
Pixel based Semantic Labelling and Image Parsing using Intelligent Learning Techniques

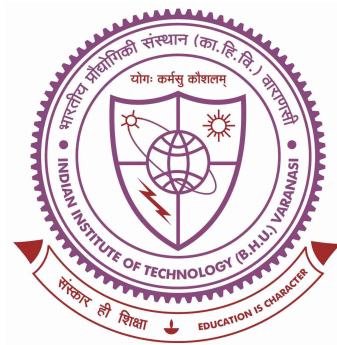
Thesis submitted in partial fulfillment

for the Award of Degree

DOCTOR OF PHILOSOPHY

by

Vishal Srivastava



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY

(BANARAS HINDU UNIVERSITY),

VARANASI-221 005

Roll No: 17071022

October 2020

Certificate

It is certified that the work contained in this thesis titled "Pixel based Semantic Labelling and Image Parsing using Intelligent Learning Techniques" by "Vishal Srivastava" has been carried out under my supervision and that it has not been submitted elsewhere for a degree.

It is further certified that the student has fulfilled all the requirements of Comprehensive Examination, Candidacy and SOTA for the award of Ph.D. Degree.

Dr. Bhaskar Biswas

Associate Professor

Department of Computer Science and Engineering

Indian Institute of Technology (Banaras Hindu University),

Varanasi-221 005

DECLARATION BY THE CANDIDATE

I, Vishal Srivastava, certify that the work embodied in this thesis is my own bonafide work and carried out by me under the supervision of Dr. Bhaskar Biswas from July-2017 to October , 2020, at the Department of Computer Science and Engineering, Indian Institute of Technology (BHU), Varanasi. The matter embodied in this thesis has not been submitted for the award of any other degree/diploma. I declare that I have faithfully acknowledged and given credits to the research workers wherever their works have been cited in my work in this thesis. I further declare that I have not willfully copied any other's work, paragraphs, text, data, results, etc., reported in journals, books, magazines, reports dissertations, theses, etc., or available at websites and have not included them in this thesis and have not cited as my own work.

Date :

Signature of the Student

Place :

(Vishal Srivastava)

CERTIFICATE BY THE SUPERVISOR

It is certified that the above statement made by the student is correct to the best of my/our knowledge.

Signature of Supervisor

(Dr. Bhaskar Biswas)

Signature of Head of Department

COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis : **Pixel based Semantic Labelling and Image Parsing using Intelligent Learning Techniques**

Name of the Student : **Vishal Srivastava**

Copyright Transfer

The undersigned hereby assigns to the Indian Institute of Technology (Banaras Hindu University) Varanasi all rights under copyright that may exist in and for the above thesis submitted for the award of the "DOCTOR OF PHILOSOPHY".

Date :

Signature of the Student

Place :

(Vishal Srivastava)

Note: However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provided that the source and the Institute's copyright notice are indicated.

Acknowledgements

Undertaking this Ph.D. has been a truly life-changing experience for me, and it would not have been possible to do without the support and guidance that I received from many people. I take this opportunity to extend my sincere gratitude and appreciation to all those who made this Ph.D. thesis possible. First and foremost, I would like to express my sincere gratitude to my research guide, Dr. Bhaskar Biswas, for introducing me to this exciting field of Computer Science & Engineering and for his dedicated help, advice, inspiration, encouragement, and continuous support my Ph.D. His enthusiasm, integral view on research, and his mission to provide high-quality work have made a deep impression on me. During our course of interaction during the last three years, I have learned extensively from him, including how to raise new possibilities, regard an old question from a new perspective, approach a problem through systematic thinking, data-driven decision making, and exploiting serendipity. I owe him lots of gratitude for having me shown this way of research. My special words of thanks should also go to Prof. Rajeev Srivastava, Head, Computer Science & Engineering Department, for his kindness and valuable support for facilitating all the requirements. I express my heartfelt gratitude to faculty members and staff of the department. I especially thank the members of my Research Program Evaluation committee, Dr. Bhaskar Biswas, Dr. Tanima Dutta, Dr. Kalpana Choudhary and DPGC Convener Dr. Pratik Chattopadhyay for their valuable suggestions regarding the thesis. I have a special mention of thanks to my lab mates in Research Lab, Mr. Navin, Mr. Shivansh, and Ms. Sneha. I wish to extend my gratitude to all my seniors. Their advice, encouragements, and critics were always the source of inspiration. I would want to thank all my family members who continuously support me throughout my work. My parents did hard work to support the logistics for my studies and encouraged me every decision that I took. Finally, I am grateful to my almighty god shiv for blessings and for giving me the strength to persevere throughout the long and arduous journey.

- Vishal Srivastava

PREFACE

Semantic image labelling, also known as pixel-based classification, is a task of segmenting the objects within an image by pixel-level spectral similarity. Semantic image labelling has also referred to as image parsing, which is a process of decomposing the image in different regions and constructing a structured input. In this thesis, we summarise the four aspects of research in the semantic labelling, i.e., classical machine learning(ML), feature engineering, deep learning(DL), and relaxation labelling(RL).

The above mentioned four aspects of semantic labelling lead to realizing that it is not a separate domain but a natural step in moving from coarse to fine interpretation. The original procedure could have been derived from a classification scheme, which predicts the label for a complete input. This process is also known as image category classification in the literature. This process has various applications such as bio-metric image classification, classification of tumor in a different grade, classification of different classes of species, digit classification using digit databases, emotion detection using face databases, and category classification using features from CBIR systems. A vast set of image databases, such as MNIST, CIFAR, ORL, YALE, etc., are available to validate the category classification methods.

In the previous classification methods, i.e., coarse grain, the classifiers predict the objects or provide a rank list in case of many objects. The next step is to localize and detect the objects within the image, which is a fine grain inference. The main aim of such inference is not only to provide classes but also some specific information such as the spatial location of classes, centroids, and bounding boxes. These kinds of fine-grain classification processes have performed on pixel-level, not image level. Therefore, they are computationally costly. Such inferences in various literature have denoted semantic labelling, pixel-based semantic classification/segmentation, or semantic image parsing. Consequently, we can summarise the semantic parsing as an image-based method to achieve fine-grain predictions. The goal is to make a dense prediction for the label of each pixel in such a way that each pixel is labelled with its class of enclosed regions. These kinds of fine-grained studies are the main interest of this thesis.

Another important aspect is to select the complex data-sets for semantic labelling. We have used low dimensional RGB images such as sift-flow, pascal-voc data-sets, and high

dimensional hyperspectral(HD-HSI) data-sets to perform the semantic labelling based experiments. The facial expression based images in ORL, YALE-A, and B, COIL data-sets have been used for category prediction also.

We have discussed the classical machine learning methods, some feature engineering, and knowledge embedding techniques to develop accurate and efficient frameworks. The salient and CNN feature-based approaches have been adapted to achieve effective and robust features from raw data. Feature selection and dimension embedding have also played a pivotal role in proposed frameworks. Deep learning and CNN based approaches have been used to design the custom CNN architecture and exploited the significant results from the image. Some relaxation labelling-based methods have also detailed to improve the CNN and classical ML-based probabilistic outcomes significantly. The advantages and drawbacks of the proposed frameworks have discussed. The comparative analysis for benchmark image sets and evaluation matrices have also been performed. Finally, some encouraging future works have drawn out, and the conclusion has drawn for pixel-wise semantic scene labelling or image parsing.

Contents

Certificate	ii
Declaration by the Candidate	iii
Copyright Transfer Certificate	iv
Acknowledgements	v
Preface	vi
Contents	viii
List of Figures	xvii
List of Tables	xxi
Abbreviations	xxiii
Symbols	xxv
1 Introduction	1
1.1 Introduction	1
1.1.1 Major Issues	2
1.2 Objectives	4
1.2.1 Exploring the DNN feature for scene labelling	4
Objective-1	4
Objective-2	5
1.2.2 Exploring the Label relaxation and feature expansion based information generation	5

	Objective-3	5
	Objective-4	5
1.2.3	Develop a deep learning architecture, study of manifold learning and convex optimization	6
	Objective-5	6
	Objective-6	6
1.2.4	Exploring the multi-scale morphological features and predicate based feature merging	6
	Objective-7	7
1.3	Contribution	7
1.3.1	Deep CNN feature and manifold reduction based scene labelling	7
1.3.2	Extracting the salient features from Deep CNN and investigate the feature expansion	8
1.3.3	Label relaxation for scene labelling	8
1.3.4	CNN architecture design and relaxation labelling	9
1.3.5	Morphological features and predicate based feature merging for scene labelling	9
1.4	Thesis Organization	10
2	Literature and Related work	11
2.1	Survey on Scene labelling	11
2.2	Scene Labelling/Parsing based techniques	12
2.2.1	Machine Learning based techniques for Scene Labelling	12
	Machine learning based methods on Hyper-spectral images (HSI)	13
2.2.2	Deep Learning based techniques for Scene Labelling	13
2.2.3	Relaxation based techniques for Scene Labelling	14
2.2.4	Evaluation Matrices and Validation	15
2.2.4.1	Overall Accuracy (OA)	15
2.2.4.2	Average Accuracy (AA)	16
2.2.4.3	Kappa value (K)	16
2.2.4.4	Intersection Over Union (IOU)	16
2.2.4.5	Precision, Recall, and Fscore	17
3	Deep CNN Feature Fusion with Manifold Learning and Regression for semantic pixel classification	19
3.1	Introduction	19
3.2	Proposed work	20
3.2.1	Deep CNN Spatial Feature by off the self networks	20
3.2.2	Estimation of Global probability using Manifold learning	22
3.2.3	Estimation of Regional probability by mixing of classes in pixel using Subspace Projection	24

3.2.3.1	Subspace projection based MLR method:subspaceMLR	25
3.2.3.2	Regional probability calculation	26
3.2.3.3	Fusion of regional and global Probabilities using Regularizer	27
3.3	Experimental Result Analysis	27
3.3.1	Dataset detail	27
3.3.2	Deep spatial features extraction	28
3.3.3	Global probability estimate with manifold learning	28
3.3.4	Regional Probability estimate with multinomial regression	32
3.3.5	Fusion of Global and Regional information	33
3.3.6	Effect of λ_2 on OA	36
3.3.7	Computational Complexity	37
3.3.8	Comparison with other methods	38
3.4	Conclusion	39
4	Convolution Neural Network(CNN) and correlation based Salient Features	41
4.1	Introduction	41
4.2	Proposed Method	42
4.2.1	Outline of the Proposed scheme	42
4.2.2	Method-1 for Feature extraction: Deep CNN Spatial Feature by 'off-the-self' networks	42
4.2.3	Method-2 for Feature extraction: Expansion of weakly Correlated Features	43
4.2.4	Salient Features selection using Scale identification for maximum object(MO) classification	45
4.2.4.1	Significant feature selection	45
4.2.5	Fusion of spatial and spectral Information	47
4.3	Experimental Analysis-1	48
4.3.1	Experimental Parameters	48
4.3.2	Experimental Results for Deep Spatial Feature	48
4.3.3	Experimental Results for spatial-spectral features using salient features	54
4.3.4	Impact of training samples on OA	56
4.3.5	Running Time	57
4.4	Experimental Analysis-2	58
4.4.1	Setting up the Experiment	58
4.4.2	Experiment 1:On Indiana Pines Image	59
4.4.3	Experiment 2:On Salinas Valley Image	60
4.4.4	Experiment 3:On Pavia University Image	61
4.4.5	Comparative analysis	62
4.4.6	Impact of μ, c_{min} and selected channels on Overall Accuracy	63

4.5	MO Signals	64
4.6	Conclusion	65
5	A subspace regression, two phase label optimization, and efficient edge preservation scheme(EPS), for relaxation labelling	67
5.1	Introduction	67
5.2	Proposed Method-1	70
5.2.1	Step-1:Subspace Projection and Logistic Regression method:SubspaceMLR	71
5.2.2	Step-2:Two phase optimization method: paraKERNALGC	72
5.2.2.1	Functional description	72
5.2.2.2	Phase-1: Class Label update using gradient descent	73
5.2.2.3	phase-2:Partition Update with Graph Cut	74
5.2.3	Fusion of Datacosts using Regularizer (λ)	75
5.3	Experimental Result Analysis-1	75
5.3.1	Experiment 1: Application on synthetic image	76
5.3.1.1	Experiment 1:Impact of Training Sample Size on OA	77
5.3.1.2	Experiment 2:Impact of Parameter σ_1 on OA :	78
5.3.2	Experiment 2:Experiments with real images data set	79
5.3.2.1	Observation 1:On Indiana Pines Dataset	79
5.3.2.2	Observation 2:On Salinas Valley Image Dataset	81
5.3.3	Comparative Analysis	84
5.3.4	Parameter Analysis	85
5.3.5	Computational Time	86
5.4	Proposed Method-2	87
5.4.1	Posterior probability based pixel Learning Method:subspaceMLR	88
5.4.2	Edge Preservation Scheme(EPS) based relaxation	88
5.4.3	Multilevel logistic prior diffusion in Posterior with Graph cut (GC)	91
5.5	Experimental Result Analysis-2	92
5.5.1	Setting up the Experiment for Simulated Dataset	92
5.5.2	Experiment 1:Modelling on synthetic image dataset	93
5.5.2.1	Exercise 1:Effect of Regularizer λ on OA	94
5.5.2.2	Exercise 2:Effect of error (Err) on iterations(itr):	95
5.5.2.3	Exercise 3:Effect of noise term sigma on OA:	96
5.5.3	Experiment 2:Modelling with Real Images Data Set	96
5.5.3.1	Observation 1:On Indiana Pines(IP) Dataset	97
5.5.3.2	Observation 2:On Pavia University(PU) Rosis Sensor dataset	98
5.5.3.3	Observation 3:On Salinas Valley (SV) Dataset	100
5.5.3.4	Overall performance on IP, PU, and SV datasets	102

5.5.4	Contrast Experiments	104
5.6	Conclusion	105
6	CNN-EFF: CNN based Edge Feature Fusion in Semantic Class Prediction	107
6.1	Introduction	107
6.2	Proposed Method:(CNN-EFF)	108
6.2.1	Deep CNN based training and soft-max probability computation	109
6.2.1.1	First convolution and maximum pooling Layer(conv-1 and pool-1)	109
6.2.1.2	Second convolution and maximum pooling Layer(conv-2 and pool-2)	110
6.2.1.3	Third Convolution and pooling Layer(conv-3 and pool-3)	110
6.2.1.4	Data in fully connected (fc) Layers	111
6.2.2	Edge Feature Fusion based post processing(EFF-Jacobian optimization)	112
6.3	Experimental Result Analysis	116
6.3.1	Data-set details	117
6.3.2	Evaluation Metrics	117
6.3.3	Experiment on Sift-Flow dataset	118
6.3.3.1	Highway dataset	118
6.3.3.2	House dataset	121
6.3.4	Experiment on Pascal-Voc dataset	122
6.3.4.1	sheep dataset	122
6.3.4.2	Horse rider dataset	124
6.3.4.3	Horse keeper dataset	125
6.3.5	Comparative Analysis	126
6.3.5.1	Comparison with other CNN based approaches	126
6.3.5.2	Conditional random field(CRF) and total variations(TV) based analysis	128
6.3.6	Effect of CNN parameters on Accuracy	129
6.3.7	Effect of CNN parameters on Computation Time	130
6.3.8	Impact of EFF parameters	131
6.3.9	Statistical Test	133
6.4	Conclusions and Future Work	135
7	LM-MFP: Large Scale Morphology and Multi-criteria Feature pooling based Image Labelling	137
7.1	Introduction	137
7.2	Proposed Method:(LM-MFP)	139
7.2.1	Morphological Feature Extraction	140
7.2.2	LM-MFP based feature pooling method	144

7.3	Experimental Result Analysis	148
7.3.1	Dataset detail	148
7.3.2	Evaluation Metrics	148
7.3.3	Experiment 1: Semantic prediction on Sift-Flow dataset	149
7.3.3.1	highway dataset	149
7.3.3.2	house dataset	152
7.3.4	Experiment 2: Semantic prediction on Pascal-Voc dataset	153
7.3.4.1	Sheep dataset	153
7.3.4.2	Horse Rider dataset	154
7.3.4.3	Horse Keeper dataset	155
7.3.5	Experiment 3: Category prediction on YALE and ORL dataset	156
7.3.6	Comparative Analysis	157
7.3.7	Effect of correlation, SSIM, Mutual information, and Training samples on Overall accuracy (OA)	159
7.3.8	Complexity, Computation Time	159
7.3.9	Accuracy, Precision, Recall, Kappa, F-score, computation time for LM and LM-MFP	160
7.4	Conclusions and Future Work	162
8	Concluding Remarks and Future Directions	165
8.1	Contribution Summary	165
	Deep CNN Feature Fusion with Manifold Learning and Regression	165
	CNN and correlation based Salient Features for Image Labelling	166
	Two phase label optimization, and efficient edge preservation scheme(EPS) for Relaxation Labelling	166
	CNN-EFF: CNN based Edge Feature Fusion	166
	LM-MFP: Large Scale Morphology and Multi-criteria Feature pooling	167
8.2	Scope for Future Work	167
	Graph based embedding and matching techniques for informative knowledge fusion	167
	Application of CNN based metric learning, extreme learning machines for better feature search	167
	Variational optimization by using convex methods and classical techniques	168
	Brain disease classification on MRI, fMRI and medical hyper-spectral images by using classical and deep learning techniques	168

Application on other semantically labelled data-sets for camouflage detection and object detection in images	168
A Variational Optimization	169
A.1 Energy function	169
A.2 Variations by Minimizing Least Square Regression	170
B List of Publications	173
 Bibliography	 175
Bibliography	175

List of Figures

3.1	Deep feature extraction framework using principle components followed by convolution map generation	21
3.2	False colour image and Ground truth for (a)Indian Pines(IP)(b)Pavia University(PU) (c)Salinas valley(SV) datasets	24
3.3	original image and class mixing image for I=2	25
3.4	Deep spatial features for IP,PU and SV in AlexNet	29
3.5	Data distribution on principal components(pc) with spectral datacube(a, b, c) and Deep spatial-spectral datacube(d, e, f)	29
3.6	Laplace and spatial potential graphs of IP, PU,SV for AlexNet architecture	30
3.7	Data distribution on first three Embedded Manifold for spatial-spectral combined datacube obtained from graph based manifold embedding of IP, PU, SV datasets	31
3.8	(a,d)-class mixing images for I=2, (b,e)- classification results obtained from Regional probability with (OA=72.95 %) and (OA=73.03 %), (c,f)- Gobal probability for spatial-spectral datacube of IP and PU respectively .	32
3.9	Class-wise Regional probability images for spectral datacube of IP	32
3.10	class mixing image(a) and Regional probability images(b to j) for spectral datacube of PU	33
3.11	Regional and global probability fusion for IP,PU,SV datasets for AlexNet, VGG-16 and VGG-19	35
3.12	OA vs λ_2 for Regional-global probability fusion of spatial-spectral IP, PU, SV datasets	37
3.13	Computation Time for IP, PU, SV datasets	38
3.14	Overall Accuracies for different models in (a) Indian Pines(b)Pavia University(c) Salinas Valley Dataset	39
4.1	Deep spatial features for IP,PU and SV in AlexNet	50
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	51
4.3	The Salient feature obtained from the combined Deep spatial and spectral datacube	54
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	56

4.5	Overall Accuracies for different models for (a) Indian Pines(b)Pavia University(c) Salinas Valley Dataset	57
4.6	Classification Results on selected bands with SVM and LDA predictors for Indiana Pines Data Set	60
4.7	Classification Results on selected bands with SVM and LDA predictors for Salinas Valley Data Set	61
4.8	Classification Results with selected bands and SVM and LDA predictors for Pavia University Data Set	61
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	63
4.10	Salient Dimensions selection by using MO(Maximum Objects) in function peaks	65
5.1	Process flow of proposed method	70
5.2	Classification and Clustering with spectral and parametric kernel method in projected subspaces	77
5.3	Overall Accuracy(OA) Vs Training samples	78
5.4	Overall Accuracy(OA) Vs σ_1	78
5.5	Cluster energy minimization	81
5.6	Cluster,Contours and Label mapping on image by paraKERNELGC method on Indaian pines data	82
5.7	Classification and Clustering with subspaceMLR and for Indaian pines data	83
5.8	Cluster energy, Cluster labels,Contours and Label mapping on image by paraKERNELGC method on Salinas valley data	84
5.9	Classification and Clustering results(in%) with spectral spatial attribute,spectral data,LDA and QDA in projected subspaces for Salinas valley Data Set	85
5.10	86
5.11	Process flow diagram for EPS optimization of central pixel	87
5.12	simulated image with spectral signatures	92
5.13	Classification and class relaxation outcome	93
5.14	Overall Accuracy Vs Regularizer(λ)	95
5.15	Error term Vs Number of Iterations	95
5.16	Overall Accuracy Vs Error term	96
5.17	(a)Ground truth (b)False color image (c)Feature map (d)One class map, for Indian Pines Image	99
5.18	Class probability image for Corn-no till class in 4 cases for Indian Pines Image	99
5.19	Classification Results and Accuracy(in%), obtained For Indian Pines	100
5.20	Class probability image for Bare-soil class of Pavia university dataset	101
5.21	Classification results for Pavia university dataset	102
5.22	(a)-(f):Class probability images for class2 of Salinas Data, and (g)-(l):Classification results for Salinas valley Dataset	103

5.23	Overall and Average Accuracy for IP, PU, SV Datasets	104
5.24	Hysime+LDA and Hysime+QDA results on IP, PU, SV Datasets [(a)to(f)] and MSE, Projection error and noise power in Hysime projection [(g)to(i)]	105
6.1	Process flow for CNN-EFF method	109
6.2	Pixel-ordering in 8-Neighbour computation for pixels-5(central pixel) . . .	115
6.3	Confusion matrix, Training Loss, MIOU of CNN-EFF method for highway and house image from sift-flow dataset	119
6.4	Probability images and classification results for Highway data(sift-flow) .	120
6.5	Probability images and classification results for House data(siftflow) . . .	121
6.6	Confusion matrix, Training Loss, MIOU of CNN-EFF method for sheep, horse rider and horse keeper image from pascal-voc dataset	123
6.7	Probability images and classification results for sheep data(pascal-voc) . .	124
6.8	Probability images and classification results for Horse-rider data(pascal-voc)	125
6.9	Probability images and classification results for horse-keeper data(pascal-voc)	126
6.10	Results of Comparative approaches in sift-flow, pascal-voc datasets	128
6.11	Results of Comparative approaches for PCNN+CRF from 11-(a) to (e) and for $PCNN + TV_{lag}$ from 11-(f) to (j)	130
6.12	Classification Accuracy vs CNN parameters for images of sift-flow and pascal-voc datasets	131
6.13	Computation Time vs CNN parameters for images of sift-flow and pascal-voc datasets	132
6.14	Accuracies vs iterations and Accuracy vs μ for sift-flow and pascal-voc datasets	133
7.1	Methodology for LM-MFP method	138
7.2	Process flow for LM-MFP method	139
7.3	Pooled features and classification results for highway data(siftflow)	151
7.4	Pooled features and classification results for highway data(siftflow)	152
7.5	Pooled features and classification results for sheep data(pascal-voc)	154
7.6	Pooled features and classification results (in %) for horse-rider data(pascal-voc)	155
7.7	Pooled features and classification results for horse-keeper data(pascal-voc)	156
7.8	Category prediction results for YALE(8(a) to 8(j)) and ORL(8(k) to 8(t)) datasets	157
7.9	Results of Comparative approaches in sift-flow, pascal-voc datasets	158
7.10	Classification Accuracy vs LM-MFP parameters for images of sift-flow and pascal-voc datasets	160
7.11	Computation Time vs LM-MFP parameters for images of sift-flow and pascal-voc datasets	161
7.12	Accuracies, Kappa, Precision, Recall, F-score, and Time	162

List of Tables

3.1	Spatial features extracted for different architecture	28
3.2	Result of spatial-spectral data on Indian Pines Image	36
3.3	Result of spatial-spectral data on Pavia University Dataset	36
3.4	Result of spatial-spectral data on Salinas valley Image	36
4.1	training and test samples for Indian Pines Data	49
4.2	training and test samples for Pavia university Data	49
4.3	training and test samples for Salinas Valley Dataset	49
4.4	Numbers of spatial features Extracted from Different CNN Architectures .	51
4.5	Result on Indian Pines Image(in %) with spatial features	52
4.6	Result for Pavia University Dataset(in %) using spatial features	52
4.7	Result on Salinas Valley Image(in %) with spatial features	53
4.8	Result of spatial spectral data on Indian Pines Image with salient features	53
4.9	Result of spatial spectral data on Pavia University Dataset	57
4.10	Result of spatial spectral data on Salinas valley Image	58
4.11	Running time for proposed(CNN trainig time+feature learning time for 500 features) and past methods	58
4.12	$\lambda=0.1$ and $c_{\min}=0.8$, Accuracies(in %) for Indian Pines Data	59
4.13	$\lambda=0.1$ and $c_{\min}=0.8$,Accuracies(in %) for Salinas Data	60
4.14	$\lambda=0.1$ and $c_{\min}=0.8$,Accuracies(in %) for Pavia University Data	61
4.15	Performance (in %) in state-of-art methods	62
5.1	Class wise and Overall performance of proposed method in Indiana Pines Dataset	81
5.2	Class wise and Overall performance of proposed method in Salinas Valley Dataset	84
5.3	Overall accuracy(in %) for recently proposed methods	85
5.4	Class-wise and Overall performance of proposed method in Indiana Pines(IP) Dataset	97
5.5	Class wise and Overall performance of proposed method in Pavia University (PU) Dataset	100
5.6	Class wise and Overall performance of proposed method in Salinas Valley (SV) Dataset	103
5.7	Overall accuracy for Inidain pines and pavia dataset for past methods . . .	105

6.1	Testing results for sift-flow dataset	119
6.2	performance of CNN-EFF and CNN model on pascal-voc dataset	122
6.3	Comparative Analysis	128
6.4	<i>PCNN + CRF</i> Results with Smoothness factor $\beta=2$ and Accuracies(in %)	129
6.5	<i>PCNN + TV_{lag}</i> method with $\lambda_{Reg}=0.1$ and Accuracies(in %)	129
6.6	McNemar's Test with 5% Level of Significance	134
7.1	Testing results for sift-flow dataset	150
7.2	performance of LM and LM-MFP model on pascal-voc dataset	153
7.3	Comparative Analysis	158

Abbreviations

DNN	Deep Neural Network
CNN	Convolution Neural Network
ML	Machine Learning
RGB	Red Green Blue Image
MLP	Multi Layer Perceptron
SVM	Support Vector Machine
OA	Overall Accuracy
AA	Average Accuracy
K	Kappa Value
SVD	Singular value Decomposition
PCA	Principal Component Analysis
HD	High Dimensional
RL	Relaxation Labelling
MRF	Markov Random Field
CRF	Conditional Random Field
IOU	Intersection Over Union
MLR	Multinomial Regression
MO	Maximum Object Signal
GC	Graph Cut
EPS	Edge Preservation Scheme
RBF	Radial Basis Function
MAP	Maximum A Posterior

paraKERNAL	parametric Kernal
EFF	Edge Feature Fusion
FC	Fully Connected
LM	Large scale Morphology
MFP	Multi criteria Feature Pooling
SSIM	Structural Similarity
MI	Mutual Information

Symbols

exp	exponential
γ	$\frac{spectral}{spatial}$
λ	eigenvectors or regularizers
α	regularizer
$\phi_{(k)}$	length k subspaces
λ_1, λ_2	multiregularizers
λ_d	step length for new dimensions
μ	regularizer
c_{min}	minimum correlation
σ	standard deviation
ϕ	relaxation prior
\oplus_N^j	Fusion operator on nearest neighbour and features
Δ	gradient
η	Learning Rate