Contents

List of Figures **xiii**

List of Tables **xv**

Abbreviations **xvii**

Symbols **xix**

Preface **xxi**

1 Introduction **1**
1.1 Background . . . 1
1.2 Motivation . . . 4
1.3 Problem Statement and Thesis Objectives . . . 6
1.4 Contribution to the Thesis . . . 8
1.5 Outline of the Thesis . . . 9

2 Theoretical Foundation and Literature Survey **11**
2.1 Background . . . 12
2.1.1 Salient Object Detection . . . 12
2.1.2 Automatic Image Annotation . . . 12
2.2 Literature Review . . . 13
2.2.1 Salient Object Detection . . . 13
2.2.1.1 Statistical and Machine Learning Methods . . . 14
2.2.1.2 Deep Learning Methods . . . 21
2.2.2 Image Annotation . . . 29
2.2.2.1 Machine Learning Methods . . . 30
2.2.2.2 Deep Learning Methods . . . 34
2.2.3 Colon Tumour Localization . . . 37
2.3 Issues and Challenges . . . 41
2.4 Benchmark Databases . . . 46
2.5 Performance Metrics . . . 48

2.6 Conclusion 51

3 Design and Implementation of Salient Object Detection Model using Active Contours and Gradient Vector Flow 53
3.1 Background 54
3.2 Literature of segmentation and saliency computation methods for proposed models 55
3.2.1 Otsu's Thresholding Method 56
3.2.2 Active Contour Model 60
3.2.3 Saliency Computation 61
3.2.4 Gradient Vector Flow 62
3.3 Methods and Models 66
3.3.1 Salient Object Detection using Threshold and Region-Based Segmentation 66
3.3.1.1 Threshold-Based Component 67
3.3.1.2 Region-Based Component 68
3.3.2 An Improved Active Contour Model for Salient Object Detection using Edge Cues 70
3.3.2.1 Generating Initial Contour using Sobel Operator . . 72
3.3.2.2 Pseudocode 74
3.3.3 An efficient modification of Gradient Vector Flow using Directional Contrast for Salient Object Detection and Intelligent Scene Analysis 74
3.3.3.1 Edge Map Formation 75
3.3.3.2 Computation of the gradient of MDC for Saliency . . 75
3.3.3.3 Proposed Energy Functional for Salient Object Detection 76
3.3.3.4 The algorithm of the proposed model 76
3.4 Results 78
3.4.1 Result analysis of threshold based model 78
3.4.2 Result analysis of edge based model 81
3.4.3 Result analysis of GVF based model 82
3.4.4 Setting Parameter Values 84
3.4.4.1 Comparison with GVF and MDC algorithm 85
3.4.4.2 Comparison with other algorithms 86
3.4.4.3 Salient Object and Scene Analysis 89
3.5 Conclusion 91

4 Salient Object Detection using background subtraction, Gabor filters, objectness and directional backgroundness 93
4.1 Background 94

4.2 Literature of backgroundness, Gabor filters, objectness and lazy snapping . . . 94
4.3 Methods and Models . . . 98
4.3.1 Background Subtraction . . . 98
4.3.2 Saliency Map using Minimum Directional Backgroundness . . . 100
4.3.3 Texture segmentation using Gabor Filters . . . 103
4.3.4 Objectness . . . 104
4.3.5 Morphological Operations . . . 106
4.3.6 Boundary correction using Lazy Snapping . . . 107
4.3.7 Algorithm . . . 107
4.4 Results . . . 109
4.4.1 Saliency Results . . . 110
4.4.2 Parameter Change Results . . . 114
4.4.3 Failure Results . . . 115
4.5 Conclusion . . . 116

5 Salient Object Detection using YOLOv2 and Faster R-CNN 117
5.1 Background . . . 118
5.2 Literature of YOLOv2 and Faster R-CNN . . . 119
5.3 Methods and Models . . . 122
5.3.1 User Interactive Salient Object Detection using Yolov2, Lazy Snapping and Gabor Filters . . . 122
5.3.1.1 Selection of pre-trained deep network . . . 122
5.3.1.2 Training . . . 123
5.3.1.3 Procedure of locating salient object . . . 125
5.3.2 Colon Tumour Localization using Three Input Variants to Faster R-CNN and Lazy Snapping . . . 128
5.3.2.1 Block Diagram . . . 128
5.3.2.2 Object Detection on Faster R-CNN using Resnet50 . . . 128
5.3.2.3 Input Variants . . . 130
5.3.2.4 Lazy Snapping . . . 131
5.4 Results . . . 131
5.4.1 Result of YOLOv2 networks for general images . . . 132
5.4.1.1 Parameter Selection . . . 132
5.4.1.2 Ablation Study . . . 132
5.4.1.3 Comparison with state-of-the-art methods . . . 134
5.4.1.4 Failure Results . . . 135
5.4.2 Result of YOLOv2 network for medical images . . . 135
5.4.2.1 Environment Setup . . . 136
5.4.2.2 Exploring Combinations of output from a network trained using three input variants . . . 136
5.4.2.3 Evaluation Parameters . . . 138

5.4.2.4 Choice of Deep Network . . . 138
5.4.2.5 Choice of Input Variants . . . 138
5.4.2.6 Ablation Study Results . . . 139
5.4.2.7 Combination Study Results . . . 141
5.4.2.8 Discussions . . . 141
5.5 Conclusion . . . 143

6 Automatic Image Annotation using Texture Features and Sparse Reconstruction 147
6.1 Background . . . 148
6.2 Survey of salient object detection, feature selection and image annotation methods . . . 148
6.2.1 Salient Object Detection . . . 149
6.2.2 Feature Selection . . . 153
6.2.3 Image Annotation . . . 155
6.3 Method and Model . . . 157
6.3.1 Extraction of LBP, LDP and Color Features . . . 158
6.3.2 Winner of Feature Selection Methods . . . 160
6.3.3 Proposed subset of features: Double Sort Feature Selection . . . 161
6.3.4 Image Annotation . . . 162
6.3.5 Algorithm of the proposed image annotation model . . . 164
6.4 Results . . . 165
6.4.1 Salient Object Detection Results . . . 165
6.4.2 Scene Analysis Results . . . 166
6.4.3 Feature Selection Results . . . 166
6.4.4 Double Sort Feature Selection Results . . . 171
6.4.5 Image Annotation Results . . . 172
6.4.6 Failure Cases . . . 173
6.5 Conclusion . . . 175

7 Conclusion and Future Work 179
7.1 Conclusions . . . 179
7.2 Future Work . . . 181

References 183

List of Publications 215