Declaration

I hereby declare that the work presented in this Thesis titled *Some Methods for Annotation Localization and Writer Identification for Processing Annotated Documents* submitted to the Indian Institute of Technology Jodhpur in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy, is a bonafide record of the research work carried out under the supervision of Dr. Gaurav Harit. The contents of this thesis in full or in parts, have not been submitted to, and will not be submitted by me to, any other Institute or University in India or abroad for the award of any degree or diploma.

Shilpa Pandey
PG201282501

Certificate

This is to certify that the thesis titled *Some Methods for Annotation Localization and Writer Identification for Processing Annotated Documents*, submitted by *Shilpa Pandey (PG201282501)* to the Indian Institute of Technology Jodhpur for the award of the degree of *Doctor of Philosophy*, is a bonafide record of the research work done by her under my supervision. To the best of my knowledge, the contents of this report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Gaurav Harit
Ph.D.Thesis Supervisor

Acknowledgments

It is actually difficult to remember all the experiences, persons and incidents, which have contributed with their roles to shape me as the present individual. However, none of these is insignificant, as these are the lessons only associated with each of those experiences, persons and incidents which have helped me to stand still and keep my focus undeviated.

Firstly, I would like to express my sincere gratitude to my advisor, *Dr. Gaurav Harit*, for providing necessary infrastructure and resources to accomplish my research work. I deeply extend my sincere thanks for his patience, motivation, and immense knowledge. His guidance helped me for the entire time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study. In particular, I am grateful to him for enlightening me the first glance of research. Besides my advisor, I would like to thank the rest of my thesis committee: *Dr. Anil Kumar Tiwari, Dr. Gaurav Bhatnagar*, and *Dr. Venkata Ramana Badarla*, for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

I am grateful to my fellow lab mates *Ms. Hiteshi Jain, Mrs. Ridhi Agarwal, Ms. Divya Shrivastav* and *Ms. Divya Sharma* for providing a stimulating and fun-filled work environment. They have been an endless source of support and motivation. My thanks go in particular to *Ms. Hiteshi Jain* with whom I had many rounds of discussions on various concepts of the course. I admire her distinguished helping nature. I voluminously thank my friends *Dr. Deepak Bharti, Mr. Ishan Varun, Mr. Parveen* and *Ms. Vibha.* They have been the reasons, which have made me smile amidst tough situations. Recalling ever this wonderful time, which I have spent with them at *IIT Jodhpur*, will always be soul-soothing. Lots of memories to cherish…

I would also like to extend huge, warm thanks to my roommates, *Ms. Priya Malpani, Ms. Sana Maidulla, Mrs. Vandana Dahiya* and *Ms. Shivanjali Saxena* for providing their lively company during the research. I also thank them for taking care of me in a very affectionate and homely way. It was their affection and concern which never made me feel away from home. I thank them for their endless motivation through all the sleepless nights I had working before deadlines. With them, for the last five years, I had a lot of fun.

I expand my gratitude to my grandparents *Late. Mrs. Revati Pandey* and *Late. Mr. R.D.Pandey* and parents *Late. Mrs. Asha Pandey* and *Mr. D.N.Pandey* for their unprecedented support and blessings. I acknowledge the trust which they have kept in my capabilities. I also value the commendable levels of patience of my parents for withholding their plans for long for sake of my career. I am also grateful to my brother *Mr. Prateek Pandey* and sister-in-law *Mrs. Ruchi Pandey* for their sincere encouragement and inspiration throughout my research work and lifting me uphill this phase of life. I owe so much to them.

Above all, I owe it all to Almighty God for granting me the wisdom, health, and strength to undertake this research task and enabling me to its completion. I'm blessed and I thank God for each day and everything that happened to me.

" The true sign of intelligence is not knowledge but imagination... "

Shilpa Pandey
Ph.D. Student

List of Figures

Figure	*Title*	*page*
1.1	Structural Diversities Among Documents	3
1.2	Unconstrained Annotations for Non-controlled Environment. Writing within the margins, between the paragraphs, multi-oriented text-lines, overlapping with the printed text, and presence of symbolic annotations like arrows, underlines, cuts, and encirclement are the examples of the unconstrained annotations.	4
1.3	Organization of the Thesis	6
2.1	Examples of such documents [warrelics.eu, 2010; Japanese Military Postcards, 1931] in real environment where the text is broken to characters and encourages segmentation of handwritten and printed content at character level.	11
2.2	Examples of such documents in real environment where large chunks of text persists and encourages segmentation of handwritten and printed content at Block Level.	13
2.3	Examples of such documents [Genealogy Photos, 2010; Birth Certificate, 2018] in real environment where mixed types of text persists and encourages segmentation of handwritten and printed content at word/connected component Level.	15
2.4	Multi-oriented Handwritten Dataset used by the work [Pandey and Harit, 2015].	20
2.5	Examples of such documents [Peng *et al.*, 2013; Pandey and Harit, 2015] in real environment where the text is of overlapping nature and encourages segmentation of handwritten and printed content at Pixel Level.	22
2.6	Examples of such documents [Lin Yangchen, 2015; Nathan Raab, 2017] where long independent text-lines of single category persist and encourages segmentation of handwritten and printed content at text-line Level.	24
3.1	Discriminating Elements in Handwriting.	30
4.1	Kinds of Handwritten Annotations.	56
4.2	Contrast among Controlled and Non-controlled Annotation Documents.	57
4.3	Type of documents processed/included in Our Dataset for extracting annotations. The blue bounding boxes represent hand annotated pseudo words while the red bounding boxes represent the printed pseudo-words.	59
4.4	Schematic diagram of the proposed scheme for Handwritten and Printed Segmentation	61
4.5	Extraction of pseudo-words by morphological analysis. At first the boundaries of foreground (dark) regions in the original image is enlarged by dilation mechanism using a structuring element. Then the foreground region of the dilated image is shrunk by erosion mechanism using the same structuring element. The resulting image is then smoothed by a median filter and finally connected blocks of pixels are boxed as patches using boundary extraction algorithm.	61
4.6	Maximum Horizontal Runlength for Printed and Handwritten Pseudo-word.	62
4.7	Maximum Height of the Connected Components within a Printed and Handwritten Pseudo-word.	62
4.8	Overlap-area of the connected components within a pseudo-word/patch.	63
4.9	Vertical Projection Profile is drawn by Counting the Foreground Pixels column wise.	63
4.10	Selection of Sample Points.	64
4.11	Envelop Straightness Feature Value Computation.	64
4.12	Envelop Straightness	65

4.13 Sample result of annotations and printed text extraction with Graph Cuts framework on one(two) image(s) from **our dataset** 68
4.14 Sample result of annotations and printed text extraction with Graph Cuts framework on one(two) image(s) from **IAM dataset**. 69
4.15 Results of Feature Sets mentioned by [Peng *et al.*, 2013] and [Benjlaiel *et al.*, 2014] in contrast to proposed Feature Set on Our Dataset using the Graph Cut Framework. 70
4.16 Results of Feature Sets mentioned by [Peng *et al.*, 2013] and [Benjlaiel *et al.*, 2014] in contrast to proposed Feature Set on IAM Dataset using the Graph Cut Framework. 70

5.1 Salient object and its Neighborhood. 72
5.2 Types of visual saliency. 72
5.3 Types of annotations used in our work. 76
5.4 Set 1: All kinds of annotation localization in documents using CRF. 83
5.5 Set 2: Underlined region localization in multi-annotated images using CRF 85
5.6 Set 2: Marginal annotation region localization in multi-annotated images using CRF. 86
5.7 Set 2: Encircled region localization in multi-annotated images using CRF 87
5.8 Set 2: Inline annotation region localization in multi-annotated images using CRF 88
5.9 Set 3: Textual region localization in documents comprising both Textual and Symbolic annotations using CRF 88
5.10 Set 3: Symbolic Region Localization in Documents comprising both Textual and Symbolic Annotations using CRF 89
5.11 Set 4: Underlined annotation localization in single-class annotated images using CRF 90
5.12 Set 4: Marginal annotation localization in single-class annotated images using CRF 91
5.13 Set 4: Encircled annotation localization in single-class annotated images using CRF 91
5.14 Set 4: Inline annotation localization in single-class annotated images using CRF. 92
5.15 Textual region localization in IAM images using CRF. 93
5.16 Textual region localization in PRImA images using CRF. 93
5.17 All kinds of annotation localization in documents using SVM. 94

6.1 Set 1: All kinds of Annotation Localization in Documents using DS. 104
6.2 Set 2: Underlined Region Localization in Multi-annotated Images using DS 105
6.3 Set 2: Marginal Annotation Region Localization in Multi-annotated Images using DS. 106
6.4 Set 2: Encircled Region Localization in Multi-annotated Images using DS 107
6.5 Set 2: Inline Annotation Region Localization in multi-annotated Images using DS 108
6.6 Set 3: Textual Region Localization in Documents comprising both Textual and Symbolic Annotations using DS 109
6.7 Set 3: Symbolic Region Localization in Documents comprising both Textual and Symbolic Annotations using DS 110
6.8 Set 4: Underlined region localization in single-class annotated images using DS 111
6.9 Set 4: Marginal annotations localization in single-class annotated images using DS 112
6.10 Set 4: Encircled annotation localization in single-class annotated images using DS 113
6.11 Set 4: Inline annotation localization in single-class annotated images using DS 114
6.12 Textual region localization in IAM images using DS. 115
6.13 Textual region localization in PRImA images using DS. 115

7.1 Variety of handwriting challenges for core-region extraction. 117
7.2 A word divided into its corresponding zones. 118
7.3 Zoning by using Horizontal Projection Profile Method 119
7.4 Erroneous detection of reference lines due to longer horizontal strokes by Horizontal Projection Profile method. 119
7.5 The peaks of the horizontal stroke are discarded in the HPP derivative while the peaks of the core-region are preserved [Bozinovic and Srihari, 1989]. 119

7.6 Contiguous set of pixels are drawn on the upper envelop. The pixel marked as **U** belongs to the upper envelop. 122
7.7 The pixels numbered from **1 --- 8** represents the sampled sequence of pixels on the upper envelop. 122
7.9 The pixels **2, 4, 5** and **7** are the maximal points selected when matched with the Topological Patterns for the upper reference line. 123
7.11 Partitioning the computed maxima points into two sets of core and non-core region maximal points. The first and the last maximal points i.e. **1** and **8** are also included. 124
7.12 Histogram indicating the percentage of word images having core-region area mismatch with its corresponding ground truth core-region for specified percentage of pixels mismatch. 125
7.13 A few sample results of our proposed method to identify core regions in handwritten words. The region between blue lines indicates the core region obtained by our proposed method while the region between red lines indicates the core region marked as ground truth. 127

8.1 The three types of handwritten Word Allographs from CVL Dataset 130
8.2 The four topological masks with P as the central pixel, the dark region as the foreground pixel and the filled pattern as *do not care* care region. 134
8.3 Results for grapheme based features with non overlapping and overlapping window. 136
8.4 Results for Character based features with non overlapping and overlapping window. 137
8.5 Performance comparison between grapheme and character feature clusters. 137
8.6 Identification rate improvement between [Slimane and Margner, 2014] and our method using graphemes. 138

List of Tables

Table	*Title*	*page*
2.1	Visual Differences Between Machine Printed and Handwritten Text.	10
2.2	Overview of dataset, performance and methods applied for handwritten character classification from printed text.	12
2.3	Overview of Dataset, Performance and Methods applied for Classification of Handwritten Text Blocks from Printed Text Blocks.	14
2.4	Overview of Dataset, Performance and Methods applied for Handwritten Words Classification from Printed Text.	17
2.5	Overview of dataset, performance and methods applied for classification of handwritten connected components from printed connected components.	21
2.6	Overview of Dataset, Performance and Methods applied for Classifying Handwritten Pixels from Printed Pixels.	23
2.7	Overview of Dataset, Performance and Methods applied for Handwritten Text-line Classification from Printed Text-line.	25
2.8	Overview of Post Processing Methods applied for Classification of Handwritten Text Printed Text.	27
3.1	Writer Identification: Dependent vs Independent Text.	31
3.2	Distinction Between Writer Identification and Writer Verification.	32
3.3	Overview of dataset, performance and methods applied for writer identification by means of feature based character shape analysis.	34
3.4	Overview of dataset, performance and methods applied for writer identification by means of feature based grapheme/connected component shape analysis.	35
3.5	Overview of dataset, performance and methods applied for writer identification by means of codebook based character shape analysis.	36
3.6	Overview of dataset, performance and methods applied for writer identification by means of codebook based grapheme/connected Component shape analysis.	38
3.7	Overview of dataset, performance and methods applied for writer identification by means hybrid approaches.	39
3.8	Overview of dataset, performance and methods applied for writer identification of a document by means of local and global features.	43
3.9	Overview of dataset, performance and methods applied for writer identification of a document by means of word property analysis.	48
3.10	Overview of dataset, performance and methods applied for writer identification of a document by means of text-line property analysis.	51
4.1	Types of annotations present in dataset	60
4.2	Comparison in Graph Cut, K-means and Agglomerative Clustering on **our Dataset**	60
4.3	Comparison in Graph Cut, K-means and Agglomerative Clustering on **IAM Dataset**	60
4.4	*Variation of Envelop Straightness (ES)* Feature Values for Printed Text and Annotations.	65
4.5	ES performance	66
4.6	Description of The Feature Sets Used in our Experiment	66
4.7	Annotations and printed text extraction rates with Graph Cuts framework on **Our Dataset**	67

4.8 Annotations and printed text extraction rates with Graph Cuts framework on **IAM Dataset** 67

5.1 Set 1: Annotation localization when the dictionary is trained on images comprising all annotations and the testing is performed on similar images. 83
5.2 Set 2: Annotation localization when the dictionary is trained on individual annotations and the testing is performed on the images containing all the annotations. 84
5.3 Set 3: Annotation localization when the dictionary is trained on images comprising only textual annotations and the testing is performed on similar images, and vice versa. 84
5.4 Set 4: Annotation localization when the dictionary is trained on individual annotations and the testing is performed on the images containing individual annotations. 84
5.5 Annotation localization on IAM and PRImA dataset by CRF textual annotation detection model. 86
5.6 Comparison of SVM with CRF on multi-annotated documents. 89
5.7 Comparison of SVM with CRF on single-annotated documents. 89

6.1 Set 1: Annotation localization when the dictionary is trained on images comprising all annotations and the testing is performed on similar images. 104
6.2 Set 2: Annotation localization when the dictionary is trained on individual annotations and the testing is performed on the images containing all the annotations. 105
6.3 Set 3: Annotation localization when the dictionary is trained on images comprising only textual annotations and the testing is performed on similar images, and vice versa. 108
6.4 Set 4 : Annotation localization when the dictionary is trained on individual annotations and the testing is performed on the images containing individual annotations. 108
6.5 Annotation localization on IAM and PRImA dataset by DS textual annotation detection model 109
6.6 Comparison of SVM with DS on multi-annotated documents. 109
6.7 Comparison of SVM with DS on single-class annotated documents. 110

7.1 Composition of Dataset along with the achieved accuracies. 124
7.2 Performance accuracies of our core-region detection method for handwritten words in comparison with the *state-of-the-art* methods on different datasets. 126

8.1 Writer identification Rate for an overlapping window at grapheme level 135
8.2 Writer identification rate for non-overlapping window at grapheme level 135
8.3 Writer Identification Rate for an overlapping window at Character Level 135
8.4 Writer Identification Rate for non-overlapping window at Character Level 135
8.5 Writer Identification rate for an overlapping window based on [Slimane and Margner, 2014] 136
8.6 Writer Identification rate for non-overlapping window based on [Slimane and Margner, 2014] 136

List of Symbols

Symbol	*Description*
$\mathbf{B}$	Bases vector dictionary
$\mathbf{X}$	Grid of patches
$\mathbf{Y}$	Label set for grid of patches
D	Dictionary for image patches
$\mathbf{y}$	Label for a single patch
$\mathbf{x}$	Target patch
λ	Regularization Term
$\mathbf{S}$	Sparse latent variable
$\mathscr{G}$	Graph
υ	Nodes
ε	Edges
$\mathbf{w}$	Weight vector
$\mathscr{X}$	Set of training instances
$\mathscr{Y}$	Set of labels for training instances
γ	Weight regularization term
$\mathbf{O}$	Saliency Map
$\mathbf{A}$	an input descriptor matrix
$\mathbf{s}_i$	sparse latent variable
sim	similarity function
dis_i	distance function
p	pseudo-word
σ	neighborhood width control parameter

List of Abbreviations

Symbol	*Description*
BoVW	Bag of Visual Words
CC	Connected Component
CLPSD	Character Lowermost Point Standard Deviation
CNN	Convolution Neural Network
CRF	Conditional Random Field
DCM	Dissimilarity Counter Method
GSC	Gradient, Structural and Concavity
GSCM	Grey-Scale Co-Occurrence Matrix
HMM	Hidden Markov Model
HMT	Hidden Markov Tree
ICR	Intelligent Character Recognition
IFK	Improved Fisher Kernels
KLD	Kullback-Leibler Distance
LLC	Locality-constrained Linear Coding
MGF	Multi-channel Gabor Filtering
MLP	Multi-Layer Perceptron
MRF	Markov Random Field
OCR	Optical Character Recognition
PCA	Principal Component Analysis
PDF	Probability Density Function
RLSA	Run Length Smoothing Algorithm
SDLFD	Spectrum-Domain Local Fluctuation Detection
SGONG	Self-Growing and Self-Organized Neural Gas
SIFT	Scale-Invariant Feature Transform
SVM	Support Vector Machine
SOF	Sorted Order of Features
SOM	Self-Organizing Map
TAS	Triple-Adjacent-Segment
UBM	Universal Background Model
VLAD	Vectors of Locally Aggregated Descriptors
WED	Weighted Euclidean Distance
WEKA	Waikato Environment for Knowledge Analysis