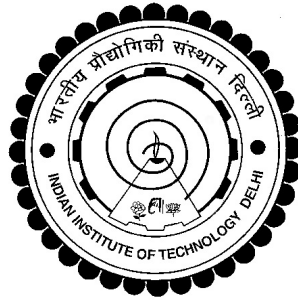


**LEVERAGING MACHINE LEARNING FOR ENHANCED
PERFORMANCE OF OPTICAL AND WIRELESS
COMMUNICATION SYSTEMS**

NAVEENTA GAUTAM



**BHARTI SCHOOL OF TELECOMMUNICATION TECHNOLOGY
AND MANAGEMENT
INDIAN INSTITUTE OF TECHNOLOGY DELHI
JANUARY 2024**

©Indian Institute of Technology Delhi (IITD), New Delhi 2024

LEVERAGING MACHINE LEARNING FOR ENHANCED
PERFORMANCE OF OPTICAL AND WIRELESS
COMMUNICATION SYSTEMS

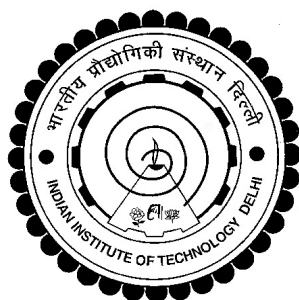
by

NAVEENTA GAUTAM

BHARTI SCHOOL OF TELECOMMUNICATION TECHNOLOGY AND
MANAGEMENT

Submitted

*in fulfillment of the requirements of the degree of Doctor of Philosophy
to the*



INDIAN INSTITUTE OF TECHNOLOGY DELHI
JANUARY 2024

To my parents and brother,
for being my biggest cheerleaders.

Certificate

I hereby certify that the work which is being presented in the thesis entitled “**LEVERAGING MACHINE LEARNING FOR ENHANCED PERFORMANCE OF OPTICAL AND WIRELESS COMMUNICATION SYSTEMS**” except Chapter 3 in the fulfillment of the requirements for the award of the degree of **Doctor of Philosophy** and submitted to the Bharti School of Telecommunication Technology and Management, Indian Institute of Technology Delhi, is an authentic record of my own work carried out during the time period from November 2019 to September 2023 under the supervision of Prof. Brejesh Lall and Prof. Amol Choudhary from the Department of Electrical Engineering, Indian Institute of Technology Delhi, India. The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute and any collaboration with other students with their contributions has been clearly highlighted. Any work that has been published is mentioned under the “List of Publications”.

Naveenta Gautam

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

Prof. Brejesh Lall

Professor

Department of Electrical Engineering

Indian Institute of Technology Delhi

Prof. Amol Choudhary

Associate Professor

Department of Electrical Engineering

Indian Institute of Technology Delhi



1st September 2023

To whom it may concern

This is to certify that Chapter 3 entitled "*Leveraging Transfer Learning for 5G Slicing in Industrial Networks*" contained in the thesis entitled "*LEVERAGING MACHINE LEARNING FOR ENHANCED PERFORMANCE OF OPTICAL AND WIRELESS COMMUNICATION SYSTEMS*" being submitted by NAVEENTA GAUTAM to the Bharti School of Telecommunication Technology and Management, Indian Institute of Technology Delhi, for the award of the degree of Doctor of Philosophy is the record of the bona-fide research work carried out by her during the internship at Nokia Bell Labs, Stuttgart, Germany during June-September 2022 under the supervision of myself and Nokia colleagues.

The results contained in this thesis have not been submitted either in part or in full to any other university or institute for the award of any degree or diploma.

Best regards,

Dr. Ilaria Malanchini.

A handwritten signature in black ink that reads "Ilaria Malanchini".

Dr. Ilaria Malanchini
Senior Research Engineer
Nokia Bell Labs
Direct mobile: +49 175 7266815
Ilaria.Malanchini@nokia-bell-labs.com
Nokia Solutions and Networks
GmbH & Co. KG
Magirusstraße 8
70469 Stuttgart

Acknowledgments

I would like to express my deepest gratitude to my supervisors, Prof. Brejesh Lall and Prof. Amol Choudhary, for their invaluable guidance, unwavering support, and continuous encouragement throughout my doctoral journey. Prof. Lall, your mentorship and guidance have been invaluable to me. Your depth of knowledge in the field and insightful feedback have greatly enriched my research. You believed in me before I believed in myself. Your encouragement to explore new ideas and approaches pushed me to achieve more than I thought possible. Prof. Choudhary, I am incredibly grateful for your insightful perspectives and advice throughout this journey. Your experience and innovative thinking have broadened my understanding of optics and encouraged me to explore new avenues. Your guidance in refining my research methodology and analyzing the results has been invaluable. I sincerely appreciate your patience, constant support, and dedication to my academic growth. Thank you for entertaining my half baked ideas and always being up for discussions.

I would like to extend my gratitude to the entire academic committee, Prof. Abhishek Dixit, Prof. Monika Aggarwal and Prof. Sumantra Dutta Roy for their valuable insights, constructive criticism, and fruitful discussions that have shaped the direction of my research. I thank Prof. S.D. Joshi for his continuous support and words of encouragement.

I am grateful to Dr. Ilaria Malanchini, Dr. Qi Liao, Dr. Alessandro Leito and Dr. Markus Gruber at Nokia Bell Labs, Stuttgart, Germany for welcoming me to their

team and supporting me outside of our collaboration. The internship experience was truly enriching and helped me to grow both personally and professionally. Special thanks to Ilaria for going out of her way to organise events for the interns. My heartfelt thanks go to my colleagues at Nokia — Marco Barletta, Emre Cinar, Maximilian Anderlohr, Jessica Greger, Deborah Djon, Patrick Watters, Anthony and Tianlun Hu for a great time inside and outside of the office.

I want to thank Dr. Prakriti Saxena, Dr. Manishika Rawat and Dr. Richa Priyadarshini for sharing this journey with me and to IIT Delhi for giving me three beautiful lifelong friendships. I extend my thanks to my fellow researchers and colleagues - Harsh Vaid, Megha Kataria, Neeraj Kumar, Anushikha Singh, Col. Gurjant Singh, Utkarsh Agarwal, Tayeba Qazi, Sanhita Pathak, Sadia Hussain who have provided a supportive environment for research and have inspired me with their dedication and passion for knowledge. Special thanks to Dr. Vinay Kaushik for useful discussions on machine learning and Dr. Rajveer Dhawan for providing valuable support on the coherent optical communications testbed. I want to thank Major Arshdeep Sodhi, Sai Vikranth Pendem and Harigovind P S for choosing to do their masters thesis with me. I have enjoyed our discussions and the time we worked together as a team.

I am indebted to Bharti School of Telecommunication Technology and Management, IIT Delhi, for providing the essential resources, facilities, and financial support that made this research possible. My sincere thanks to the academic staff— Anita Bisht, Dilip Jana, Savita Sharma and Vinita Yadav — for their seamless administrative support.

I am thankful to Aditya Bansal for being the calm to my storm. Your ability to nod attentively while I rambled is truly commendable.

Lastly, I would like to express my heartfelt appreciation to my parents for their unwavering support, love, and understanding throughout my academic pursuit. Your

encouragement and belief in my abilities have been the pillars of my strength. To my brother, Prince, Thanks for keeping me sane and grounded and for listening to all my rants without any judgement.

To everyone who has contributed to my academic and personal growth, I offer my deepest gratitude. Without your support, this journey would not have been possible. Thank you for being an integral part of my life and for shaping me into the researcher I have become today.

Naveenta Gautam

Abstract

Telecommunication systems, integral to modern society, facilitate seamless information transmission across vast distances via wired and wireless networks. The surge in data demand, fueled by smartphone adoption, IoT devices, and streaming services, has expanded the role of these systems beyond communication to encompass critical information access, business operations, and entertainment [1].

As data demands soar, telecommunication providers confront challenges meeting user requirements. This compels substantial investments in network capacity, 5G technology deployment, and existing infrastructure optimization for enhanced data transmission efficiency [1]. Amid these challenges, machine learning (ML) emerges as a potent tool to transform telecommunication performance. ML techniques, applicable in signal processing, network optimization, security enhancement, and personalized services [2], are reshaping telecommunication systems. This ongoing integration promises continuous technological advancement in communication.

We harness the power of machine learning to address critical challenges spanning wireless and optical communications domain. We explore innovative methodologies, algorithms, and techniques that can elevate the performance and signal quality of communication networks. In the first part of our research we explore the use of transfer learning and reinforcement learning for network slicing in industrial scenarios. Here, we seek to optimize resource allocation and network management, ensuring

efficient and adaptable communication systems for diverse industrial applications [3]. Transitioning our focus to optical fiber communication systems, which serve as the fundamental backbone for the burgeoning 5G networks, our primary objective is to tackle the formidable challenge posed by nonlinear effects. These nonlinear phenomena have emerged as a critical bottleneck impeding the realization of high-speed optical communications [4]. Through machine learning techniques, we aim to mitigate these nonlinearities, unlocking the potential for faster and more reliable data transmission. As we delve deeper into the intricacies of optical fiber systems, we explore the dynamic behavior of ultrashort pulses within nonlinear fibers [5]. Leveraging neural networks, we develop models capable of predicting and reconstructing these intricate pulse dynamics. This breakthrough not only enhances our understanding of pulse dynamics but also facilitates the design of more robust optical communication systems [6]. In the final phase of our research journey, we tackle the challenge of deploying deep neural networks within real-time systems. To this end, we propose a model compression method utilizing knowledge distillation. This innovative approach allows us to distill the vast knowledge contained within our deep neural networks into more compact architectures, enabling their seamless integration into real-world latency-sensitive communication systems [7].

In summary, the combination of telecommunication systems and machine learning has the potential to reshape communication technology fundamentally. As ML techniques continue to advance and become more seamlessly integrated into telecommunication networks, we can expect improved performance, enhanced user experiences, and a transformative impact on how information is transmitted, received, and utilized in our increasingly connected world. Our work explores some facets of the same and the findings are encouraging.

Contents

Certificate	i
Acknowledgments	iii
Abstract	vi
List of Figures	xviii
List of Tables	xx
List of Abbreviations	xxii
1 Introduction	1
1.1 Overview	1
1.2 Motivation and Scope	6
1.2.1 Introduction	6
1.2.2 5G Network Slicing	6
1.2.3 Nonlinear Compensation in Coherent Communication Systems	8
1.2.4 Ultrashort Pulse Propagation	9
1.2.5 Knowledge Distillation for Nonlinear Schrodinger Equation . .	12
1.3 Challenges & Objectives	13
1.4 Major Contributions of the Thesis	16

1.5	Layout of the Thesis	19
2	Technological Background and Literature Review	21
2.1	Machine Learning	21
2.1.1	Supervised Learning	23
2.1.2	Training Process of Neural Networks	28
2.1.3	Reinforcement Learning	31
2.1.4	Transfer Learning	37
2.2	5G Network Slicing	40
2.2.1	Introduction	40
2.2.2	Categories of services in 5G Slicing	42
2.2.3	ML based approach for network slicing	44
2.3	Nonlinear compensation in Coherent Optical Communication Systems	46
2.3.1	Introduction	46
2.3.2	DSP for Coherent Optical Communication Systems	51
2.3.3	ML techniques for Nonlinear Compensation	57
2.4	Ultrashort Pulse Propagation	59
2.4.1	Introduction	59
2.4.2	Measuring Ultrashort Pulses	61
2.4.3	ML for characterizing Ultrashort Pulses	62
2.5	Knowledge Distillation	64
2.5.1	Introduction	65
2.5.2	Training the student-teacher model	65
2.5.3	Recent developments in Knowledge Distillation	67
3	Leveraging Transfer Learning for 5G Slicing in Industrial Networks	70
3.1	Introduction	70

3.1.1	Proposed Work	72
3.2	System Model	73
3.3	Problem Formulation	76
3.3.1	The Reinforcement Learning Problem	77
3.3.2	The Transfer Learning Problem	79
3.4	Proposed Solution	80
3.4.1	Algorithm design for Deep Reinforcement Learning	80
3.4.2	Proposed Transfer Learning Approach	81
3.4.3	Summarized Transfer learning (TL) aided deep reinforcement learning (DRL) algorithm	81
3.5	Numerical Results	82
3.5.1	Simulation setup	82
3.5.2	Baseline scenario with DRL	85
3.5.3	TL aided approach	87
3.6	Summary	90
4	Transformers for Nonlinear Equalization in Coherent Optical Com- munication Systems	93
4.1	Introduction	93
4.2	Proposed Work	95
4.3	System Architecture and Methodology	96
4.3.1	System Setup	96
4.3.2	Transformer Architecture	99
4.3.3	Training Data Preparation	101
4.4	Results	103
4.4.1	Training Data and Feature Extraction methods	104
4.4.2	Transformer Architecture	109

4.4.3	Simulation Results	112
4.4.4	Computational Complexity Analysis	119
4.5	Conclusions and Future Outlook	123
5	Neural Network architectures for modelling Nonlinear Optical Pulse	
	Propagation	125
5.1	Introduction	125
5.2	Proposed work	127
5.3	Methodology	129
5.3.1	Forward mapping of the NLSE	130
5.3.2	Reverse mapping of the NLSE	133
5.3.3	β_2 and γ parameter prediction	134
5.4	Results	134
5.4.1	NN Architectures Analysis	134
5.4.2	Simulation Results	143
5.4.3	β_2 and γ parameter prediction	157
5.4.4	Training FCNN for different parameters	158
5.5	Discussions	161
5.6	Summary	167
6	OptiDistillNet: Learning nonlinear pulse propagation using the student-teacher model	169
6.1	Introduction	169
6.2	Proposed Work	171
6.3	Methodology	172
6.3.1	Training Data	175
6.3.2	Student Training loss functions	177

6.3.3	Network architecture	178
6.4	Results	179
6.5	Summary	190
7	Conclusions and Future work	192
7.1	Conclusions of this research	192
7.2	Future work	195
	List of publications	199
	Bibliography	200
	Technical Biography of the Author	220

List of Figures

1.1	A simplified overview of 5G network with optical fiber as the backhaul	3
2.1	The three types of learning in ML	22
2.2	Architecture of a multilayer perceptron with 3 hidden layers	24
2.3	The Transformer- model architecture [56]	26
2.4	Illustration of the Neural Network Training Process	30
2.5	Actor-Critic network reinforcement learning algorithm.	33
2.6	Comparison of Model Transfer and Feature Transfer Techniques	38
2.7	Network Slicing in 5G	42
2.8	Transfer learning technique : NN pretrained on one factory environment is leveraged to enhance performance of a different factory environment.	45
2.9	Overview of a coherent optical communications system setup	48
2.10	Schematic of digital back propagation	56
2.11	Trained NN used for prediction of temporal and spectral intensity points of the pulses ; I : Temporal Intensity , S : Spectral Intensity . .	64
2.12	Knowledge distillation framework	66
3.1	An example of temporal evolution of the production activities and their mutual dependencies.	75

3.2	DRL agent optimizing spectrum allocation for network slices by exploiting the relationship between network KPIs and production status	77
3.3	Production state evolution	84
3.4	Traffic pattern for all slices	84
3.5	Comparison between activation functions and exploration strategies .	86
3.6	Comparison among memory buffer sizes for production state history .	86
3.7	Traffic demand of the TL scenario compared with the baseline scenario	88
3.8	Cumulative reward obtained by the model trained from scratch and TL models during training	89
3.9	Performance of slice Robots during testing in terms of average values and confidence intervals	91
3.10	Performance of slice AGVs during testing in terms of average values and confidence intervals	91
3.11	Performance of slice Containers during testing in terms of average values and confidence intervals	91
4.1	System Setup of a DP-16QAM coherent optical communication system	97
4.2	Transformer architecture and training data preparation	102
4.3	Q factor Vs. Fiber length for different hyper-parameters; d : dropout, ff: feedforward	107
4.4	Training curves for different feature extractors	108
4.5	SER vs window showing the optimum window length for our experiments	109
4.6	Training curves for different window lengths	110
4.7	Training the Transformer model	111
4.8	Testing the trained Transformer model	111
4.9	Q factor Vs. Fiber length for DP-16QAM single channel system . . .	113

4.10	Q factor Vs. LOP for a distance of 640km for DP-16QAM single channel system. BiLSTM gives 0 symbol errors for a LOP of -3 dBm while transformer gives 0 symbol errors corresponding to LOP in the range of -3 to -1 dBm	114
4.11	EVM comparison for CDC, DBP, LSTM and transformer	115
4.12	Constellations before (blue) and after NLE using Bi-LSTM and Transformer for a LOP of 0dBm and 1200 km fiber length.	117
4.13	Visualizing EVM calculation	118
4.14	Computational complexity analysis of DBP, FCNN, BiLSTM and Transformer	122
5.1	System diagram for training a NN using error backpropagation for learning the forward and reverse mapping. Arrows on the top show the input and target data for forward mapping while arrows at the bottom are for reverse mapping; I: Temporal Intensity, S: Spectrum, OSA: Optical Spectrum Analyser	131
5.2	FCNN (all nodes of one layer connected to all nodes of next layer) and CaNN (includes connections between alternate layers)	137
5.3	Convolutional NN : Convolutional + ReLu layer: Feature extraction, Output layer: Regression	139
5.4	LSTM (LSTM layer 1 only): red connections and BiLSTM (LSTM layer1 + LSTM layer2): dashed connections, Common: solid black	141
5.5	Estimating the temporal characteristics of the O/P pulse for a known fiber.	147
5.6	Estimating the spectral characteristics of the O/P pulse for a known fiber.	147

5.7	Estimating the temporal characteristics of the O/P pulse for an unknown fiber.	148
5.8	Estimating the spectral characteristics of the O/P pulse for an unknown fiber.	148
5.9	Estimating the temporal characteristics of the O/P pulse for anomalous and normal dispersion regime (Case 4,5).	149
5.10	Error Histogram of the FCNN for the forward mapping (Case 2 in Table 5.1).	149
5.11	Reconstruction of the temporal characteristics of an I/P pulse using RNNs.	153
5.12	Reconstruction of the temporal characteristics of an I/P pulse.	154
5.13	Reconstruction of the spectral characteristics of an I/P pulse.	154
5.14	Estimating the temporal characteristics of the FCNN trained for learning the forward and reverse mapping for an unchirped pulse.	155
5.15	Error Histogram of the FCNN for the reverse mapping (Case 2)	156
5.16	Comparing the γ and β_2 values of 10 commonly used HNLF [130] with NN predicted values.	157
5.17	MSE vs length of the fiber for learning the forward and reverse mapping for a pulse of width 1ps, $C = -2$, $\beta_2 = ps^2/km$ and $\gamma = 30W^{-1}/km$	159
5.18	Estimating the temporal characteristics of the FCNNs trained for learning the forward and reverse mapping of β_3	159
5.19	Estimating the spectral characteristics of the FCNNs trained for learning the forward and reverse mapping of β_3	160
5.20	Reconstructing the temporal profile of a pulse of width 0.5 ps transmitted through a SSMF.	160
5.21	Training curves of different NNs for learning the forward mapping (Case 2)	162

5.22	Training curves of different NNs for learning the reverse mapping (Case 2)	162
5.23	MSE vs dispersion of the fiber for forward and reverse mapping for a pulse of width 1ps, $C = -2$ and $\gamma = 10W^{-1}/km$	163
5.24	MSE vs γ of the fiber for forward and reverse mapping for a pulse of width 1ps, $C = -2$ and $\beta_2 = 17ps^2/km$	164
5.25	MSE vs pulse width of the fiber for forward and reverse mapping for $C = -2$, $\beta_2 = 17ps^2/km$ and $\gamma = 10W^{-1}/km$	164
6.1	Training a shallow student network using knowledge transfer from the teacher network	175
6.2	Details of network architecture and loss function for teacher (top) and student (bottom) network with 91.23% distillation rate. The values inside rectangular boxes indicate the spatial dimension of the convolutional feature.	176
6.3	Comparison of teacher and student networks in terms of the trainable parameters and MSE	182
6.4	Training curves of the teacher, student trained independently and proposed OptiDistillNet	183
6.5	Comparison between the teacher, student (trained independently) and OptiDistillNet for the reconstruction of temporal domain curves of the pulses after propagation	186
6.6	Comparison between the teacher, student (trained independently) and OptiDistillNet for the reconstruction of spectral domain curves of the pulses after propagation	186

6.7	Comparison between the teacher, student (trained independently) and OptiDistillNet for the reconstruction of temporal domain curves of the pulses after propagation	188
6.8	Comparison between the teacher, student (trained independently) and OptiDistillNet for the reconstruction of spectral domain curves of the pulses after propagation	188
7.1	Adapting the Transformer nonlinear equalizer for WDM systems . . .	196
7.2	RBM being used as a feature extractor for the Transformer nonlinear equalizer	197

List of Tables

2.1	Summary of ML techniques addressing nonlinearities categorized by nonlinearity type and system setup	57
2.2	ML techniques employed for NLC along with the system configurations and limitations	58
3.1	System Setup	83
4.1	System Parameters	98
4.2	Transformer Architecture	110
4.3	BiLSTM architecture	112
4.4	EVM values for difference techniques at 1200 km	117
5.1	Pulse and Fiber parameters used for creating the training dataset. . .	132
5.2	Hyper-parameters selection process of FCNN, CaNN and CNN for the forward mapping	144
5.3	Hyper-parameters selection process of RNNs for the FORWARD MAPPING	145
5.4	Hyper-parameter selection process of FCNN, CaNN and CNN for the REVERSE MAPPING	151
5.5	Hyper-parameter selection process of RNNs for the REVERSE MAPPING	152

5.6	Training parameters for testing the generalization performance of FCNN.	158
6.1	Comparison of different NN architectures. The second column describes the architecture of the network in the following order: [Size of input layer, hidden layer neurons, size of output layer]	179
6.2	Performance comparison of different CNN models for a training dataset of $7 \cdot 10^5$. The second column represents the size of the input channels in each convolution layer.	180
6.3	Performance comparison of student networks trained using CNN5 as the teacher	181
6.4	Comparison of the teacher network, student network trained independently and OptiDistillNet in terms of the size of training data, MSE and trainable parameters.	189

List of Abbreviations

BiLSTM Bidirectional long short term memory. 9, 95, 128

CaNN Cascade Neural Network. 128

CNN Convolutional Neural Network(s). 18, 24, 128

DBP Digital Backpropagation. 8, 55, 95

DDPG Deep Deterministic Policy Gradient. 32, 34, 72, 85

DNN Deep Neural Network(s). 169

DRL Deep Reinforcement Learning. 7, 80, 92

DSP Digital Signal Processing. 8, 14, 47, 51

EVM Error Vector Magnitude. 114

FCNN Fully connected neural network. 9, 128

FPGA Field-Programmable Gate Array. 69, 195

FWM Four Wave Mixing. 4, 49

GRU Gated Recurrent Unit. 128

GVD Group Velocity Dispersion. 129

KD Knowledge Distillation. 18, 170

LSTM Long short term memory. 128

ML Machine Learning. vi, 21, 125

NLC Nonlinear compensation. 54

NLSE Nonlinear Schrodinger equation. 15, 128, 170

NN Neural Network. 15

QAM Quadrature Amplitude Modulation. 3

RNN Recurrent Neural Network. 24, 140

SPM Self Phase Modulation. 4, 48, 99

TL Transfer Learning. 7, 37, 72, 89

WDM Wavelength Division Multiplexing. 3, 98, 99

XPM Cross Phase Modulation. 4, 49, 99