

# CONTENTS

<b>List of Figures</b>	<b>v</b>
<b>List of Tables</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The segmentation problem . . . . .	1
1.2 Image representation . . . . .	2
1.3 Trends in image segmentation . . . . .	5
1.3.1 Active contour . . . . .	5
1.3.2 Graph-Partitioning approaches . . . . .	7
1.3.3 Classification based approaches . . . . .	8
1.3.4 Structured prediction models . . . . .	9
1.4 Medical imaging . . . . .	10
1.5 Contributions of this thesis . . . . .	11
<b>2 Background on graphical models</b>	<b>17</b>
2.1 Markov random fields . . . . .	18
2.1.1 Example: a 2D MRF for images . . . . .	20
2.1.2 Labeling problem . . . . .	21
2.1.2.1 MAP inference . . . . .	22
2.1.2.2 Parameter learning . . . . .	22
2.1.3 Conditional random fields . . . . .	23
2.2 Inference . . . . .	23
2.2.1 Variational methods . . . . .	24

# CONTENTS

---

2.2.2	Mean field . . . . .	26
2.2.2.1	Mean field approximation . . . . .	26
2.2.2.2	Image denoising with mean field . . . . .	28
2.2.3	Bethe approximation . . . . .	29
2.2.4	Belief propagation . . . . .	30
2.2.4.1	Sum-product algorithm . . . . .	30
2.2.4.2	Max-Product algorithm . . . . .	31
2.2.5	More advanced variational algorithms . . . . .	32
2.2.6	Graph-cuts . . . . .	32
2.3	Parameter Learning . . . . .	33
2.3.1	Maximum Likelihood Estimation . . . . .	33
2.3.2	Approximation methods . . . . .	35
2.3.2.1	Pseudo-likelihood . . . . .	35
2.3.2.2	Contrastive divergence . . . . .	36
2.3.3	Loss-sensitive training . . . . .	36
2.3.4	Maximum-margin learning . . . . .	37
<b>3</b>	<b>Supervoxel-Based Segmentation of Mitochondria in EM Image Stacks with Learned Shape Features</b>	<b>39</b>
3.1	Introduction . . . . .	39
3.2	Related Work . . . . .	41
3.2.1	Mitochondria Segmentation . . . . .	41
3.2.2	Machine Learning in EM Imagery . . . . .	42
3.2.3	Segmentation by Graph-Partitioning . . . . .	43
3.3	Method . . . . .	44
3.3.1	Supervoxel Over-segmentation . . . . .	44
3.3.2	Feature Vector Extraction . . . . .	48
3.3.2.1	Ray Descriptors . . . . .	49
3.3.2.2	Histogram Features . . . . .	52
3.3.3	Graph Cuts with Learned Potentials . . . . .	52
3.3.3.1	Energy Function . . . . .	52
3.3.3.2	Learned Shape Cues in the Unary Term . . . . .	53
3.3.3.3	Learned Boundary Appearance in the Pairwise Term . . . . .	55

3.4	Results . . . . .	56
3.4.1	Experimental Setup . . . . .	56
3.4.2	Parameters and Implementation Details . . . . .	57
3.4.3	Experiments and Evaluation . . . . .	57
3.4.4	Discussion . . . . .	62
3.4.4.1	Computational Advantage of SLIC Supervoxels . . . . .	62
3.4.4.2	Benefits of Ray Descriptors . . . . .	63
3.4.4.3	Learning the Pairwise Term . . . . .	64
3.4.4.4	Comparing against a state-of-the-art method . . . . .	65
3.4.4.5	Failure modes . . . . .	66
3.5	Conclusion . . . . .	67
<b>4</b>	<b>Structured Image Segmentation using Kernelized Features</b>	<b>69</b>
4.1	Introduction . . . . .	69
4.2	Related Work . . . . .	71
4.3	Learning a CRF with Kernelized Features . . . . .	73
4.3.1	CRF for Segmentation . . . . .	73
4.3.2	Learning the CRF Using SSVM . . . . .	74
4.3.3	Linearizing the CRF Energy Function . . . . .	76
4.4	Kernel-transformed Features . . . . .	77
4.5	Results . . . . .	80
4.5.1	Competing Methods . . . . .	80
4.5.2	MSRC Dataset . . . . .	81
4.5.3	Synaptic Gap Dataset . . . . .	84
4.5.4	Mitochondria Dataset . . . . .	85
4.5.5	Discussion . . . . .	86
4.6	Conclusion . . . . .	87
<b>5</b>	<b>Learning for Structured Prediction Using Stochastic Descent with Working Sets</b>	<b>89</b>
5.1	Introduction . . . . .	89
5.2	Related work . . . . .	90
5.3	Formulation . . . . .	92
5.4	Max-margin Learning of CRFs . . . . .	92

# CONTENTS

---

5.4.1	Discriminative Learning . . . . .	92
5.4.2	Max-margin Formulation . . . . .	93
5.4.3	Stochastic Subgradient Descent . . . . .	94
5.5	Estimating Subgradient Using Working Sets . . . . .	95
5.6	Experimental Results . . . . .	98
5.6.1	CRF for Image Segmentation . . . . .	98
5.6.2	Methods . . . . .	99
5.6.3	MSRC Dataset . . . . .	99
5.6.4	Electron Microscopy Dataset . . . . .	102
5.6.5	Time analysis . . . . .	104
5.7	Conclusion . . . . .	106
<b>6</b>	<b>Enhanced techniques for the segmentation of mitochondria</b>	<b>107</b>
6.1	Introduction . . . . .	107
6.2	Encoding geometric interactions . . . . .	108
6.3	Results . . . . .	109
6.4	Semi-automated splitting of merged regions . . . . .	115
6.5	Conclusion . . . . .	116
<b>7</b>	<b>Conclusion</b>	<b>119</b>
<b>8</b>	<b>Appendix</b>	<b>123</b>
8.1	Quadratic complexity of non-linear Structured SVM . . . . .	123
8.2	Supplementary: Learning for Structured Prediction Using Approximate Subgradient Descent with Working Sets . . . . .	125
8.2.1	Convergence properties of the $t^{\text{th}}$ parameter vector . . . . .	125
	<b>Bibliography</b>	<b>129</b>