



Gradient-Free Optimization of Artificial and Biological Networks using Learning to Learn

Alper Yeğenoğlu

IAS Series

Band / Volume 55

ISBN 978-3-95806-719-6

Mitglied der Helmholtz-Gemeinschaft

 **JÜLICH**
Forschungszentrum

Forschungszentrum Jülich GmbH
Institute for Advanced Simulation (IAS)
Jülich Supercomputing Centre (JSC)

Gradient-Free Optimization of Artificial and Biological Networks using Learning to Learn

Alper Yeğenoğlu

Schriften des Forschungszentrums Jülich
IAS Series

Band / Volume 55

ISSN 1868-8489

ISBN 978-3-95806-719-6

Contents

Abstract	1
Zusammenfassung	3
1 Introduction and Motivation	11
1.1 Motivation	11
1.2 Contribution of the Thesis	14
1.3 Scope and Structure of this Thesis	15
2 Artificial and Biological Neural Networks	17
2.1 Overview of Artificial Neural Networks	17
2.1.1 Multilayer Perceptron	18
2.1.2 Activation Functions	20
2.1.3 Convolutional Neural Networks	21
2.2 Introduction into Neurons & Spiking Neural Networks	23
2.2.1 Structure of a Neuron	23
2.2.2 Properties of a Neuron	24
2.3 Recurrent Neural Networks	26
2.3.1 The Basics of Recurrent Networks	26
2.3.2 Reservoir Computing	27
2.4 Simulation Tools for Neural Networks	29
2.5 Neuro-inspired Learning	30
2.5.1 Visual Cortex as Inspiration for Convolutional Networks	30
2.5.2 Other Neuro-Inspired Methods	32
2.6 Summary	34
3 Optimization methods applied on Neural Networks	37
3.1 Introduction into Optimization	38
3.1.1 Optimization algorithms	39
3.2 Gradient Descent and Backpropagation	39
3.2.1 Backpropagation Through Time	40
3.2.2 Adaptive Moment Estimation	41

Contents

3.2.3	The Problem of Exploding and Vanishing Gradients	43
3.2.4	Biological Plausibility of Gradient Descent and Backpropagation in Spiking Neural Networks	44
3.3	Metaheuristics	46
3.3.1	Evolutionary Strategies	47
3.3.2	Ant Colony Optimization	48
3.3.3	Ensemble Kalman Filter	49
3.4	Summary	51
4	Deep Neural Networks Optimized by the Ensemble Kalman Filter	53
4.1	The Effects of Vanishing Gradients in Deep Neural Networks	54
4.2	Experimental Setup and Network Optimization	55
4.2.1	Ensemble Kalman Filter Optimization	55
4.2.2	Training with Gradient Descent	56
4.3	Numerical Results with Gradient Descent Optimizers	57
4.3.1	Optimization with SGD: Non evolving Gradients and Activation Values	57
4.3.2	Optimization with Adam: Slowly evolving Gradients and Activation Values	60
4.4	Optimization Results with the EnKF	63
4.4.1	The Effect of Different Activation Functions	63
4.4.2	Varying Number of the Ensemble Members	64
4.4.3	Performance on the Letters Dataset	65
4.4.4	Adapting the Hyper-Parameters	65
4.4.5	Convergence of the EnKF	67
4.4.6	Benchmarking the EnKF	69
4.5	Discussion	70
5	Learning to Learn	73
5.1	From the Historical Context to the State of the Art	74
5.2	Basic concepts of Learning to Learn	75
5.3	Technical description of L2L	77
5.3.1	Executing an L2L run	78
5.4	Optimizing with L2L	80
5.4.1	MNIST Classification with a Liquid State Machine	80
5.4.2	Fitness function of the Reservoir Network	82
5.4.3	Optimizing the Reservoir with the Ensemble Kalman Filter	82
5.4.4	Classification Performance of the Reservoir	83
5.5	Analysis of the Parameters	85
5.5.1	Weights Analysis	85

5.5.2	Covariance Matrix Analysis	87
5.6	Hyper-Parameter Optimization with L2L	92
5.6.1	Hyper-Parameter Optimization Workflow	92
5.6.2	Hyper-Parameter Optimization Results	96
5.7	Discussion	99
6	Optimizing Spiking Neural Networks enhances Foraging Behaviour of Swarm-Agents	101
6.1	Modeling the Ant Colony	102
6.1.1	Environment and Task	103
6.1.2	Network Architecture	104
6.2	Optimizing the Ant's Network with L2L	105
6.2.1	L2L Simulation Details	106
6.2.2	Ant Foraging Performance	107
6.3	Analysis of the Ant Foraging Behaviour	109
6.3.1	Comparing the Performance to a rule-based Model	109
6.3.2	Correlating the Input and Output Activity	110
6.4	Discussion	113
7	Conclusions and Outlook	115
7.1	Conclusion and Discussion	115
7.2	Outlook	119
	List of Figures	122
	Bibliography	125

IAS Series
Band / Volume 55
ISBN 978-3-95806-719-6