

Contents

1	INTRODUCTION	1
1.1	Key Contributions	4
1.2	Publications	5
2	BACKGROUND ON DEEP LEARNING	7
2.1	The Argument for Many-Layered Neural Networks	7
2.2	Notation	9
2.3	Difficulties in Learning Deep Architectures	11
2.4	Greedy Layer-wise Training with RBMs	12
2.4.1	Greedy Training with Auto-Encoders	14
2.5	Regularization in Unsupervised Learning	15
2.5.1	Limiting Representation Size	15
2.5.2	Enforcing Representation Sparsity	15
2.5.3	Infinite Training Data	16
2.6	Convolutional Neural Networks	16
2.6.1	Sparse Parameters	17
2.6.2	Weight Sharing	18
2.6.3	Pooling	19
2.6.4	Margin Handling	19
2.7	Cross-Dataset Transfer Learning	19
2.8	Conclusion	21
3	GPU-BASED MACHINE LEARNING	23
3.1	Background on the Backpropagation Algorithm	23
3.1.1	Forward Pass	24
3.1.2	Backward Pass	24
3.1.3	Weight Update	25
3.2	GPU Considerations	26
3.3	CUV: N-Dimensional Arrays in CUDA	27

3.4	Generic Model Optimization with CUVNET	31
3.4.1	Requirements of Deep Learning Algorithms	31
3.4.2	Fast GPU Implementation	31
3.4.3	User Interface	32
3.4.4	Complex Models	33
3.4.5	Differentiation Graph	35
3.4.6	Special Copy-on-write Semantics	35
3.4.7	Monitoring and Visualization	37
3.4.8	Testing	38
3.4.9	Other Major Features	39
3.5	Comparison with other Neural Network Software	39
3.6	Acknowledgments	40
3.7	Applications	41
4	CURFIL: FAST RANDOM FORESTS IN CUDA	43
4.1	Related Work	45
4.2	Random Forests	46
4.3	Visual Features for Node Tests	47
4.4	CURFIL Software Package	49
4.4.1	GPU Kernels	53
4.4.2	Global Memory Limitations	56
4.4.3	Extensions	57
4.5	Experimental Results	58
4.5.1	Datasets	58
4.5.2	Training and Prediction Time	59
4.5.3	Classification Accuracy	60
4.5.4	Incorporating Novel Features	62
4.5.5	Random Forest Parameters	63
4.6	Conclusion	63
5	UNSUPERVISED METHODS FOR IMAGE CATEGORIZATION	65
5.1	Exploiting Local Structure in Boltzmann Machines	65
5.1.1	Background on Boltzmann Machines	66
5.1.2	Local Impact Semi-Restricted Boltzmann Machines ⁺	67
5.1.3	Related Work	69
5.1.4	Experimental Results	70
5.1.5	Conclusions	73

5.2	Investigating Convergence of RBM Learning	73
5.2.1	Background on Restricted Boltzmann Machines	76
5.2.2	Experimental Setup	79
5.2.3	Results	83
5.2.4	Conclusions	85
5.3	Two-Layer Encodings for Semi-supervised Learning	86
5.3.1	Related Work	88
5.3.2	Background	89
5.3.3	Where Pre-Training of One-Layer Encoders Fails	93
5.3.4	Two-Layer Encoders and Contractive Regularization	95
5.3.5	Two-Layer Encoders and Shortcut Connections	96
5.3.6	Experiments	96
5.3.7	Discussion	105
5.3.8	Conclusions	106
6	LEARNING OBJECT-CLASS SEGMENTATION	109
6.1	Learning Object Class Segmentation with CNN	110
6.1.1	Methods	110
6.1.2	Results	114
6.1.3	Related Work	117
6.1.4	Conclusion	119
6.2	Encoding Depth Information for CNN	119
6.2.1	Methods	120
6.2.2	Experiments	121
6.2.3	Related Work	122
6.2.4	Conclusion	124
6.3	Depth and Height Aware Semantic Perception	125
6.3.1	Related Work	127
6.3.2	Methods	128
6.3.3	Experiments	130
6.3.4	Conclusion	133
6.4	Recurrent Networks for Depth Perception	133
6.4.1	Model Description	135
6.4.2	Related Work	139
6.4.3	Experiments	141
6.4.4	Conclusion	151
6.5	Chapter Summary	152

7	STRUCTURED PREDICTION FOR OBJECT DETECTION	153
7.1	Structured Prediction for Object Detection	154
7.2	Experiments	158
7.3	Related Work	160
7.4	Conclusion	162
8	TRANSFER LEARNING FOR RGB-D OBJECT RECOGNITION	165
8.1	Related Work	167
8.2	CNN Feature Extraction Pipeline	169
8.2.1	RGB Image Preprocessing	169
8.2.2	Depth Image Preprocessing	170
8.2.3	Image Feature Extraction	174
8.3	Learning Method	175
8.3.1	Object Classification	175
8.3.2	Object Pose Estimation	175
8.4	Evaluation	176
8.4.1	Evaluation Protocol	176
8.4.2	Results	177
8.5	Conclusion	183
9	CONCLUSION	185
9.1	Future Directions	187
A	ACRONYMS AND SYMBOLS	189
	BIBLIOGRAPHY	195