# Contents

## List of Figures

<table>
<thead>
<tr>
<th>1. INTRODUCTION</th>
<th>1-27</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Optical Character Recognition</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Bi-lingual Document Image Processing</td>
<td>8</td>
</tr>
<tr>
<td>1.4 Review of Literature</td>
<td>10</td>
</tr>
<tr>
<td>1.5 Objectives of the Research Problem</td>
<td>12</td>
</tr>
<tr>
<td>1.6 Methodology and Issues of Bi-lingual Character Recognition</td>
<td>13</td>
</tr>
<tr>
<td>1.7 Proposed System</td>
<td>21</td>
</tr>
<tr>
<td>1.7.1 Pre-processing</td>
<td>22</td>
</tr>
<tr>
<td>1.7.2 Feature Extraction</td>
<td>24</td>
</tr>
<tr>
<td>1.7.3 Classification</td>
<td>24</td>
</tr>
<tr>
<td>1.7.3.1 Nearest Neighbor</td>
<td>25</td>
</tr>
<tr>
<td>1.7.3.2 K-Nearest Neighbor</td>
<td>25</td>
</tr>
<tr>
<td>1.7.3.3 Support Vector Machine</td>
<td>25</td>
</tr>
<tr>
<td>1.7.4 Evaluation Protocol</td>
<td>26</td>
</tr>
</tbody>
</table>

## Feature Extraction Techniques

<table>
<thead>
<tr>
<th>2. FEATURE EXTRACTION TECHNIQUES</th>
<th>28-38</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Grid Based Statistical Features</td>
<td>28</td>
</tr>
<tr>
<td>2.2 Grid Based Wavelet Energy Features</td>
<td>31</td>
</tr>
<tr>
<td>2.3 Comparative Analysis of Feature Extraction Techniques</td>
<td>36</td>
</tr>
<tr>
<td>2.4 Summary</td>
<td>38</td>
</tr>
</tbody>
</table>
3. KANNADA CHARACTER RECOGNITION SYSTEM  39-76
  3.1 Properties of Kannada Characters  39
  3.2 Data set  41
  3.3 Experimental Results on the Grid Based Statistical Features  42-57
    3.3.1 Hand printed Kannada Numerals  42
    3.3.2 Hand printed Kannada Vowels  44
    3.3.3 Hand printed Kannada Consonants  47
    3.3.4 Hand printed Kannada Characters  49
    3.3.5 Printed and Hand printed Kannada Numerals  53
    3.3.6 Analysis of the Grid Based Statistical Features  56
  3.4 Experimental Results on the Grid Based Wavelet Energy Features  58-75
    3.4.1 Hand printed Kannada Numerals  58
    3.4.2 Hand printed Kannada Vowels  61
    3.4.3 Hand printed Kannada Consonants  64
    3.4.4 Hand printed Kannada Characters  68
    3.4.5 Printed and Hand printed Kannada Numerals  71
    3.4.6 Analysis of the Grid Based Wavelet Energy Features  74
  3.5 Comparative Analysis  75
  3.6 Summary  76

4. ENGLISH CHARACTER RECOGNITION SYSTEM  77-106
  4.1 Properties of English Characters  77
  4.2 Data set  79
  4.3 Experimental Results on the Grid Based Statistical Features  80-91
    4.3.1 Hand printed English Numerals  80
    4.3.2 Hand printed English Uppercase Characters  82
    4.3.3 Hand printed English Lowercase Characters  84
    4.3.4 Hand printed English Characters  85
    4.3.5 Printed and Hand printed English Numerals  88
    4.3.6 Analysis of the Grid Based Statistical Features  90
4.4 Experimental Results on the Grid Based Wavelet Energy Features
   4.4.1 Hand printed English Numerals 92
   4.4.2 Hand printed English Uppercase Characters 94
   4.4.3 Hand printed English Lowercase Characters 97
   4.4.4 Hand printed English Characters 99
   4.4.5 Printed and Hand printed English Numerals 102
   4.4.6 Analysis of the Grid Based Wavelet Energy Features 104

4.5 Comparative Analysis 105

4.6 Summary 106

5. BI-LINGUAL KANNADA AND ENGLISH CHARACTER RECOGNITION SYSTEM 107 - 140

5.1 Data set 107

5.2 Experimental Results on the Grid Based Statistical Features 108-123
   5.2.1 Printed and Hand printed Kannada and English Numerals 108
   5.2.2 Hand printed Kannada Vowels and English Uppercase Characters 110
   5.2.3 Hand printed Kannada Vowels and English Lowercase Characters 112
   5.2.4 Hand printed Kannada Consonants and English Uppercase Characters 113
   5.2.5 Hand printed Kannada Consonants and English Lowercase Characters 116
   5.2.6 Hand printed Kannada and English Characters 118
   5.2.7 Analysis of the Grid Based Statistical Features 122
5.3 Experimental Results on the Grid Based Wavelet Energy Features

5.3.1 Printed and Hand printed Kannada and English Numerals 124

5.3.2 Hand printed Kannada Vowels and English Uppercase Characters 126

5.3.3 Hand printed Kannada Vowels and English Lowercase Characters 128

5.3.4 Hand printed Kannada Consonants and English Uppercase Characters 130

5.3.5 Hand printed Kannada Consonants and English Lowercase Characters 132

5.3.6 Hand printed Kannada and English Characters 134

5.3.7 Analysis of the Grid Based Wavelet Energy Features 138

5.4 Comparative Analysis 140

5.5 Summary 140

6. CONCLUSION AND FUTURE RESEARCH 141-142

6.1 Conclusion 141

6.2 Future Research 142

BIBLIOGRAPHY 143

APPENDIX 150