Chapter 4

Convolution Neural Network(CNN) and correlation based Salient Features

This chapter focuses on two methods for feature extraction followed by a salient feature selection for image labelling, as described in objective-2. We have explored the 'off-the-shelf' CNN features and the expansion of low correlated features in experimental analysis-1 and 2. Salient features are selected based on the MO signal that we have created by computing the signal peaks. We have applied the SVM classifier to train and predict the semantic labels for images.

4.1 Introduction

Deep networks have escalated the computational performance in the sensor-based high dimensional(HD) imaging such as Hyperspectral images(HSI), due to their informative feature extraction competency. Therefore in this work, we have extracted the informative features from different CNN models for the benchmark HD image datasets. The deep features have concatenated with spectral features to increase the informative knowledge in the image datacube. The feature concatenation has massively increased the size of datacube. Therefore, we have applied an unsupervised maximum object identification based salient feature selection on identifying the most informative features of datacube and discarding the less informative features. This leads to reducing the computational

time without compromising the accuracy. An unsupervised feature selection approach that transforms the data into scale space and achieves robust and strong features. In the previous CNN based methods, raw features have directly fed to the MLP(multilayer perception) layers for target prediction whereas we have provided our salient features into a multi-core SVM based set-up and have achieved high accuracy with low computational time as compared to the previous state-of-art

4.2 Proposed Method

4.2.1 Outline of the Proposed scheme

In the first step, 'off-the-shelf' CNN has been used for extracting the spatial features. Section-4.2.2 describes the first step. In the second step, scale-space saliency learning has applied to extracted spatial features. Section-4.2.3 describes the step-2. The step-2 has followed by an SVM based classifier training with minimal training samples to achieve the maximum accuracy and minimum computation cost.

4.2.2 Method-1 for Feature extraction: Deep CNN Spatial Feature by 'off-the-self' networks

In the forward propagation of Deep CNN architecture, the previous layer's features maintain the higher spatial resolution for contextual information extraction. In contrast, subsequent layers occupy more semantic information and less spatial knowledge. A fully connected layer exhibits the small spatial resolution, and we are concerned with exact spatial details. For CNN based spatial feature extraction, we have used the approach adapted in [49]. A hierarchical level of up-sampling on the convolution layers has been performed to extract the spatial features. Up-sampling has performed by applying the activation on convolution layers. Deep spatial features for an HSI pixel has been described as an

array of activation for all CNN convolution unit at one pixel. To determine the spatial features of a given image, the principal components of all spectral channels have extracted. Subsequently, their hierarchical convolution features are drawn by 'off-the-shelf' CNN's namely AlexNet [57], VGG-16 [47],VGG-19 [47, 64], trained on ImageNet data for 1000 classes. Up-sampling and pooling operation reduces the spatial resolution of a feature map as the depth of the convolution layer increases. To resolve this issue, we have up-sampled the feature maps by using bi-cubic interpolation. Finally, up-sampled feature maps are concatenated from each convolution layer in a long array of pixel vector, which we have described as deep spatial features for that position. For example, convolution layer-1 to 5 in AlexNet [[57]] architecture exhibits 96, 256, 384 and 256 features that generates 1376 cumulative spatial features for every pixel.

4.2.3 Method-2 for Feature extraction: Expansion of weakly Correlated Features

Weakly correlated bands are a great source of information for accurate classification. Therefore, we have generated low correlated intermediate features and reduced the highly correlated features to obtain more informative image datacube. The idea of feature expansion in equation-4.1 to 4.11 is taken and modified from [72]. Let, we have an HD image

$$\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_D]^{\mathrm{T}} \varepsilon \mathrm{R}^{D \times I} \tag{4.1}$$

having I pixels and D feature dimensions. $\mathbf{p}_d = [\mathbf{p}_{d,1}, ..., p_{d,i}]$, is the image corresponding to the d^{th} feature dimensions having wavelength λ_d, where $[\mathbf{p}_{d,1}, ..\mathbf{p}_{d,i}]$ are the pixel values i in d dimensions. At first, we examine the weakly correlated pairs of dimensions which are adjacent to each other. If d and $d+1$ are two weakly correlated adjacent dimensions, then two intermediate dimensions have been generated which contains a similar correlation as d and $d+1$. Such virtually created dimensions are simply concatenated with our HD image to obtain more informative datacube. The inclusion of intermediate dimensions in HD image will lead to increase the size, therefore, in the next iteration of our expansion algorithm, we have transferred only one feature from a pair of two highly correlated features. Let $\mathbf{p}^{(s)}$ is the image after s^{th} iteration and two adjacent dimensions are $\mathbf{p}_d^{(s)}$ and $\mathbf{p}_{d+1}^{(s)}$ with wavelength $\lambda_d^{(s)}$ and $\lambda_{d+1}^{(s)}$ respectively. $\mathbf{p}_d^{(s)}$ and $\mathbf{p}_{d+1}^{(s)}$ have generated the intermediate dimensions, if correlation coefficient is less than a

minimum threshold .i.e. $c_{\min}$=0.8, in our experiment.

$$c_{d,d+1}^{(s)} < c_{\min} \tag{4.2}$$

where as correlation $c_{d,d+1}^{(s)}$ is calculated as:

$$c_{d,d+1}^{(s)} = \frac{\sum_{i=1}^{I}(\mathbf{p}_{d,i}^{(s)} - \hat{\mathbf{p}}_{d}^{(s)})(\mathbf{p}_{d+1,i}^{s} - \hat{\mathbf{p}}_{d+1}^{(s)})}{\sqrt{\sum_{i=1}^{I}(\mathbf{p}_{d,i}^{(s)} - \hat{\mathbf{p}}_{d}^{(s)})^2}\sqrt{\sum_{i=1}^{I}(\mathbf{p}_{d,i}^{(s)} - \hat{\mathbf{p}}_{d+1}^{(s)})^2}} \tag{4.3}$$

where i is the index of pixels in HD image and $\hat{\mathbf{p}}_d$ is mean pixel value in d^{th} dimension .i.e.

$$\hat{\mathbf{p}}_d = \frac{1}{I}\sum_{i=1}^{I}\mathbf{p}_{d,i} \tag{4.4}$$

two intermediate images .i.e $\mathbf{p}_{d+1}^{(s+1)}$ and $\mathbf{p}_{d+2}^{(s+1)}$ have been generated in $(s+1)_{th}$ iteration. The image channel which is already present in $(d+1)_{th}$ dimension has been shifted to $(d+3)^{rd}$ dimension. Therefore, for $\mathbf{p}_{d+1}^{(s+1)}$ and $\mathbf{p}_{d+2}^{(s+1)}$ feature in $(s+1)^{th}$ iteration, we have updated the HD image. The updation rules for new HD image are given as:

$$\mathbf{p}_{d}^{(s+1)} = \mathbf{p}_{d}^{(s)} \tag{4.5}$$

$$\mathbf{p}_{d+1}^{(s+1)} = \mu(\mathbf{p}_{d}^{(s)} + \mathbf{p}_{d+1}^{(s)}) \tag{4.6}$$

$$\mathbf{p}_{d+2}^{(s+1)} = \left(\frac{1}{10\mu}\right)(\mathbf{p}_{d}^{(s)} + \mathbf{p}_{d+1}^{(s)}) \tag{4.7}$$

$$\mathbf{p}_{d+3}^{(s+1)} = \mathbf{p}_{d+1}^{(s)} \tag{4.8}$$

$\mu\varepsilon(0,1)$ is a weight parameter for new dimensions. The updation rules have resulted in new intermediate dimensions as $(\mathbf{p}_{d+1}^{(s+1)}, \mathbf{p}_{d+2}^{(s+1)})$, which have similar correlation as $\mathbf{p}_{d}^{(s)}$ and $\mathbf{p}_{d+1}^{(s)}$. For a new channels, the wavelength has increased by a step length of $\triangle\lambda^{(s)}$. The wavelengths of intermediate channels are $\lambda_{d+1}^{(s+1)}$ and $\lambda_{d+2}^{(s+1)}$, which lies between $\lambda_{d}^{(s)}$ and $\lambda_{d+1}^{(s)}$ as:

$$\lambda_{d+1}^{(s+1)} = \lambda_{d}^{(s)} + \triangle\lambda^{(s)} \tag{4.9}$$

$$\lambda_{d+2}^{(s+1)} = \lambda_{d+1}^{(s)} - \triangle\lambda^{(s)} \tag{4.10}$$

The image pair, .i.e. $(\mathbf{p}_d^{(s)}, \mathbf{p}_{d+1}^{(s)})$ will not split if correlation $c_{k,k+1}^{(s)} > \mathrm{c}_{\min}$. In this case, $\mathbf{p}_d^{(s)}$ will transfer to next state as: $\mathbf{p}_d^{(s+1)} = \mathbf{p}_d^{(s)}$, where $\triangle\lambda^{(s)}$ is updated as:

$$\triangle \lambda^{(s+1)} = \mu \triangle \lambda^{(s)} \tag{4.11}$$

such update will terminate when $\triangle\lambda^{(s+1)}$ falls down to a minimum step wavelength.

4.2.4 Salient Features selection using Scale identification for maximum object(MO) classification

Salient channels are more significant than other channels for classification. It is not possible to conclude which band or channel is carrying more information, other than noisy channels. The author [55] have analytically observed the salient features from scale identification and found that the image descriptors and scales are fundamentally related [63, 71]. The salient features reduce the computation requirement and increases the classification accuracy. The ability to classify an image object [61] is a remarkable aspect of scale selection in salient features. Therefore these features contain maximum possible information in an image.

4.2.4.1 Significant feature selection

In this section, we have used two mathematical methods, which are represented as functions f1 and f2, to identify interesting local patterns in HD images. Such methods are the outcome of the Hessian matrix. We have adapted and modified the Hessian-based operations from the [63] for local pattern extraction; therefore, we refer the readers to [63] for better understanding of equation-4.12 to equation-4.20. The Hessian matrix operations are used as an auxiliary method to determine the most interesting features in neighbourhood that have the expertise to identify the local information. We have defined the f1 and f2 operations for HD image P, where p,q are x and y coordinate directions of HD-HSI, as:

$$f1(P) = P_{pp} + P_{qq} \tag{4.12}$$

$$f2(P) = P_{pp}P_{qq} + P_{pq}^2 \tag{4.13}$$

Applying the sequence of f1 and f2 operations as:

$$f1(f2(P)) = (P_{pp}P_{qq} + P_{pq}^2)_{pp} + (P_{pp}P_{qq} + P_{pq}^2)_{qq} \tag{4.14}$$

The f2 of f1 is

$$f2(f1(P)) = (P_{pp} + P_{qq})_{pp} * (P_{pp} + P_{qq})_{qq} + (P_{pp} + P_{qq})_{pq}^2 \tag{4.15}$$

The f1 of f1 is

$$f1(f1(P)) = (P_{pp} + P_{qq})_{pp} + (P_{pp} + P_{qq})_{qq} \tag{4.16}$$

The f2 of f2 is

$$f2(f2(P)) = (P_{pp}P_{qq} + P_{pq}^2)_{pp} * (P_{pp}P_{qq} + P_{pq}^2)_{qq} + (P_{pp}P_{qq} + P_{pq}^2)_{pq}^2 \tag{4.17}$$

Similarly more such sequences can be generated to Select further saliency in Image datacube where P_{pp},P_{qq} and P_{pq} are second order difference of every band in HSI P as:

$$P_{pp} = \nabla_p^G(\nabla_p^G P) \tag{4.18}$$

$$P_{qq} = \nabla_q^G(\nabla_q^G P) \tag{4.19}$$

$$P_{pq} = \nabla_q^G(\nabla_p^G P) \tag{4.20}$$

In the above equations ∇_p^G and ∇_p^G are the Gaussian derivatives of the image P in x and y direction, respectively. The optimal points of local structure for an image have calculated concerning the number of channels and their spatial-spectral dimensions. For such points $f\varepsilon z^3$, the local neighbourhood S of the connected region is:

$$S = (\delta\varepsilon z^3 : ||f - \delta|| \leq 1)$$

We have concluded that $f\varepsilon z^3$ in the image is a locally optimal point of hd(hessian derivative) if hd is the HD image derivative. The algorithm-1 shows the f2 of f1 operation on spatial-spectral image datacube.

Algorithm 1 Significant band selection method for f1 of f2

procedure 1($For\ Fused\ Datacube\ Image\ \ HD-HSI\ P'$,$Z \leftarrow\ Bands\ (\overline{P})$)
    Output is selected image bands with salient information
    Begin
    Calculate $f1(\overline{P})$ *and* $f2(\overline{P})$ *from eq.*1 *and* 2
    Calculate $f1(f2(\overline{P}'))$ *from eq.* 3
    for do$(i \leftarrow Z)$
        $num = 0$
        for do$(Each\ pixel\ (px)\ \leftarrow\ i^{th}\ band)$
            if $f1((f2(\overline{p})))(px) > f1((f2(\overline{p})(\delta)))\ for\ \delta \varepsilon N(px)$ **then**
                $Num \leftarrow Num + 1$
                $t_{num} = f1((f2(\overline{p})(px)))$
            end if
        end for
    end for
    $MO(i) = \sum_{s=1}^{num} t_s$
    MO is the maximum identified objects
    $Salient\ dimensions,\ (i) \leftarrow \underset{i}{argmax}\ [Peak_{signal}\ (MO(i))]$
end procedure

4.2.5 Fusion of spatial and spectral Information

In Section-4.3.2, only spatial features have been used to obtain the salient features used in classification. In section-4.3.3, the spatial and spectral channels have been concatenated to extract the salient features of datacube. The number of channels has reduced to salient features by using the above-mentioned saliency selection method. The features obtained in this method are called 'spat-spec salient features' for short. To further examine the proposed scheme, our method has been distinguished with various recently proposed HD-HSI classification methods including local binary patterns (LBP)-extreme learning machine (ELM) [58], SVM-random feature selection (RFS) [65], 1-D CNN [53], and CNN-PPF [59]. In the Last-3 columns of table 4.5-4.7 and the last four columns of table 4.8-4.10, the sequence of Hessian (f1, f2) operations followed by several selected salient features are mentioned at next to its CNN model. Subsequently, 'spat-spec salient features' method is compared with the recently proposed schemes mentioned above. The results obtained from recently proposed methods are mentioned in the first three columns of table 4.5-4.7 and the first four columns of table 4.8-4.10 along with their reference. In subsequent sections, we have performed the salient feature selection method on Deep features as well

as another correlation-based feature extraction method. Both methods are divided into experimental analysis-1, and 2.

4.3 Experimental Analysis-1

In this section, we have applied the saliency selection method on deep CNN features. We have also applied salient feature selection on expanded features based on correlation in Experimental analysis-2.

4.3.1 Experimental Parameters

To conduct the experiments, 'off-the-shelf' networks are trained on 3 HSIs. Based on the learning of HSIs, all three CNN models have been modified. In modified architecture, the bias learning rate and weight learning factor have fixed as 20, which is adopted by many classical CNN. We have replaced the fully connected layer with a new softmax layer and a classification layer that belongs to Indian pines (IP), Pavia University(PU), and Salinas Valley(SV) datasets. Parameters for training have described as follows:
The training method is 'stochastic gradient descent with momentum(sgdm)',learning rate schedule is 'piecewise',drop factor = 0.2,drop period=5,maximum Epochs= 10, mini-batch size=64. The layer activation function is used to extract features from the convolution layer, and 'bicubic' interpolation is applied to transform the data in the original size. In subsequent steps, two different filters such as Gaussian filters are applied to obtain a channel-wise Hessian matrix. The f1 and f2 matrix has been obtained by using the Hessian matrix. The f1 and f2 of the Hessian matrix have extracted the local extremums of each feature.

4.3.2 Experimental Results for Deep Spatial Feature

To compare our proposed method, which is based on the spatial feature using saliency channels, an experiment has conducted for deep spatial features only. In this experiment, 3 CNN architectures have been used namely AlexNet, VGG-16[47] and VGG-19[47, 64].

TABLE 4.1: training and test samples for Indian Pines Data

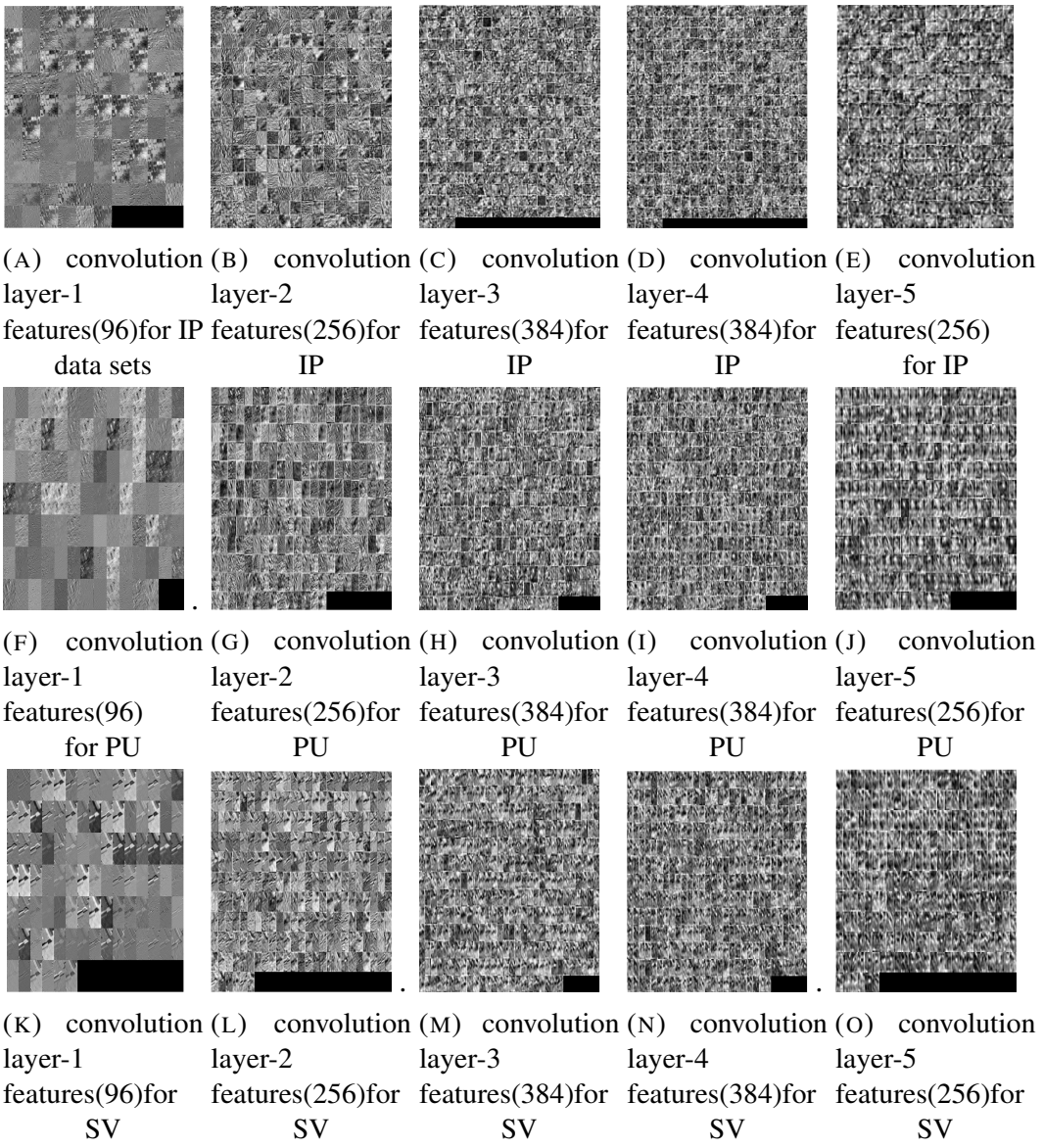
NO.	Classes	Training samples	Test samples
1	alfalfa	10%	54
2	corn_no till	10%	1434
3	corn_min till	10%	834
4	corn	10%	234
5	grass/pasture	10%	497
6	grass/tree	10%	747
7	grass/pasture-mowed	10%	26
8	Hay-windrowed	10%	489
9	oats	10%	20
10	soyabean_no till	10%	968
11	soyabean_min till	10%	2468
12	soyabean_clean till	10%	614
13	wheat	10%	212
14	woods	10%	1294
15	bldg-grass tree	10%	380
16	stone-steel tower	10%	95
Total	-	1018	10249

TABLE 4.2: training and test samples for Pavia university Data

NO.	Classes	Training samples	Test samples
1	Asphalt	4%	6631
2	Meadows till	4%	18649
3	Gravel	4%	2099
4	Trees	4%	3064
5	Painted metal sheets	4%	1345
6	Bare Soil	4%	5029
7	Bitumen	4%	1330
8	Self-Blocking Bricks	4%	3682
9	Shadows	4%	947
Total	-	1706	42770

TABLE 4.3: training and test samples for Salinas Valley Dataset

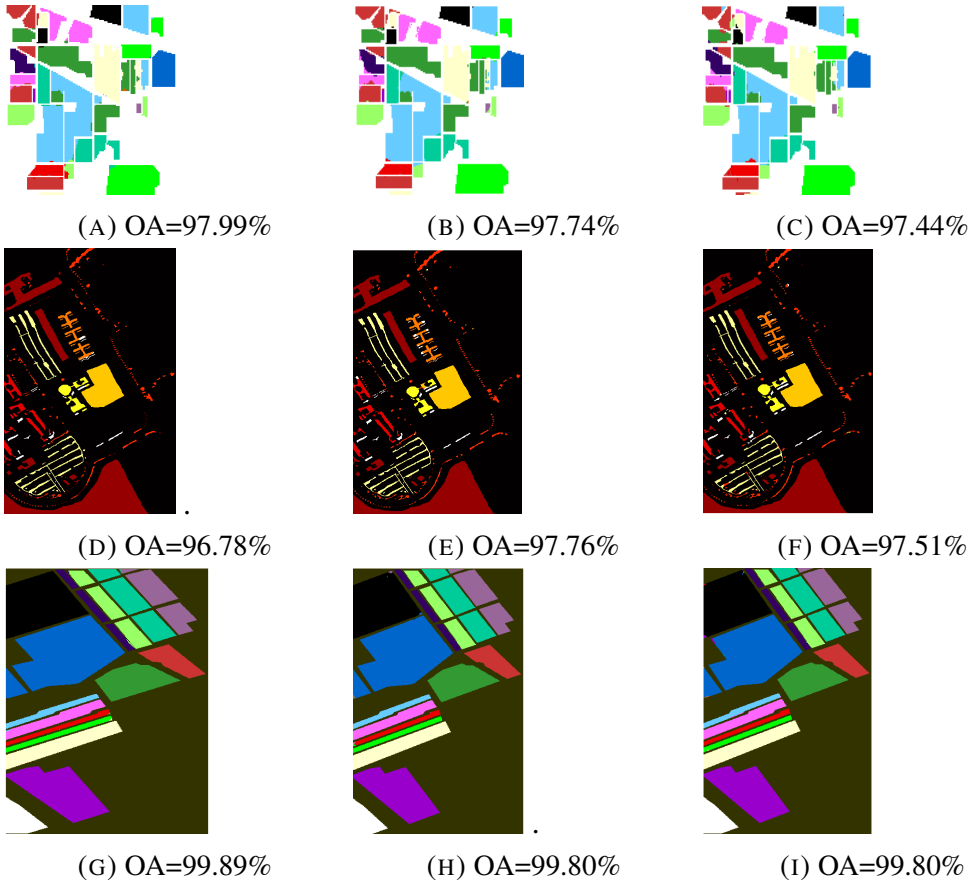
NO.	Classes	Training samples	Test samples
1	Brocoli_green_weeds_1	6%	2009
2	Brocoli_green_weeds_2	6%	3726
3	Fallow	6%	1976
4	Fallow_rough_plow	6%	1394
5	Fallow_smooth	6%	2678
6	Stubble	6%	3959
7	Celery	6%	3579
8	Grapes_untrained	6%	11271
9	Soil_vinyard_develop	6%	6203
10	Corn_senesced_green_weeds till	6%	3278
11	Lettuce_romaine_4wk till	6%	1068
12	Lettuce_romaine_5wk till	6%	1927
13	Lettuce_romaine_6wk	6%	916
14	Lettuce_romaine_7wk	6%	1070
15	Vinyard_untrained	6%	7268
16	Vinyard_vertical_trellis	6%	1807
Total	-	3240	54129



(A) convolution layer-1 features(96)for IP data sets
(B) convolution layer-2 features(256)for IP
(C) convolution layer-3 features(384)for IP
(D) convolution layer-4 features(384)for IP
(E) convolution layer-5 features(256) for IP
(F) convolution layer-1 features(96) for PU
(G) convolution layer-2 features(256)for PU
(H) convolution layer-3 features(384)for PU
(I) convolution layer-4 features(384)for PU
(J) convolution layer-5 features(256)for PU
(K) convolution layer-1 features(96)for SV
(L) convolution layer-2 features(256)for SV
(M) convolution layer-3 features(384)for SV
(N) convolution layer-4 features(384)for SV
(O) convolution layer-5 features(256)for SV

FIGURE 4.1: Deep spatial features for IP,PU and SV in AlexNet

These CNN models were initially trained on the ImageNet dataset. To obtain deep CNN features, the network architecture of the CNNs has modified as described in the previous section. As CNN Models have initially trained for 1000 class categories, the modification has been done to adapt the number of classes to three. The last layer of the modified network is updated using transfer learning. For example, the deep CNN features of IP, PU, and SV datasets for AlexNet architecture are shown in figure-4.1. In Table-4, the layers and number of spatial features extracted from three 'off the shelf' networks have shown. In VGG-16 and VGG-19, in case of more than one convolution layer at a single



(A) OA=97.99% (B) OA=97.74% (C) OA=97.44%

(D) OA=96.78% (E) OA=97.76% (F) OA=97.51%

(G) OA=99.89% (H) OA=99.80% (I) OA=99.80%

FIGURE 4.2: Classification results of spatial features:(1)IP[a,b,c] (2)PU[d,e,f](3)SV[g,h,i]for AlexNet,VGG-16 and19

TABLE 4.4: Numbers of spatial features Extracted from Different CNN Architectures

CNN	Layers	Features(IP)	Features(Pavia)	Features(Salinas)
AlexNet	conv1-conv5	1376	1372	1380
VGG-16	conv1-conv5	1472	1372	1476
VGG-19	conv1-conv5	1472	1472	1476

point, we have considered the features obtained from the last layer only. Table 4.1-4.3 shows the training sample obtained for each class in training. Further SVM has applied for classification. We have adapted 'one-vs-all' SVM, and RBF kernel with box constrain '10' for the Linear soft margin SVM model. The spatial features extracted from AlexNet, VGG-16, and VGG-19 are used to train the SVM. The entire set of spatial features are used to train the SVM for Indian pines dataset and 200 spatial features for Pavia university and Salinas Valley dataset. Training of the model has performed on 4 'cores' of CPU. Figures-4.2(a),(b),(c), shows the classification result obtained for Indian Pines Dataset with features extracted from Alexnet, VGG-16, and VGG-19, respectively. Figures-4.2(d),(e),(f), shows the classification results for the Pavia university dataset. Figure-4.2(g),(h),(i), depict

TABLE 4.5: Result on Indian Pines Image(in %) with spatial features

Classes	Alexnet [49]	VGG-CNN[49]	GoogleNet [49]	Alexnet(1376)	vgg16(1472)	vgg19(1472)
alfalfa	92.83	91.53	86.07	100	95.65	86.96
corn_no till	97.94	98.89	99.21	97.69	99.09	97.13
corn_min till	96.47	96.47	96.47	97.95	97.47	94.34
corn	100	100	100	95.78	89.03	95.78
grass/pasture	100	100	100	92.34	93.17	96.07
grass/tree	91.71	93.01	93.91	98.08	99.32	99.32
grass/pasture-mowed	89.40	89.45	88.69	100	75.00	100
Hay-windrowed	98.22	99.49	98.98	100	100	100
oats	99.34	99.62	99.44	100	95.00	90.00
soyabean_no till	-	-	-	98.97	97.12	94.14
soyabean_min till	-	-	-	98.25	98.74	98.33
soyabean_clean till	-	-	-	98.99	92.75	98.31
wheat	-	-	-	95.12	97.07	93.66
woods	-	-	-	98.66	99.53	99.92
bldg-grass tree	-	-	-	100	100	97.93
stone-steel tower	-	-	-	90.32	89.25	97.85
OA	94.24	94.36	93.30	97.99	97.74	97.44
AA	96.21	96.49	95.89	97.63	94.89	96.23

TABLE 4.6: Result for Pavia University Dataset(in %) using spatial features

Classes	Alexnet [49]	VGG-CNN- [49]	GoogleNet [49]	Alexnet(100)	vgg16(100)	vgg19(100)
alfalfa	77.42	83.86	76.35	97.04	98.05	97.51
corn_no till	88.87	83.06	76.35	99.65	99.95	99.99
corn_min till	96.52	96.37	96.37	98.52	97.67	98.48
corn	67.77	63.93	62.50	78.23	89.46	85.51
grass/pasture	100	99.91	100	93.01	91.52	98.36
grass/tree	98.55	99.03	98.70	100	99.82	99.74
grass/pasture-mowed	97.35	97.96	99.73	100	90.08	96.69
Hay-windrowed	97.96	94.77	98.39	99.57	99.19	96.28
oats	53.68	57.83	52.21	67.37	82.68	78.04
OA	87.77	85.76	88.86	96.78	97.76	97.51
AA	86.46	86.30	86.28	92.60	94.27	94.51

the Salinas Valley results. Table 4.5-4.7 has depicted the classifier outcomes achieved by using deep spatial features and a comparison of the results with the recently proposed method [64]. It has observed that overall accuracy with derived spatial features is heading up to 97.99%,97.74%, and 97.44% for Indian Pines dataset, which are 3-4 % greater than the past method [64].

TABLE 4.7: Result on Salinas Valley Image(in %) with spatial features

Classes	Alexnet [49]	VGG-CNN-S [49]	GoogleNet [49]	Alexnet(100)	vgg16(100)	vgg19(100)
alfalfa	94.14	94.75	95.63	100	100	100
corn_no till	98.81	97.65	98.07	100	100	100
corn_min till	97.75	97.13	98.03	100	100	100
corn	91.71	88.36	94.97	98.78	99.86	99.93
grass/pasture	94.51	86.48	93.42	99.74	99.89	99.93
grass/tree	94.12	93.83	92.63	100	100	100
grass/pasture-mowed	97.69	97.78	96.98	99.69	99.89	100
Hay-windrowed	86.28	81.98	87.00	100	99.79	99.65
oats	98.07	95.97	98.87	100	100	100
soyabean_no till	87.52	81.03	92.56	99.94	99.94	99.94
soyabean_min till	81.68	73.04	80.76	100	100	97.66
soyabean_clean till	66.59	57.67	61.32	99.43	99.38	100
wheat	62.71	67.04	62.57	99.34	96.94	100
woods	66.21	58.62	63.91	99.25	100	99.53
bldg-grass tree	97.38	97.45	97.98	100	99.53	99.56
stone-steel tower	100	100	100	100	100	100
OA	91.71	89.06	91.99	99.89	99.80	99.80
AA	88.45	85.55	88.42	99.76	99.70	99.76

TABLE 4.8: Result of spatial spectral data on Indian Pines Image with salient features

Classes	**LBP-ELM [58]**	**SVM-RFS [65]**	**1-D-CNN [53]**	**CNN-[59]**	**Alexnet(f1(f1)-200)**	**vgg16(f1(f1)206)**	**vgg19(f1(f1)209)**
alfalfa	**86.06**	**88.73**	**78.58**	**92.99**	**100**	**100**	**100**
corn_no till	**88.19**	**91.20**	**85.23**	**96.66**	**98.88**	**98.74**	**97.20**
corn_min till	**96.07**	**97.52**	**95.75**	**98.58**	**99.16**	**97.47**	**94.82**
corn	**99.73**	**99.86**	**99.81**	**100**	**98.73**	**96.62**	**99.16**
grass/pasture	**100**	**100**	**99.64**	**100**	**96.27**	**96.89**	**97.31**
grass/tree	**90.02**	**91.67**	**89.63**	**96.24**	**100**	**100**	**99.86**
grass/pasture-mowed	**71.00**	**78.79**	**81.55**	**87.80**	**100**	**100**	**100**
Hay-windrowed	**95.62**	**93.76**	**95.42**	**98.98**	**100**	**100**	**100**
oats	**98.66**	**98.74**	**98.59**	**99.81**	**70.00**	**55.00**	**80.00**
soyabean_no till	**-**	**-**	**-**	**-**	**95.58**	**99.90**	**94.55**
soyabean_min till	**-**	**-**	**-**	**-**	**99.47**	**99.59**	**99.31**
soyabean_clean till	**-**	**-**	**-**	**-**	**96.63**	**96.63**	**92.58**
wheat	**-**	**-**	**-**	**-**	**99.02**	**96.10**	**98.05**
woods	**-**	**-**	**-**	**-**	**100**	**100**	**100**
bldg-grass tree	**-**	**-**	**-**	**-**	**100**	**99.74**	**99.22**
stone-steel tower	**-**	**-**	**-**	**-**	**94.62**	**98.92**	**98.92**
OA	**87.33**	**89.83**	**86.44**	**94.34**	**98.70**	**98.91**	**97.80**
AA	**-**	**-**	**-**	**-**	**96.77**	**95.98**	**96.94**

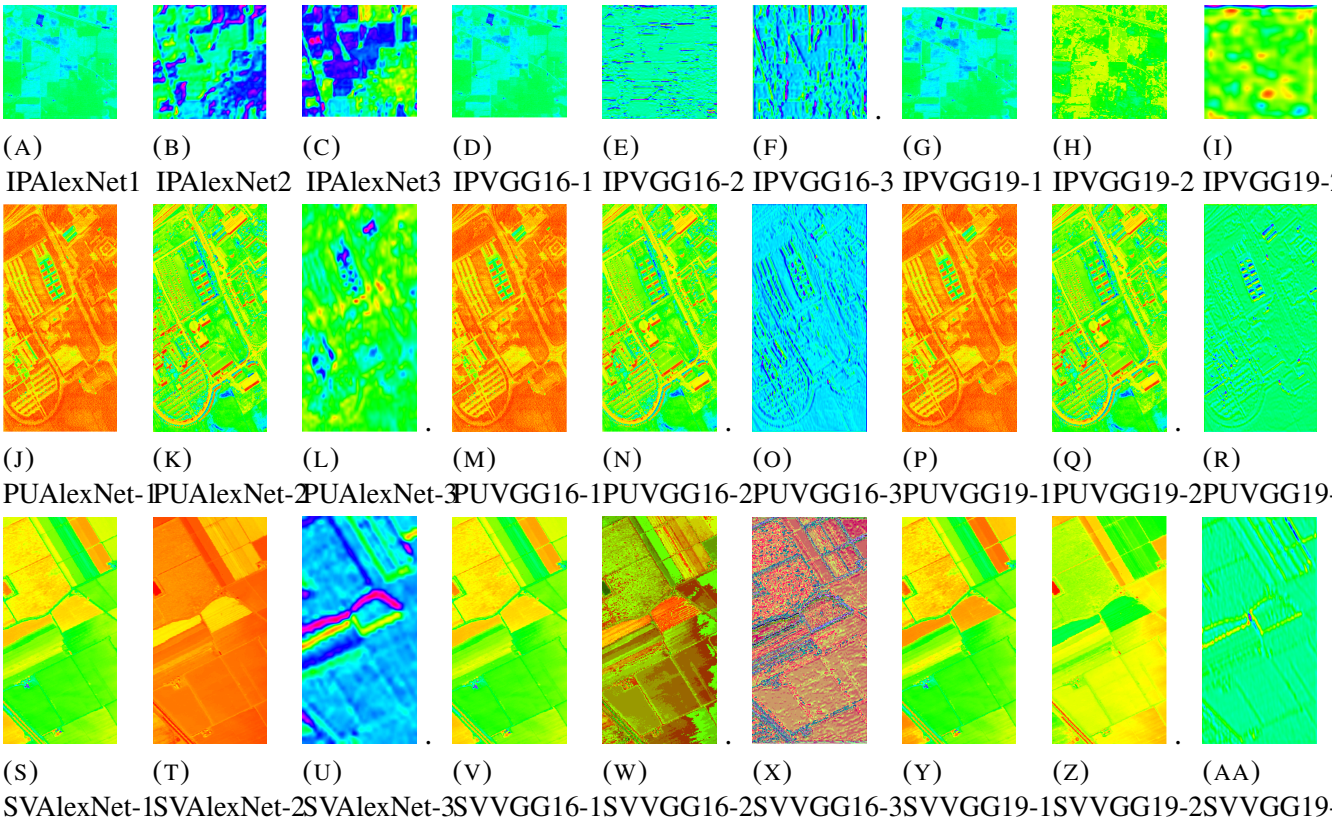


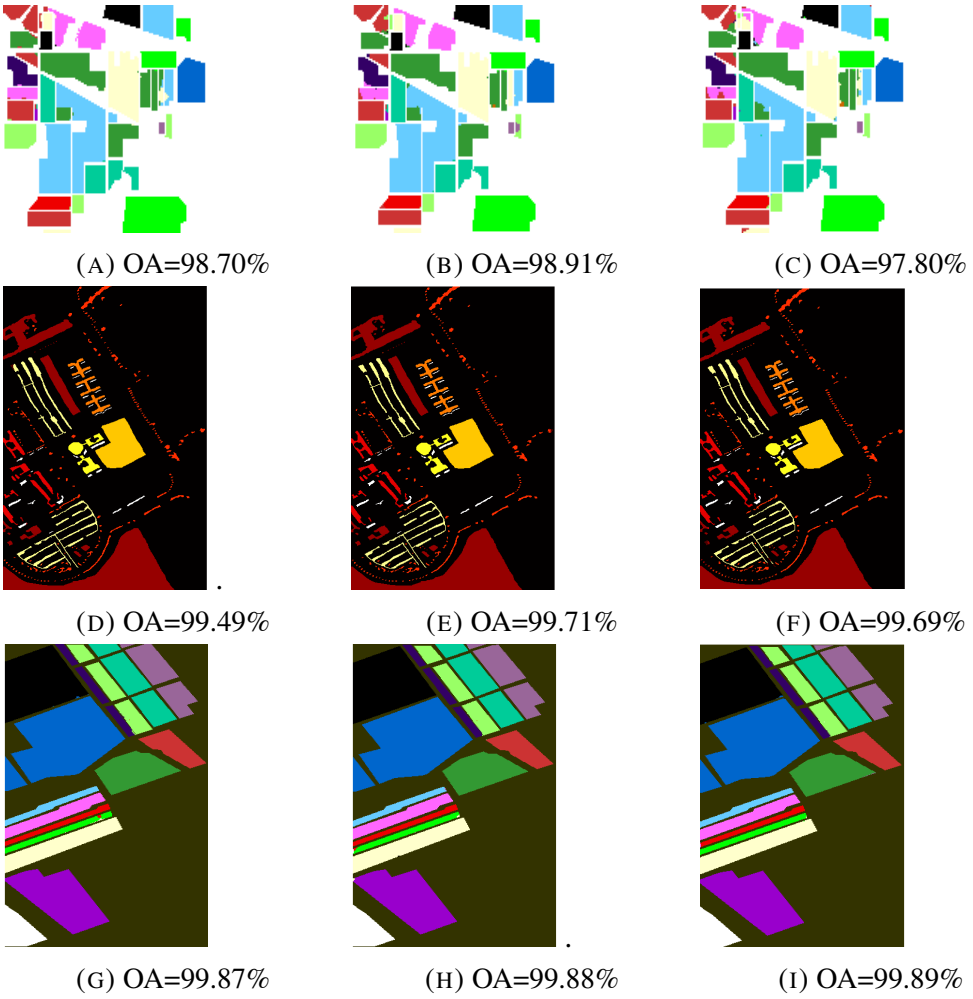
FIGURE 4.3: The Salient feature obtained from the combined Deep spatial and spectral datacube

4.3.3 Experimental Results for spatial-spectral features using salient features

In this section, the spatial and spectral features have fused to integrate the information. The features are concatenated to form a huge datacube of more than 1000 channels. In this section, the saliency selection method has become essential to reduce datacube by selecting salient features. From combined spatial-spectral datacube, we have shown the selected salient bands obtained by our saliency selection method in figure-4.3. Salient band selection followed by an SVM based training has performed on fused datacube. The results obtained from SVM based testing on the complete dataset has compared with some previously proposed methods. The practical aspect of the previously proposed method is that they have used the limited training samples .i.e 200 per class for just nine classes in

Indian pines(IP) data and all available classes in Pavia and Salinas datasets. Past methods have used 24% training set for Indian pines, 4-5% for Pavia university, and 6-7% for Salinas valley dataset. Therefore, for a fair comparison, lesser training samples have been used in our experiment .i.e. 10% per class for Indian pines, 4% for Pavia, and 6% for Salinas valley dataset in 16, 9 and 16 classes respectively. The obtained training sets are lower than previously proposed methods. On the other hand, the outcome of past methods .i.e [58, 65, 53, 59], are precisely taken from respective references and can be observed in the table 4.8-4.10, which reports the proposed and previous method results on three HSIs. Our results are achieving remarkably better accuracy than contemporary methods. The first four columns of the tables-4.8,4.9,4.10 are reporting the accuracy of past methods, and the last four columns report the proposed methods. In the tables mentioned above, The number in parenthesis next to the previous methods shows the reference. Next to the proposed methods indicates the number of channels retained after saliency selection. We have achieved the overall accuracy 98.70%, 98.91%, 97.80% for IP, PU, SV for AlexNet, which is 3-4% greater than the CNN-PPF [59] result, 11-12% higher than 1-D-CNN [53], 10-11% higher than the LBP-ELM [58] and SVM-RFS [65] for IP dataset by choosing the similar SVM parameters as in previous section . Similarly, a remarkable increase in accuracy can also be notice in PU and SV datasets. In PU dataset, overall accuracy is 99.49%, 99.71% and 99.69% for AlexNet,VGG-16 and VGG-19 respectively, which is 3-4% higher than CNN-PPF [59], 7-8% higher than 1-D-CNN[53], 8-9% higher than the LBP-ELM [58] and 9-10% higher than SVM-RFS [65]. In Salinas valley(SV) datasets, Overall accuracy is achieved up to 99.87%, 99.88% and 99.89% for AlexNet, VGG-16 and VGG-19 CNNs, that is 5-6% higher than CNN-PPF [59], 10-11% higher than 1-D-CNN [53], 7-8% higher than the LBP-ELM [58] and 6-7% higher than SVM-RFS [65]. Classified image results for 'spat-spec salient' method has described as follows:

The figure-4.4(a),(b),(c) shows the classification measures obtained for Indian Pines Dataset with AlexNet, VGG-16, and VGG-19 features. Figure-4.4(d), (e), (f), shows the Pavia university results, and figure-4.4(g), (h), (i) represents the Salinas Valley results for respective CNN models.



(A) OA=98.70% (B) OA=98.91% (C) OA=97.80%

(D) OA=99.49% (E) OA=99.71% (F) OA=99.69%

(G) OA=99.87% (H) OA=99.88% (I) OA=99.89%

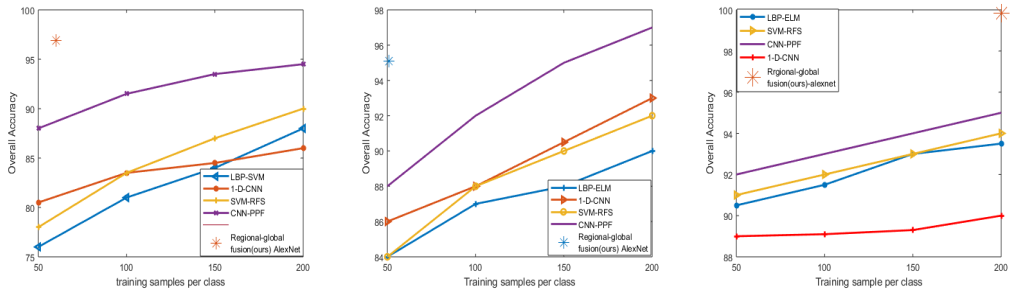
FIGURE 4.4: Classification results of spatial-spectral salient features:(1)IP[a,b,c](2)PU[d,e,f] (3)SV[g,h,i]for AlexNet,VGG-16 and 19

4.3.4 Impact of training samples on OA

To explore the impact of training samples on the overall performance, we have altered the number of training samples from 50 to 200 with an interval of 50. Figure-4.5(a),(b), and (c) shows that our proposed salient feature method has outperformed the previously proposed methods with a significant margin. Notably, our salient feature-based approach is gaining over past methods. The proposed 'spect-spat salient' approach, denoted by a single point in all the graphs, achieves significantly higher accuracy than any other case. The reason is that the saliency-based method embeds the local spatial extremes with spectral features, which reduces the computation cost and enhances the overall accuracy.

TABLE 4.9: Result of spatial spectral data on Pavia University Dataset

Classes	LBP-ELM [58]	SVM-RFS [65]	1-D-CNN [53]	CNN-[59]	Alexnet(f2(f1)-193)	vgg16(f2(f2)-214)	vgg19(f2(f1)-212)
alfalfa	81.32	87.95	88.38	97.42	99.38	99.73	100
corn_no till	90.91	91.17	91.27	95.76	99.97	100	99.75
corn_min till	85.09	86.99	85.88	94.05	99.52	100	100
corn	96.61	95.50	97.24	97.52	95.95	96.54	98.11
grass/pasture	99.63	99.85	99.91	100	100	100	100
grass/tree	94.33	94.31	96.41	99.13	100	100	100
grass/pasture-mowed	95.94	94.74	93.62	96.19	100	100	100
Hay-windrowed	82.65	85.89	87.45	93.62	99.70	99.97	99.27
oats	99.79	99.89	99.57	99.60	97.04	100	100
OA	89.86	91.10	92.27	96.48	99.49	99.71	99.69
AA	-	-	-	-	99.06	99.58	99.68



(A) (B) (C)

FIGURE 4.5: Overall Accuracies for different models for (a) Indian Pines(b)Pavia University(c) Salinas Valley Dataset

4.3.5 Running Time

All the experiments have performed on a machine with 14 core CPU and 1 TB memory. The CNN based image training has conducted on NVIDIA Tesla P4 with compute capability 6.1 GPU acceleration. Table-4.11 shows the running time of our proposed CNN and past CNN based schemes [64, 53] on the standard HSI's. In table-4.11, the first number in parenthesis denotes the CNN training time, and the second number indicates the salient feature extraction time in the above GPU accelerated set-up. In contrast, the SVM classification task has performed on the multi-core CPU. As shown in table-4.11, our method is taking reasonable computation time but obtained better accuracy than older methods. The computation aspect affirms that our scheme is robust and efficient for HD images.

TABLE 4.10: Result of spatial spectral data on Salinas valley Image

Classes	LBP-ELM [58]	SVM-RFS [65]	1-D-CNN [53]	CNN-[59]	Alexnet(f2(f2)-202)	vgg16(f2(f1)-234)	vgg19(f2(f1)-220)
alfalfa	**99.75**	**99.55**	**97.34**	**100**	**100**	**99.65**	**100**
corn_no till	**99.87**	**99.92**	**99.29**	**99.88**	**100**	**100**	**99.73**
corn_min till	**99.60**	**99.44**	**96.51**	**99.60**	**100**	**99.95**	**100**
corn	**99.64**	**99.86**	**99.66**	**99.49**	**99.28**	**99.86**	**99.93**
grass/pasture	**98.81**	**98.02**	**96.97**	**98.34**	**99.66**	**99.70**	**99.48**
grass/tree	**99.67**	**99.70**	**99.60**	**99.97**	**100**	**100**	**100**
grass/pasture-mowed	**99.66**	**99.69**	**99.49**	**100**	**100**	**100**	**100**
Hay-windrowed	**84.04**	**84.85**	**72.25**	**88.68**	**100**	**100**	**99.99**
oats	**99.89**	**99.58**	**97.53**	**98.33**	**100**	**99.87**	**100**
soyabean_no till	**95.03**	**96.49**	**91.29**	**98.60**	**99.97**	**99.91**	**99.94**
soyabean_min till	**96.82**	**98.78**	**97.58**	**99.54**	**100**	**99.72**	**100**
soyabean_clean till	**100**	**100**	**100**	**100**	**99.95**	**100**	**100**
wheat	**98.25**	**99.13**	**99.02**	**99.44**	**99.78**	**100**	**98.47**
woods	**97.94**	**98.97**	**95.05**	**98.96**	**96.73**	**97.10**	**99.72**
bldg-grass tree	**72.96**	**76.38**	**76.83**	**83.53**	**99.83**	**100**	**99.81**
stone-steel tower	**99.06**	**99.56**	**98.94**	**99.31**	**100**	**100**	**100**
OA	**92.42**	**93.15**	**89.28**	**94.80**	**99.87**	**99.88**	**99.89**
AA	**-**	**-**	**-**	**-**	**99.70**	**99.74**	**99.82**

TABLE 4.11: Running time for proposed(CNN trainig time+feature learning time for 500 features) and past methods

Methods	IP	PU	SV
spat-spect salient	1.54+0.35	1.54+0.978	1.54+0.904
Spatial features[[49]]	1.0+.11	1.68+0.55	1.16+0.77
1-D-CNN[[53]]	0.21	0.37	0.26
CNN-PPF[[59]]	4.76	16.92	20.97

4.4 Experimental Analysis-2

In this section, we have performed the salient feature selection method on low correlation-based feature expansion. The idea is to rely on the fact that low correlated features are informative and lead to better prediction accuracy than highly correlated features. Therefore, we have generated more features that are low correlated and discard the highly correlated one.

4.4.1 Setting up the Experiment

This section communicates the experimental results organized using three previously used images, available as benchmarks. We have denoted the proposed approach in this section as a divide-and-select channel selection(DSCS) approach. The experiments have conducted using LDA and SVM classifiers to authenticate the efficiency of the proposed

method. A MAP based Bayesian α expansion graph cut method has applied to the posterior probability of LDA to improve classifier accuracy. A graph cut method has denoted as GC in the suffix of LDA as LDAGC. The wavelength of extended channels λ and a minimum correlation $c_{\min}$ is 0.1 and 0.8 respectively in each experiment to generate low correlated channels in narrow wavelength. In Experiment-4.4.2, 4.4.3, and 4.4.4, only 10% of data samples have used as the training set for IP(Indian Pines), PU(Pavia University), and SV(Salinas Valley) datasets in classifier training. Subsequently, the approach has evaluated to obtain performance metrics like OA(Overall Accuracy), AA(Average Accuracy), Kappa Coefficient(k).

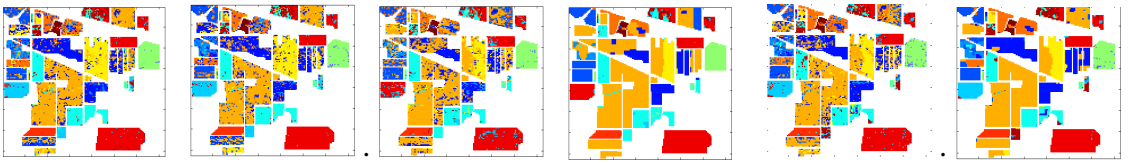
TABLE 4.12: λ=0.1 and $c_{\min}$=0.8, Accuracies(in %) for Indian Pines Data

Band	DSCS+LDA			DSCS+LDAGC			DSCS+SVM		
	OA	AA	k	OA	AA	k	OA	AA	k
Band_deter_deter(16)	78.92	77.55	0.7674	82.45	80.87	0.7906	73.74	72.10	0.7124
Band_trace_deter(18)	79.68	79.18	0.7891	84.79	83.04	0.8233	89.50	88.54	0.8739
Band_trace_trace(17)	73.29	72.10	0.7124	84.24	83.04	0.8233	89.73	88.11	0.8703
Band_deter_trace(16)	81.02	80.73	0.7910	85.13	84.41	0.8346	77.08	76.86	0.7590
Band_deter(61)	85.61	82.59	0.8301	90.70	88.67	0.8852	77.57	76.73	0.7573
Band_trace(55)	85.07	84.57	0.8326	89.96	88.78	0.8619	76.16	75.47	0.7484

4.4.2 Experiment 1:On Indiana Pines Image

This experiment also explores the HD Indiana Pines(IP) image obtained by the AVIRIS sensor, which contains 145×145 pixels in 200 spectral bands. We have 16 labelled classes in the validation dataset. In Table-4.12, the term 'Band_deter_deter(16)' shows that the determinant of determinant operation has selected the 16 feature channels for classification. The abbreviations for other annotations can be inferred similarly. The number of selected channel subsets followed by the classifier performance evaluation have shown in Table-4.12. Analytically the DSCS+LDAGC method is achieving remarkable OA of 89.96%, 90.70%, in the case of 55 and 61 selected channels. These channels have obtained by applying the trace(f1) and determinant(f2) operations in the proposed DSCS scheme, as described in equation-4.12 and 4.13. For subsequent reduction of channels, we have asked a different combination of trace and determinant in DSCS method on the extended image, as described in equation-4.14, 4.15, 4.16, and 4.17. Trace followed by a determinant operation is resulting in 18 significant channel combinations and OA of 84.79%. By applying trace followed by a determinant operation on the extended image, 18 channels have selected, and our DSCS+SVM method is achieving OA of 79.68%. Classified image result of DSCS+SVM

method has shown in figure-4.6(a) and 2(b), while DSCS+LDA and DSCS+LDAGC results have shown in figure-4.6(c),(d),(e) and (f). Figure-4.9(g) shows the affiliation between OA and increasing band selection by various combinations for the Indiana pines channel set.



(A) DSCS+SVM trace_deter(18) (OA = 89.50%) (B) DSCS+SVM deter(61) (OA = 77.57%) (C) DSCS+LDA trace_deter(18) (OA = 79.68%) (D) DSCS+LDAGC trace_deter(18) (OA = 85.61%) (E) DSCS+LDA deter(61) (OA = 84.79%) (F) DSCS+LDAGC deter(61) (OA = 90.70%)

FIGURE 4.6: Classification Results on selected bands with SVM and LDA predictors for Indiana Pines Data Set

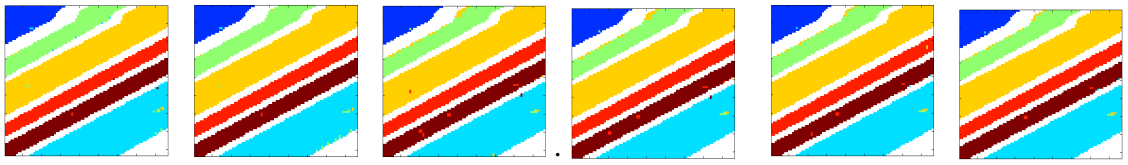
4.4.3 Experiment 2:On Salinas Valley Image

In this section, the experimental evaluation on the Salinas Valley(SV) dataset has

TABLE 4.13: λ=0.1 and $c_{\min}$=0.8,Accuracies(in %) for Salinas Data

Band	DSCS+LDA			DSCS+LDAGC			DSCS+SVM		
	OA	AA	k	OA	AA	k	OA	AA	k
Band_deter_deter(20)	98.87	98.36	0.9874	98.98	98.38	0.9874	99.48	99.24	0.9909
Band_trace_deter(21)	99.05	98.61	0.9886	99.13	98.69	0.9898	99.31	99.18	0.9911
Band_trace_trace(21)	99.12	98.88	0.9908	99.19	98.95	0.9916	99.27	99.10	0.9909
Band_deter_trace(23)	99.08	98.68	0.9898	99.16	98.72	0.9902	99.50	99.30	0.9909
Band_deter(67)	99.34	98.68	0.9902	99.43	98.77	0.9910	99.44	99.36	0.9930
Band_trace(72)	99.22	99.05	0.9924	99.33	99.07	0.9928	99.27	99.13	0.9892

conducted. This HSI contains 83×86 pixel image in 204 spectral channels. Table-4.13 depicts the Overall accuracy(OA), Average accuracy(AA) and Kappa Coefficient(k) for DSCS+LDA, DSCS+LDAGC and DSCS+SVM methods. Overall accuracy has increased when the number of selected channels are increasing in every case, which has shown in figure-4.9(h). Figure-4.7 is representing the output images after classification with the proposed DSCS algorithm.

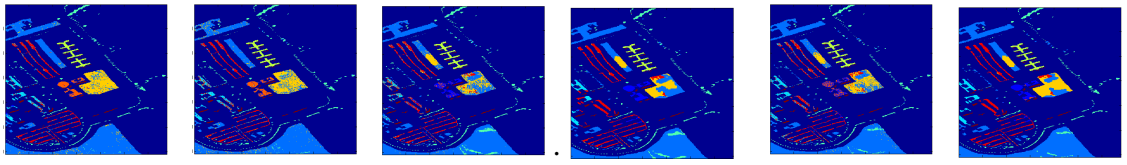


(A) DSCS+SVM trace_deter(21) (OA = 99.31%) (B) DSCS+SVM deter(67) (OA = 99.44%) (C) DSCS+LDA trace_deter(21) (OA = 99.05%) (D) DSCS+LDAGC trace_deter(21) (OA = 99.13%) (E) DSCS+LDA deter(67) (OA = 99.34%) (F) DSCS+LDAGC deter(67) (OA = 99.43%)

FIGURE 4.7: Classification Results on selected bands with SVM and LDA predictors for Salinas Valley Data Set

4.4.4 Experiment 3:On Pavia University Image

This experiment estimates the classifier metrics on the Pavia University(PU) image scene



(A) DSCS+SVM deter(35) (OA = 77.57%) (B) DSCS+SVM trace_deter(10) (OA = 79.68%) (C) DSCS+LDA trace_deter(10) (OA = 92.87%) (D) DSCS+LDAGC trace_deter(10) (OA = 93.24%) (E) DSCS+LDA deter(35) (OA = 93.75%) (F) DSCS+LDAGC deter(35) (OA = 94.13%)

FIGURE 4.8: Classification Results with selected bands and SVM and LDA predictors for Pavia University Data Set

TABLE 4.14: λ=0.1 and $c_{\min}$=0.8,Accuracies(in %) for Pavia University Data

Band	DSCS+LDA			DSCS+LDAGC			DSCS+SVM		
	OA	AA	k	OA	AA	k	OA	AA	k
Band_deter_deter(12)	92.89	91.90	0.9090	93.00	92.14	0.9191	89.92	88.63	0.8655
Band_deter_trace(8)	92.75	90.42	0.8950	92.85	91.94	0.9091	89.85	87.73	0.8735
Band_trace_trace(8)	92.89	91.46	0.9034	92.94	90.46	0.8948	89.48	87.78	0.8597
Band_trace_deter(10)	92.87	91.65	0.9071	93.24	91.13	0.9045	90.54	88.18	0.8646
Band_trace(27)	93.60	91.83	0.9102	93.96	93.37	0.9218	92.87	90.82	0.9154
Band_deter(35)	93.75	91.33	0.9054	94.13	93.28	0.9176	93.64	91.82	0.9050

obtained from the Rosis sensor of Pavia University. It is a 610×340 pixels image with 103 spectral bands. The DSCS+LDA, DSCS+LDAGC, and DSCS+SVM methods are achieving OA of 93.75%, 94.13%, 93.64% for 67 selected channels by using determinant operation as feature selector. The determinant based feature selector method results in 35 features and achieves maximum overall accuracy .i.e 94.13% for DSCS+LDAGC method, as shown

in Table-4.14. Trace followed by a determinant and trace only operations are resulting in 10 and 27 selected features. The OA obtained for LDA, LDAGC and SVM classifiers are 92.87%, 93.24%, 90.54% and 93.60%, 93.96%, 92.87% respectively for 10 and 27 selected features. In Table-4.14, it has observed that salient channels are reducing up to 8 channels, but accuracy is not decreasing that much by our proposed approach. Table-4.14 shows the Overall accuracy, Average accuracy, and Kappa coefficient for the different combinations of channels in proposed methods. Figure-4.8(a), 4.8(b) shows the DSCS+SVM results when a determinant[deter(35)] and a trace followed by a determinant[trace(deter(10))] operations are applied. These operations select the 35 and 10 channels, respectively. Figures-4.8(c), 4.8(d), 4.8(e), and 4.8(f), shows the DSCS+LDA and DSCS+LDAGC results when the number of selected features are 10 and 35. Classifier outcomes obtained from 8, 10, and 12 selected features have shared in Table-4.14.

4.4.5 Comparative analysis

In table-4.15, the OA of previously proposed methods has shared for IP, PU, and SV datasets. In the FRSR method, the OA is coming out to be 92.38% and 86.96% for SVM and KNN, which is lower than our DSCS+LDA and DSCS+LDAGC methods, having accuracy in a range of 92.75% to 94.13%. Proposed DSCS+LDA, DSCS+LDAGC, and DSCS+SVM methods are achieving higher accuracy than DWSSR(KNN) method with a significant margin of 2-20% for IP dataset. For the PU dataset, the proposed method is achieving up to 94.13% accuracy in the DSCS+LDAGC approach, which is higher than DWSSR(KNN) and DWSSR(SVM) methods. The Proposed scheme is achieving better accuracy than SSR, *k*-means, RPCA+*k*-Means for IP, PU, and SV datasets. For the SV dataset, overall accuracy is above 99% for both proposed and past methods, but the proposed scheme is gaining maximum accuracy up to 99.50% and 99.43% for DSCS+SVM and DSCS+LDAGC which is greater than FRPCAGL(SVM), *k*-Means and RPCA+*k*-Means methods.

TABLE 4.15: Performance (in %) in state-of-art methods

Dataset	FRSR(SVM)	FRSR(KNN)	FRPCAGL(SVM)	DWSSR(SVM)	DWSSR(KNN)	SSR(SVM)	SSR(KNN)	*k*-means	RPCA+*k*-means
IP	-	-	84.58	79.89	73.80	71.86	63.53	80.84	82.41
PU	92.38	86.96	93.78	93.68	88.68	89.58	83.07	87.61	89.83
SV	-	-	99.37	-	-	-	-	99.21	99.26

4.4.6 Impact of μ, c_{min} and selected channels on Overall Accuracy

To investigate the impact of the parameter μ and minimum correlation $c_{\min}$ on overall accuracy, we have varied the values of μ and $c_{\min}$ from 0.2 to 0.8, as shown in Figure-4.9. By assumption, the μ and $c_{\min}$ is lying between 0 and 1, but the value of μ cannot be equal to 0 since it is indeterminate for second intermediate feature generation. It has observed that OA increases with an increase of μ for SVM as well as an LDA classifier. OA increases with an increase of $c_{\min}$ on IP and SV dataset for both SVM and LDA learners. In PU dataset, OA increases with $c_{\min}$ for LDA and decreases for SVM classifier. The effect of selected channels on overall accuracy has shown in figure-4.9(g), (h), (i) for IP, PU, SV datasets.

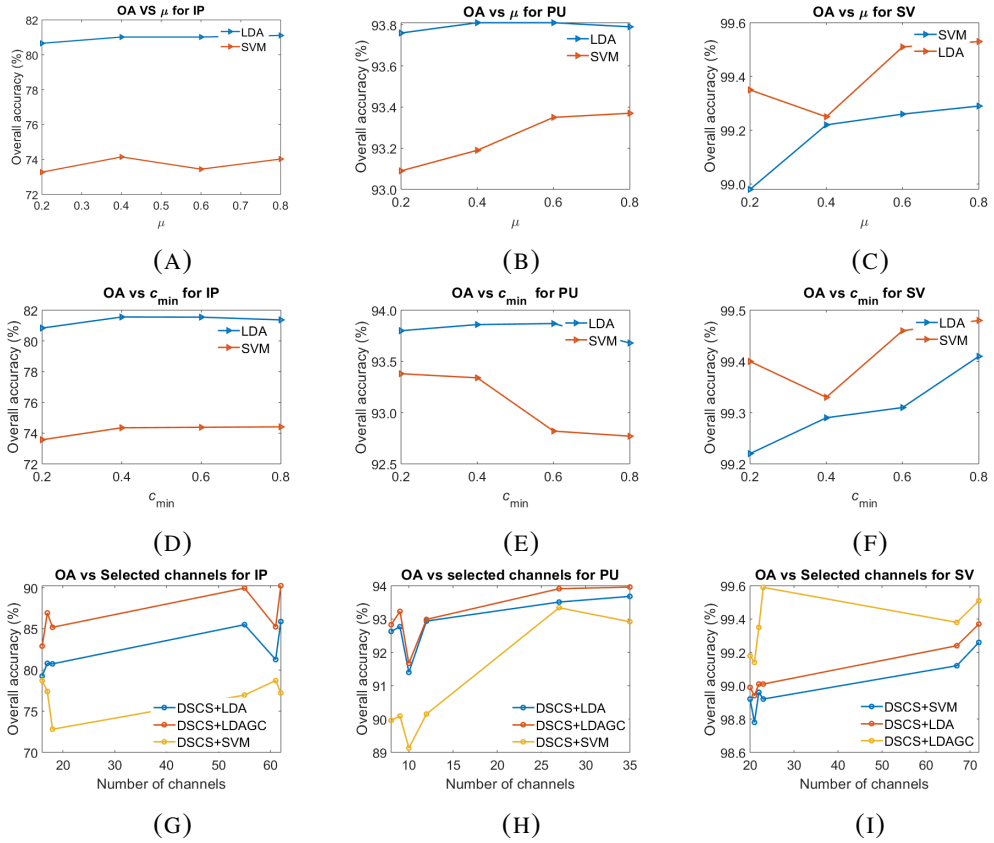
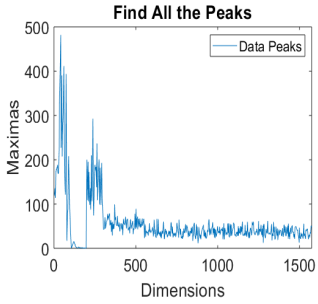


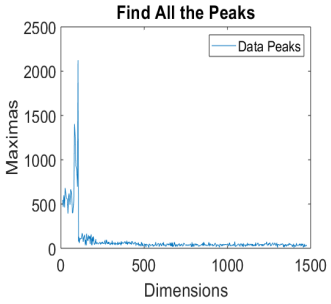
FIGURE 4.9: OA(in %) vs μ(a to c), OA(in %) vs minimum correlation $c_{\min}$ (d to f) and OA(in %) vs Selected channels(g to i) for IP, PU, SV datasets

4.5 MO Signals

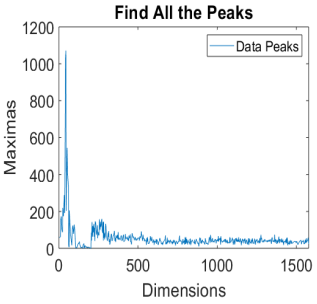
In this section, we have shown the maximum objects(MO) identified in IP, PU, and SV datasets. We have selected the salient features based on the maximum peaks of these signals. These peaks have resulted in the salient feature channels as shown in the step-15, 16, and 17 of algorithm-1. We have obtained the HSI datasets from ”Grupo de Inteligencia Computacional (GIC) Buscar.” Figure-4.10(a), (b), (c) have denoted the number of objects detected in each dimension for IP(Indian Pines). Figure-4.10(d), (e), (f) are the same signal function for PU(Pavia Univerity) dataset. Figure-4.10(g), (h), and (i) have shown the object count function for SV(Salinas Valley) dataset. The peak values in these functions have denoted the salient feature dimensions that we have used for classification in SVM.



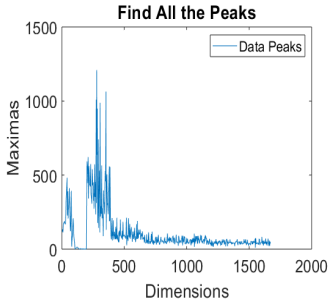
(A) MO(Maximum Objects) in function peaks of IP-AlexNet



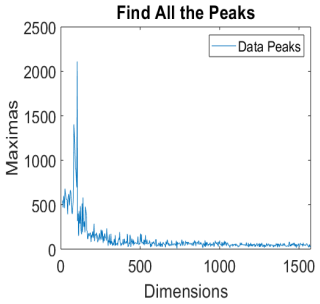
(B) MO(Maximum Objects) in function peaks of PU-AlexNet



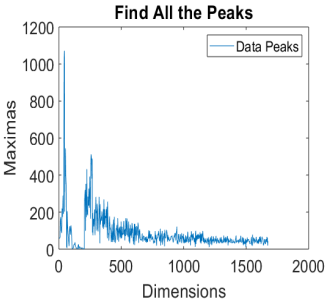
(C) MO(Maximum Objects) in function peaks of SV-AlexNet



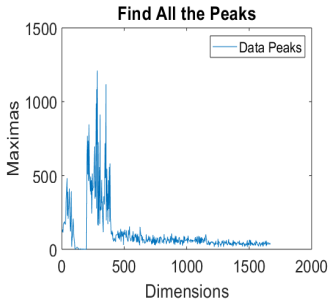
(D) MO(Maximum Objects) in function peaks of IP-VGG16



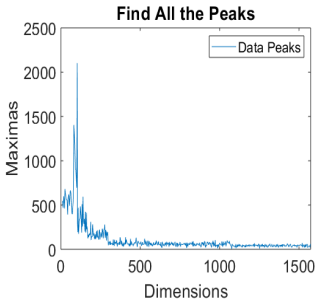
(E) MO(Maximum Objects) in function peaks of PU-VGG16



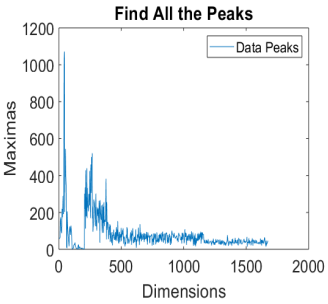
(F) MO(Maximum Objects) in function peaks of SV-VGG16



(G) MO(Maximum Objects) in function peaks of IP-VGG19



(H) MO(Maximum Objects) in function peaks of PU-VGG19



(I) MO(Maximum Objects) in function peaks of SV-VGG19

FIGURE 4.10: Salient Dimensions selection by using MO(Maximum Objects) in function peaks

4.6 Conclusion

In this section, we present a robust and effective deep feature extraction method by using different CNN architecture and an unsupervised dimension expansion approach. To achieve better classification accuracies and lesser computation cost, we have applied a

saliency-based feature learning to extract more capable spatial and spectral features by embedding only selected features in the SVM classifier. These features have selected based on the latent relation between saliency and image descriptors. The salient features have been selected by emulating the scale selection method. Experimental results, performed on three benchmark Images, are demonstrating the outstanding performance of “spat–spec saliency method” as compared to recently proposed state-of-art methods.