

Contents

1	Introduction	1
1.1	Handwritten Character Recognition (HCR)	1
1.1.1	Applications of HCR	2
1.1.2	Challenges in development of HCR system	3
1.2	Handwritten character recognition approaches	4
1.2.1	Template matching technique	5
1.2.2	Techniques based on handcrafted features	5
1.2.3	Techniques based on deep learned features	5
1.2.4	Techniques based on combination of handcrafted and deep-learned features	6
1.3	Historical background and evolution of present HCR systems . . .	6
1.4	Meitei Mayek	8
1.4.1	Origin and history of Meitei Mayek	9
1.4.2	Present day Meitei Mayek	10
1.5	Existing works in HCR of Meitei Mayek	12
1.6	Motivation	13
1.7	Objective	14
1.8	Dissertation contributions	15

1.8.1	Creation of a Meitei Mayek handwritten character dataset (TUMMHCD) and performance analysis of existing features with state-of-the-art classifiers	15
1.8.2	Development of a Convolutional Neural Network (CNN)-based recognition for the developed dataset.	16
1.8.3	A multilevel recognition of Meitei Mayek handwritten characters using fusion of feature strategy.	16
1.8.4	Identification of script-specific properties for achieving enhanced recognition accuracy.	17
1.8.5	Incorporation of language model in the recognition system.	17
1.9	Organisation of the thesis	18
2	Dataset creation and performance analysis of state-of-the-art classifiers with existing features	20
2.1	Existing Datasets	20
2.2	Creation of TUMMHCD dataset	22
2.2.1	Data acquisition	23
2.2.2	Scanning	23
2.2.3	Pre-processing	24
2.2.4	Resizing	25
2.3	Performance analysis of existing techniques on TUMMHCD	26
2.3.1	Recognition using handcrafted feature descriptors	28
2.3.2	Recognition using image pixel intensity (IPI) values	31
2.4	Discussion and Conclusion	31
3	CNN-based recognition of TUMMHCD	34
3.1	Convolutional neural network (CNN)	34

3.2	Related work	38
3.2.1	CNN for HCR of Indic scripts	39
3.2.2	CNN for HCR of Meitei Mayek	40
3.3	Performance analysis of state-of-the-art CNN models on TUMMHCD	40
3.4	Building CNN from scratch for recognition of TUMMHCD	44
3.4.1	Base CNN architecture	44
3.4.2	Proposed CNN architecture	47
3.5	Experimental results and discussion	49
3.5.1	Analysis w.r.t feature-classifier combinations	50
3.5.2	Analysis w.r.t pre-trained CNN models	51
3.5.3	Performance analysis of TUMMHCD with noise	52
3.6	Conclusion	54
4	A multilevel recognition of TUMMHCD using fusion of features strategy	55
4.1	Introduction	55
4.2	Related work	57
4.3	Proposed methodology	58
4.3.1	First level recognition module	59
4.3.2	Softmax module	60
4.3.3	Filtering module	62
4.3.4	Second level recognition module	64
4.4	Experimental results and discussion	67
4.4.1	Overall system performance	67

4.4.2	Performance with different training scenarios	68
4.4.3	Performance with different filtering threshold values	69
4.4.4	Performance on MNIST, DIDA and CArDIS	69
4.4.5	Discussion	70
4.5	Conclusion	71
5	Zone and rule assisted recognition of Meitei Mayek handwritten characters	72
5.1	Introduction	72
5.2	Related work	73
5.3	Features of Meitei Mayek	78
5.4	Proposed methodology	81
5.4.1	Training CNN	83
5.4.2	Finding vulnerable set	84
5.4.3	Zone identification	87
5.4.4	Rules used	91
5.5	Experiments	91
5.5.1	Dataset details	91
5.5.2	Experimental results on zone identification	92
5.5.3	Evaluation metrics	92
5.5.4	Overall experimental results	93
5.5.5	Error analysis	95
5.5.6	Comparison with other existing works	98
5.6	Limitation	98

5.7	Conclusion	99
6	Using LSTM language model with CNN for handwritten character recognition	101
6.1	Introduction	101
6.2	Related work	103
6.3	Proposed methodology	105
6.3.1	CNN training	106
6.3.2	LSTM language model	106
6.3.3	Testing of test images	109
6.4	Experimental results	111
6.5	Conclusion	115
7	Conclusion & Future Direction	116
7.1	Conclusion	116
7.2	Future Work	118
	Publications based on the Thesis Works	143