
Contents

CONTENTS

CHAPTER No.	TITLE	PAGE No.
1	INTRODUCTION	1
1.1	Overview	1
1.2	Optical Character Recognition (OCR)	2
1.3	Tamil Script	3
1.4	Machine Learning	6
1.5	General Approach to Classification	8
1.6	Image Classification	12
1.7	Multi-Class Classification	13
1.8	Fusion Classifier	15
1.9	Motivation	16
1.10	Problem Statement and Research Objectives	18
1.11	Layout of the Thesis	19
1.12	Conclusion	20
2	REVIEW OF LITERATURE	21
2.1	Tamil Character Recognition System	21
2.2	General Classification Techniques	31
2.3	Multi-Learner Systems	33
2.4	Fusion or Multiple Classification Approaches	34
2.5	Advanced Classification Approaches	37
2.6	Evaluation of Classification Performance	38
2.7	Conclusion	42

3	METHODOLOGY	43
3.1	Classifiers	44
3.2	Image Features	52
3.3	Techniques used during Fusion Classification	54
3.4	Proposed Classification Fusion Systems	57
3.5	Conclusion	61
4	RESULTS AND DISCUSSION	62
4.1	Dataset Used	62
4.2	Performance Metrics	62
4.3	Experimental Results	63
4.4	Conclusion	74
5	SUMMARY AND CONCLUSION	75
	BIBLIOGRAPHY	
	PUBLICATION	

LIST OF TABLES

TABLE No.	TITLE	PAGE No.
2.1	Consolidation of Single and Multiple Classification Techniques	39
4.1	Error Rates of the Classifiers	65
4.2	Accuracy of the Classifiers	68
4.3	Average Efficiency Gain of KNN + SVM Classification Fusion System	70
4.4	Classification Time	71

LIST OF FIGURES

FIGURE No.	TITLE	PAGE No.
1.1	Tamil Scripts	5
1.2	Combinant Tamil Letters	5
1.3	Classification Model	9
1.4	Process of Classification	11
1.5	Paradigms in Image Classification	12
1.6	Fusion Classifier Model	16
1.7	Problem Statement	19
3.1	Feature Vector Data Structure	53
3.2	Class Label Code for each Tamil Character	54
3.3	Hold-out Method	55
3.4	Dataset Split into training and testing set	55
3.5	Proposed Methodology	60
4.1	Error Rate of Classifiers	67
4.2	Average Error Rate of Classifiers	67
4.3	Accuracy of Classifiers	69
4.4	Average Accuracy	69
4.5	Comparison of Single Base classifiers with Fusion Classifiers	72