classification performance are improved, computational obstacles arise when applying edge-preserving filters to hyperspectral images with many spectral bands, especially when it is necessary to preserve small features.

This requires separate processing for each band or spectral layer, which increases the computing burden. On the Indian dataset, the CNN+AL+AVG method took 165.46 seconds to process.

However, when the WLS filter was combined with the CNN+AL+AVG approach, the processing time increased to 179.53 seconds. The rise in computational requirements is a result of the use of different algorithms across many spectral levels, resulting in longer processing times compared to simpler methods.

Managing spatial spectral images is a major problem as a result of requiring a delicate balance between computational complexity and the need to preserve fine edge information. Edge-preserving filters effectively enhance image features and reduce noise.

They significantly increase processing times. Therefore, it is important to carefully consider the balance between computational resources and the need to maintain high-quality spectral detail. Therefore, processing hyperspectral images efficiently while maintaining balance is a constant challenge. This requires careful consideration in choosing appropriate processing methods that are compatible with computational constraints and accuracy needs.

4.4.3.1 Classification Map of SA

The visual representation of Salinas data maps is illustrated in these figures.4.7,4.8,4.9. For remote sensing applications, map classifications help in making decisions about the characteristics or configurations of the scene and providing information about it, such as the categories to which it belongs.

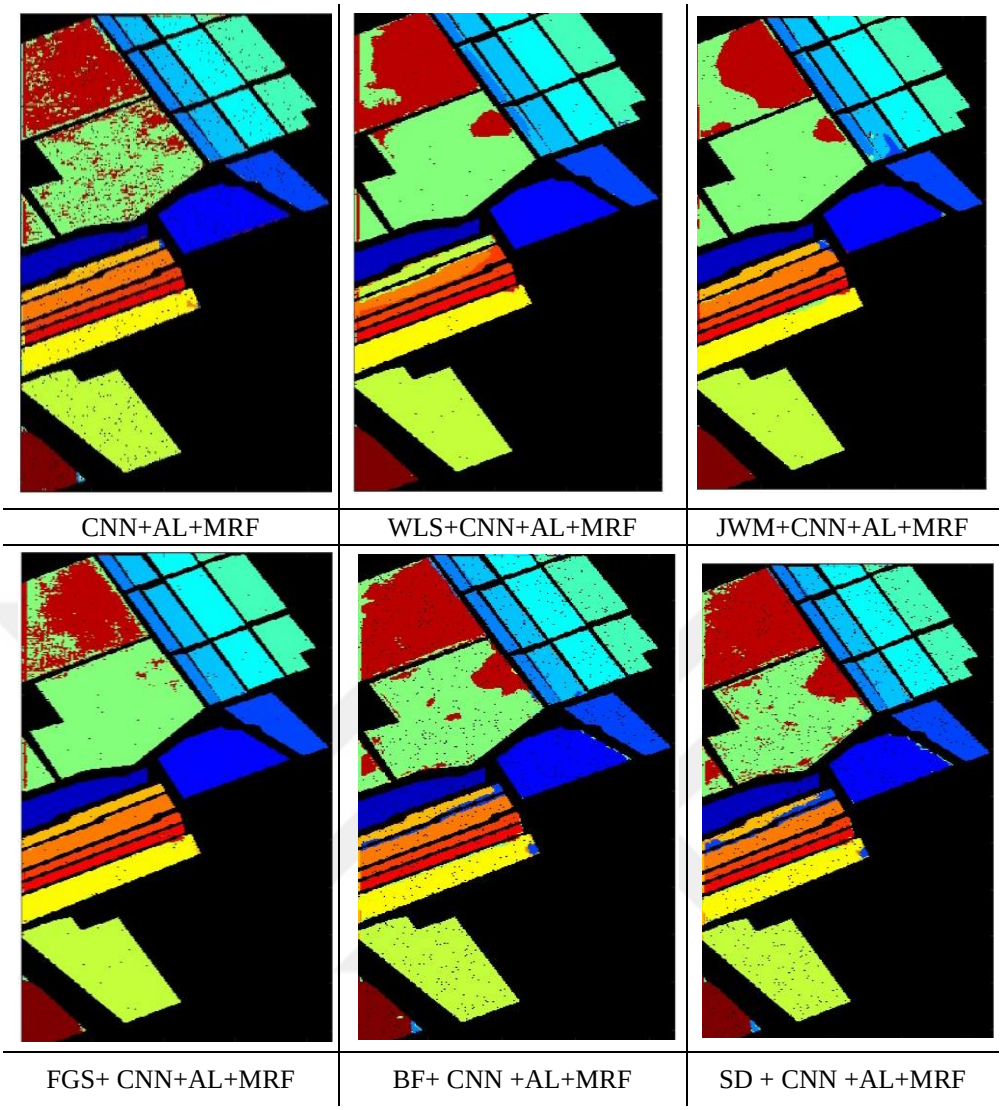


Figure 4.7 Classification maps for SA dataset using 2% training samples of the methods without dimension reduction

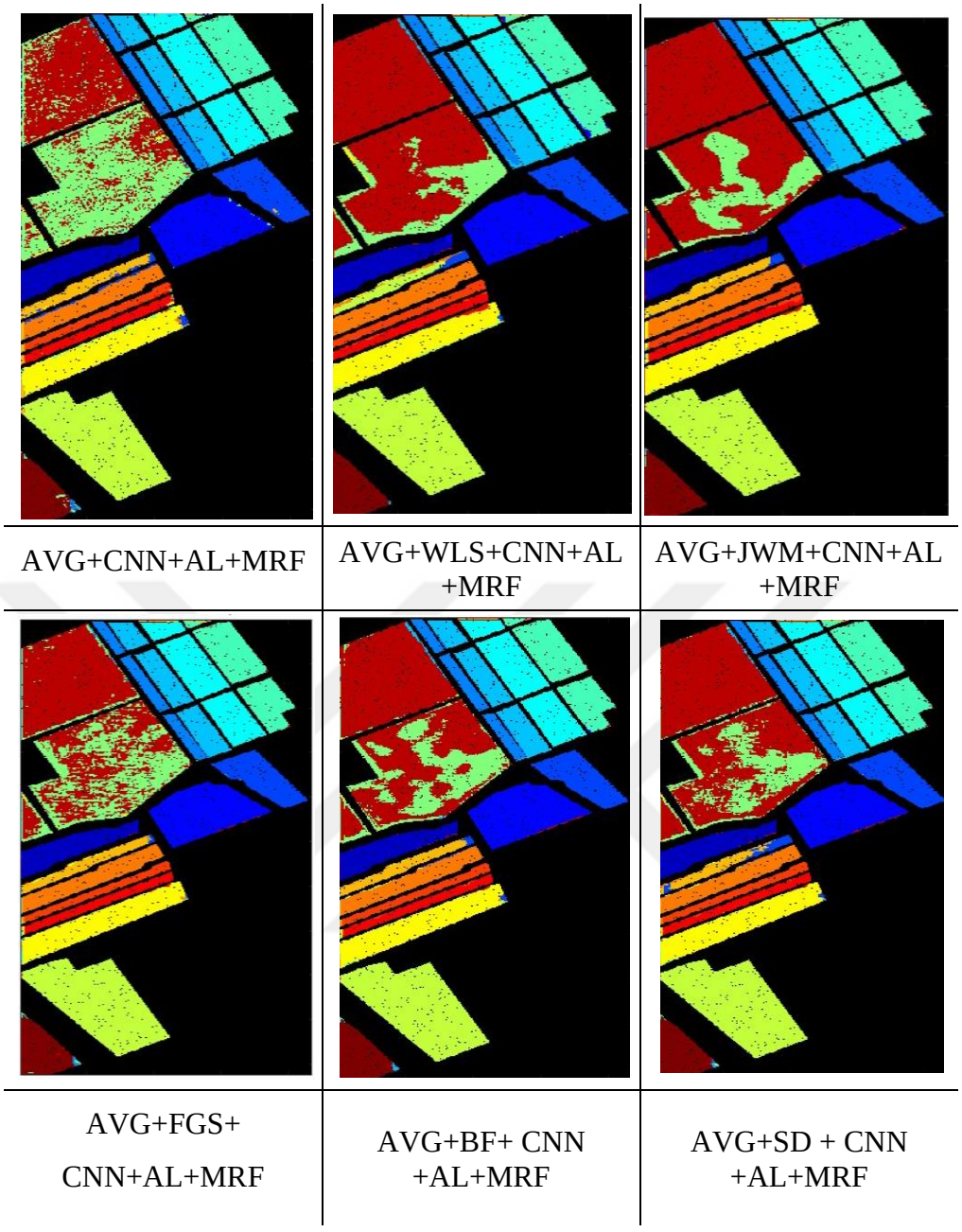


Figure 4.8 Classification maps for SA dataset using 2% training samples of the methods with AVG and EPF

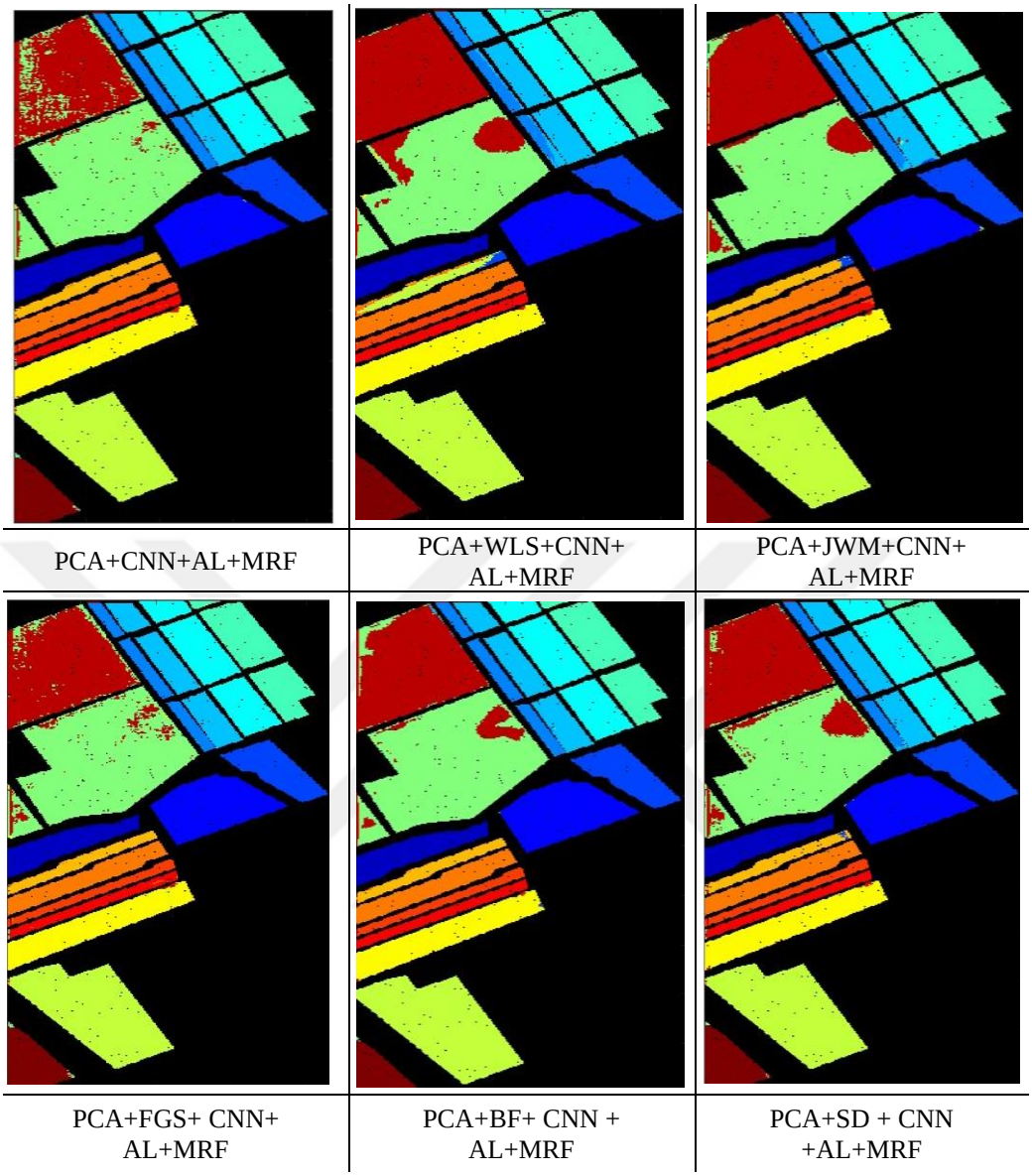


Figure 4.9 Classification maps for SA dataset using 2% training samples of the methods with PCA and EPF

4.5 Overall Comparison

The proposed method gave superior results; this is what we concluded from comparing the results of CNN+AL+MRF method [26] with the results obtained from applying dimension reduction techniques and the five filter techniques. During the experiments we don't consider CNN without AL, because in [26] 3D_CNN show poor results compared to CNN+AL+MRF. This proves the effectiveness of CNN with AL, as shown by the results for the CNN+AL+MRF method for the IP

dataset were 94.28%, while the training was 5%, and the results for the PU data were 96.40% with a training rate of 1%. While that Table 4.10 shows a summary of the best two techniques that gave the highest results for the three data used.

Table 4.12 Summarized the Best Results (%) for Used Techniques

No.	Data set	1st Best Technique	OA	2nd Best Technique	OA
1.	IP	AVG+SD+CNN+AL+MRF	99.03	PCA+SD+CNN+AL+MRF	99.02
2.	PU	AVG+SD+CNN+AL+MRF	98.00	PCA+SD+CNN+AL+MRF	97.00
3.	SA	AVG+FGS+CNN+AL+MRF	98.05	PCA+WLS+CNN+AL+MRF	99.33

In summary, the results show the effectiveness of pre-processing techniques when applied to the data set, achieving superior performance in classifying hyperspectral images, and one of the best techniques that contributed effectively was the static/dynamic filter technique, which had the highest score in most of the data results.

5
CONCLUSION

The enhancement of the classification achievement of hyperspectral images (HSIs) is critical for many uses, such as agricultural applications, environmental monitoring, and geological research. Active learning frameworks, particularly the CNN+AL approach, have gained recognition for addressing the issues caused by limited labeled samples. Nevertheless, the CNN+AL approach does not often perform processing operations on the spectral signatures of HSIs during the process. In this study, we present a novel strategy that incorporates pre-processing methods before utilizing the CNN+AL method. These pre-processing methods include a variety of aspect reduction approaches and edge-preserving filters. When it comes to preventing the dimensionality reduction process and removing redundant information from the data, dimensionality reduction methods are an extremely important resource.

Simultaneously, edge-preserving filters lessen the amount of noise, improve the smoothness of the image, and maintain the sharpness of the class borders inside the HSI. Additionally, edge-preserving filtering algorithms help make interclass variances higher and intraclass variances lower. This lowers the amount of spectral variability that exists within each class in the end. Three sets of reference data, in particular Indian Pines, Pavia University, and Salinas, are used to conduct a thorough evaluation of the methodology.

These datasets reflect different scenarios and characteristics. The architecture incorporates several alternative edge-preserving filters and applies PCA and AVG as approaches for dimension reduction during the process. In terms of classwise OA and AA, the results suggest that the pre-processing stage in active learning is far superior to the CNN+AL approach.

This describes the major significance of the pre-processing step in active learning methodology. Filtering algorithms lead to an increase in the total processing time of the CNN+AL framework, but dimension reduction methods contribute to a drop in the training time of CNNs. Dimension reduction methods reduce the training time of CNNs. When AVG technology was combined with SD Filter, Indian Pines

and Pavia University achieved the highest percentages of (99.03%, 98.00%) respectively, while Salinas achieved the highest result of 99.33% when PCA was combined with WLS. This all-encompassing approach shows how important pre-processing techniques are for improving the active learning process for HSI classification, which leads to better accuracy in classification and higher operational efficiency. The results clearly demonstrate that applying our meticulously selected filter technique and dimension reduction greatly enhances the classification achievement. for example, "Grasstrees" at Indian Pines, " Painted metal sheets" at Pavia University, "Brocoli green weeds 2" and " Lettuce romaine 4wk" at Salinas, and nearly all the land cover classes reach near-perfect accuracy, proving that our approach is resilient.

Future work in hyperspectral image classification (HSI) will concentrate on refining methods for pre-processing images. Making improvements to pipelines, investigating novel dimensionality reduction methods, and fine-tuning edge-preserving filters are all part of this. Enhancing feature extraction while simultaneously decreasing noise and preserving data integrity is the objective. Integration of spatio-temporal data into HSI is critical for temporal computation, which is gaining attention because it improves classification accuracy and enriches analytical depth.

REFERENCES

[1] G. M. Foddy, “Natalia Kosol, Mykola Lavryniuk, Sergey Skakon, and Andrey Shelestov. Deep learning classification of land cover and crop types using remote sensing data,” IEEE Geoscience and Remote Sensing Letters, vol. 80, no. 1, 2002.

[2] M. S. Navin and L. Agilandeeswari, “Multispectral and hyperspectral images-based land use / land cover change prediction analysis: an extensive review,” Multimed. Tools Appl., vol. 79, no. 39–40, pp. 29751–29774, 2020.

[3] P. Wang, S. Xu, Y. Li, J. Wang, and S. Liu, “Hyperspectral image classification based on joint sparsity model with low-dimensional spectral-spatial features,” Journal of Applied Remote Sensing, vol. 11, no. 1, pp. 15010–015010, 2017.

[4] D.-W. Sun, “Hyperspectral Imaging for Food Quality Analysis and Control,” Elsevier, 2010.

[5] M. José, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, “Hyperspectral remote sensing data analysis and future challenges,” IEEE Geoscience and remote sensing magazine, vol. 1, no. 2, pp. 6–36, 2013.

[6] ” Raytheon.com. [Online]. Available: http://www.raytheon.com. [Accessed: 14-Jan-2024].

[7] Y.-Q. Wan, Y.-H. Fan, and M.-S. Jin, “Application of hyperspectral remote sensing for supplementary investigation of polymetallic deposits in Huaniushan ore region, northwestern China,” Sci. Rep., vol. 11, no. 1, p. 440, 2021.

[8] Hyperspectral imaging,” Mathworks.com. [Online]. Available: https://www.mathworks.com/discovery/hyperspectral-imaging.html. [Accessed: 14-Jan-2024].

[9] “Hyperspectral Imaging,” Fas.org. [Online]. Available: https://irp.fas.org/imint/hyper.htm. [Accessed: 12-Jan-2024.

[10] H. Peng and K. B. Zhang, “Active deep learning for classification of hyperspectral images,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 10, no. 2, pp. 712–724, 2016.

[11] F. Pacifici, M. Chini, and W. J. Emery, “A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification,” Remote Sens. Environ., vol. 113, no. 6, pp. 1276–1292, 2009.

[12] E. Izquierdo-Verdiguier, L. Gomez-Chova, L. Bruzzone, and G. Camps-Valls, “Semisupervised kernel feature extraction for remote sensing image analysis,” IEEE Trans. Geosci. Remote Sens., vol. 52, no. 9, pp. 5567–5578, 2014.

[13] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, “Deep feature extraction and classification of hyperspectral images based on convolutional neural

networks," IEEE Trans. Geosci. Remote Sens., vol. 54, no. 10, pp. 6232–6251, 2016.

[14] L. Hu, X. Luo, and Y. Wei, "Hyperspectral image classification of convolutional neural network combined with valuable samples," J. Phys. Conf. Ser., vol. 1549, no. 5, p. 052011, 2020.

[15] D. Tuia, M. Volpi, L. Copa, M. Kanevski, and J. Munoz-Mari, "A survey of active learning algorithms for supervised remote sensing image classification," IEEE J. Sel. Top. Signal Process., vol. 5, no. 3, pp. 606–617, 2011

[16] M. Chen, Q. Wang, and X. Li, "Discriminant analysis with graph learning for hyperspectral image classification," Remote Sens. (Basel), vol. 10, no. 6, p. 836, 2018.

[17] 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., vol. 56, no. 11, pp. 6440–6461, 2018.

[18] D. Zhang, J. Kang, L. Xun, and Y. Huang, "Hyperspectral image classification using spatial and edge features based on deep learning," Intern. J. Pattern Recognit. Artif. Intell., vol. 33, no. 09, p. 1954027, 2019.

[19] P. Ren et al., "A Survey of Deep Active Learning," ACM Computing Surveys, vol. 54, no. 9, pp. 1–40, Oct. 2021, doi: 10.1145/3472291.

[20] D. Tuia, F. Pacifici, M. Kanevski, and W. J. Emery, "Classification of very high spatial resolution imagery using mathematical morphology and support vector machines," IEEE Trans. Geosci. Remote Sens., vol. 47, no. 11, pp. 3866–3879, 2009.

[21] F. Shi, Z. Wang, M. Hu, and G. Zhai, "Active learning plus deep learning can establish cost-effective and robust model for multichannel image: A case on hyperspectral image classification," Sensors (Basel), vol. 20, no. 17, p. 4975, 2020.

[22] H. Bi, F. Xu, Z. Wei, Y. Xue, and Z. Xu, "An active deep learning approach for minimally supervised PolSAR image classification," IEEE Trans. Geosci. Remote Sens., vol. 57, no. 11, pp. 9378–9395, 2019.

[23] M. E. Ieee, S. Paoletti, and S. Junli, Active Learning with Convolutional Neural Networks for Hyperspectral Image Classification Using a New Bayesian Approach Juan Mario Haut Javier Plaza, Student Member. IEEE.

[24] P. Liu, H. Zhang, and K. B. Eom, "Active deep learning for classification of hyperspectral images," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 10, no. 2, pp. 712–724, 2017.

[25] Q. Wang, M. Chen, J. Zhang, S. Kang, and Y. Wang, "Improved active deep learning for semi-supervised classification of hyperspectral image," Remote Sens. (Basel), vol. 14, no. 1, p. 171, 2021.

[26] X. Cao, F. Zhou, L. Xu, D. Meng, Z. Xu, and J. Paisley, "Hyperspectral image classification with Markov random fields and a convolutional neural network," IEEE Transactions on image processing, vol. 27, no. 5, pp. 2354–2367, 2018.

[27] V. Gupta, S. Sastry, and S. K. Mitra, “Hyperspectral image classification using trilateral filter and deep learning,” in 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), 2020.

[28] X. Kang, P. Duan, and S. Li, “Hyperspectral image visualization with edge-preserving filtering and principal component analysis,” Inf. Fusion, vol. 57, pp. 130–143, 2020.

[29] Y. Shi et al., “Spectral-spatial attention rotation-invariant classification network for airborne hyperspectral images,” Drones, vol. 7, no. 4, p. 240, 2023.

[30] J. Jiang, J. Ma, C. Chen, Z. Wang, Z. Cai, and L. Wang, “SuperPCA: A Super pixelwise PCA Approach for Unsupervised Feature Extraction of Hyperspectral Imagery,” IEEE Trans. Geosci. Remote Sens., vol. 56, no. 8, pp. 4581–4593, 2018.

[31] L. Fang, S. Li, X. Kang, and J. A. Benediktsson, “Spectral–spatial classification of hyperspectral images with a super pixel-based discriminative sparse model,” IEEE Trans. Geosci. Remote Sens., vol. 53, no. 8, pp. 4186–4201, 2015.

[32] F. Zhu, Z. Liang, X. Jia, L. Zhang, and Y. Yu, “A benchmark for edge-preserving image smoothing,” IEEE Trans. Image Process., vol. 28, no. 7, pp. 3556–3570, 2019.

[33] C. Pan, X. Jia, J. Li, and X. Gao, “Adaptive edge preserving maps in Markov random fields for hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 10, pp. 8568–8583, 2021.

[34] X. Kang, S. Li, and J. A. Benediktsson, “Spectral–spatial hyperspectral image classification with edge-preserving filtering,” IEEE Trans. Geosci. Remote Sens., vol. 52, no. 5, pp. 2666–2677, 2014.

[35] S. Zhong, C. Chang, and Y. Zhang, “Iterative Edge Preserving Filtering Approach to Hyperspectral Image Classification,” IEEE Geoscience and Remote Sensing Letters, vol. 16, no. 1, pp. 90–94, Jan. 2019, doi: 10.1109/lgrs.2018.2868841.

[36] Z. Zhang, “Semi-supervised hyperspectral image classification algorithm based on graph embedding and discriminative spatial information,” Microprocessors and Microsystems, vol. 75, 2020.

[37] Z. Xie and M. Li, “Semi-Supervised AUC Optimization Without Guessing Labels of Unlabeled Data,” Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32, no. 1, Apr. 2018, doi: 10.1609/aaai. v32i1.11812.

[38] J. Xiangyong, Z. Yao, and D. Xu, “Hyperspectral image classification with convolutional neural network and active learning,” IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 7, pp. 4604–4616, 2020.

[39] F. Melgani and L. Bruzzone, “Classification of hyperspectral remote sensing images with support vector machines,” IEEE Trans. Geosci. Remote Sens., vol. 42, no. 8, pp. 1778–1790, 2004.

[40] L. Sun, Z. Wu, J. Liu, L. Xiao, and Z. Wei, “Supervised spectral–spatial hyperspectral image classification with weighted Markov random fields,” IEEE Trans. Geosci. Remote Sens., vol. 53, no. 3, pp. 1490–1503, 2015.

[41] R. Dianat and S. Kasaei, “Dimension reduction of optical remote sensing images via minimum change rate deviation method,” IEEE Trans. Geosci. Remote Sens., vol. 48, no. 1, pp. 198–206, 2010.

[42] C. Guilloteau, T. Oberlin, O. Berné, and N. Dobigeon, “Hyperspectral and multispectral image fusion under spectrally varying spatial blurs-Application to high dimensional infrared astronomical imaging,” IEEE Transactions on Computational Imaging, vol. 6, pp. 1362–1374, 2020.

[43] S. Ayesha, M. K. Hanif, and R. Talib, “Overview and comparative study of dimensionality reduction techniques for high dimensional data,” Inf. Fusion, vol. 59, pp. 44–58, 2020.

[44] X. Cao, J. Yao, X. Fu, H. Bi, and D. Hong, “An enhanced 3-D discrete wavelet transform for hyperspectral image classification,” IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 6, pp. 1104–1108, 2020.

[45] J. Ieee, L. Liu, and Z. Xiao, Supervised Spectral-Spatial Hyperspectral Image Classification with Weighted Markov Random Fields Le Sun, Student Member.

[46] J. Khodr and R. Younes, “Dimensionality reduction on hyperspectral images: A comparative review based on artificial datas,” in 2011 4th International Congress on Image and Signal Processing, 2011.

[47] K. Thilagavathi, R. Nagendran, and I. T. B. Mary, “Hyperspectral image classification using ensemble average method,” in 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), IEEE, 2021, pp. 1–6.

[48] Y. Cai, X. Liu, and Z. Cai, “BS-nets: An end-to-end framework for band selection of hyperspectral image,” IEEE Trans. Geosci. Remote Sens., vol. 58, no. 3, pp. 1969–1984, 2020.

[49] L. O. Jimenez and D. A. Landgrebe, “aSupervised classification in high dimensional space: Geometrical, statistical, and asymptotical proper ties of multivariate data,” IEEE Trans. Syst.

[50] A. Paul, P. Sahoo, and N. Chaki, “Dimensionality reduction of hyperspectral images: A data-driven approach for band selection,” in Advances in Intelligent Systems and Computing, Singapore: Springer Singapore, 2020, pp. 11–27.

[51] L. Lebart and G. Saporta, “Historical elements of correspondence analysis and multiple correspondence analysis,” in Visualization and verbalization of data, 2014, pp. 31–44.

[52] R. R. Arvind, “Principal component analysis,” Devopedia, 12-Jan-2019. [Online]. Available: https://devopedia.org/principal-component-analysis. [Accessed: 14-Jan-2024].

[53] J. Nalepa, M. Myller, and M. Kawulok, “Training- and Test-Time Data Augmentation for Hyperspectral Image Segmentation,” IEEE Geosci. Remote Sens. Lett., vol. 17, no. 2, pp. 292–296, 2020.

[54] C. Wang, L. Zhang, W. Wei, and Y. Zhang, “Hyperspectral Image Classification with Data Augmentation and Classifier Fusion,” IEEE Geosci. Remote Sens. Lett., vol. 17, no. 8, pp. 1420–1424, 2020.

[55] L. Perez and J. Wang, “The effectiveness of data augmentation in image classification using deep learning,” arXiv [cs.CV], 2017.

[56] Y. Yang, H. Hui, L. Zeng, Y. Zhao, Y. Zhan, and T. Yan, “Edge-preserving image filtering based on soft clustering,” IEEE Trans. Circuits Syst. Video Technol., vol. 32, no. 7, pp. 4150–4162, 2022.

[57] L. Xu, C. Lu, Y. Xu, and J. Jia, “Image smoothing via L 0 gradient minimization,” in Proceedings of the 2011 SIGGRAPH Asia conference, 2011, pp. 1–12.

[58] W. Liu, P. Zhang, X. Chen, C. Shen, X. Huang, and J. Yang, “Embedding bilateral filter in least squares for efficient edge-preserving image smoothing,” IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 1, pp. 23–35, 2020

[59] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson, “Advances in hyperspectral image classification: Earth monitoring with statistical learning methods,” IEEE Signal Process. Mag., vol. 31, no. 1, pp. 45–54, 2014.

[60] J. Ren, J. Zabalza, S. Marshall, and J. Zheng, “Effective feature extraction and data reduction in remote sensing using hyperspectral imaging [applications corner],” IEEE Signal Process. Mag., vol. 31, no. 4, pp. 149–154, 2014.

[61] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, “Digital photography with flash and no-flash image pairs,” ACM Trans. Graph., vol. 23, no. 3, pp. 664–672, 2004.

[62] X. Cao, “Hyperspectral/Multispectral Imaging,” in Computer Vision: A Reference Guide, Cham: Springer International Publishing, 2021, pp. 592–598.

[63] C. O. Ancuti and C. Ancuti, “Single Image Dehazing by Multi-Scale Fusion,” IEEE Trans. Image Process., vol. 22, no. 8, pp. 3271–3282, 2013.

[64] X. Quan and W. He, “Image reconstruction algorithm based on guided image filtering in parallel-beam CT,” in 2018 14th International Conference on Computational Intelligence and Security (CIS), 2018.

[65] C. Zhou, H. Zhang, X. Shen, and J. Jia, “Unsupervised learning of stereo matching,” in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 1567–1575.

[66] X. Pan, F. Xie, Z. Jiang, and J. Yin, “Haze removal for a single remote sensing image based on deformed haze imaging model,” IEEE Signal Process. Lett., vol. 22, no. 10, pp. 1806–1810, 2015.

[67] L. Zhang, Q. Yan, Y. Zhu, X. Zhang, and C. Xiao, “Effective shadow removal via multi-scale image decomposition,” Vis. Comput., vol. 35, no. 6–8, pp. 1091–1104, 2019.

[68] X. Wan and S. Chen, “Hyperspectral image classification using improved multi-scale block local binary pattern and bi-exponential edge-preserving smoother,” European Journal of Remote Sensing, vol. 56, no. 1, 2023.

[69] H. Lv, Z. Wang, and H. Zhang, “Edge protection filtering and convolutional neural network for hyperspectral remote sensing image classification,” Infrared Physics & Technology, vol. 122, 2022

[70] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, “Edge-preserving decompositions for multi-scale tone and detail manipulation,” ACM Trans. Graph., vol. 27, no. 3, pp. 1–10, 2008.

[71] Y. Yang, D. Wu, L. Zeng, and Z. Li, “Weighted least square filter via deep unsupervised learning,” Multimedia Tools and Applications, pp. 1–17, 2023.

[72] Tappen and Freeman, “Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters,” in Proceedings Ninth IEEE International Conference on Computer Vision, 2003.

[73] I. Hassan and A. Hussain, “Image Segmentation using Weighted Average Local Histogram,” International Journal of Computer Applications, vol. 64, no. 19, pp. 37–41, Feb. 2013, doi: 10.5120/10745-5598.

[74] J. Van de Weijer and R. Van den Boomgaard, “Local mode filtering,” in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 2005.

[75] Y. Shao, S. Du, and H. Tang, “An extended bi-dimensional empirical wavelet transforms based filtering approach for engineering surface separation using high-definition metrology,” Measurement, vol. 178, p. 109259, Jun. 2021, doi: 10.1016/j.measurement.2021.109259.

[76] R. Kilik, “Histogram-based weighted median filtering used for noise reduction of digital elevation model data,” Acta Geod. Geophys., vol. 56, no. 4, pp. 743–764, 2021.

[77] C. Duan, C. Xing, Y. Liu, and Z. Wang, “Fusion of infrared and visible images using fast global smoothing decomposition and target-enhanced parallel gaussian fuzzy logic,” Sensors, vol. 22, no. 1, 2021.

[78] D. Min, S. Choi, J. Lu, B. Ham, K. Sohn, and M. N. Do, “Fast global image smoothing based on weighted least squares,” IEEE Trans. Image Process., vol. 23, no. 12, pp. 5638–5653, 2014.

[79] S. Reich, A. Abramov, J. Papon, F. Wörgötter, and B. Dellen, A novel real-time edge-preserving smoothing filter. 2013.

[80] K. Kotwal and S. Chaudhuri, “Visualization of hyperspectral images using bilateral filtering,” IEEE Trans. Geosci. Remote Sens., vol. 48, no. 5, pp. 2308–2316, 2010.

[81] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), 2002.

[82] L. Dai, L. Tang, and J. Tang, “Speed up bilateral filtering via sparse approximation on a learned cosine dictionary,” IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 3, pp. 603–617, 2020.

[83] W. Liu, P. Zhang, X. Chen, C. Shen, X. Huang, and J. Yang, “Embedding bilateral filter in least squares for efficient edge-preserving image smoothing,” arXiv [cs.GR], 2018.

[84] M. Zhang, Bilateral filter in image processing. 2009.

[85] C. Zhao, X. Wan, and Y. Yan, “Spectral-spatial classification of hyperspectral images based on joint bilateral filter and stacked sparse autoencoder,” in 2017

First International Conference on Electronics Instrumentation & Information Systems (EIIS), 2017.

[86] B. Ham, M. Cho, and J. Ponce, “Robust guided image filtering using nonconvex potentials,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 1, pp. 192–207, 2018.

[87] L. Pompilio, M. Pepe, and G. Candiani, “Hyperdimensional data exploitation through parametric reduction,” in 2014 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2014.

[88] B. Ham, M. Cho, and J. Ponce, “Robust image filtering using joint static and dynamic guidance,” in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[89] B. Kumar, O. Dikshit, A. Gupta, and M. K. Singh, “Feature extraction for hyperspectral image classification: a review,” Int. J. Remote Sens., vol. 41, no. 16, pp. 6248–6287, 2020.

[90] E. Myasnikov and Samara National Research University, “Hyperspectral image segmentation using dimensionality reduction and classical segmentation approaches,” Comput. Opt., vol. 41, no. 4, pp. 564–572, 2017.

[91] X. Hu, X. Chen, B. Ni, T. Li, and Y. Liu, “Bi-volution: A static and dynamic coupled filter,” Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36, pp. 960–968, 2022.

[92] A review: Data pre-processing and data augmentation techniques Kiran Maharana ∗, Surajit Mondal, Bhushankumar Nemade Department of Information Technology. India.

[93] H. Li, J. Cui, X. Zhang, Y. Han, and L. Cao, “Dimensionality Reduction and Classification of Hyperspectral Remote Sensing Image Feature Extraction,” Remote Sensing, vol. 14, no. 18, 2022.

[94] I. T. Jolliffe and J. Cadima, “Principal component analysis: a review and recent developments,” Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences, vol. 374, 2016.

[95] C. Wang et al., “A review of deep learning used in the hyperspectral image analysis for agriculture,” Artif. Intell. Rev., vol. 54, no. 7, pp. 5205–5253, 2021.

PUBLICATIONS FROM THE THESIS

Conference Papers

1. Z. D. AL-SAMMARRAIE and A. C. Karaca, “Modified active learning framework for hyperspectral image classification with convolutions neural networks,” in 2023 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2023.