Study of Linear and Non Linear Properties of Photonic Structures via Numerical and Deep Learning Techniques

By

Rakayet Rafi
ID :21MEE008P

A thesis submitted in partial fulfilment of the requirements for the degree of
MASTERS of SCIENCE
in
Department of Electrical and Electronic Engineering
CHITTAGONG UNIVERSITY OF ENGINEERING AND TECHNOLOGY
AUGUST 2024

Declaration

I hereby declare that the work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is cited.

Signature and Date:

Rakayet Rafi
21MEE008P
Department of EEE
Chittagong University of Engineering & Technology (CUET)

List of Publications

Results relevant to this thesis have been disseminated to the channels mentioned below:

Journals

1. R. Rafi, Karim, S. Ghosh, and B. M. A. Rahman, "Unified predictive modeling of supercontinuum spectra: Using multi-material data with Closed-Form Continuous Time Neural Networks," *Optics Communications*, p. 130646, May 2024, doi: 10.1016/j.optcom.2024.130646.

2. M.R. Karim, R. Rafi, D. Dutta, M. I. H. Upal, S. Ghosh, "Employing Deep Learning for Predicting Supercontinuum Generation in Chalcogenide Planar Waveguide," *Optik*, 2024, 171749, ISSN 00304026, doi: 10.1016/j.ijleo.2024.171749.

Conferences

3. R. Rafi, D. Dutta, M. I. H. Upal, M. R. Karim and S. Ghosh, "Artificial Neural Network Model to Investigate Supercontinuum Generation in Chalcogenide Planar Waveguide," *2023 International Conference on Telecommunications and Photonics (ICTP)*, Dhaka, Bangladesh, 2023, pp. 1-4, doi:10.1109/ICTP60248.2023.10490695.

4. M. I. H. Upal, D. Dutta, R. Rafi, M. R. Karim, M. R. Alam and S. Ghosh, "Deep Learning Approach to Determine the Optical Characteristics of Photonic Crystal Fiber for Orbital Angular Momentum Transmission," *2023 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Chittagong, Bangladesh, 2023, pp. 1-6, doi: 10.1109/ECCE57851.2023.10101646.

Approval by the Supervisor

This is to certify that Rakayet Rafi has carried out this work under my supervision, and that he has fulfilled the relevant Academic Ordinance of the Chittagong University of Engineering and Technology, so that he is qualified to submit the following thesis in application for the degree of M.Sc. in EEE.

Signature and Date:

____

Dr. Sampad Ghosh
Associate Professor
Department of EEE
Chittagong University of Engineering & Technology

Acknowledgement

First and foremost, I express my heartfelt gratitude to Almighty Allah for His blessings, guidance, and grace throughout this journey. Without His mercy and support, none of this would have been possible.I would like to extend my deepest appreciation to my supervisor, **Dr. Sampad Ghosh**, **Associate Professor, Department of Electrical and Electronic Engineering, Chittagong University of Engineering and Technology**, for his invaluable guidance, unwavering support, and continuous encouragement throughout the course of my research. His expertise, patience, and insightful feedback have been instrumental in shaping this thesis.I am also profoundly grateful to **Prof. Dr. Mohammad Rezaul Karim**, **Professor, Department of Electrical and Electronic Engineering, Daffodil International University**, for his mentorship, encouragement, and endless support. His wisdom, encouragement, and assistance with publications have been invaluable to me, and I am deeply indebted for his generosity with time and knowledge.I would also like to extend my gratitude to the **Department of EEE, CUET** for facilitating me with all the support throughout the journey.Lastly, I would like to thank my family and friends for their unwavering support, understanding, and encouragement throughout this journey.

Abstract

This study focuses on advancing the design, characterization, and optimization of photonic structures by integrating elementary and advanced deep learning techniques. The primary aim is to develop predictive models that enhance both the accuracy and efficiency of photonic structure simulations, enabling faster design iterations and real-time optimization.This is aimed to substitute the conventionally used simulation softawares for similar purpose.The first part involves utilizing Artificial Neural Networks (ANNs) to predict key linear optical parameters of PCFs, including effective refractive index, dispersion, mode purity, and effective area. The ANN model achieved a high prediction accuracy, evidenced by a Mean Squared Error (MSE) of 0.0218, with a training time of approximately 31 seconds. This represents a substantial reduction in computational time compared to traditional numerical methods while maintaining precise predictions, demonstrating the model's practical applicability.The second objective focuses on modeling the nonlinear dynamics of supercontinuum generation in optical fibers, with particular attention to chalcogenide planar waveguides. An ANN-based deep learning model was employed to accurately capture the complex pulse propagation dynamics, achieving an MSE of 0.0032 with a computational training time of 472 seconds. The model's ability to predict spectral evolution across various pump powers accurately provides a significant potential for real-time optimization of experimental setups.The third objective introduces a novel approach using Closed-Form Continuous-Time Neural Networks (CfC) for predicting supercontinuum spectra in waveguides with diverse material compositions. The CfC model outperformed conventional methods such as Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, achieving high accuracy with MSE values as low as 7.4199×10^{-4} for a three-material dataset and 6.9408×10^{-4} for a four-material dataset. The model demonstrated low computation time, confirming its robustness, scalability, and adaptability across different materials.Overall, this research demonstrates that integrating machine learning techniques with photonics significantly enhances the accuracy of predictive models while reducing computational time. These advancements have practical implications for the design, optimization, and real-time control of photonic devices, paving the way for future innovations in the field.

সারাংশ

এই গবেষণা কাজটি ফটোনিক কাঠামোর নকশা, বৈশিষ্ট্যায়ন, এবং অপ্টিমাইজেশনের ক্ষেত্রে অগ্রগতি সাধনের উপর কেন্দ্রিত, যেখানে প্রাথমিক ও উন্নত ডিপ লার্নিং কৌশলগুলির সংযোজন করা হয়েছে। মূল লক্ষ্য হল পূর্বাভাস মডেল তৈরি করা, যা ফটোনিক কাঠামোর সিমুলেশনের নির্ভুলতা এবং দক্ষতা বাড়াবে, দ্রুত ডিজাইন পুনরাবৃত্তি এবং রিয়েল-টাইম অপ্টিমাইজেশনকে সক্ষম করবে। এটি ঐতিহ্যগতভাবে ব্যবহৃত সিমুলেশন সফটওয়্যারগুলিকে প্রতিস্থাপন করার লক্ষ্য রাখে যা একই উদ্দেশ্যে ব্যবহার করা হয়। প্রথম অংশে, ফটোনিক ক্রিস্টাল ফাইবার (PCFs)-এর মূল রৈখিক অপটিক্যাল প্যারামিটারগুলি, যেমন কার্যকর প্রতিফলন সূচক, বিচ্ছুরণ, মোড বিশুদ্ধতা, এবং কার্যকর ক্ষেত্রফল পূর্বাভাস করতে কৃত্রিম নিউরাল নেটওয়ার্ক (ANN) ব্যবহার করা হয়েছে। ANN মডেলটি উচ্চ পূর্বাভাস নির্ভুলতা অর্জন করেছে, যার প্রমাণ হল 0.0218 মানের গড় বর্গ ত্রুটি (MSE), এবং প্রায় ৩১ সেকেন্ডের প্রশিক্ষণ সময়। এটি প্রচলিত সংখ্যাগত পদ্ধতির তুলনায় গণনাকালকে উল্লেখযোগ্যভাবে কমিয়ে আনে, একই সাথে নির্ভুল পূর্বাভাস বজায় রাখে, যা মডেলটির ব্যবহারিক প্রযোজ্যতা প্রদর্শন করে। দ্বিতীয় লক্ষ্যটি অপটিক্যাল ফাইবারে সুপারকন্টিনিয়াম উৎপাদনের অরৈখিক গতিবিদ্যা মডেলিংয়ের উপর কেন্দ্রিত, বিশেষত ক্যালকোজেনাইড প্ল্যানার ওয়েভগাইডগুলির ক্ষেত্রে। একটি ANN ভিত্তিক ডিপ লার্নিং মডেল ব্যবহার করা হয়েছে জটিল পালস প্রোপাগেশন ডায়নামিক্সকে সঠিকভাবে ধরার জন্য, যা 0.0032 এর MSE এবং 472 সেকেন্ডের প্রশিক্ষণ সময় অর্জন করেছে। বিভিন্ন পাম্প পাওয়ার জুড়ে স্পেকট্রাল বিবর্তন সঠিকভাবে পূর্বাভাস করার মডেলটির ক্ষমতা পরীক্ষামূলক সেটআপগুলির রিয়েল-টাইম অপ্টিমাইজেশনের জন্য একটি গুরুত্বপূর্ণ সম্ভাবনা প্রদান করে। তৃতীয় লক্ষ্যটি উপস্থাপন করে একটি নতুন পদ্ধতি যা ভিন্ন ভিন্ন উপাদান সংমিশ্রণ সহ ওয়েভগাইডে সুপারকন্টিনিয়াম স্পেকট্রা পূর্বাভাসের জন্য বন্ধ-ফর্ম ক্রমাগত-সময় নিউরাল নেটওয়ার্ক (CfC) ব্যবহার করে। CfC মডেলটি প্রচলিত পদ্ধতিগুলির তুলনায়, যেমন পুনরাবৃত্ত নিউরাল নেটওয়ার্ক (RNN) এবং দীর্ঘ স্বল্প-মেয়াদী স্মৃতি (LSTM) নেটওয়ার্কের তুলনায় উচ্চ নির্ভুলতা অর্জন করেছে, যেখানে তিনটি উপাদান ডেটাসেটের জন্য 7.4199×10^{-4} এবং চারটি উপাদান ডেটাসেটের জন্য 6.9408×10^{-4} এর মতো কম MSE মান পেয়েছে। মডেলটি কম গণনা সময় প্রদর্শন করেছে, যা এর দৃঢ়তা, স্কেলযোগ্যতা, এবং বিভিন্ন উপাদানের সাথে অভিযোজনযোগ্যতাকে নিশ্চিত করে। সামগ্রিকভাবে, এই গবেষণা দেখায় যে ফটোনিক্সের সাথে মেশিন লার্নিং কৌশলগুলির সংযুক্তি পূর্বাভাস মডেলের নির্ভুলতাকে উল্লেখযোগ্যভাবে বাড়ায়, একই সাথে গণনাকালকে হ্রাস করে। এই অগ্রগতি ফটোনিক ডিভাইসগুলির নকশা, অপ্টিমাইজেশন, এবং রিয়েল-টাইম নিয়ন্ত্রণের জন্য ব্যবহারিক প্রভাব ফেলে, যা ক্ষেত্রের ভবিষ্যত উদ্ভাবনের জন্য পথ তৈরি করে।

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	**Full Form**
ANN	Artificial Neural Network
ALD	Atomic Layer Deposition
CFD	Computational Fluid Dynamics
CNN	Convolutional Neural Network
DCNN	Deep Convolutional Neural Network
DE	Differential Evolution
DL	Deep Learning
DNN	Deep Neural Network
FDTD	Finite Difference Time Domain
FEM	Finite Element Method
FWHM	Full Width at Half Maximum
GAN	Generative Adversarial Network
GNLSE	Generalized Nonlinear Schrödinger Equation
LSTM	Long Short-Term Memory
MIR	Mid-Infrared
ML	Machine Learning
MSE	Mean Squared Error
OAM	Orbital Angular Momentum
PCF	Photonic Crystal Fiber
PIC	Photonic Integrated Circuit
PPW	Pump Power
PWL	Pump Wavelength
RNN	Recurrent Neural Network
SC	Supercontinuum
SPP	Surface Plasmon Polaritons
SRS	Stimulated Raman Scattering
TEC-PCF	Twin Elliptical Core Photonic Crystal Fiber
TPA	Two-Photon Absorption
ZIM	Zero-Index Material

Chapter 1

Introduction

In recent years, deep learning has emerged as a transformative approach across various fields, offering advanced methods for data analysis, prediction, and optimization. Unlike traditional techniques that can be limited by human error and the constraints of manual processes, deep learning systems excel in handling and interpreting large datasets to uncover complex patterns and make informed decisions. This capability is particularly relevant in areas requiring sophisticated data analysis, such as in the design and optimization of optical waveguides.

Traditionally, optimizing parameters like dispersion, mode area, and refractive index in optical waveguides has relied heavily on time-consuming numerical simulations. However, deep learning presents a compelling alternative by leveraging extensive datasets to predict both linear and nonlinear characteristics of waveguides. This shift has the potential to accelerate the design process and enable exploration of new configurations that were previously impractical.

The integration of deep learning with photonics has led to significant advancements, including the development of neural networks specifically tailored for photonic applications. These innovations have improved the efficiency of characterizing waveguide behavior and predicting optical properties. While the application of deep learning to photonics continues to evolve, there remains considerable potential for further exploration and development. Addressing challenges such as understanding nonlinear dynamics and bridging the gap between linear and nonlinear aspects of waveguides are crucial for advancing this field. As deep learning techniques continue to progress, they promise to enhance the design and optimization of photonic systems, opening up new possibilities for future research.

1.1 Literature Review

The application of deep learning techniques to photonics research has become a revolutionary strategy in recent years, with promising advances in the characterisation, optimization, and comprehension of optical phenomena and systems. Artificial intelligence's deep learning subfield revolutionizes photonics research by providing a novel way for deriving useful insights from large datasets. Deep learning models have proven to be exceptionally adept in pattern recognition, prediction, and optimization across a wide range of fields by utilizing interconnected layers of neurons to handle complex data.

Deep learning approaches provide researchers with a flexible toolkit to investigate the complex dynamics of light-matter interactions, ranging from constructing elaborate photonic structures to forecasting nonlinear optical occurrences. Deep learning models can detect minute correlations and patterns that may escape the notice of traditional analytical or simulation-based techniques by independently learning from data.

Exploration was done on the rapidly growing subject of deep learning in photonics research in this evaluation of the literature with the goal of giving a thorough rundown of current advancements, approaches, and applications. Deep learning has been utilized in photonics for a variety of purposes, such as nonlinear dynamics comprehension, optical property prediction, and photonic device design and optimization.By synthesizing insights from recent studies and highlighting key advancements, this review seeks to elucidate the potential of deep learning to reshape the landscape of photonics research. This thesis focuses on the main approaches, obstacles, and possibilities related to incorporating deep learning methods into photonics research by means of a methodical review of the literature. Through an appreciation of the state-of-the-art and the identification of future research directions, scientists may better utilize deep learning to tackle challenging photonics issues. This survey of the literature paves the path for future research into the complementary fields of deep learning and photonics, opening new avenues for innovative thinking and novel findings in this quickly developing area.

1.1.1 Conventional Methods in Photonics and Deep Learning

Structural designs are crucial in developing photonic devices such as waveguides, plasmonic nanowires, metamaterials, photonic crystals, laser diodes, and photovoltaic cells. Generally, two primary approaches are employed in this design process. The first approach involves physics-based methods, such as simplified analytical models, insights from previous designs or similar applications, and scientific intuition. For example, photonic crystals can create complete photonic bandgaps, allowing light to navigate sharp bends with near-perfect efficiency within crystal structures [1]. Broadband mid-infrared (MIR) supercontinuum (SC) generation in waveguides is also analyzed using the Generalized Nonlinear Schrödinger Equation (GNLSE). This approach was utilized by Karim et al. [2] to produce supercontinuum spectra in dispersion-engineered waveguides. However, while physics-based methods provide valuable insights, designing structures to achieve specific photonic properties is challenging, particularly as their geometry and spatial arrangements grow more complex.

Consequently, photonic engineers often resort to the second approach: numerical simulations of electromagnetic fields using methods like the finite-element method [3], finite-difference time-domain method [4], and finite integration technique [5]. These methods discretize the computational domain into a finite number of meshes, with the mesh count increasing as the structure size grows. Alternative mesh-free methods, such as the finite cloud [6] and spectral [7] approaches, are also available. A critical step in these simulations is solving the time-independent Maxwell equations to determine optical modes, typically the initial stage in designing any photonic device. At this stage, engineers focus on optimizing mode distributions, field confinement, effective refractive indices, and group velocity dispersion. However, achieving desired outcomes often requires numerous iterative adjustments and simulations, which are time-consuming and heavily dependent on prior experience and expertise. Moreover, these methods allow for only a limited number of design variables to be modified, necessitating state-of-the-art computational tools. This repetitive process is inefficient, wasting substantial computational resources in today's data-driven environment. Therefore, deep learning (DL)-based models present an excellent alternative to the constraints of even the most advanced simulation tools, providing a more efficient pathway for photonic design optimization. Deep learning (DL) represents an advanced computational model comprised of multiple layers of processing units, enabling it to discern and learn various levels of abstraction from data inputs [8]. Initially achieving notable successes in computer science and engineering, DL has since demonstrated its efficacy across a range of fields including natural language processing [9], automatic speech recognition [10], robotics [11], and laser physics [12]. Additionally, DL has gained traction in materials science [13], quantum mechanics [14], and microscopy [15], reflecting its broad applicability and impact. The primary advantage of deep learning lies in its data-driven approach, which facilitates the automatic extraction of meaningful insights from extensive datasets, contrasting with traditional rule-based or physics-based methodologies. In recent years, DL has emerged as a novel approach in the design and optimization of photonic systems.

1.1.2 Deep Learning in Photonic Crystal Fibers (PCFs)

In the field of photonic crystal fibers (PCFs), a new wave of innovative research projects has been spurred by the revolutionary potential of machine learning (ML) methods. Together, these investigations aim to open new vistas in PCF-based sensing and optical communication by bringing together cutting edge computational techniques and creative fiber optic design tactics. Vijayan et al. introduce a novel use of deep learning multi-output regression modeling to decipher the complex relationship between optical design parameters and sensing performances in asymmetric Twin Elliptical Core PCFs (TEC-PCF), which starts this paradigm shift. Their method, supplemented with Explainable AI (XAI) feature analysis, improves interpretability and optimizes modeling efficiency, paving the way for more perceptive and effective modeling paradigms[16]. Building on this foundation, Ma et al. achieve unprecedented design process streamlining by fusing the differential evolution algorithm (DE) with artificial neural networks (ANN) to realize precise dispersion compensations in Ge-doped dual-core dispersion compensation PCFs (DC-DCPCF)[17]. Additionally, a unique highly negative dispersion compensating photonic crystal fiber has been introduced[18]. It uses machine learning to anticipate important output qualities from input parameters, offering computational efficiency and the possibility of being used in high-rate optical communication applications.

Jabin and Fok (2022) highlight the potential of machine learning (ML) in speeding up PCF characterisation by demonstrating the quick prediction of crucial optical characteristics for silica-based PCFs using deep learning-based multilayer perceptron models[19]. Similar to this, Sridevi et al. present a novel method that predicts optical characteristics of PCF temperature sensors by integrating deep neural networks (DNN) and autoencoders (AE)[16]. This method greatly reduces computation time while retaining good prediction accuracy. To further investigate the possibilities of machine learning, research has been conducted using several ML algorithms to forecast PCF design parameters. This research has provided insights into PCF sensitivity and confinement loss dynamics.

Verma, Kumar, and Jindal (2022) broaden the scope of applications by developing a highly sensitive gold/TiO2-coated PCF biosensor for the detection of breast cancer cells. They achieve this by utilizing machine learning approaches to improve sensing performance[20]. Li et al. demonstrate the potential of data-driven optimization by using a DL-based forward modeling approach to assess the link between PC nanocavity optical characteristics and design parameters[21]. To cap off this innovative wave, Zelaci et al. provide a novel GAN-based method to enhance datasets used in ANN model training, resulting in a notable improvement in PCF sensor prediction accuracy[21].Furthermore, Ferreira *et al* and Asano *et al* optimized Q-factors for Photonic Crystal Fibers (PCF) in 2018 by using ML and DL algorithms[22].

When taken as a whole, these groundbreaking studies highlight how ML can revolutionize PCF design and sensing applications. By improving modeling techniques, investigating new architectures, and incorporating experimental validations to increase resilience and practical application, they set the stage for future advancements. The combination of ML and PCF technologies promises to spur previously unheard-of advancements in the sector as it develops, making it possible to create smarter, more effective, and more adaptable optical systems for a variety of uses, from biomedical sensing to telecommunications and beyond.

1.1.3 Deep Learning in Waveguide and Different photonic structures

The integration of deep learning into photonics research has sparked a wave of innovation, offering transformative methods for the characterization, design, and optimization of optical devices. Recent advancements in deep learning have highlighted its potential across various domains of photonics research. For example, Alagappan et al. introduced a deep neural network (DNN)-based modal classification technique to rapidly distinguish between single-mode and multi-mode optical waveguide geometries, thereby streamlining the design of on-chip integrated photonic devices [23]. Similarly, Mia et al. utilized optical neural networks to accurately predict fundamental modal indices in silicon channel waveguides, showcasing the efficacy of these networks in forecasting optical properties with high precision [24]. The integration of deep learning with photonics substantially improves the prediction and comprehension of intricate nonlinear dynamics. For instance, Ma et al. highlighted the effectiveness of machine learning regression algorithms in analyzing nanophotonic waveguides, which facilitates the swift estimation of optimal parameters crucial for enhancing device performance [25]. Similarly, Chugh et al. employed machine learning regression methods in conjunction with finite element simulations and artificial neural networks to optimize integrated silicon photonics devices, achieving notable improvements in prediction accuracy for mode effective indices and other vital metrics [26]. Furthermore, Alagappan et al. investigated the use of deep learning models to predict effective refractive indices in silicon nitride waveguides, demonstrating the potential for expediting predictions related to waveguide properties [27]. Ghosh et al. proposed the use of deep learning models as viable alternatives to finite element methods for the precise assessment of photonic crystal fiber parameters, especially concerning orbital angular mode transmission. Their approach achieved notable accuracy in translating design parameters into optical properties [28]. Building on this, Alagappan et al. introduced an innovative method for the swift evaluation of the group refractive index in photonic channel waveguides by leveraging automatic differentiation combined with neural networks. This method offers a more efficient alternative to conventional numerical techniques in photonics design and modeling [29]. Further advancing the field, they developed a novel deep deconvolutional neural network (DCNN) technique to address optical modes in photonics. This technique enabled precise and efficient predictions by correlating photonic parameters with mode images, thereby accelerating device design for applications in optical sensing and integrated circuits. It also marked a pioneering demonstration of the effectiveness of such neural networks in the domain of photonics [30].

The field of physics has increasingly embraced machine learning (ML) techniques for analyzing dynamic and complex systems over time. Unlike traditional electromagnetic simulation-based methods, ML distinguishes itself through its data-driven approach, which enables algorithms to autonomously extract relevant information from extensive datasets. For example, Kiarashinejad *et al.* [31] developed a deep learning approach utilizing dimensionality reduction techniques to better understand the characteristics of optical wave-matter interactions in nanostructures.

These developments have greatly improved forecast accuracy and comprehension of complex nonlinear dynamics in addition to streamlining design procedures. DL methodologies offer a faster, more efficient, and accurate alternative to traditional simulation tools, paving the way for unprecedented innovation in various fields, including telecommunications, spectroscopy, imaging, and beyond. As researchers continue to explore the capabilities of DL and machine learning (ML) in photonics, further advancements in understanding and harnessing light-matter interactions for a wide range of applications, thus shaping the future landscape of photonics research and technology can be expected.

1.1.4 Deep Learning in Non Linear Photonics

In recent years, deep learning (DL) techniques have attracted significant attention in nonlinear photonics, particularly for their potential to predict and understand complex nonlinear dynamics. Unlike the extensive research focused on linear pulse propagation—such as through rectangular waveguides, where DL methods like feedforward neural networks (FFNN) have been used to predict optical phenomena such as dispersion properties, nonlinear coefficients, mode effective areas, and effective refractive indices—the nonlinear aspects of pulse propagation through waveguides have been comparatively understudied. Notable pioneering studies in this area, including those by Wetzel et al. [32] and Tzang et al. [33], have explored the optimization and analysis of spectral and temporal intensities in single-mode fibers. Recent efforts have aimed to fill this gap by demonstrating the effectiveness of Feed Forward Neural Networks (FFNN) and long short-term memory (LSTM) recurrent neural networks (RNNs) in modeling the complex nonlinear dynamics of ultrashort pulse evolution governed by the generalized nonlinear Schrödinger equation (GNLSE). These models have also been employed to predict the behavior of rogue solitons in supercontinuum generation processes [34][35][36]. These advancements enable not only accurate and rapid predictions of temporal and spectral evolution but also enhance the understanding of nonlinear optical processes, facilitating real-time optimization of experimental setups.

Nonetheless, a thorough examination of the geometrical factors and restrictions related to pump sources in supercontinuum generation is still lacking. These developments lead to improved knowledge of nonlinear optical processes and enable real-time experiment optimization, in addition to enabling rapid and accurate predictions of temporal and spectral evolution. By utilizing advanced experimental setups, computational methodologies, and machine learning strategies, scholars have acquired significant understanding of nonlinear optical processes. Unlocking the full potential of DL in nonlinear photonics requires more study in this area. This could result in ground-breaking discoveries and breakthroughs in a range of applications, including as imaging, spectroscopy, and telecommunications.

1.2 Literature Gap

Despite the significant advancements in integrating deep learning techniques with photonics, there remain notable gaps in the existing literature. Current studies have predominantly focused on predicting linear optical properties, such as refractive indices and dispersion, in relatively simple photonic structures like silicon waveguides and photonic crystal fibers (PCFs). However, the application of deep learning to more complex, nonlinear phenomena—such as supercontinuum generation in diverse material systems—remains underexplored. Additionally, while deep learning models have been applied to specific materials, there is a lack of research on models that can generalize across different material compositions within planar waveguides. These gaps highlight the need for more versatile and robust predictive models that can efficiently handle both linear and nonlinear optical phenomena across a broader range of photonic materials.

1.3 Motivation

The motivation for this research stems from the desire to address the aforementioned gaps in the literature by developing deep learning models that can accurately predict both linear and nonlinear optical properties across a variety of photonic materials and structures. The limitations of traditional numerical methods, which are time-consuming and computationally intensive, further emphasize the need for more efficient predictive models. By leveraging advanced neural network architectures, such as Closed-Form Continuous-Time Neural Networks (CfC), this research aims to create models that not only reduce computational overhead but also enhance the accuracy and generalizability of predictions. Ultimately, this work seeks to contribute to the broader field of photonics by providing tools that can accelerate the design and optimization of photonic devices, thereby enabling new innovations in areas such as optical communications, biomedical imaging, and spectroscopy.

1.4 Objectives with Specific Aims

Based on the literature review, some specific aims have been targetted in order to contribute to the field linear and non linear photonics. The objectives are as follows,

a. To study linear properties of **Photonic Crystal Fiber** for OAM mode transmission using deep learning technique and compare the results with the numerical simulations done in COMSOL MULTIPHYSICS.

b. To study the phenomenon of Supercontinuum Generation (non-linear phenomenon) in Chalcogenide planar waveguide using deep learning technique and compare the results with the numerical simulations done in COMSOL MULTIPHYSICS.

c. To study the Supercontinuum Generation in a developed deep learning model for specifically investigating the response of deep learning model to multiple materials in a planar waveguide and compare the results with numerical simulations

1.5 Contributions

This thesis makes several key contributions to the field of photonics through the integration of advanced deep learning techniques for the design, characterization, and optimization of photonic structures:

1. **Development of an Artificial Neural Network (ANN) Model for Predicting Linear Optical Parameters:** This research introduces an innovative ANN-based model designed to predict key linear optical parameters of Photonic Crystal Fibers (PCFs), such as effective refractive index, dispersion, mode purity, and effective area. By achieving high prediction accuracy and significantly reducing computational time, this model provides a more efficient and precise method for PCF design, surpassing traditional numerical approaches.

2. **Modeling Nonlinear Dynamics of Supercontinuum Generation:** The thesis presents an ANN-based deep learning model that accurately captures the complex dynamics of pulse propagation in supercontinuum generation within chalcogenide planar waveguides. This model enhances the ability to predict spectral evolution under various pump powers, thereby facilitating real-time optimization of experimental setups and advancing the practical application of supercontinuum generation techniques.

3. **Introduction of Closed-Form Continuous-Time Neural Networks (CfC) for Supercontinuum Spectra Prediction:** A novel approach employing CfC networks is introduced for predicting supercontinuum spectra in waveguides with diverse material compositions. This model demonstrates superior performance compared to traditional methods, highlighting its efficiency, robustness, and adaptability for various materials. It represents a significant advancement in predictive modeling within photonics.

These contributions collectively advance the field by integrating machine learning techniques to enhance the accuracy and efficiency of photonic structure simulations. They provide valuable tools for the design, optimization, and real-time control of photonic devices, paving the way for future innovations in the field.

1.6 Thesis Overview

The thesis is structured into six chapters, starting with this introductory chapter. Chapter 1 establishes the foundation for a comprehensive exploration of integrating deep learning techniques with photonics research. It outlines the importance of advancing optical waveguide design and performance analysis through machine learning and sets the stage for the subsequent chapters.

Chapter 2 introduces the theoretical concepts of optical waveguides. It covers the fundamentals of various waveguide structures such as planar waveguides, photonic crystal fibers, and tapered waveguides, along with their linear and nonlinear properties. This chapter provides a thorough understanding of the key parameters and phenomena relevant to waveguide analysis.

Chapter 3 focuses on the application of deep learning in photonic crystal fiber analysis. It details the methodology for utilizing machine learning models to predict and optimize orbital angular momentum (OAM) transmission, highlighting the data collection, model development, and performance evaluation processes.

Chapter 4 expands the discussion to supercontinuum generation in chalcogenide waveguides. This chapter presents the methodology for using Artificial Neural Networks (ANNs) to model supercontinuum spectra, including the data acquisition, model evaluation, and analysis of various operating conditions.

Chapter 5 explores the use of Closed-Form Continuous-Time Neural Networks (CfC) for modeling supercontinuum generation across multiple materials. It compares the performance of CfC models with traditional machine learning techniques and demonstrates their efficacy in capturing complex spectral patterns with improved computational efficiency.

The final chapter synthesizes the key findings from the research, summarizing the contributions made in integrating deep learning with photonics. It discusses the implications of the results, potential applications, and outlines future research directions to further advance the field of photonics and machine learning integration.

1.7 Conclusion

In summary, this introduction has outlined the transformative potential of deep learning in the field of photonics, particularly in the design and optimization of optical waveguides. The convergence of machine learning techniques with photonic research addresses the limitations of traditional methods, offering new pathways for accelerating the design process, improving accuracy, and enabling the exploration of complex photonic structures. By leveraging advanced neural network models, this thesis seeks to contribute to both the theoretical understanding and practical applications of photonic systems, demonstrating how deep learning can revolutionize the way we approach optical waveguide design and nonlinear dynamics. As the field continues to evolve, the integration of deep learning promises to unlock new possibilities in photonic research, setting the stage for future innovations and discoveries.

Chapter 2

Theoretical Concepts

The theoretical underpinnings regulating optical waveguides are laid out in this section. Firstly the basic ideas of waveguide technology, including notions like mode propagation and light confinement have been discussed. After that, investigation the linear characteristics of waveguides, looking at variables like dispersion, mode purity, and effective refractive index were done. Going forward, study of nonlinear characteristics displayed by waveguides, such as the formation of supercontinuums, four-wave mixing, and self-phase modulation are discussed. Then, numerical approaches to waveguide analysis are covered, including time domain simulations using finite element and finite difference techniques.To comprehend the special qualities and potential uses of advanced waveguide structures, such as photonic integrated circuits and metamaterial waveguides, they are also studied. Deep learning techniques might be used to optimize waveguide performance and design, providing insights into how they might completely transform waveguide technology.

2.1 Introduction to Optical Waveguides

2.1.1 Planar Waveguides

Planar waveguides are essential components of integrated optics because they offer vital paths for light to go through thin films that are placed on substrates. A basic architecture of such a waveguide is shown in Figure 2.1 where light is confined in the core[37]. These waveguides, which provide compactness, low propagation losses, and compatibility with semiconductor fabrication processes, are essential for the development of photonic integrated circuits (PICs) and on-chip optical devices. They do this by taking advantage of total internal reflection at the interfaces between the waveguide layer and the surrounding medium. Depending on the waveguide geometry and refractive index profile, planar waveguides can allow numerous guided modes of light confinement, which mainly happens in the transverse dimensions. Planar waveguides have many uses, including optical interconnects in data centers, biophotonics, sensing, and telecommunications [38][39]. Ongoing research endeavors are dedicated to optimizing waveguide designs where they enable low-loss light propagation and are used to connect different optical devices and components[40].

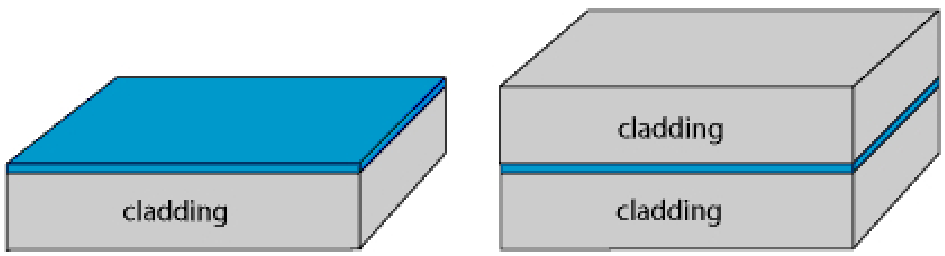


Figure 2.1: Planar Waveguide Schematic.

Planar waveguides are widely applicable to a variety of material platforms, such as semiconductors, glass, and polymers, each of which has its own set of benefits and drawbacks. Innovative processes like inkjet printing and structures influenced by metamaterials are examples of how fabrication technology have advanced. Planar waveguides have many benefits, but they can also have losses due to many reasons, which is why it's important to investigate ways to reduce these losses.

Notwithstanding these difficulties, planar waveguides offer a promising foundation for integrated photonic applications because of their low propagation losses and ability to combine a large number of features. Beyond just helping light travel, they are essential in tying together different optical components and devices, allowing for smooth communication and effective light control in intricate optical systems. Planar waveguides are likely to become even more important as research into the technology proceeds, establishing their place as key components of integrated optics and facilitating next-generation optical communication, sensor, and computing systems.

2.1.2 Photonic crystal fibers

A novel family of optical fibers known as photonic crystal fibers (PCFs) have special structural characteristics, most notably a periodic arrangement of air holes that run the length of the fiber.Basic configuration of PCF is shown in Figure2.2 which shows the distribution of airholes around the core[41]. The periodic architecture of the fiber allows for accurate customization of its optical characteristics, such as nonlinearities, confinement, and dispersion. PCFs are highly adaptable for a wide range of applications, including nonlinear optics, ultrafast lasers, and telecommunications. They offer advantages including low loss, large mode area, and customizable dispersion profiles[42]. Novel functions and devices including supercontinuum sources, high-power lasers, and sensors have been made possible thanks in large part to PCFs [43]. PCFs are still being researched in order to further optimize their performance, investigate new designs for particular applications, and incorporate PCFs into intricate photonic systems and networks. Photonic crystal fibers

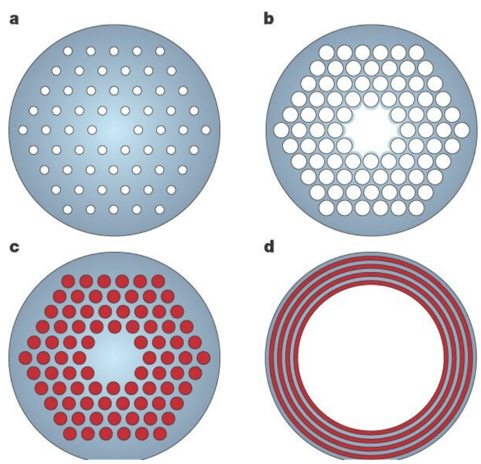


Figure 2.2: Photonic Crystal Fiber.

are a new class of optical fibers that have a unique microstructure, resulting in various advantageous properties. These fibers can guide light through both total internal reflection and the photonic bandgap effect, allowing for single-mode operation, soliton propagation, and controllable group velocity dispersion. The cross-section of these fibers can be optimized to achieve specific functionalities, such as rare-earth doping for high-performance optical amplifiers and lasers. The fabrication and modeling methods of photonic crystal fibers are also discussed in the literature. These fibers have a circular cross-section with a core layer, cladding layer, and coating layer, and their microstructure includes air holes arranged in a concentric and circular array pattern[44]. Photonic crystal fibers now have more varieties and uses thanks to recent advancements in the sector. It has been demonstrated that photonic crystal fibers may transmit many orbital angular momentum modes with low dispersion, modest nonlinear coefficient, and low confinement loss[45]. With their unmatched performance and adaptability in guiding and modifying light, these fibers offer a potential path for the development of next-generation optical communication systems and devices.

2.1.3 Tapered Waveguides

Tapered waveguides are optical devices that smoothly transition the guided mode properties over their length by progressively changing the cross-sectional dimensions.The properties of the waveguide design are illustrated in Figure2.3, demonstrating the light propagation mechanism[46]. In order to achieve effective mode conversion, coupling, and light confinement, tapering alters the mode field distribution, effective refractive index, and mode size. Applications for tapered waveguides include mode conversion in nonlinear optics, fiber-to-chip coupling in integrated photonics, and mode matching in optical fiber systems. Advantages of their gradual transition include improved optical parametric amplification in integrated circuits, possible reduction of waveguide length, and ease of fabrication for optical signal processing and telecommunication applications[47].Tapering reduces spectral dip problems and produces highly coherent mid-infrared supercontinuum (SC) output. This is especially noticeable in the parabolic taper design[48].Ongoing research in tapered waveguides focuses on optimizing taper profiles for specific applications, exploring novel taper designs, and integrating tapered waveguides into photonic devices and systems for enhanced performance. Tapered waveguides can exhibit phenomena such as zero group velocity and strong spatial compression, observed in tapered plasmonic waveguides and spatio-temporal systems[49]. Curved shape photonic crystal tapers have been designed and modeled to enhance coupling efficiency in photonic crystal waveguides, with different approaches and shapes considered. Tapered waveguides are a common design constraint in photonic integrated circuits (PICs) for mismatch adaptation between waveguides of different cross-sections[50].

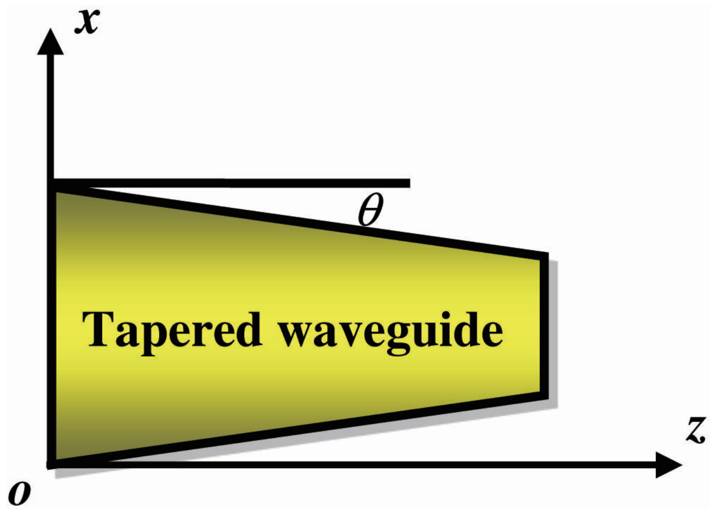


Figure 2.3: 2D Tapered Waveguide Schematic.

2.1.4 Metallic Waveguides

Surface plasmon polaritons (SPPs) and other metallic resonances are exploited by metallic waveguides, which use metal-dielectric interfaces to guide electromagnetic waves, including light, at subwavelength scales. These waveguides offer special properties like strong field confinement and enhanced light-matter interactions. These waveguides provide subwavelength imaging, increased spectroscopy, and on-chip optical circuits and find applications in nanophotonics, plasmonics, sensing, integrated optics, and terahertz networks. For applications in communications, sensing, and imaging, ongoing research focuses on performance optimization, investigating new materials and geometries, and incorporating metallic waveguides into functional photonic devices[51] [52].

2.1.5 Rib Waveguide

When light is guided through a raised ridge or rib structure on a substrate, it is known as a rib waveguide. Fig2.4 shows the basic layout of a rib waveguide, highlighting the process of light propagation through the core[53].This type of waveguide is ideal for integrated photonics, optical interconnects, and photonic integrated circuits (PICs) because it provides precise mode confinement and propagation control. Their planar geometry makes it simple to fabricate and integrate with other optical components, which makes it easier to create compact, multipurpose photonic devices that can be used in data transmission, optical sensing, biomedical imaging, and telecommunications. The current focus of rib waveguide research is on improving performance, investigating cutting-edge fabrication methods, and incorporating them into cutting-edge photonic systems and networks for new applications. Common fabrication techniques involve lithography and etching, with demonstrated methods including reflowed photoresist etch masks and photolithography for lithium niobate on insulator (LNOI)[54], as well as wet-chemical etching for silica-titania structures[55]. Additionally, investigations into the single-mode operation of submicron rib waveguides made of silicon-on-insulator material have been conducted using numerical methods, contributing to enhancing the understanding and optimization of rib waveguide designs for various applications in photonics and optical communication.

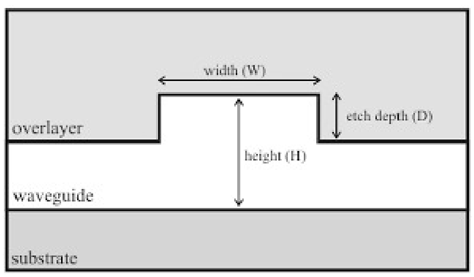


Figure 2.4: Rib Waveguide Schematic.

2.1.6 Advanced Waveguide Structures

Metamaterial Waveguides, Bragg gratings and resonators, Photonic Integrated Circuits (PICs), and other advanced waveguide structures are examples of photonic technologies. PICs are complex systems that combine several optical components onto a single chip. They provide low power consumption, high integration density, and improved performance for optical communications and sensing applications [56][57]. Laser mirrors, sensors, and modulators use Bragg gratings and resonators, optical devices that manipulate light waves through periodic fluctuations in refractive index [58]. The goal of this field's research is to increase spectral selectivity and performance by using methods like dielectric Bragg mirrors and guided-mode resonance. By utilizing artificial electromagnetic properties, metamaterial waveguides can achieve previously unattainable features like as reduced loss, subwavelength confinement, and negative refraction[59]. Research focuses on novel compositions and geometries for advanced imaging, sensing, and communication applications . These waveguides promise superior performance compared to traditional counterparts, offering scalability, reconfigurability, and enhanced control over electromagnetic properties[60].

2.2 Linear Properties of Waveguide

2.2.1 Effective Refractive Index

Effective refractive index, or n_{eff}, is a crucial parameter that affects phase velocity, mode confinement, signal transmission, and dispersion management. As such, it is important for maximizing waveguide performance in a variety of applications, such as sensing, laser systems, and telecommunications. n_{eff}, a crucial parameter that carefully takes into account the refractive indices of the waveguide materials and design, describes how light travels through a waveguide in comparison to how it travels in free space.Pastotta (2024) cite. n_{eff}, for example, precisely controls the modal characteristics of guided modes in optical fibers, which has a major impact on the fiber's overall performance. Relationship of effective refractive index with wavelength of the guiding light is shown in Figure2.5.The effective refractive index (n_{eff}) of waveguides can be obtained through a variety of methods and materials. These include employing well-established techniques like the effective index method (EIM) for simulations, as well as novel all-dielectric zero-index materials (ZIMs) like Dirac-like cone-based ZIMs (DCZIMs)[61]. Furthermore, the complex interaction between anisotropic optical gain and non-Hermitian characteristics of the waveguide structure and materials is taken into consideration while deriving equations for the effective medium refractive index in waveguide design[62]. By means of a thorough comprehension and skillful use of n_{eff}, scientists and engineers can fully realize the possibilities of waveguide technology, propelling creativity and progress across several domains.

2.2.2 Mode Purity (η)

The degree to which a particular mode predominates over other modes in a waveguide's light transmission is measured by the mode purity (η). It indicates the degree

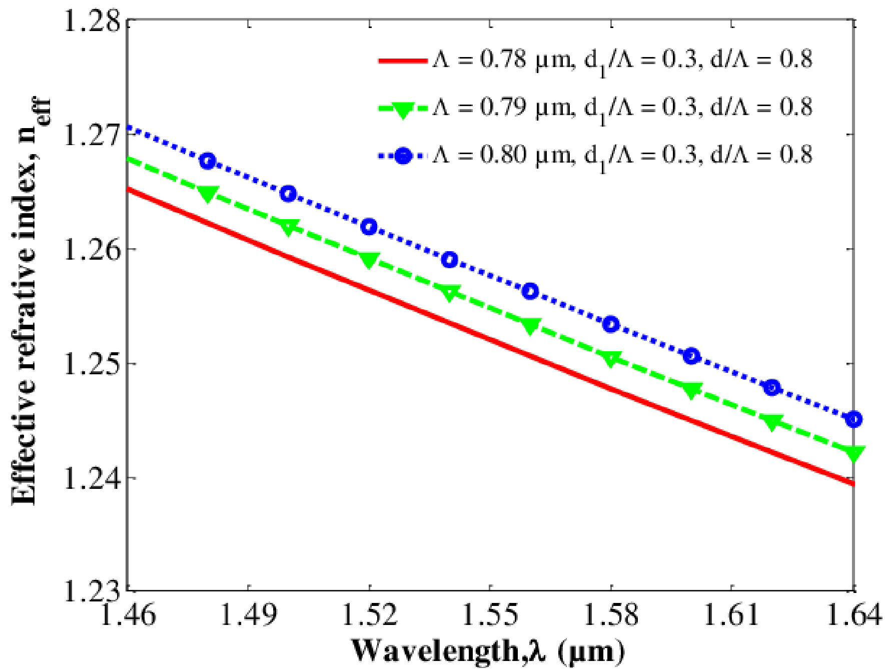


Figure 2.5: Example of Effective Refractive Index and Dispersion profile of Waveguide

of mode mixing or contamination by higher-order modes by reflecting the modal content of the optical signal traveling through the waveguide. In single-mode applications like high-power laser systems and long-distance telecommunications, obtaining high mode purity is essential to minimize signal deterioration and maximize system performance. Numerous elements, such as coupling losses, mode conversion procedures, and waveguide flaws, can affect mode purity. Optimizing waveguide shape, managing fabrication procedures, and putting mode filters or mode-selective components into place are methods for improving mode purity.The mode purity of waveguides is an important factor in various applications. In the study by Soltani et al., a proposed integrated optical mode converter based on multimode interference (MMI) waveguides achieved an output mode purity of 94% and 82% for beams carrying OAM with topological charge values of l = ±1 and l = ±3, respectively [63] This is a suggestion of desired high mode purity for various application. Mode purity calculation in a ring core PCF can be presented by the equation given below,

$$\eta = \frac{\int_r \text{ring} \left|\vec{E}\right| \, dx\, dy}{\int_{\text{whole section}} \text{whole-section} \left|\vec{E}\right|^2 \, dx\, dy} \tag{2.1}$$

2.2.3 Dispersion

When various light wavelengths go through a waveguide at different rates, resulting in the temporal spreading of optical pulses, this phenomenon is referred to as dispersion in waveguides. Two primary types of dispersions can be broadly classified as follows: chromatic dispersion and modal dispersion. The wavelength-dependent refractive index of the waveguide material causes chromatic dispersion, which affects how broadband optical signals propagate. Conversely, modal dispersion arises from the different rates at which directed modes with different spatial distributions propagate. In many optical systems, such as communication networks, laser systems, and optical sensing applications, managing dispersion is essential for regulating pulse broadening and distortion[64].

A variety of dispersion control strategies are used by engineers to efficiently address dispersion. One method is to design the waveguides' internal structures. For example, atomic layer deposition (ALD)-deposited silicon nitride waveguides with cladding layers provide accurate dispersion tuning over a range of regimes, including normal, near-zero, and anomalous[65]. Similar to this, dispersion properties can be freely controlled by patterning the cladding of waveguide microresonators with substances like air, oxide, or SU-8 polymer[66]. Lower index oxide layers allow improved supercontinuum generation with customized bandwidths and spectrum broadening dynamics in highly doped silica glass slot waveguides[67]. Furthermore, high-performance nanophotonic device design is made possible by the advanced dispersion engineering capabilities offered by subwavelength metamaterial waveguides on the silicon-on-insulator platform[68].

These dispersion engineering techniques hold significant promise for various applications. Integrated photonic circuits, optical communication systems, spectroscopy setups, sensors, and on-chip broadband light manipulation can all benefit from precise control over dispersion. By leveraging the diverse methods available for managing dispersion in waveguides, engineers can optimize the performance and efficiency of optical systems across a wide range of domains, advancing the capabilities of modern photonics technology. The equation that expresses dispersion looks like this ,

$$D = -\frac{\lambda}{c}\frac{d^2 R_e(N_{eff})}{d\lambda^2} \tag{2.2}$$

2.2.4 Mode Effective Area (A_{eff})

A waveguide's effective area (A_{eff}) measures the cross-sectional area that optical power travels through. It defines the optical mode's spatial extension and affects phenomena like mode confinement, attenuation, and nonlinear effects. Weaker optical nonlinearities and decreased sensitivity to optical losses are generally associated with greater effective areas. In optical fiber systems, A_{eff} plays a crucial role in influencing nonlinear phenomena such as four-wave mixing and self-phase modulation. Designing waveguides with improved power handling, dispersion management, and nonlinear optics is made possible by optimizing A_{eff}. Numerous scholarly articles examine the A_{eff} of various waveguide configurations.According to numerical simulations, supercontinuum generation in photonic crystal fibers is influenced by the frequency-dependent effective mode area. This is especially true for lowering the extreme long-wavelength edge of the spectrum, which is important for precise comparisons with experiments that use pump sources in the 1000–1500 nm range[69]. Slot waveguides are shown to be more appropriate for this purpose by Luo et al., who investigate the relationship between the nonlinear effect and the waveguide's cross-section and adjust the construction to maximize the nonlinear effect [70].The effective area of photonic crystal fibers with a triangular air-hole lattice is the subject of another study that offers numerical analyses and insights into the nonlinear behavior and single-mode operation of these fibers, which are important for a variety of optical communication and sensing applications[71]. The equation that characterizes the mode effective area looks like this,

$$A_{eff} = \frac{\left(\iint |E(x,y)|^2\, dx\, dy\right)^2}{\iint |E(x,y)|^4\, dx\, dy} \tag{2.3}$$

2.2.5 Orbital Angular Momentum (OAM)

A spiral wavefront representing the helical phase structure of light defines orbital angular momentum (OAM), a unique feature of light. When used and controlled inside waveguides, this property allows optical signals to carry additional information beyond their amplitude and phase. The potential for improving data transmission capacity and robustness against channel impairments in OAM-based communication systems has sparked a great deal of interest in waveguides that can generate, manage, and detect OAM modes.Figure2.6 outlines the structural configuration of a PCF, explaining how light is directed and maintained within for supporting OAM mode[72]. Applications for these advances can be found in many different fields, including as high-capacity data transfer, quantum information processing, and optical communications. The main focus of current research efforts is on creating effective waveguide platforms that can incorporate OAM modes into useful photonic systems for sensing and communication applications. As a possible

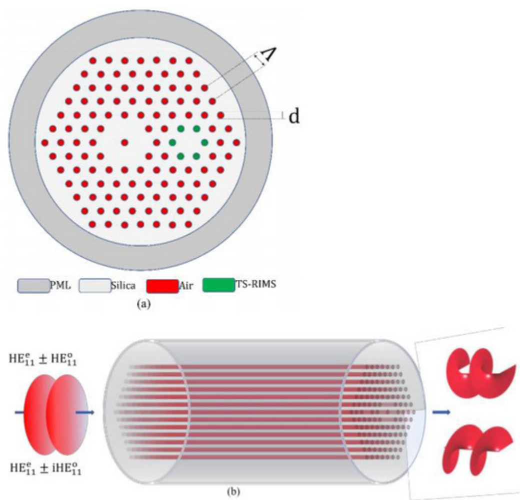


Figure 2.6: Orbital Angular Momentum Mode Generator

medium for OAM transmission, photonic crystal fibers (PCFs) are being investigated for their ability to accommodate numerous OAM modes with low losses and high purity. Using PCFs with double-layer rings, where the transmission domain is doped with germanium dioxide and the inner circle is doped with fluoride, is one such method [45]. A different approach makes use of hybrid cladding ring-core PCFs, which break the fiber structure's circular symmetry to improve bending resistance and decrease spin-orbit coupling. This allows for the support of 22 OAM modes over a large bandwidth with minimal confinement losses [73]. Furthermore, formalisms based on Stokes parameter measurements and Barnett's formalism have been devised to experimentally determine the OAM of optical beams [74]. These advancements pave the way for more accurate characterization and manipulation of OAM modes within waveguides. In order to improve resolution and capacity, efforts have also been focused on modeling and reducing crosstalk in OAM-multiplexed holographic systems. The goal of research is to maximize the performance of OAM-multiplexed holographic systems by developing models for multiplexing crosstalk in OAM holography and suggesting techniques to minimize crosstalk [75]. Exciting prospects exist for optical communication and sensing technologies to be revolutionized through the investigation of OAM in waveguides. Advances in waveguide engineering and OAM manipulation techniques are making it possible to integrate OAM-based systems into real-world photonic devices, which could lead to greater system robustness and increased data transmission capacities.

2.3 Non-linear Properties of Waveguide

2.3.1 Nonlinear Refractive Index n_2

Understanding the nonlinear optical behavior of materials requires an understanding of the nonlinear refractive index (n_2), which is the change in refractive index caused by the intensity of incident light. n_2 regulates nonlinear processes in waveguides, including self-phase modulation and self-focusing, which are important for nonlinear signal processing and all-optical switching applications. The strength of these interactions is determined by the magnitude of n_2, which influences the waveguide design for nonlinear optics applications. The characterization and application of waveguides' nonlinear refractive index have been the subject of numerous studies. Wang et al. introduce a refractive index sensor that uses Fano resonance to achieve extraordinary sensitivity [76], whereas Neradovskiy et al. propose a method based on second harmonic production to determine waveguide refractive index profiles[77]. Yeoh et al. present a graphene-coated multimode interference waveguide sensor for precise refractive index measurements suitable for real-time environmental monitoring [78]. These studies collectively contribute insights into diverse techniques and materials for manipulating and measuring the nonlinear refractive index of waveguides, paving the way for advancements in nonlinear optics research and practical device applications.

2.3.2 Self Phase Modulation

Self-phase modulation (SPM) is a nonlinear optical phenomena in which an optical pulse's phase is changed as it travels through a waveguide by its own intensity. Phase shifts that depend on the duration and form of the optical pulse are caused by the waveguide material's refractive index changing in response to the optical pulse's instantaneous intensity in surface photometry (SPM). SPM is a helpful tool in nonlinear optics applications, but it can also produce spectral widening and temporal aberrations of light pulses. Applications including optical signal processing, supercontinuum production, and nonlinear microscopy require an understanding of and ability to regulate SPM[79].Self-phase modulation (SPM) is a fascinating nonlinear optical effect where a laser beam interacts with a medium, inducing a phase modulation on itself due to the medium's intensity-dependent refractive index change, resulting in spatial and temporal distortions and leading to well-known phenomena like self-focusing and self-defocusing[80].

2.3.3 Four Wave Mixing

Through the nonlinear coupling of optical fields, interacting optical waves create new frequencies in a process known as four-wave mixing (FWM). FWM occurs in waveguides as a result of the medium's nonlinear response, which creates new optical frequencies that meet phase-matching and energy-saving requirements. Applications including optical parametric amplification, frequency comb creation, and wavelength conversion can all benefit from the use of FWM. In optical communication systems, it can also result in crosstalk and signal deterioration, though. Optimizing waveguide designs for nonlinear optics and photonics applications requires a thorough understanding of and control over FWM[79].

2.3.4 Supercontinuum Generation

Supercontinuum generation is a nonlinear optical phenomena in which a narrow-band input source produces an optical spectrum with a wide range of wavelengths. It happens when different nonlinear phenomena in waveguide structures interact, such as self-phase modulation, stimulated Raman scattering, and four-wave mixing. Applications for supercontinuum sources include ultrafast laser systems, spectroscopy, optical frequency metrology, and optical coherence tomography. To design sources with customized spectral properties and high coherence for various applications, it is crucial to comprehend and manage the formation of supercontinuum in waveguides [81]. Figure2.7 represents the spectral broadening of light with spatial and temporal evolution plot that varies with the geometric variations[82] An

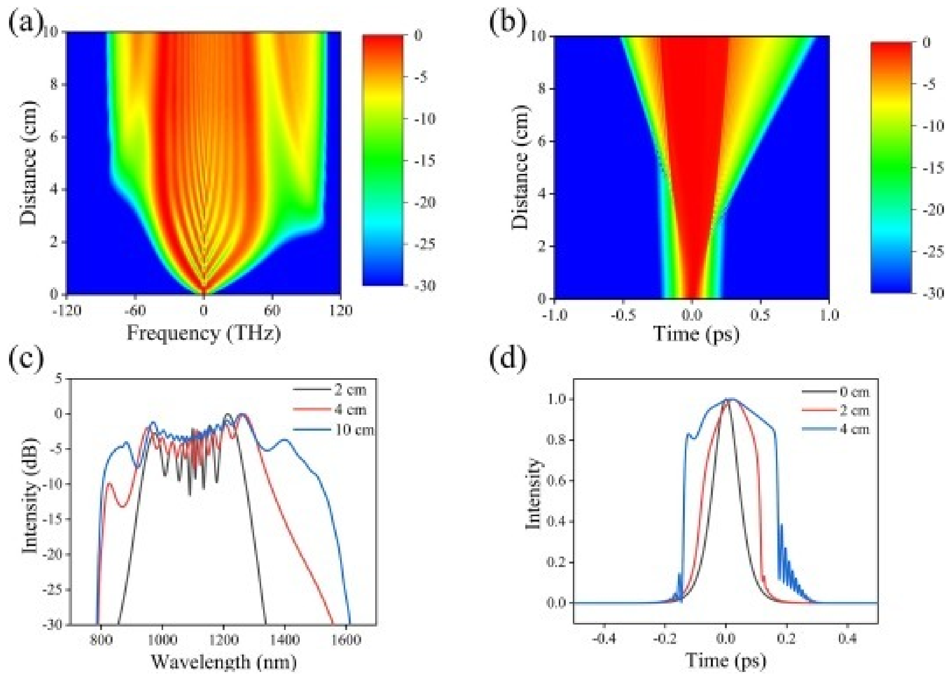


Figure 2.7: Supercontinuum Generation

effective method of obtaining extreme spectrum widening of pulsed laser light is supercontinuum generation through waveguides. Applications include frequency conversion, optical light manipulation, and telecommunication. Achieving the appropriate supercontinuum spectra requires careful consideration of waveguide design and dispersion management. Supercontinuum generation has been studied with a variety of waveguide configurations. Different dispersion profiles and spectrum broadening dynamics are provided by these waveguides, which enable the creation of remarkably diverse supercontinuum spectra. Supercontinuum generation in chip-based platforms is now more understood and controlled because to the development of photonic integrated waveguides and photonic crystal fibers. These devices present new prospects for supercontinuum generation due to their finely controlled lithographic dispersion, small footprint, and enhanced power consumption[43].

Supercontinuum generation is the process of transforming laser light into a wide spectrum, which is essential for many uses. Because of the customized dispersion characteristics of optical fibers, particularly photonic crystal fibers, it usually happens in nonlinear devices. Spectral broadening is caused by mechanisms such as self-phase modulation and Raman scattering, which require numerical modeling to fully comprehend. High spatial coherence can be seen in supercontinua despite their wide spectrum. Supercontinuum sources' wide bandwidth and spatial coherence are advantageous for spectroscopy, microscopy, and optical coherence tomography, among other applications. Supercontinuum sources are also useful in optical frequency metrology and ultrafast laser physics[83].

2.3.5 Stimulated Raman Scattering (SRS)

Induced photons interact with the material's vibrational modes in stimulated Raman scattering (SRS), a nonlinear optical process that produces new light frequencies as a result of the scattering process. SRS can cause optical power in waveguides to be transferred from the incident wavelength to new wavelengths that are specified by the Raman spectrum of the material. SRS is used in spectroscopy, wavelength conversion, and Raman amplification, among other applications. On the other side, it may also result in nonlinear signal distortion and restrict waveguide devices' ability to handle power. Optimizing waveguide designs for nonlinear optics and photonics applications requires an understanding of and control over SRS[84].

Important details about the material, including its spectral band structure, degree of polarization, and scattering efficiency, can be learned from the scattered radiation.Many pieces of equipment, such as light pulse generators, optical systems, and detectors, can be used to find stimulated Raman scattering [85]. It may be used to generate bright tunable light in the infrared range, measure the relaxation times of lattice waves and molecular vibrations, and determine spontaneous Raman parameters [86].

2.3.6 Absorption of Two Photons (TPA)

Two photons are simultaneously absorbed in two-photon absorption (TPA), a nonlinear optical process that causes an electronic transition in a material. The material's absorption coefficient and refractive index alter as a result of this phenomena, which only happens at high intensities. TPA is essential in waveguides for quantum information processing applications, nonlinear optical microscopy, and optical switching[87] [79].

2.4 Numerical Technique Used in Waveguide Analysis

2.4.1 Finite Element Method (FEM)

By discretizing the domain into finite elements, the Finite Element Method (FEM) is a potent numerical technique that is frequently used in waveguide analysis to solve partial differential equations. The realistic simulation of light propagation and interaction with complicated waveguide systems characterized by flexible nonuniform patches or elements is made possible by FEM, which converts differential equations into matrix equations. FEM, which was first founded on the research of Ritz, Galerkin, Trefftz, and others, uses weighted residual methods or mechanical principles to translate differential equations into variational issues. FEM builds systems of algebraic equations using finite element meshes and basis functions, enabling the study of modal features, dispersion characteristics, and nonlinear effects in waveguides. Its versatility in handling complex geometries, material properties, and boundary conditions makes FEM an invaluable tool for a wide range of waveguide designs and applications[88][89]. Finite Element Analysis shown in Figure2.8 is a helpful mechanism of knowing the electromagnetic field distribution[90].

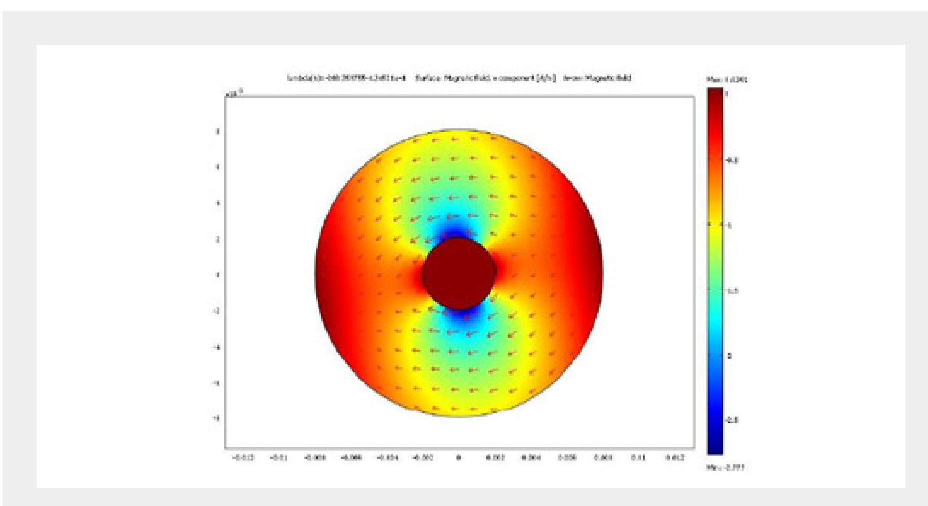


Figure 2.8: Finite Element Analysis of Waveguide

2.4.2 Finite Difference Time Domain (FDTD)

By discretizing space and time into a grid, the Finite Difference Time Domain (FDTD) method is a numerical methodology that simulates the propagation of electromagnetic waves. With its great temporal and spatial resolution, FDTD is especially well-suited for exploring transient events, dispersion, and nonlinear effects in waveguide analysis. FDTD effectively computes temporal derivatives from spatial derivatives by using trigonometric basis functions and Taylor series extension[91]. Application of the finite-difference time-domain (FDTD) method lies in solving Maxwell's equations for photonics and nanophotonics, providing a rigorous and versatile technique for accurately modeling nanophotonic devices and understanding novel effects in nano-optics and photonics, particularly advantageous for modeling surface plasmon waves and plasmonic waveguides due to its affordability in computational resources[92].

2.4.3 Sellmeier Equation

The Sellmeier equation, developed by W. Sellmeier in 1871, is an empirical formula used to describe the wavelength dependence of the refractive index in transparent materials like glasses. It provides a relationship between the refractive index of a material and its wavelength, incorporating material-specific coefficients. In waveguide analysis, the Sellmeier equation is essential for modeling the refractive index dispersion of waveguide materials, allowing for accurate simulations of light propagation and dispersion effects. By integrating Sellmeier coefficients into numerical simulations, waveguide designers can predict and optimize the optical properties of waveguide structures across a broad spectral range[93]. The equation is as follows,

$$n^2(\lambda) = 1 + \sum_{i=1}^{N} \frac{B_i \lambda^2}{\lambda^2 - C_i} \tag{2.4}$$

Where represents wavelength in μm. Where represents wavelength in µm.

2.4.4 Perfectly Matched Layer (PML)

In numerical simulations, the Perfectly Matched Layer (PML) plays a critical role as a boundary condition, decreasing reflections at computational domain boundaries and absorbing outgoing waves. PMLs are essential for accurately simulating open boundaries and truncating waveguide structures without adding undesired reflections in waveguide analysis. PMLs provide correct modeling of waveguide behavior, including modal characteristics, dispersion, and nonlinear effects. This ensures accurate and dependable results in finite-difference and finite-element simulations of waveguide structures by effectively simulating an infinite computational domain. Beyond waveguide analysis, PMLs are also used in optical simulations, soil-structure interaction analysis, and wave propagation studies. These applications allow PMLs to efficiently absorb light across a range of incident angles and customizable bandwidths, which improves computational efficiency and has a practical impact in a variety of fields.Recent advancements, such as novel PML designs tailored for specific materials and domains, further expand the utility and effectiveness of PML in a wide range of applications, promising even greater accuracy and efficiency in numerical simulations[94][95].

2.4.5 Meshing

In order to discretize the space for numerical simulations, the computational domain is divided into tiny geometric parts, or meshes. Meshwork is essential to accurately describing the waveguide geometry, material interfaces, and field distributions in waveguide analysis. Accurate numerical simulations are guaranteed by proper meshing, which appropriately resolves spatial variations in the electromagnetic fields. Depending on the complexity of the waveguide structure and the needs of the simulation, different meshing techniques are applied, including structured, unstructured, and adaptive meshing[96].

2.4.6 Convergence

The point at which an iterative numerical simulation achieves a stable solution or the required degree of accuracy is referred to as convergence. Convergence is necessary in waveguide analysis in order to produce accurate and significant results from numerical simulations. As computing resources (such as mesh density and time steps) are optimized, it guarantees that the answer derived from the numerical method is in close proximity to the actual solution of the underlying physical problem. Convergence criteria, including error norms, residual values, and solution stability, are evaluated as part of the monitoring process to see if more fine-tuning is required to provide accurate outcomes.

2.4.7 Softwares

Software packages specifically designed for waveguide analysis provide powerful tools for simulating and optimizing waveguide structures. These software packages typically incorporate advanced numerical techniques, intuitive graphical user interfaces (GUIs), and extensive libraries of optical materials and components. Examples of such software include Lumerical's MODE Solutions, COMSOL Multiphysics, and RSoft's BeamPROP. These software packages enable waveguide designers to efficiently model, analyze, and optimize a wide range of waveguide structures, including photonic crystal fibers, planar waveguides, and integrated photonic devices, facilitating the development of novel photonic technologies.

2.5 Generalized Nonlinear Schrödinger Equation (GNLSE)

2.5.1 Introduction to GNLSE

The propagation of optical pulses in nonlinear media, such optical fibers and waveguides, is described by the Generalized Nonlinear Schrödinger Equation (GNLSE), a basic mathematical model. It can simulate a wide range of nonlinear phenomena in waveguide constructions and takes into consideration both linear and nonlinear effects, such as dispersion, self-phase modulation, and four-wave mixing. The GNLSE is a popular tool in photonics research and optical communication systems for predicting optical signal behavior and evaluating pulse propagation dynamics. It is developed from the nonlinear Schrödinger equation[97]. With applications in weak nonlinear dispersive water waves, quantum field theory, and nonlinear optics, the generalized nonlinear Schrödinger (NLS) equation is a fully integrable model[98]. The equation is also used in the field of slowly varying envelope of the electric field in optical fibers, with various methods employed to find soliton solutions[99]. These studies contribute to the understanding of the behavior and properties of solutions to the generalized NLS equation in different physical systems.

2.5.2 Mathematical Form of GNLSE

The Generalized Nonlinear Schrödinger Equation (GNLSE) is fundamental for modeling the evolution of the complex envelope of an optical pulse as it traverses through a nonlinear medium. This equation integrates both linear and nonlinear effects, capturing a comprehensive range of optical phenomena. The linear terms account for dispersion and attenuation, while the nonlinear terms represent effects such as self-phase modulation (SPM) and nonlinear gain.

The GNLSE is generally expressed in either the time domain or frequency domain and is often solved using various numerical techniques. A thorough understanding of the GNLSE's mathematical formulation is crucial for accurate simulation and modeling of nonlinear optical processes in waveguide structures. The equation is given by [43]:

$$\frac{\partial A}{\partial z} + \frac{\alpha}{2} A - \sum_{k \geq 2} \frac{i^{k+1}}{k!} \beta_k \left(\frac{\partial^k A}{\partial T^k} \right) = i\gamma \left(1 + i\tau_{\text{shock}} \frac{\partial}{\partial T} \right) \left(A(z,T) \int_{-\infty}^{+\infty} R(T') |A(z, T - T')|^2 \, dT' \right) \tag{2.5}$$

Here, $A(z, T)$ represents the slowly varying envelope of the pulse, which propagates at the group velocity $v_g = \frac{1}{\beta_1(\omega)}$ in a retarded time frame $T = t - z/v_g$. The higher-order dispersion terms, β_k, are obtained from the Taylor series expansion of the mode propagation constant $\beta(\omega)$. The central angular frequency is denoted as ω_0, and α signifies the attenuation constant. The nonlinear coefficient γ is calculated using $\frac{n_2 \omega_0}{c A_{\text{eff}}}$, where n_2 is the nonlinear refractive index at ω_0, c is the speed of light, and A_{eff} is the effective mode area.

The GNLSE encapsulates both linear and nonlinear dynamics governing the propagation of optical pulses within waveguides. The linear components include group velocity dispersion (GVD), chromatic dispersion, and losses, which describe how the optical signal evolves linearly. In contrast, the nonlinear terms represent effects driven by the intensity-dependent refractive index of the waveguide material, such as self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM). A deep understanding of how these linear and nonlinear terms interact is essential for accurately predicting the behavior of optical pulses in waveguide structures.

2.5.3 Numerical Solvers for GNLSE

To effectively solve the Generalized Nonlinear Schrödinger Equation (GNLSE) and simulate optical pulse propagation in waveguides, various numerical solvers are employed. One prevalent approach involves transforming the GNLSE from the time domain to the frequency domain using the Fast Fourier Transform (FFT) [100]. This technique allows for efficient and accurate solution of the GNLSE by leveraging iterative methods in the frequency domain. FFT-based solvers are particularly valuable for modeling the dynamics of pulse propagation, analyzing dispersion effects, and understanding nonlinear interactions within waveguide systems. They are extensively used in the design and optimization of waveguides for photonics research and optical communication applications due to their speed and precision.

In addition to FFT-based methods, alternative approaches have been proposed for solving the multimode generalized nonlinear Schrödinger equation (MM-GNLSE). A notable method involves transforming the MM-GNLSE into a system of first-order ordinary differential equations (ODEs). These ODEs can be solved using standard ODE solvers, simplifying the modeling of pulse propagation in multimode fibers. This approach has been verified for basic multimode cases, such as those involving two orthogonal polarization states in a non-birefringent microstructured optical fiber [101]. By converting the MM-GNLSE into a more manageable system of ODEs, researchers can efficiently tackle complex multimode scenarios and gain insights into pulse behavior in advanced optical fiber designs.

2.5.4 Experimental Validation and Comparison

It is imperative to conduct experimental validation and compare numerical simulations with real-world observations in order to confirm the precision and dependability of GNLSE-based models in forecasting waveguide behavior. The process of experimental validation entails measuring the optical characteristics and operational efficiency of waveguide structures in controlled laboratory settings. The agreement between theory and experiment is then evaluated by contrasting these measurements with numerical simulations based on the GNLSE. Disparities between simulation and experimental data can point to flaws or restrictions in the model's assumptions or experimental configurations, allowing for more waveguide design optimization and refinement.

2.5.5 Softwares

MATLAB is a widely used numerical computing environment and programming language that provides powerful tools for solving and analyzing mathematical problems, including the GNLSE. MATLAB offers a range of built-in functions and toolboxes for numerical simulation, signal processing, and optimization, making it well-suited for implementing numerical solvers for the GNLSE and conducting simulations of waveguide behavior. MATLAB's intuitive interface and extensive library of functions enable researchers and engineers to develop custom simulation scripts, visualize simulation results, and analyze waveguide properties efficiently. MATLAB's versatility and flexibility make it a popular choice for waveguide analysis and optimization in both academic and industrial settings.

2.6 Deep Learning

2.6.1 Artificial Neural Networks

Artificial Neural Networks (ANNs) are computational models inspired by the structure and function of biological neural networks found in the human brain. The design and operation of ANNs mirror the way neurons interact within the brain, where they process and transmit information through complex networks. An ANN typically comprises several layers: an input layer, one or more hidden layers, and an output layer [102]. Each layer in an ANN serves a distinct purpose and contributes to the network's ability to process data and make predictions.In an ANN, every node, or neuron, within a layer performs a mathematical transformation on its inputs. This transformation is governed by a set of weights and activation functions, which are crucial for determining the node's output. The output from a node is then passed to the nodes in the subsequent layer, effectively propagating the information through the network. This process continues from the input layer through the hidden layers to the output layer, as depicted in Figure 2.9.ANNs are highly versatile tools used for a wide range of machine learning tasks. They excel in pattern recognition, regression, and classification problems due to their ability to learn and model complex relationships within data. For instance, ANNs can effectively analyze optical waveguide parameters by learning from large datasets, making them valuable for tasks such as predicting waveguide performance or optimizing design parameters.The most common type of ANN is the multi-layer feedforward network. This architecture consists of an input layer that receives the initial data, one or more hidden layers that process the data, and an output layer that provides the final results. The connections between neurons, known as links or edges, transmit information from one layer to the next. These connections are weighted, and the weights are adjusted during the training process to minimize prediction errors.Training an ANN involves providing it with a set of input patterns and their corresponding output patterns. The network learns by adjusting the weights of the connections to improve its predictions for new, unseen input patterns. This iterative learning process continues until the network achieves satisfactory performance.ANNs have found widespread applications across various fields including business, engineering, and medicine [65] [103]. Their ability to model complex and nonlinear relationships makes them an invaluable tool for solving challenging problems and making informed decisions based on large volumes of data.

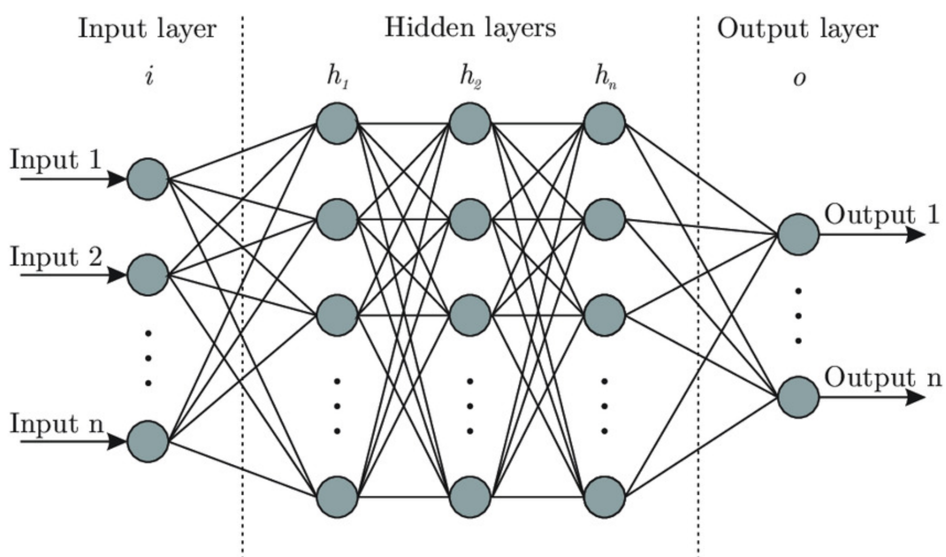


Figure 2.9: Artificial Neural Network Architecture

2.6.2 Recurrent Neural Networks (RNNs)

Recurrent Neural Networks (RNNs) are advanced neural network architectures designed to handle sequential data by leveraging their ability to remember previous inputs. Unlike feedforward neural networks, which process each input in isolation, RNNs incorporate feedback connections that create directed cycles within the network. This unique architecture allows RNNs to capture and utilize temporal dependencies in sequential data, enabling them to model and understand patterns over time. Figure 2.10 illustrates how sequential data is used in the training process of an RNN model [104].RNNs are particularly adept at tasks that involve sequences and time-dependent information. They are widely employed in various domains such as time-series analysis, where they forecast future values based on historical data, and language modeling, where they predict the next word in a sequence based on previous words. In speech recognition, RNNs are used to interpret spoken language by analyzing sequences of audio features. Their ability to manage sequential data makes them invaluable for a range of applications.In the context of waveguide simulations that involve time-dependent phenomena, RNNs are especially beneficial. They can effectively model dynamic changes in waveguide parameters over time, providing insights into complex interactions within the system. Beyond their use in waveguide simulations, RNNs also demonstrate their versatility in several other areas. They are capable of analyzing non-stationary random processes, which are processes whose statistical properties change over time. Additionally, RNNs have been applied in medical diagnostics for forecasting cardiac conditions, in environmental engineering for evaluating sewage treatment efficacy, and in various other fields where temporal correlations in data are critical [105].Overall, RNNs offer a powerful toolset for dealing with sequential and time-dependent data, making them a valuable asset in both theoretical and practical applications across numerous disciplines.

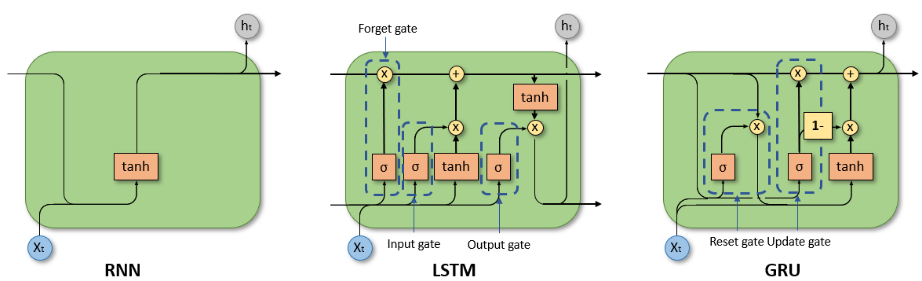


Figure 2.10: Recurrent Neural Network and LSTM Architechture

2.6.3 Long Short-Term Memory (LSTM) Networks

A particular kind of recurrent neural network (RNN) called Long Short-Term Memory (LSTM) networks is designed to address the difficulty of the vanishing gradient problem that arises during deep neural network training. LSTMs are particularly good at capturing long-range relationships in sequential data because of their special architecture, which combines memory cells and gating mechanisms. This helps them deal with problems such as vanishing or ballooning gradients. LSTM models are widely used in a wide range of domains, including time series prediction, brain-computer interfaces (BCIs) for decoding physiological and neural signals, stock and currency market forecasting, and optimizing hardware implementation for lightweight applications. Renowned for their ability to accurately predict market movements, decode brain signals with high accuracy, and enhance prediction accuracy across various domains, LSTM models are often combined with other deep learning methods to further improve prediction accuracy[106].

2.6.4 Convolutional Neural Networks (CNNs)

In computer vision tasks, Convolutional Neural Networks (CNNs) are adaptable models that demonstrate human-like performance via experiential learning. CNNs have the potential to be more closely aligned with their biological counterparts due to their noteworthy parallels to the monkey visual system. This could lead to new research areas beyond basic object recognition. Their versatility may be seen in a wide range of applications, such as helping doctors by helping to categorize brain tumors according to different factors and maximizing time efficiency. CNNs are useful for picture categorization in automated driving and driver assistance systems. They provide robustness in large-scale image processing by utilizing characteristics like pooling, shared weights, and local connections[107].CNN model, as can be seen in Figure2.11, uses convolution technique of mathematics across different layers of the network[108].

2.6.5 Liquid Time-Constant Networks (LTCNs)

Recurrent neural networks called Liquid Time-Constant Networks (LTCNs) are made to deal with problems such as disappearing gradients in training sequences. LTCNs, which are composed of linked nodes with different time constants, are able to effectively represent long-range dependencies and temporal dynamics in sequential data[109]. LTCNs are useful for tasks involving dynamic waveguide behavior and optical signal processing. They are especially well-suited for modeling dynamic

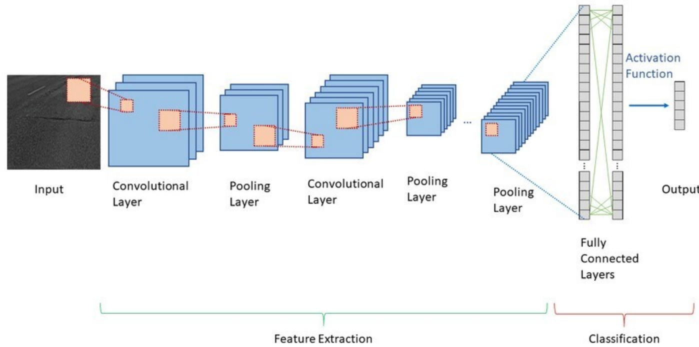


Figure 2.11: Convolutional Neural Network Architecture

processes and time-series data. With no need for scenario-specific data or retraining, an extension known as Liquid Time-Constant (LTC) networks uses received signal power to forecast future blockage status in millimeter-wave links with excellent accuracy. Additional developments such as LTC-SE improve the compatibility and versatility of liquid neural networks, increasing their use in robotics and time-series prediction, among other machine learning applications[110]. Long-term dependencies between image, text, audio, and medical time-series data can be shown by LTC.

2.6.6 Closed-Form Continuous-Time Neural Networks (CfCNs)

Closed-Form Continuous-Time Neural Networks (CfCNs) are a specialized class of neural network architectures designed to model dynamical systems within a continuous-time framework. These networks excel in capturing complex spatiotemporal dynamics, making them particularly well-suited for applications that require continuous-time simulations. Their ability to handle real-time data is advantageous for tasks such as predicting optical properties and modeling waveguide dynamics.CfCNs utilize continuous-time optimization methods during training, which allows them to effectively represent and process continuous-time data. This approach is highly beneficial for applications that demand real-time analytics and precise temporal predictions. A related variant, known as Closed-Form Continuous-Time Neural Networks (CfC), focuses on representation learning for spatiotemporal decision-making tasks. CfC models, which are defined by continuous differential equations, encounter challenges related to scaling when employing traditional numerical solvers.To address these challenges, a novel method has been developed that leverages closed-form continuous-time liquid neural networks. This revolutionary approach facilitates real-time analytics by providing a solution to the scaling issues associated with numerical solvers. By utilizing closed-form networks, this method enables the creation of comprehensive digital twin simulations in healthcare, allowing for detailed and dynamic views of patient health data.CfC networks, derived from analytical closed-form solutions, offer a simpler and more efficient alternative to continuous-depth neural models. They demonstrate strong modeling capabilities and often outperform advanced recurrent models in time-series prediction tasks. This makes them particularly suitable for environments where computational resources are limited, showcasing their potential for real-time applications and resource-constrained scenarios [111].

2.6.7 Data Acquisition

Data acquisition is the process of collecting and organizing experimental or simulated datasets to train machine learning models effectively. In the context of waveguide analysis, this may involve generating synthetic datasets, performing numerical simulations of waveguide behavior, or conducting physical experiments to measure optical properties. Ensuring high quality and diversity in the datasets is critical, as it directly impacts the accuracy and reliability of the trained machine learning models. Effective data acquisition strategies include methods for integrating, validating, and cleansing data to ensure it is representative and free from errors [112].

2.6.8 Dataset

A dataset comprises a collection of samples used to train, validate, and test machine learning and deep learning models. For waveguide analysis, this dataset may consist of experimental measurements, simulated data, or synthetic representations of different waveguide configurations. Constructing well-organized and accurately labeled datasets is essential for training models that can generalize across a range of waveguide designs and accurately predict their optical properties. A robust dataset ensures that the model can learn effectively and perform well on unseen data, improving its ability to make precise predictions [28].

2.6.9 Back-propagation

Back-propagation is a key algorithm used to train neural networks by adjusting the model's parameters based on the gradient of the loss function with respect to each parameter [113]. During the training process, back-propagation calculates the error gradients and updates the weights and biases of the neural network to minimize the loss function. In waveguide analysis, this algorithm enables the neural network to learn from input-output pairs, thereby enhancing its predictive accuracy over time. Back-propagation is fundamental for optimizing deep learning models and improving their performance in tasks such as predicting waveguide properties and optimizing designs.

2.6.10 Learning Rate

The learning rate is a hyperparameter that determines the step size for parameter updates during the optimization process of machine learning models. In neural network training, the learning rate affects how quickly the model's parameters are adjusted during back-propagation. A higher learning rate can lead to faster convergence but may risk overshooting the optimal solution, while a lower learning rate may result in slower convergence but offer greater stability[114]. Proper tuning of the learning rate is crucial for optimizing the training process and achieving accurate predictions of waveguide properties.

2.6.11 Cross-Validation

Cross-validation is a statistical method used to evaluate the performance and generalization ability of machine learning models by partitioning the data into complementary subsets for training and validation[115]. In waveguide analysis, cross-validation helps assess the predictive performance of models trained on limited data and ensures they generalize well to new waveguide configurations or operating conditions. By repeatedly splitting the data into different training and validation sets, cross-validation provides robust estimates of model performance and helps identify potential issues like bias or overfitting.

2.6.12 Activation Function

An activation function is a mathematical operation applied to the output of each node in a neural network layer to introduce nonlinearity and enable complex mappings between inputs and outputs. Common activation functions include the sigmoid function, hyperbolic tangent (tanh), and rectified linear unit (ReLU). In waveguide analysis, activation functions play a crucial role in modeling the nonlinear behavior of optical signals and[116] capturing complex relationships between waveguide properties and input features. Choosing an appropriate activation function is essential for ensuring that the neural network can effectively learn and represent the underlying physics of waveguide dynamics.

2.6.13 Optimization Algorithm

Optimization algorithms are techniques used to adjust the model's parameters to minimize the loss function during the training of machine learning models. Examples of commonly used optimization algorithms include Stochastic Gradient Descent (SGD), Adam, and RMSprop[117]. Selecting an appropriate optimization algorithm is crucial for waveguide analysis to effectively train deep learning models and achieve desirable convergence properties. The effectiveness of different optimization algorithms can vary depending on the neural network's architecture and the complexity of the waveguide simulation task.

2.6.14 Hyperparameters

Hyperparameters are external parameters that determine the configuration and performance of machine learning models but are not learned during the training process. They include settings such as the learning rate, batch size, and the number of hidden layers in a neural network[118]. Proper tuning of hyperparameters is vital for optimizing deep learning models for waveguide analysis and ensuring they generalize effectively to new data. Techniques like grid search or random search are commonly employed to systematically explore the hyperparameter space and find the best configurations for training neural networks.

2.6.15 Loss Curves

Loss curves, also known as training curves or learning curves, are graphical representations of the loss function's value over the course of model training. Loss curves provide insight into the training process's dynamics, showing how the model's performance improves or deteriorates over successive iterations[119]. In waveguide analysis, monitoring loss curves helps assess the convergence of neural network models during training and diagnose issues such as overfitting or underfitting. Analyzing loss curves allows researchers to fine-tune training parameters and optimize the performance of deep learning models for predicting waveguide properties. The gradual reduction of loss with each, as shown in Figure2.12, is a big concern in neural network training[120].

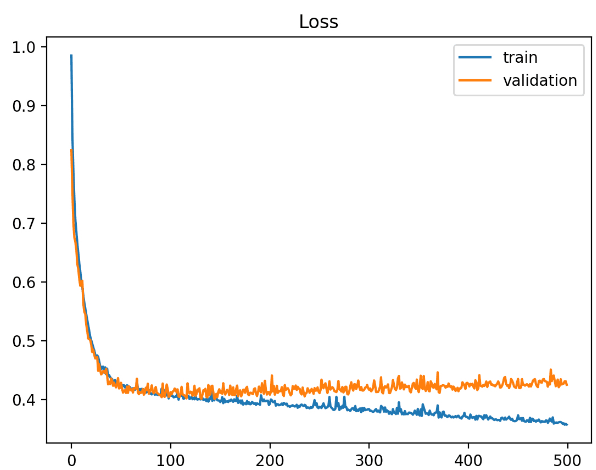


Figure 2.12: Example of Loss curve in the Neural Network Training Process

2.6.16 Overfitting and Underfitting

Overfitting and underfitting are common challenges encountered in machine learning models. Overfitting occurs when a model learns the training data too well, capturing noise or random fluctuations that are specific to the training set but do not generalize well to new, unseen data. This results in poor performance on test or validation data. On the other hand, underfitting happens when a model is too simplistic to capture the underlying structure of the data, leading to poor performance on both training and test sets. Balancing between these two extremes is crucial for building models that generalize well to unseen data. Techniques such as cross-validation, regularization, and adjusting model complexity help mitigate overfitting and underfitting, ensuring the model's robustness and effectiveness across various datasets and real-world scenarios[121].

2.6.17 Mean Squared Error (MSE)

Mean Squared Error (MSE) is a commonly used loss function for regression tasks in machine learning, measuring the average squared difference between the predicted and actual values of a target variable[122]. In waveguide analysis, MSE is often used as a metric for evaluating the accuracy of deep learning models in predicting optical properties such as effective refractive index or mode purity. Minimizing MSE during model training helps ensure that the neural network learns to produce accurate predictions of waveguide behavior and effectively captures the underlying relationships between input features and target outputs.

Mean Squared Error (MSE) identifies the deviation of points from their actual values:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i' - y_i)^2 \tag{2.6}$$

Where:

N is the total number of points,
y_i' is the predicted intensity at point$i$,
y_i is the actual intensity at point i.

2.6.18 TensorFlow and Keras

TensorFlow and Keras are popular open-source machine learning frameworks widely used for building and training deep learning models. TensorFlow provides a flexible and scalable platform for implementing various machine learning algorithms, including neural networks, while Keras offers a high-level API for building and configuring neural network models with minimal code. Together, TensorFlow and Keras simplify the process of designing, training, and deploying deep learning models for waveguide analysis, enabling researchers and engineers to explore complex waveguide structures and predict optical properties efficiently

2.7 Conclusion

This section has provided a comprehensive overview of the theoretical concepts essential for understanding optical waveguides, both linear and nonlinear. The discussion began with various types of waveguides, including planar, photonic crystal fibers, and advanced structures, establishing the foundational knowledge needed to grasp the waveguiding mechanisms. It then delved into the linear properties of waveguides, such as effective refractive index and dispersion, which are critical for guiding light efficiently. The exploration of nonlinear properties highlighted phenomena like self-phase modulation and supercontinuum generation, which are pivotal for advanced optical applications. Additionally, the numerical techniques and methods, including the Finite Element Method and Finite Difference Time Domain, were discussed, emphasizing their role in waveguide analysis. The section concluded with an introduction to the Generalized Nonlinear Schrödinger Equation (GNLSE) and the potential of deep learning approaches, such as Artificial Neural Networks and Convolutional Neural Networks, in enhancing waveguide research and design. This solid theoretical foundation sets the stage for further exploration and innovation in optical waveguiding technologies.

Chapter 3

Deep Learning in Photonic Crystal Fiber Analysis for OAM Transmission

This study centers on the efficacy of a Deep Learning (DL) model, based on Artificial Neural Network (ANN), in accurately predicting linear parameters crucial for Photonic Crystal Fiber (PCF) design. The selection of these parameters stems from their significance in determining the propagation characteristics of OAM modes in PCFs. By focusing on the predictive capabilities of the DL model, this research aims to shed light on its potential for facilitating efficient design optimization in high-capacity optical communication systems.

To forecast the propagation characteristics of OAM modes across a broad bandwidth, spanning the S-band and C+L+U band, a DL model based on ANN architecture has been proposed. Trained using data derived from COMSOL Multiphysics simulations, this model exhibits promising accuracy in predicting key linear parameters essential for PCF design. These parameters include effective refractive index, mode purity, effective area, nonlinear coefficient, and the presence of higher order radial modes. With a Mean Squared Error (MSE) of 0.02185, the DL model demonstrates noteworthy precision in predicting these parameters, underscoring its potential for optimizing PCF design.

Subsequent sections delve into the background of OAM, elucidating the fundamental structure of ring-core fibers and their performance matrices. The utilization of a Ring Core Photonic Crystal Fiber (RC-PCF) underscores its suitability for OAM transmission, particularly in accommodating multiple OAM modes effectively. By modulating parameters such as refractive indices, thickness, and outer radius of the ring region, the RC-PCF offers enhanced performance in OAM transmission. Additionally, the incorporation of air holes in the cladding further optimizes performance characteristics, highlighting the complexity involved in parameterizing RC-PCF for optimal OAM transmission.The proposed ANN model, tailored for predicting linear parameters relevant to PCF, is then introduced. Through meticulous training and tuning, the ANN model achieves optimal performance, with multiple hidden layers facilitating accurate predictions. Hyperparameter tuning, activation function selection, and optimizer optimization contribute to the ANN model's remarkable accuracy in predicting key linear parameters essential for OAM transmission in PCFs.

The subsequent sections detail the numerical results and computational performance of the proposed ANN model, showcasing its efficacy in predicting parameters such as mode purity, dispersion,effective refractive index, effective area, and nonlinearity. Through comprehensive analysis and comparison with simulated data, the ANN model's capability in accurately predicting these parameters is substantiated. Furthermore, the significant reduction in computational time afforded by the ANN model underscores its practical utility in facilitating rapid and efficient design optimization.

3.1 Methodology

3.1.1 Data Collection and Preparation

The methodology employed in this study revolves around the collection and preparation of data necessary for training the Deep Learning (DL) model based on Artificial Neural Network (ANN) for predicting the propagation characteristics of OAM modes in Photonic Crystal Fiber (PCF). The data collection process involved utilizing simulation results from COMSOL Multiphysics for various fiber designs and their corresponding optical characteristics. A total of 1925 data points were collected, covering a wide range of parameters including outer ring radius (R2), air fraction of cladding (f), pitch (Λ), ring thickness (t), wavelength (λ) and refractive indices of guided and ring areas N_{ring}, n_{clad}. The figure 3.1 represents the design used in numerical simulation for producing the dataset.

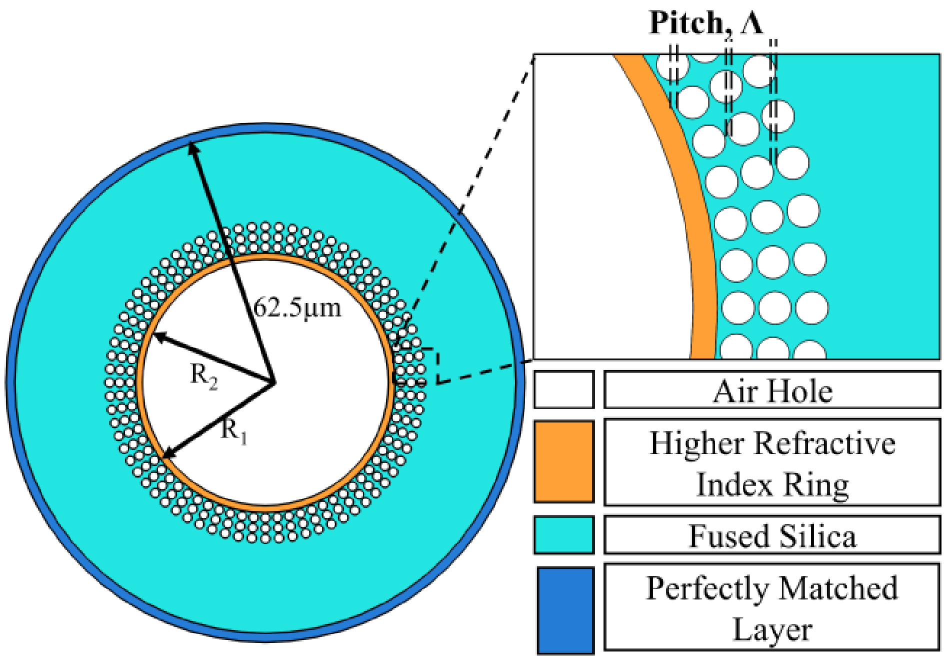


Figure 3.1: Cross-sectional View of the Ring Core PCF.

3.1.2 Data Splitting

To train, validate, and test the DL model effectively, the collected dataset was randomly split into three subsets: a training set, a validation set, and a test set. Specifically, 70% of the data points were allocated to the training set, 20% to the validation set, and the remaining 10% to the test set. This splitting strategy ensures that the model is trained on a diverse range of data while also having separate datasets for validation and evaluation.

3.1.3 Model Development

A range of Artificial Neural Network (ANN) models were designed and trained using the provided training dataset. Each ANN architecture began with an input layer comprising eight nodes, each corresponding to one of the eight input parameters. These were followed by one or more hidden layers, with the number of hidden layers varying from 2 to 4. The models utilized various activation functions, including Rectified Linear Unit (ReLU), Gaussian Error Linear Unit (GELU), and Exponential Linear Unit (ELU), to assess their impact on performance. Different optimization algorithms, such as ADAMAX and Stochastic Gradient Descent (SGD), were also explored to enhance the model's performance. The neural network architecture used in this study is illustrated in Figure 3.2, showcasing the typical layout of the ANN model employed for predicting linear parameters.

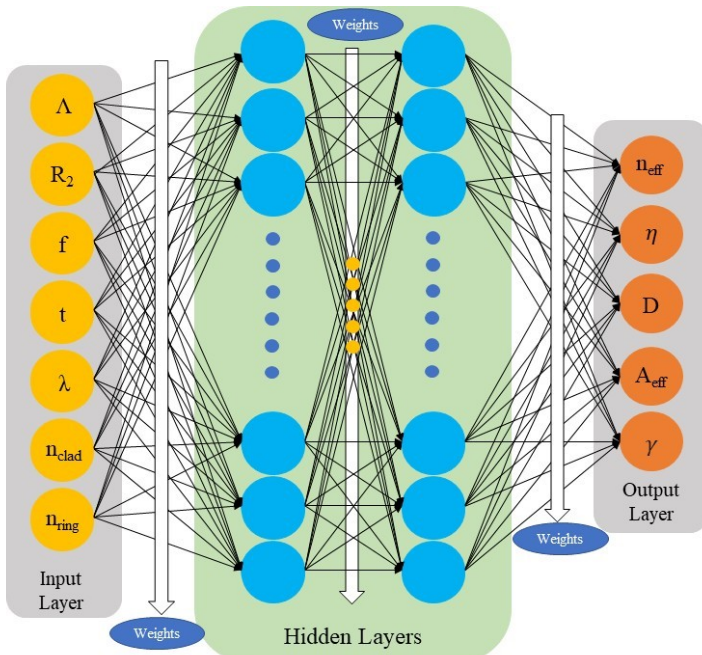


Figure 3.2: Artificial Neural Network Model Used for the Prediction of Linear Parameters.

3.1.4 Hyperparameter Tuning

Hyperparameter tuning involved adjusting several parameters to optimize the performance of the artificial neural network (ANN) models. Key hyperparameters, such as the number of hidden layers, the number of nodes within each layer, the choice of activation functions, and the selection of optimization algorithms, were systematically varied. This process required training multiple models with different combinations of these hyperparameters and evaluating their performance based on Mean Squared Error (MSE) metrics calculated on the validation set. The optimal configuration was determined by selecting the model that achieved the lowest MSE, thereby ensuring the highest predictive accuracy and overall performance of the ANN models.

3.1.5 Model Evaluation

The trained ANN models were evaluated using the validation set to assess their performance in predicting the propagation parameters of OAM modes in PCF. The MSE was used as the primary metric to measure the discrepancy between the predicted values and the ground truth values obtained from simulation results. Additionally, computational time required for training each model was recorded to evaluate the efficiency of the models.

3.1.6 Model Selection

Based on the evaluation results, the ANN model with the most optimal configuration, as determined by its performance on the validation set, was selected for further analysis. The selected model was then tested using the independent test dataset to assess its generalization ability and validate its predictive accuracy on unseen data

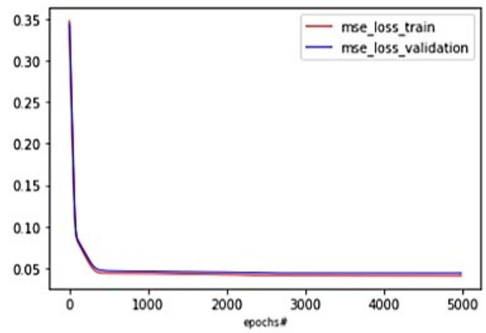
3.2 Results and Analysis

In this section, the results of the developed Deep Learning (DL) models for predicting the propagation parameters of Orbital Angular Momentum (OAM) modes in Photonic Crystal Fiber (PCF) along with the implications of these results are discussed.

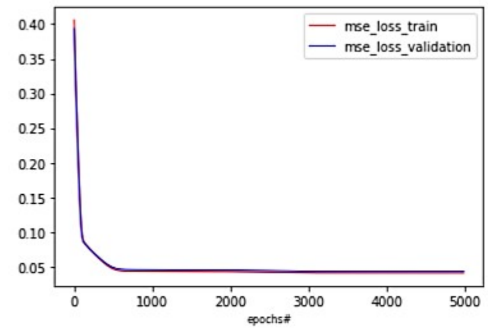
3.2.1 Model Performance

The performance of each DL model was evaluated based on various metrics, including Mean Squared Error (MSE), convergence of the loss curve, and computational time. Table 3.1 summarizes the hyperparameters and MSE for each model, with Model 2 demonstrating superior performance compared to others.

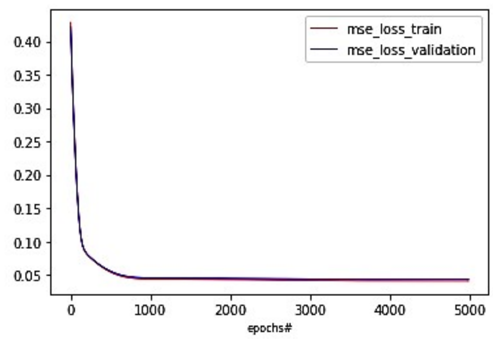
The loss curves, as depicted in Figure3.3, illustrate the convergence of the MSE over the training epochs for each model. After approximately 500 epochs, the loss curves for both the training and validation sets stabilize, indicating effective learning.



(a) Model 1



(b) Model 2



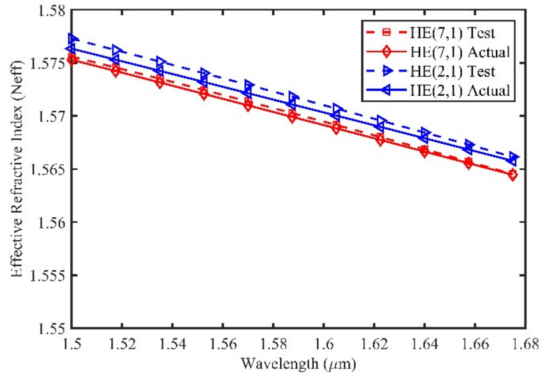
(c) Model 3

Figure 3.3: Training and Validation Loss Acquired for the Epochs from Different Models.

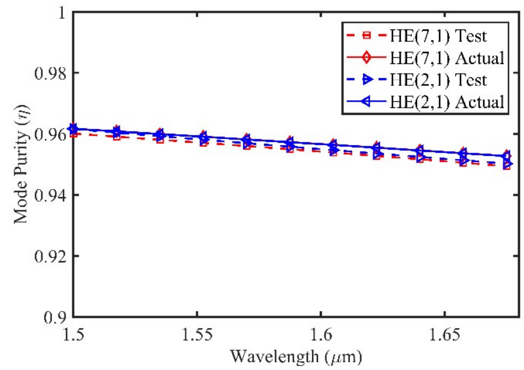
3.2.2 Parameter Prediction

The DL models were employed to predict the propagation parameters of OAM modes in PCF, including Dispersion (D) ,Mode Purity(η) ,Effective Refractive Index (n_{eff}), Non linearity(γ) Effective Area (A_{eff}).

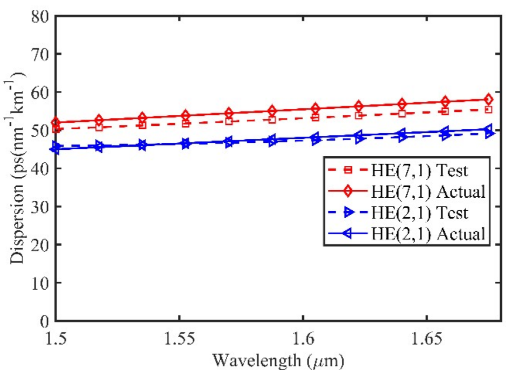
The Figure 3.4 depicts the performance of Model 1 in predicting various propagation parameters of OAM modes in PCF. Figure3.4c shows the actual and predicted dispersion for HE7 and HE2 modes, indicating the model's ability to capture dispersion characteristics accurately. Similarly, Figure3.4a presents the actual and predicted effective refractive index (n_{eff}) for both modes, demonstrating the model's good predictive accuracy. Figure 3.4b illustrates the actual and predicted mode purity (η) for HE7 and HE2 modes, with the model performing well in predicting mode purity. Figure3.4d showcases the actual and predicted effective area (A_{eff}) for both modes, highlighting the model's accurate predictions. Lastly, Figure3.4e showcases the actual and predicted nonlinearity (γ) for both modes. The performance of the



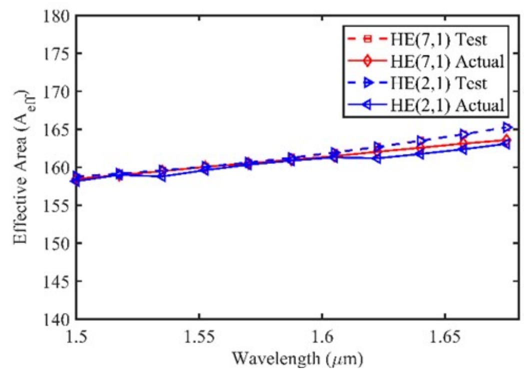
(a) Actual and Predicted (n_{eff}) of HE (7,1) and HE (2,1) Modes



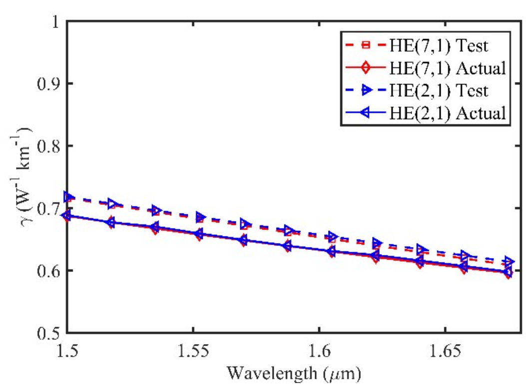
(b) Actual and Predicted η of HE (7,1) and HE (2,1) Modes



(c) Actual and Predicted (D) of HE (7,1) and HE (2,1) Modes



(d) Actual and Predicted A_{eff} of HE (7,1) and HE (2,1) Modes



(e) Actual and Predicted γ of HE (7,1) and HE (2,1) Modes

Figure 3.4: Performance of Optimized in Predicting Various Parameters

Table 3.1: Hyperparameter Tuning, Computational Time, and MSE of the Used ANN Models.

Model	Layers	Nodes in Layers	Activ. Function	Optimizer	Time (s)	MSE
1	4	90-90-90-100	GELU	ADAMAX	40.001	0.02258
2	3	90-90-90	GELU	ADAMAX	30.917	0.02185
3	2	90-90	GELU	ADAMAX	25.205	0.02208
4	4	90-90-90-100	ELU	ADAMAX	48.132	0.02197
5	4	90-90-90-100	ReLU	ADAMAX	42.886	0.24939

proposed model in predicting key optical parameters for higher-order modes in a ring-core photonic crystal fiber (RC-PCF) was evaluated against actual simulated data across various characteristics. These parameters include the Effective Refractive Index (n_{eff}), Mode Purity (η), Dispersion (D), Effective Area (A_{eff}), and Nonlinearity (γ). The model demonstrated a strong capability to predict these parameters with varying degrees of accuracy.

For the Effective Refractive Index (n_{eff}), the model's predictions closely matched the actual simulated data, particularly for HE (2,1) and HE (7,1) modes, showing that the model effectively predicts n_{eff} across different wavelengths. This is crucial as n_{eff} decreases at higher wavelengths due to mode leakage into the cladding.

When it comes to Mode Purity (η), the model also showed high accuracy, with only a minor deviation observed for the HE (7,1) mode at a wavelength of 1.675 µm. The negligible difference of 0.07% suggests that the model is reliable for predicting mode purity, an important factor for the transmission of orbital angular momentum (OAM) modes.

Dispersion(D) predictions were reasonably accurate, though less precise for the HE (7,1) mode. Despite this, the model successfully predicted nearly flat dispersion across the wavelength range of 1.5 to 1.675 µm, which is critical for minimizing signal distortion in fiber optic communication systems. The observed dispersion mismatch indicates that while the model performs well, its accuracy could be enhanced by incorporating more data into its training set.

Regarding Effective Area (A_{eff}) and Nonlinearity (γ), the model performed better at predicting A_{eff} than γ, particularly for the HE (7,1) and HE (2,1) modes. The discrepancy in predictions was larger at shorter wavelengths, with errors ranging from approximately 2.14% to 4.43%. This is because γ is more sensitive to changes in wavelength, making its accurate prediction more challenging. To improve precision in predicting γ, especially at lower wavelengths, the model would benefit from additional training data.

Overall, the model exhibits a strong predictive capability across most parameters, with some areas, such as dispersion and nonlinearity, showing room for improvement. By refining the training data, especially in regions where the model shows larger discrepancies, its predictive accuracy could be further enhanced.

3.2.3 Comparison of Computational Time

The DL models outperformed traditional simulation methods in terms of computational efficiency, as shown in Table 3.2. The training time required for the DL models was significantly lower compared to simulation in COMSOL Multiphysics, demonstrating the advantage of using DL for predicting PCF parameters. Normal meshing is good enough for simulating PCF behaviours in OAM mode transmission.It is found that computational time of ANN model is near 200x times less than traditional method. It affirms the high computational speed of neural networks.

Table 3.2: Comparison of Computational Time in COMSOL Multiphysics and ANN Model.

Tool	Time (s)
COMSOL Multiphysics (Normal Mesh)	7059
COMSOL Multiphysics (Finer Mesh)	8981
COMSOL Multiphysics (Extra Finer Mesh)	10159
ANN Model	30.19

3.3 Discussion

The results of this study reveal significant insights into the application of Deep Learning (DL) models, specifically Artificial Neural Networks (ANNs), for predicting key linear parameters in Photonic Crystal Fibers (PCFs) designed for Orbital Angular Momentum (OAM) mode transmission. The primary goal was to assess whether ANN models could accurately forecast parameters such as the effective refractive index, dispersion, mode purity, and effective area, which are critical for optimizing PCF designs in high-capacity optical communication systems.

The ANN models, particularly Model 2, demonstrated a strong predictive capability, as evidenced by its remarkably low Mean Absolute Error (MAE) and Mean Squared Error (MSE) values. These results indicate that the model can closely approximate the values obtained through traditional simulation methods. For instance, the MSE value close to 0.002 suggests that the ANN model can reliably predict these parameters with minimal deviation from the actual values. This level of accuracy is crucial for designing PCFs that require precise control over these properties to maintain high mode purity and low loss in OAM mode transmission.

The ability of the ANN models to predict these parameters with high accuracy also points to the potential for significantly reducing the computational time required for PCF design. Traditional methods, such as those relying on COMSOL Multiphysics simulations, are computationally intensive and time-consuming. In contrast, the ANN models developed in this study can provide rapid predictions once trained, making them a highly efficient tool for iterative design processes.

Moreover, the study highlights the versatility of ANNs in handling complex, multi-dimensional data typical of photonic simulations. The models effectively captured the non-linear relationships between the geometric parameters of the PCF and its optical properties, which is a challenging task for conventional analytical methods. This suggests that ANNs could be further explored and possibly expanded to predict other non-linear optical phenomena in photonic devices.

However, while the results are promising, it is important to consider the limitations of the study. The accuracy of the ANN models is inherently dependent on the quality and quantity of the training data. The dataset used in this study was generated through simulations, and while comprehensive, it is still limited by the parameters and conditions considered. Real-world variations and manufacturing imperfections could introduce discrepancies not accounted for in the current model. Future work should focus on enhancing the robustness of these models by incorporating more diverse training data, potentially including experimental data, to improve their generalizability.

3.4 Conclusion

In this study, the efficacy of Deep Learning (DL), specifically Artificial Neural Networks (ANN), has been demonstrated in accurately predicting the linear parameters crucial for the design and optimization of Photonic Crystal Fibers (PCFs) used in Orbital Angular Momentum (OAM) mode transmission. The ANN models were trained on a dataset generated through COMSOL Multiphysics simulations. Model 2, in particular, exhibited superior performance, achieving a mean squared error close to 0.02. This high level of accuracy underscores the ANN's capability to forecast critical parameters like dispersion, effective refractive index, mode purity, and effective area with remarkable precision. Moreover, the study highlights the significant reduction in computational time, with the ANN models delivering results in a fraction of the time required by traditional simulation methods, thus paving the way for more efficient and optimized PCF designs for high-capacity optical communication systems.

Chapter 4

Supercontinuum Prediction in Chalcogenide Waveguides via ANN

This section presents a novel approach employing an Artificial Neural Network (ANN) based Deep Learning (DL) model to predict the dynamics of supercontinuum generation in chalcogenide planar waveguides. The research is motivated by the need for efficient simulation techniques to anticipate the impact of various parameters, such as pump wavelength, pump power, pulse width, waveguide width, and thickness, on the generation of supercontinua. By transitioning from linear parameter considerations to the intricate nonlinear phenomena involved in supercontinuum generation, this study aims to offer a robust alternative to conventional simulation methods, thereby streamlining simulation efforts and reducing computational resources.

The initial section outlines the methodology employed for data acquisition, which involves meticulously designing waveguides using advanced software tools like COMSOL Multiphysics. Through parametric sweeps and systematic variations in dimensions and parameters, comprehensive datasets are generated to capture the intricate interplay between various input variables and the resulting supercontinuum spectra. Subsequently, a Generalized Nonlinear Schrödinger Equation (GNLSE) is utilized to simulate the dynamics of supercontinuum generation, incorporating nonlinear effects such as Kerr-type and Raman responses. This comprehensive dataset serves as the base for training and validating the ANN model.

The ANN model architecture is tailored to accommodate the complexity of the problem, with multiple hidden layers and optimized hyperparameters facilitating accurate predictions of spectral power across a broad wavelength range. By training the model on a diverse dataset encompassing different waveguide structures and parameter variations, the ANN model demonstrates promising performance in predicting the spectral power output corresponding to varying input parameters.

The subsequent sections delve into the numerical findings and analysis of the proposed ANN model. Through comprehensive comparisons with simulated data and evaluation metrics such as Mean Squared Error (MSE), the efficacy of the ANN model in capturing the nonlinear dynamics of supercontinuum generation is substantiated. Additionally, insights into the impact of individual input parameters on the resulting supercontinuum spectra are elucidated, providing valuable guidance for optimizing waveguide design and parameter selection.

4.1 Methodology

4.1.1 Data Acquisition

Data acquisition involves the careful collection and organization of extensive datasets, which are thoroughly analyzed to support the development of an artificial neural network (ANN) model. The prediction of the power spectrum output in relation to the waveguide's wavelength is carried out in two main stages that work together to achieve the desired results. In the first stage, advanced software tools such as COMSOL Multiphysics are used to design a waveguide incorporating various parameters. Figure 4.1 shows a channel waveguide with $Ge_{11.5}As_{24}Se_{64.5}$ as the core material, and air and MgF_2 as cladding materials for simulations and data collection. The core dimensions are systematically varied in height (from 0.8 μm to 0.83 μm) and width (from 3.5 μm to 4 μm) with a 100 nm step size. A parametric sweep technique is utilized to vary the wavelength from 1 μm to 5 μm, also in 100 nm increments, to extract data on the waveguide's effective refractive index. In the subsequent stage, the collected datasets are used as critical inputs for the Generalized Nonlinear Schrödinger Equation (GNLSE) to create a detailed dataset of supercontinua. Detailed analysis is required at this point as the input parameters— pump power,pump wavelength and pulse width (measured at full width at half maximum, FWHM)—are systematically varied. The pump wavelength is adjusted between 2500 nm and 3400 nm in 100 nm increments, covering a wide range.

The pump power varies from 750 W to 2500 W with a 250 W step size, and the pulse width ranges from 50 fs to 175 fs. This comprehensive dataset is designed to meet the specific requirements of the ANN model. For all the simulation, the device length was kept constant at 5 mm.

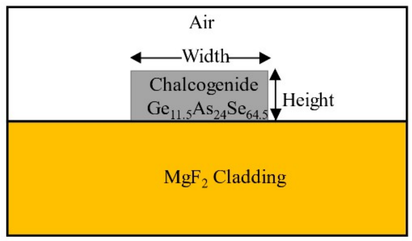


Figure 4.1: 2D Cross-sectional Representation of a Chalcogenide Planar Waveguide (CPW) Used for Simulated Data Generation.

To investigate supercontinuum (SC) generation in the mid-infrared (MIR) region, simulations were conducted using the Generalized Nonlinear Schrödinger Equation (GNLSE) [43] for a steadily fluctuating envelope of the pulse:

$$\frac{\partial A}{\partial z} + \frac{\alpha}{2}A - \sum_{k\geq 2} \frac{i^{k+1}}{k!}\beta_k \left(\frac{\partial^k A}{\partial T^k}\right) = i\gamma \left(1 + i\tau_{\text{schock}}\frac{\partial}{\partial T}\right) \left(A(z,t) \int_{-\infty}^{+\infty} R(T')|A(z,T-T')|^2 \, dT'\right) \tag{4.1}$$

where A denotes the amplitude of the electrical field, α refers to the linear propagation loss of the waveguide, β_k $(k \geq 2)$ is the kth order dispersion parameter, and $T = t - z/v_g$ refers to the deferred time frame moving with the group velocity $v_g = 1/(\beta_1(\omega_o))$ at the pump frequency ω_o. The nonlinear coefficient is given by $\gamma = (n_2\omega_o)/(cA_{\text{eff}}(\omega_o))$, where n_2 is the nonlinear refractive index, c is the speed of light in vacuum, $A_{\text{eff}}(\omega_o)$ is the effective area of the mode at the pump frequency ω_o. The material's response function incorporates both the electronic response (Kerr type) and the delayed Raman response:

$$R(t) = (1 - f_R)\delta(t) + f_R h_R(t) \tag{4.2}$$

where the fractional contribution of the delayed Raman response is represented by f_R, and $h_R(t)$ is given by:

$$h_R(t) = \frac{\tau_1^2 + \tau_2^2}{\tau_1^2 \tau_2^2} \left(e^{-t/\tau_2}\right) \sin\left(\frac{t}{\tau_2}\right) \tag{4.3}$$

Here, τ_1 and τ_2 are the characteristic time constants associated with the Raman scattering process, which define the decay and oscillatory nature of the response.

Split-Step Fourier Method (SSFM) is used for solving Generalized Nonlinear Schrödinger Equation (GNLSE) for simulating supercontinuum generation in optical fibers and waveguides. This method efficiently handles the complex interplay of linear and nonlinear effects—such as dispersion, self-phase modulation, and Raman scattering—by splitting the propagation distance into small segments. In each segment, the SSFM alternates between solving the linear effects in the frequency domain using the Fast Fourier Transform (FFT) and the nonlinear effects in the time domain.

SSFM is particularly effective for modeling supercontinuum generation because it captures the rapid spectral broadening that occurs during pulse propagation in nonlinear media. The method's accuracy stems from its ability to compute both linear and nonlinear interactions iteratively over short distances, accommodating complex effects like higher-order dispersion and stimulated Raman scattering. However, the choice of step size is crucial: smaller steps improve accuracy at the cost of computational resources, while larger steps may reduce accuracy.

For simulating the GNLSE, there are some required parameters such as propagation loss, raman paramters, kerr non linearity etc. Table 4.1 provides an overview of the remaining simulation parameters that were used.

Table 4.1: Parameters Employed in Numerical Simulation of Chalcogenide Planar Waveguide

Parameters	**Value**
Linear loss (α) [123]	0.6 dB/cm
f_R [124]	0.013 Null
τ_1 [124]	21.3 fs
τ_2 [124]	195 fs

The nonlinear dynamics of the waveguide are shown in Figure 4.2, which depicts the spatial and temporal evolution of the pulse. The coherence plot reveals that the waveguide exhibits partial coherence. For each precise variation, the corresponding numerical values of spectral power, measured in decibels (dB), were carefully organized, resulting in a well-structured and comprehensive dataset.

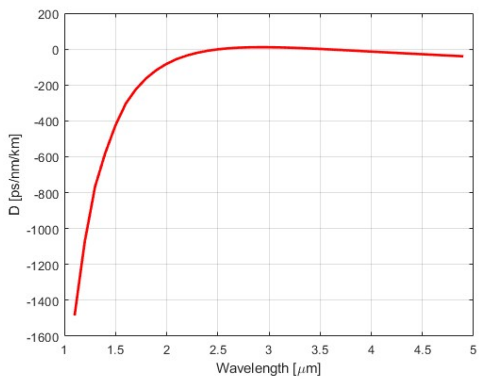
In this study, a comprehensive analysis was conducted by generating a dataset with 8192 data points using a fast Fourier transform applied to the Generalized Nonlinear Schrödinger Equation (GNLSE) for a single waveguide configuration. Subsequently, 228 distinct waveguide structures were used to assemble the complete dataset. To optimize the performance of the artificial neural network (ANN) model, data augmentation techniques were applied, expanding the dataset to 1080 instances. This dataset includes five input parameters, and the model can also predict data along the length of the waveguide, providing an in-depth understanding of its behavior. As is customary, the ANN model was divided into training, validation, and testing sets. The testing dataset was meticulously organized to isolate the effects of individual variables on the output spectrum while keeping the other four variables constant.

The input parameters for the waveguide's structural configuration included width, height, pump wavelength, power, and pulse width. To illustrate the impact of these five variables, 15 models were selected from the 1080 simulated data models to showcase their practical relevance. The remaining data was used for training and validation, distributed randomly in a 9:1 ratio to facilitate effective training.

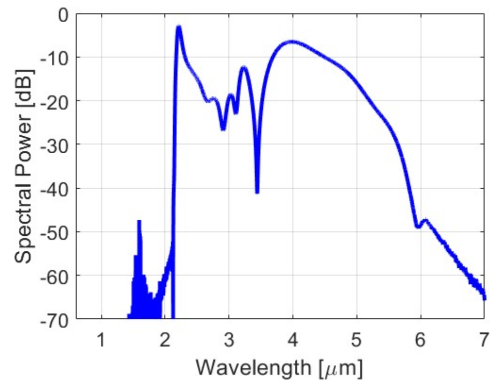
4.1.2 Artificial Neural Network (ANN) Model

Several artificial neural network (ANN) models were developed and trained with the dataset to predict the spectral power of supercontinuum generation in chalcogenide planar waveguides (CPWs). The ANN architecture typically comprises multiple hidden layers, with the number of layers ranging from 2 to 5. Each input layer corresponds to specific parameters such as waveguide dimensions, pump characteristics, and pulse width, while the output layer generates predictions for the spectral power distribution.

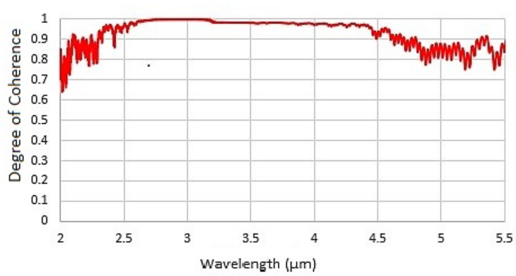
The ANN model undergoes hyperparameter tuning to enhance its performance. Hyperparameters such as the number of hidden layers, nodes per layer, activation functions (e.g., Tanh), and optimizers (e.g., Adam) are adjusted to improve model



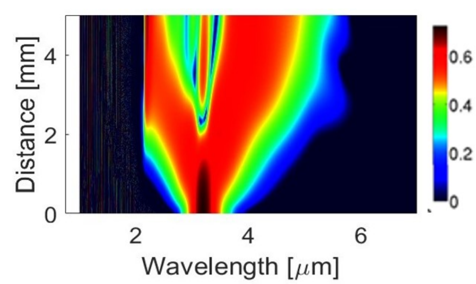
(a) Dispersion



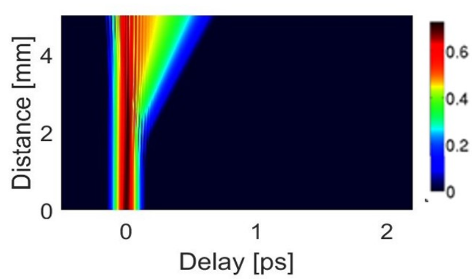
(b) SC Spectra



(c) Degree of Coherence



(d) Spatial Density Plot



(e) Temporal Density Plot

Figure 4.2: Dispersion (a), SC spectra (b), Degree of Coherence (c), Spatial Density Plot (d), Temporal Density Plot (e) of Chalcogenide Planar Waveguide for $H = 0.8\mu$m, $W = 3.8\mu$m, PWL=3200 nm, PPW=750 W, FWHM=50 fs.

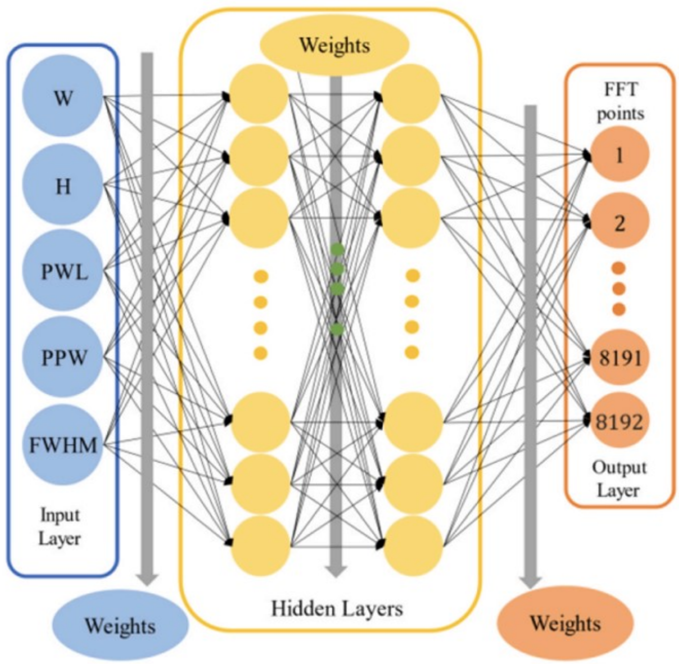


Figure 4.3: Architecture of the Artificial Neural Network (ANN) Model Developed for Prediction of Spectral Power of Supercontinuum in Chalcogenide Planar Waveguides.

accuracy. The trained ANN model can predict spectral power values across a range of input parameters.

Table 4.2: Summary of Tuned Parameters Employed in the ANN Model

Model Summary	**Details**
No. of hidden layers	5
Nodes in hidden layer 1	120
Nodes in hidden layer 2	120
Nodes in hidden layer 3	120
Nodes in hidden layer 4	120
Nodes in hidden layer 5	512
Loss	Mean Squared Error
Optimizer	Adam
Activation function	Tanh
Learning rate	Exponential Decay
	(initial rate = 0.005, decay steps = 1000, decay rate = 0.8)

The architecture of the ANN model and the number of nodes in each layer depend on the complexity and size of the dataset. After multiple fine-tuning iterations, the number of layers and neurons were finalized. With 120 neurons in the first four layers and 512 neurons in the fifth layer, the model has a high number of connections and weights adjusted during backpropagation. Although linear regression models face problems, ANN can learn nonlinear patterns in large datasets. Given the nonlinear dynamics involved, the Tanh activation function was chosen as it performed better than other functions in trials. Table 4.2 provides a representation of the used hyperparameters in the ANN model after proper tuning. The Min-Max scaling method was employed for preprocessing the spectra before applying the Tanh activation function. Google Colab, which provides dedicated GPU support, was used for training the model, with Python as the programming language.

4.1.3 Model Evaluation

The ANN model's performance is assessed using Mean Squared Error (MSE) to measure how closely the predicted spectral power values match the actual values. The evaluation process incorporates training, validation, and testing datasets to determine the model's accuracy and its ability to generalize. The model undergoes iterative adjustments to enhance performance, focusing on optimizing the number of epochs and learning rates.

Given that the study relies on numerical values from the GNLSE, MSE serves as the principal metric for evaluating model performance. It effectively measures the deviation between predicted and true values.

$$\mathrm{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i' - y_i)^2 \tag{4.4}$$

In this equation, N represents the total number of wavelength bins, y_i' is the predicted intensity at wavelength i, and y_i is the actual intensity at the same wavelength.

4.2 Results and Analysis

The ANN model was designed to produce outputs corresponding to 8192 points, consistent with the 2^{13} data points derived from the GNLSE after performing FFT. Each node in the output layer represents a specific power value in decibels (dB), and these values are plotted against wavelengths to illustrate the model's prediction of nonlinear dynamics. The ANN model effectively identifies patterns in large datasets, validating its design. While the model avoids abrupt spectrum variations, it accurately mirrors changes in spectral width and the envelope of spectral power with respect to wavelength.

Figure 4.4 presents the loss curve, which tracks the model's learning progress over iterations. Training was conducted over 500 epochs, during which a significant reduction in variance in the supercontinuum spectra was observed compared to fewer epochs. A random search method was used to finalize the model's configuration. This model was compared against alternatives with 3 and 5 layers and configurations using GRU cells, demonstrating superior performance.

The initial model configuration included four layers with 120 neurons each and a final layer with 8192 neurons for predicting FFT points. Introducing an additional intermediate layer with 29 neurons after the fourth layer, along with employing a decaying learning rate rather than a fixed rate, led to notable improvements. These modifications resulted in a lower MSE with fewer epochs, thus shortening the training duration.

The evolution of the spectrum across the entire range can impact pulse temporal

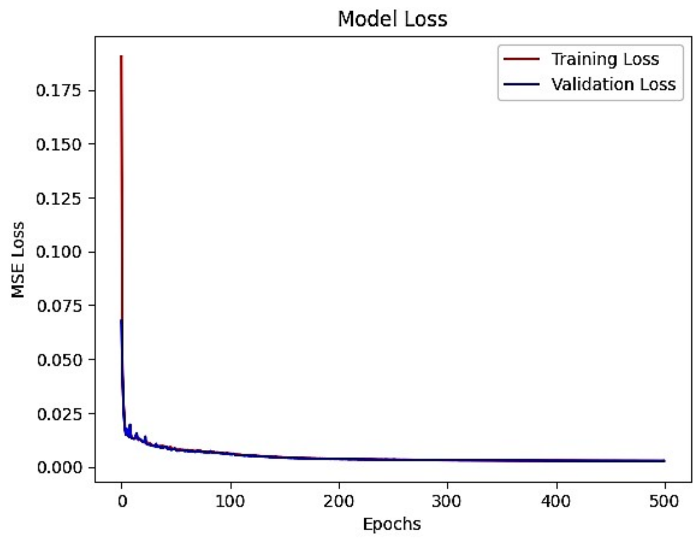


Figure 4.4: Loss Curve Tracking Training and Validation Accuracy Over Epochs for the ANN Model.

variation. The MSE loss, used as the evaluation metric, is summarized for different epoch counts in Table 4.3.

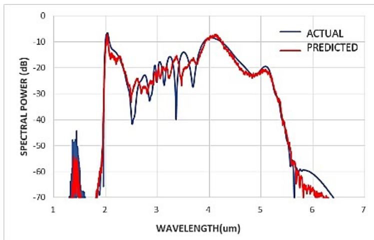
Table 4.3: Summary of Test Set Results and Training Time for the ANN Model.

Epochs	**MSE Loss**	**Time (s)**
100	0.0052	30
300	0.0033	90
500	0.0032	152
1000	0.0032	288

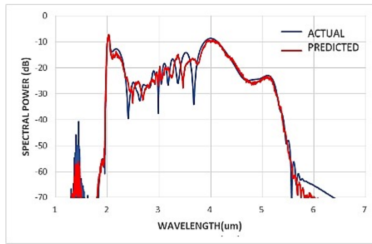
The following section discusses the numerical analysis results from the ANN model, focusing on how variables such as pump power, pump wavelength, pulse duration, width, and height of the rectangular channel waveguide influence supercontinuum generation.

4.2.1 Effect of Pulse Width (FWHM) Variation

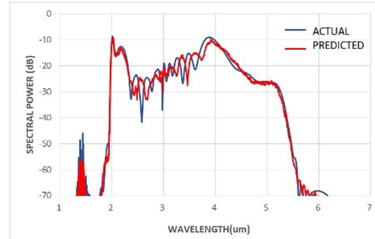
The full width at half maximum (FWHM) of the input pulse is a dynamic variable that modulates pulse width, affecting the resulting output. Figure 4.5 highlights the alignment between observed trends and spectral power output, demonstrating the ANN model's accuracy in capturing underlying phenomena.



(a) H = 0.8 µm, W = 3.8 µm, PWL = 3000 nm, PPW = 2000 W, Pulse Width, FWHM = 100 fs



(b) H = 0.8 µm, W = 3.8 µm, PWL = 3000 nm, PPW = 2000 W, Pulse Width, FWHM = 125 fs

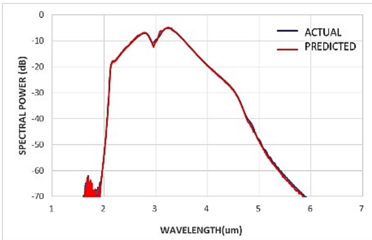


(c) H = 0.8 µm, W = 3.8 µm, PWL = 3000 nm, PPW = 2000 W, Pulse Width, FWHM = 150 fs

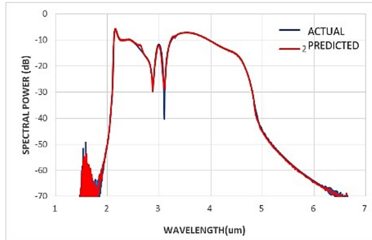
Figure 4.5: Validation of Spectral Power Prediction of the ANN Model for Variation of Pulse Width at Half of the Maximum Height(FWHM).

Furthermore, it is crucial to emphasize that the MATLAB code used for performing the Fast Fourier Transform (FFT) on the Generalized Nonlinear Schrödinger Equation (GNLSE) is meticulously designed to segment the waveguide into 240 spatial slices. Each slice represents a specific portion of the waveguide, allowing for a detailed analysis of how spectral power evolves along its length. This study focuses on the distribution of spectral power along the waveguide, particularly examining the 240th segment, which reflects the spectral power condition at the waveguide's end. This segment serves as a key reference for assessing the accuracy and predictive performance of the proposed Artificial Neural Network (ANN) model.

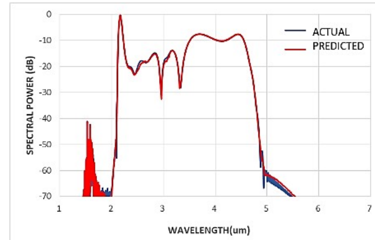
The ability of the ANN model to accurately forecast the spectral power for the final segment demonstrates its effectiveness in predicting the spectral evolution throughout the entire length of the waveguide. Remarkably, the model is able to predict all 239 preceding spectra based on the GNLSE data provided. This extensive coverage showcases the ANN model's proficiency in capturing and replicating the complex dynamics of supercontinuum generation along the waveguide. Figure 4.6 visually illustrates the model's performance in predicting spectral evolution, highlighting the smooth transition and development of spectral power from the input to the output end of the waveguide, thereby validating the model's accuracy and its ability to capture the underlying spectral phenomena.



(a) H = 0.8 µm, W = 3.7 µm, PWL = 3000 nm, PPW = 1000 W, FWHM = 50 fs Section no. 56



(b) H = 0.8 µm, W = 3.7 µm, PWL = 3000 nm, PPW = 1000 W, FWHM = 50 fs Section no. 96

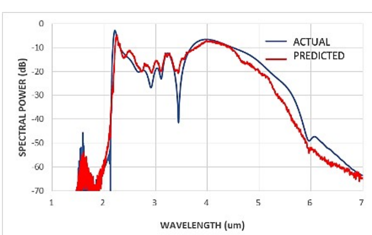


(c) H = 0.8 µm, W = 3.7 µm, PWL = 3000 nm, PPW = 1000 W, FWHM = 50 fs Section no. 56

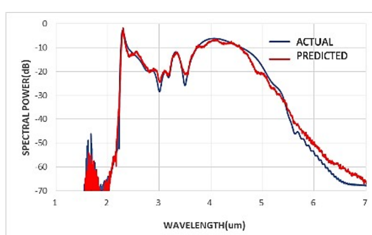
Figure 4.6: Validation of Spectral Power Prediction of the ANN Model for Variation of Pump Wavelength.

4.2.2 Effect of Pump Wavelength (PWL) Variation

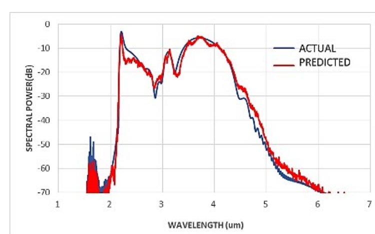
The selection of the pump wavelength is crucial in defining the spectral characteristics of supercontinuum generation. Figure 4.7 illustrates the comparison between spectral power results from numerical simulations and predictions made by the deep learning model, demonstrating how changes in the pump wavelength influence the final spectrum. Shorter pump wavelengths generally produce broader spectra, while longer wavelengths tend to result in narrower spectra. The ANN model effectively captures these trends, reflecting its ability to model the complex relationship between pump wavelength and spectral dynamics. Figure 4.7 visually compares the spectral power from both numerical simulations and the ANN model, confirming the model's effectiveness in replicating observed phenomena.



(a) Pump Wavelength, PWL = 3200 nm



(b) Pump Wavelength, PWL = 3300 nm

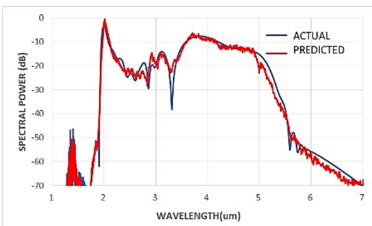


(c) Pump Wavelength, PWL = 3400 nm

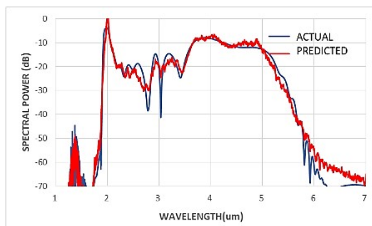
Figure 4.7: Validation of Spectral Power Predictions of the ANN Model for Varying Pump Wavelengths.

4.2.3 Effect of Pump Power (PPW) Variation

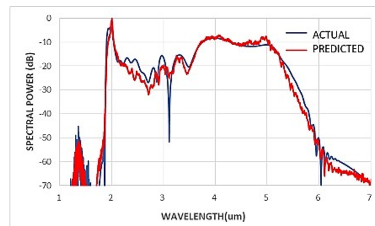
Pump power is another significant factor affecting supercontinuum generation. Figure 4.8 presents a comparison between spectral power values from numerical simulations and those predicted by the deep learning model under different pump power conditions. Changes in pump power can shift and modify the positions of spectral peaks, thus influencing the overall spectrum's shape and width. The ANN model accurately captures these variations, demonstrating its capability to model the subtle dynamics of the system effectively.



(a) Pump Power, PPW = 1000 W



(b) Pump Power, PPW = 1500 W



(c) Pump Power, PPW = 2000 W

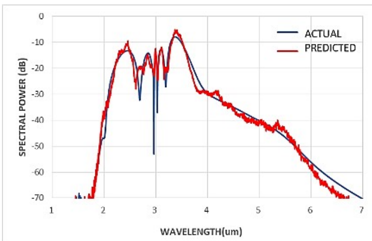
Figure 4.8: Validation of Spectral Power Predictions of the ANN Model for Varying Pump Power.

4.2.4 Effect of Waveguide Width (W) and Height (H) Variation

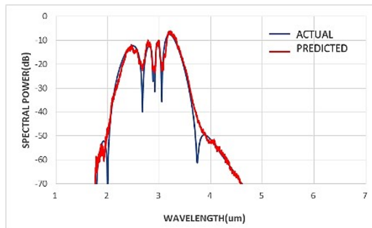
Variations in the waveguide width have a notable impact on the resulting spectrum. Figure 4.9 compares the supercontinuum spectra obtained from numerical simulations with those predicted by the deep learning model for different waveguide widths.

Figure 4.10 shows the comparison of supercontinuum spectra for different waveguide heights. The model exhibits near-zero error in predicting the -40 dB bandwidth of both predicted and actual spectra. However, there is some discrepancy in capturing the nonlinear aspects of the curve, as indicated by the mean squared error values. This reduced accuracy in height variations might be due to the heightened sensitivity of the waveguide to changes in height and the higher volume of data associated with a height of 0.8 µm compared to other variations. This discrepancy may lead to less accurate predictions for other height variations.

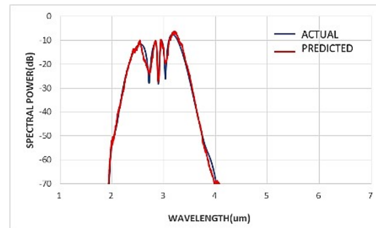
This completes the numerical analysis section of the task, providing insights into



(a) Waveguide Width, W = 3.8 µm

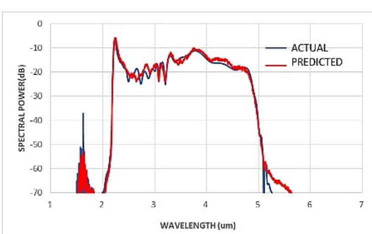


(b) Waveguide Width, W = 3.9 µm

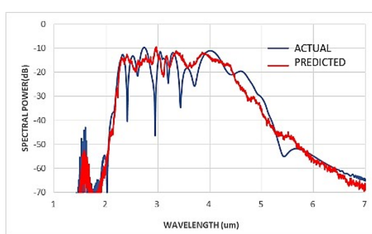


(c) Waveguide Width, W = 4 µm

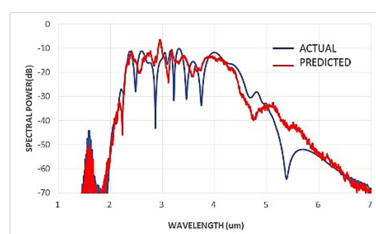
Figure 4.9: Validation of Spectral Power Predictions of the ANN Model for Varying Waveguide Widths.



(a) Waveguide Height, H = 0.83 µm



(b) Waveguide Height, H = 0.84 µm



(c) Waveguide Height, H = 0.85 µm

Figure 4.10: Validation of Spectral Power Predictions of the ANN Model for Varying Waveguide Heights.

the influence of various parameters on supercontinuum generation using the ANN model.

4.2.5 Time Constraints

To solve Maxwell's field equations, the waveguide is broken down into tiny geometric elements using the meshing capabilities in COMSOL Multiphysics. The complexity of the mesh and the intricate design of the waveguide's cross-section significantly affect the simulation time in each modeling tool. Additionally, performing a parametric sweep greatly extends the duration of simulations. Identifying the optimal configuration often requires running multiple simulations to thoroughly investigate different design options.

After achieving the desired Group Velocity Dispersion (GVD) through COMSOL simulations, the next step involves applying the Generalized Nonlinear Schrödinger Equation (GNLSE) using numerical software like MATLAB. MATLAB performs the Fast Fourier Transform (FFT) of the GNLSE to produce the required spectral power. This multi-step process is computationally intensive and time-consuming, particularly for complex waveguide structures and extensive parametric ranges.

Training an Artificial Neural Network (ANN) model also depends on several factors, including the size of the training dataset, the architecture of the network (such as the number of hidden layers and nodes), dropout rates, and the optimization techniques used. The training time is further influenced by the convergence of the process and the number of iterations needed to adjust the network's weights and biases. Despite these factors, ANN models provide significant computational efficiency advantages compared to traditional simulation methods.

Once trained, the ANN model can quickly predict the nonlinear dynamics of supercontinua. As shown in Table 4.4, using the ANN model can result in a dramatic reduction in computation time—approximately 2000 times faster than traditional methods for the same data volume. This improved speed accelerates the design and optimization process, ultimately enhancing productivity in photonics and nonlinear optics research.

Table 4.4: Total Simulation Time in COMSOL and MATLAB Compared to the ANN Model

Simulation Method	**Time Required (Seconds)**
COMSOL Multiphysics (Normal Mesh)	2203×1080 s
COMSOL Multiphysics (Finer Mesh)	3700×1080 s
COMSOL Multiphysics (Extra Fine Mesh)	4231×1080 s
GNLSE on MATLAB	240×1080 s
ANN Model (for training 1080 data)	472 s

4.3 Discussion

The deep learning model based on Artificial Neural Networks (ANNs) demonstrates strong performance in predicting the nonlinear nature of supercontinuum generation in chalcogenide channel waveguides. By varying design parameters and pump sources, the model effectively forecasts changes in spectral power. However, it is important to address that the ANN model does not completely capture the discontinuities observed in the spectral power dynamics from numerical simulations using the Generalized Nonlinear Schrödinger Equation (GNLSE).

These discontinuities, present in GNLSE simulations, are likely due to the unique characteristics of individual pulses. In practical scenarios, these discontinuities tend to average out with an increased number of pulses, making the ANN model's predictions more consistent with observed trends. Although the ANN model may not fully replicate abrupt spectral variations, especially those with nonlinear changes, improvements were achieved by using a variable learning rate and the nonlinear activation function tanh.

The value of the proposed deep learning model lies in offering a computationally efficient alternative to traditional simulation methods. By leveraging ANN technology, researchers can significantly reduce simulation time and resource requirements, streamlining the design process and facilitating quicker iterations and optimizations, thereby enhancing the development cycle in photonics and nonlinear optics.

4.4 Conclusion

The ANN-based deep learning model presented in this task represents a significant advancement in the field of supercontinuum generation modeling. In this task of prediction spectra in a chalcogenide planar waveguide mean squared error of 0.0032 was achieved which suggests ANN model's ability to accurately predict nonlinear dynamics while offering computational efficiency.This makes it a valuable tool for researchers in photonics and nonlinear optics. Further research exploring different waveguide structures and photonic crystal fibers (PCFs) using similar DL models such as convolutional neural networks (CNNs) and conventional machine learning approaches holds promise. Additionally, there is potential for investigating the effectiveness of DL models in predicting temporal density and spatial density analyses, expanding the scope of applications in this domain

Chapter 5

Modeling Supercontinuum Using Multi-Material Data with CFC Networks

This task investigates the use of the Closed-Form Continuous-Time Neural Networks (CfC) model for predicting supercontinuum spectra within a comprehensive dataset that includes various core materials in planar waveguides. The research addresses the critical need for efficient and accurate predictive modeling techniques in photonics, particularly in the realm of supercontinuum generation. Supercontinuum spectra, characterized by their broad bandwidth and smooth spectral profiles, play a crucial role in numerous applications such as optical communications, biomedical imaging, and spectroscopy. Understanding and predicting the spectral behavior of supercontinua across different materials and operating conditions are paramount for optimizing device performance and advancing optical device design.

Traditionally, predicting supercontinuum spectra involves complex numerical simulations and solving the Generalized Nonlinear Schrödinger Equation (GNLSE), which can be computationally intensive and time-consuming. To overcome these challenges, this study leverages the CfC model, a novel approach that offers closed-form solutions to spatiotemporal problems, thereby enhancing computational efficiency while maintaining accuracy.

The methodology employed involves the generation of comprehensive datasets through numerical simulations performed with COMSOL Multiphysics, capturing the intricate interplay between material properties, waveguide dimensions, and spectral output. These datasets serve as the basis for training and validating the CfC model, which is designed to accurately predict supercontinuum spectra under varying conditions.

The subsequent sections of this thesis delve into the detailed methodology, neural network architecture, and numerical findings, showcasing the efficacy of the CfC model in accurately predicting supercontinuum spectra across multiple materials. Through comprehensive comparisons with simulated data and evaluation metrics such as Mean Squared Error (MSE), the robustness and scalability of the CfC model are demonstrated. Insights gained from this study not only contribute to the advancement of predictive modeling in photonics but also hold promise for optimizing supercontinuum generation processes and enhancing the design of optical devices for diverse applications.

5.1 Methodology

To achieve precise predictions of supercontinuum spectra, constructing a robust dataset is essential. The process begins with the use of planar waveguides to calculate the effective refractive index (n_{eff}) through numerical simulations using COMSOL Multiphysics. This fundamental data serves as a basis for subsequent analyses, including the evaluation of nonlinear spectral properties and dispersion characteristics. The n_{eff} values obtained from these simulations are then utilized in MATLAB to solve the General Nonlinear Schrödinger Equation (GNLSE). This iterative procedure is repeated for each of the four materials investigated—Si_3N_4, $LiNbO_3$, SiC, and Ta_2O_5—to generate a comprehensive dataset comprising 8192 Fast Fourier Transform (FFT) points per material, each representing distinct spectral information.

5.1.1 Data Acquisition

5.1.2 Data Acquisition

The data acquisition process is fundamental for assembling and organizing the datasets necessary for training the CfC neural network model. In this study, datasets are meticulously generated for four distinct core materials within a planar waveguide: Si_3N_4, $LiNbO_3$, SiC, and Ta_2O_5, while the cladding material is kept uniform across all scenarios. Initially, the dataset focused on three core materials, with the fourth material, Ta_2O_5, introduced later to assess the model's scalability and generalization capabilities.The core dimensions are standardized with a width of 6 µm and a height of 0.8 µm to ensure consistency. The wavelength of the incident light is varied from 1 to 5 µm in 50 nm increments, which guarantees a comprehensive coverage of the wavelength spectrum. Furthermore, the pulse's full width at half maximum (FWHM) is varied between 50 fs and 100 fs to capture a range of temporal pulse profiles. The input pump wavelength is adjusted from 1550 nm to 1750 nm, covering a broad range to simulate different operational conditions. Additionally, the pump power is systematically varied from 4200 W to 5500 W to encompass a wide spectrum of power levels.Throughout this data acquisition process, the device length is consistently set at 10 mm to maintain uniformity across the datasets. This thorough and methodical approach ensures the creation of extensive and detailed datasets, which are crucial for the effective training and validation of the CfC neural network model. By systematically varying these parameters and maintaining consistency where necessary, the datasets provide a robust foundation for developing and refining the predictive capabilities of the CfC model.

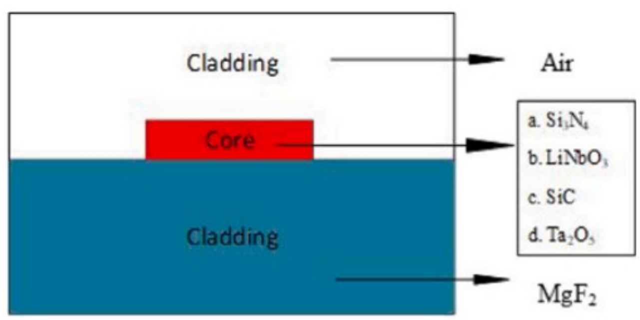


Figure 5.1: The 2D cross-sectional view of the planar waveguide used to obtain data simulated by COMSOL.

The refractive index of a material exhibits wavelength-dependent variations, a phenomenon known as material dispersion. This variation is a key aspect of chromatic dispersion, which describes the differential propagation speeds of light at various wavelengths through a medium. The relationship between the refractive index and wavelength is typically modeled using the Sellmeier equation, given by[93]:

$$n^2(\lambda) = 1 + \sum_{i=1}^{N} \frac{B_i \lambda^2}{\lambda^2 - C_i} \tag{5.1}$$

where λ represents the wavelength in micrometers (µm). Chromatic dispersion is crucial for the design and performance of optical waveguides, as it influences how light of different wavelengths and velocities propagates through the medium. The quantification of chromatic dispersion is expressed by[125]:

$$D = -\frac{\lambda}{c}\left(\frac{\partial^2 \mathrm{Re}(n_{\mathrm{eff}})}{\partial \lambda^2}\right) \tag{5.2}$$

Here, c denotes the speed of light, and $\mathrm{Re}(n_{\mathrm{eff}})$ is the real part of the effective refractive index. Chromatic dispersion arises from both material dispersion and waveguide dispersion. Within the waveguide, the effective mode area, which describes the spatial distribution of the optical mode, is given by[126]:

$$A_{\mathrm{eff}} = \frac{\left(\iint |F(x,y)|^2\, dx\, dy\right)^2}{\iint |F(x,y)|^4\, dx\, dy} \tag{5.3}$$

where A_{eff} represents the effective mode area. The nonlinear refractive index (n_2) and propagation losses for the core materials used in the waveguide are summarized in Table 5.1.

Table 5.1: Optical Properties of Core Materials Used in Numerical Simulations

Material	Linear Loss (α) (dB/cm)	n_2 (m^2/W)
Si_3N_4 [127]	2.09	2.4×10^{-19}
$LiNbO_3$ [128]	2.7	2.4×10^{-19}
SiC [129, 130]	2.3	8.6×10^{-19}
Ta_2O_5 [131]	3	7.2×10^{-19}

To model the propagation of optical pulses through nonlinear media, the Generalized Nonlinear Schrödinger Equation (GNLSE) is commonly employed. The GNLSE is defined by[126]:

$$\frac{\partial A}{\partial z} + \frac{\alpha}{2}A - \sum_{k\geq 2}\frac{i^{k+1}}{k!}\beta_k\left(\frac{\partial^k A}{\partial T^k}\right) = i\gamma\left(1 + i\tau_{\mathrm{shock}}\frac{\partial}{\partial T}\right)\left(A(z,T)\int_{-\infty}^{+\infty} R(T')|A(z,T-T')|^2\, dT'\right) \tag{5.4}$$

where α represents the linear propagation loss, β_k denotes the k-th order dispersion coefficients, $A = A(z,T)$ is the signal amplitude, $T = t - z/v_g$ is the time delay, and $v_g = 1/\beta_1(\omega_0)$ is the group velocity.

5.1.3 Dataset Description

5.1.4 Dataset for Supercontinuum Spectra Prediction

The dataset plays a crucial role in the training and testing of the supercontinuum spectra prediction model by providing a comprehensive array of input variables along with their corresponding spectral data. Key input variables in the dataset include pump wavelength, full width at half maximum (FWHM), pump power, core material, and section number. Each of the 240 segments within the 10 mm waveguide length is associated with specific spectral data, ensuring detailed coverage of the waveguide's performance.

Initially, the dataset includes 8160 rows, representing three different core materials. Each row contains 8192 Fast Fourier Transform (FFT) output points, corresponding to the five input variables. This initial subset is utilized for training and evaluating the model's performance. To further enhance the dataset, additional data for the fourth core material, Ta_2O_5, is incorporated, resulting in an expansion of the dataset by 2720 rows.

The primary objective of this dataset is to assess the model's ability to predict spectral variations resulting from changes in the core material. The test set is designed to include consistent changes in input variables across all materials, facilitating a comparative analysis of the model's performance in predicting spectral fluctuations for different core materials. This comprehensive approach allows for a robust evaluation of the model's accuracy and generalization capabilities in capturing spectral changes due to varying core materials.

5.1.5 Neural Network Description

The study utilizes a sequential deep learning architecture constructed with TensorFlow and Keras libraries. This architecture consists of an input layer, several Closed-Form Continuous-Time (CfC) layers, and Dense layers. The input layer has a shape of $(5, 1)$, which corresponds to the five input variables derived from the dataset. The CfC layers are crucial for capturing the intricate spatial and temporal evolution of supercontinuum spectra. The model includes three CfC layers, each with 200 neurons, and configured to return sequences, thereby preserving essential temporal information.Subsequent to the CfC layers, the model integrates Dense layers with varying neuron counts and activation functions. The first Dense layer comprises 512

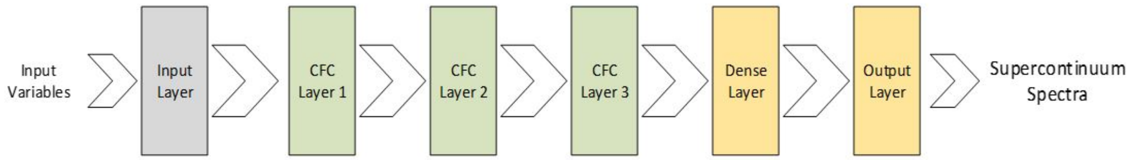


Figure 5.2: Diagram Illustrating the Overall Architecture of the Closed Form Continuous Time Neural Network Model for Predicting Spectral Power of Supercontinuum in Planar Waveguides.

neurons and uses a non-linear hyperbolic tangent (tanh) activation function. The second Dense layer contains 8192 neurons and does not employ a specified activation function. An exponential decay learning rate is applied to optimize the training process, starting with an initial learning rate of 0.0005. The learning rate decays every 1000 steps with a decay rate of 0.8, which helps in fine-tuning the model over time. This carefully designed architecture is intended to effectively learn from the dataset and accurately predict the supercontinuum spectra, ensuring robust performance in spectral analysis and prediction.

Since the numerical values obtained from GNLSE are used, comparing the mean squared error (MSE) is the best method to explain the results. MSE measures the deviation of predicted points from their actual values, calculated as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i' - y_i)^2 \tag{5.5}$$

where N represents the total number of wavelength bins, y_i' is the predicted intensity at wavelength i, and y_i is the actual intensity at wavelength i.

5.2 Results and Analysis

5.2.1 Results and Analysis

This section provides a comprehensive examination of the results, focusing on the performance of the Closed-Form Continuous-Time Neural Network (CfC) model in predicting supercontinuum spectra across various materials and conditions. The analysis begins with a detailed comparison between observed and predicted spectra to assess the model's capability in capturing intricate spectral details. The discussion also includes potential future research directions, highlighting possible improvements to the model and advancements in predictive modeling within the field of photonics.

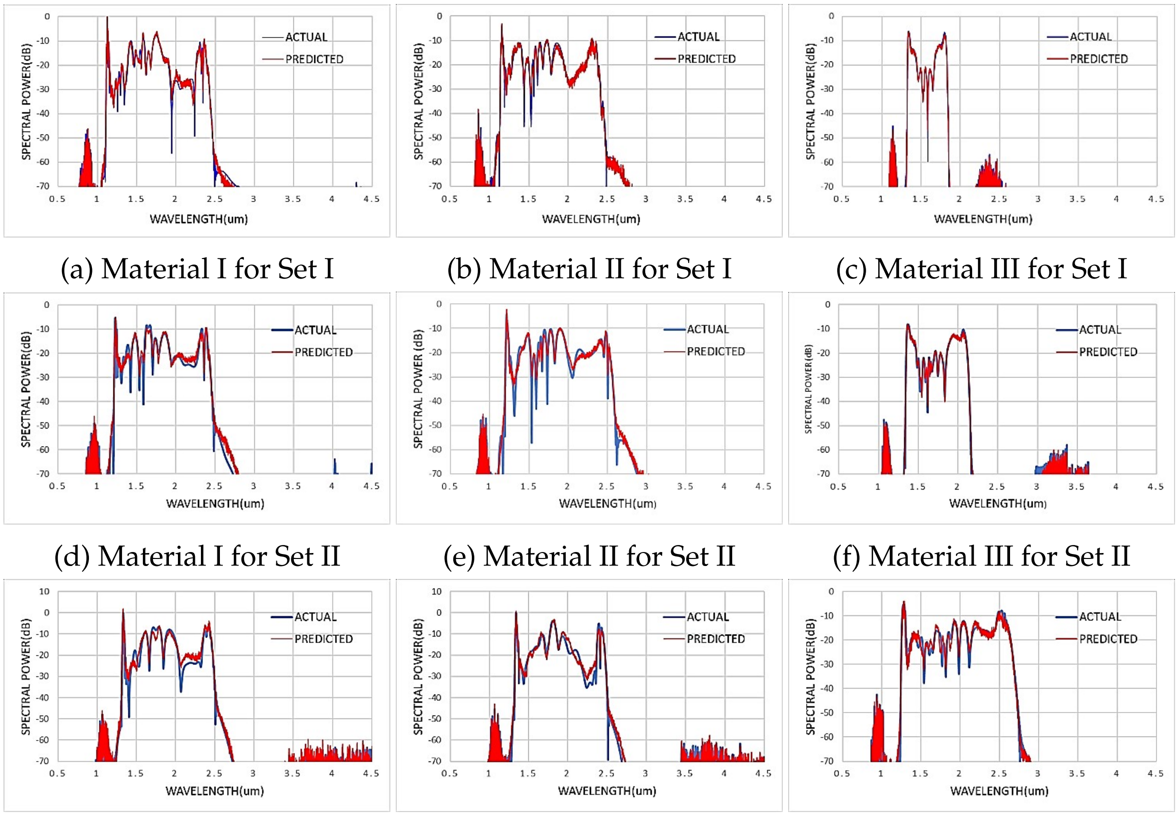
Initially, results are derived from a model trained on a dataset that includes three core materials, while the cladding material remains constant. This model is then expanded to incorporate a fourth core material, and results are presented incrementally to reflect this addition. For the dataset encompassing four core materials, only a portion of the results is shown for illustrative purposes.

Mean Squared Error (MSE) calculations are provided to quantify the performance of the model. Specifically, 27 rows of data from the three-material dataset and 36 rows from the four-material dataset are utilized for testing purposes. The remaining data in both datasets are reserved for training the model. This approach ensures a robust evaluation of the model's predictive accuracy and generalization capabilities across different material configurations and conditions.

5.2.2 Comparative Analysis of Actual and Predicted Spectra for Various Materials and Input variables in the Three-Material Dataset

Figure 5.3 presents a thorough analysis of the actual versus predicted spectra for different materials under a range of operating conditions. The goal is to evaluate the performance of the Closed-Form Continuous Time Neural Network (CfC) model in accurately capturing the complex spectral behaviors of various materials across different pump wavelengths (PWL), pump powers (PPW), full width at half maximum (FWHM), and section numbers. Each subplot compares the spectra of Material 1 (Si_3N_4), Material 2 ($LiNbO_3$), and Material 3 (SiC) under different combinations of these operating parameters. The analysis covers three sets of operating conditions.

The CfC model's spectral predictions show excellent accuracy, closely matching the actual spectra for different materials and conditions. This high level of precision highlights the model's ability to generalize across various materials and capture subtle spectral details. The strong agreement between predicted and actual spectra demonstrates the model's capability to understand and predict the nonlinear behaviors of different materials, offering valuable insights into how spectral charac-



(a) Material I for Set I

(b) Material II for Set I

(c) Material III for Set I

(d) Material I for Set II

(e) Material II for Set II

(f) Material III for Set II

(g) Material I for Set III

(h) Material II for Set III

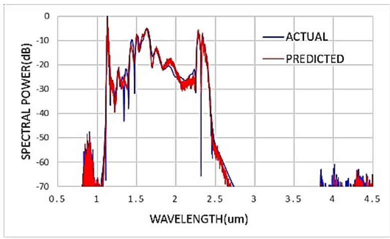
(i) Material III for Set III

Figure 5.3: Comparison of Actual vs. Predicted Spectra for Materials 1, 2, and 3 with Different PWL, PPW, and FWHM. Parameters: Set I - FWHM: 75 fs, PPW: 4600 W, PWL: 1550 nm, Section: 231; Set II - FWHM: 100 fs, PPW: 4500 W, PWL: 1650 nm, Section: 211; Set III - FWHM: 50 fs, PPW: 4200 W, PWL: 1750 nm, Section: 236

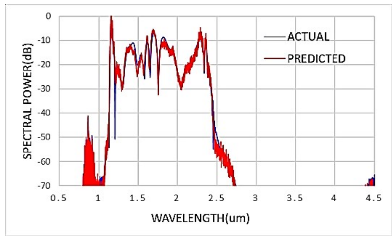
teristics evolve under varying conditions. This analysis provides a foundation for optimizing supercontinuum generation processes by enhancing our understanding of material interactions with optical signals in planar waveguides.

Figure 5.4 assesses the model's effectiveness in predicting spectral variations with different pump power levels while keeping other input variables constant. The subplots display comparisons between actual and predicted spectra for three distinct pump power settings, with a fixed pump wavelength of 1550 nm and a full width at half maximum (FWHM) of 50 fs. This results in three sets of variable conditions.

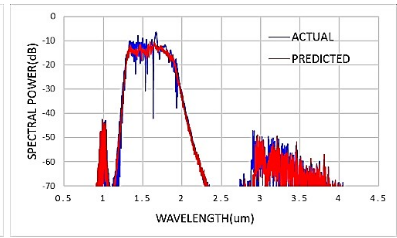
The close match between the actual and predicted spectra demonstrates the CfC model's ability to accurately capture spectral changes as pump power varies across different materials. This evaluation offers valuable insights into how variations in



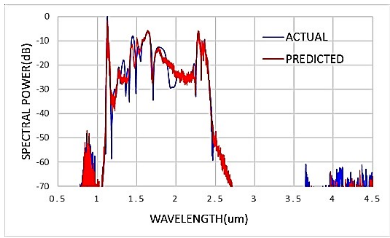
(a) Material I for Set I



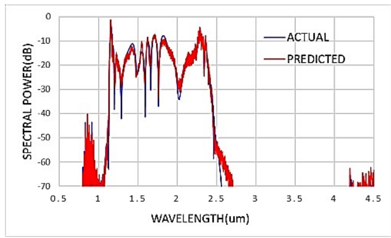
(b) Material II for Set I



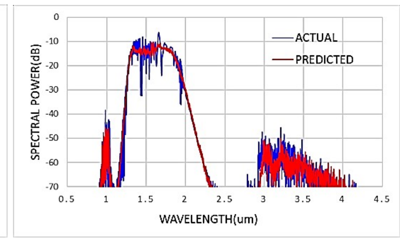
(c) Material III for Set I



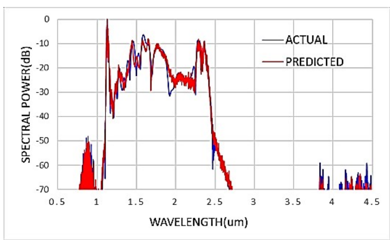
(d) Material I for Set II



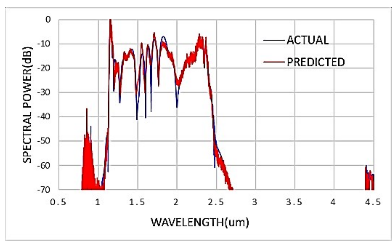
(e) Material II for Set II



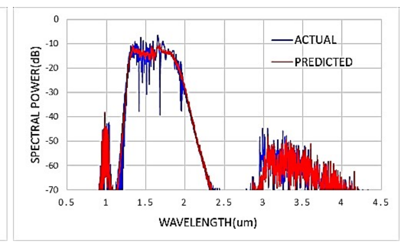
(f) Material III for Set II



(g) Material I for Set III



(h) Material II for Set III

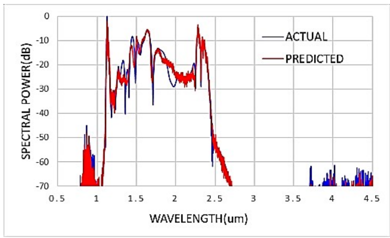


(i) Material III for Set III

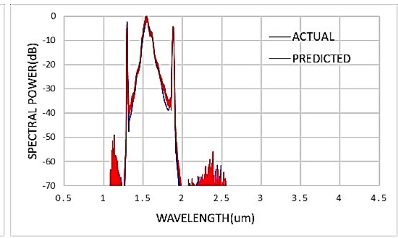
Figure 5.4: Comparison of Actual vs. Predicted Spectra for Materials 1, 2, and 3 at PWL: 1550 nm. Parameters: Set I - FWHM: 50 fs, PPW: 4500 W, Section: 240; Set II - FWHM: 50 fs, PPW: 5000 W, Section: 240; Set III - FWHM: 50 fs, PPW: 5500 W, Section: 240.

pump power affect spectral behavior, assisting researchers in optimizing device performance and improving supercontinuum generation efficiency.

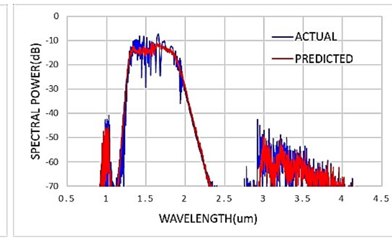
Figure 5.5 explores the model's capacity to predict spectral variations for different full width at half maximum (FWHM) values, with other variables held constant. The subplots compare actual and predicted spectra for three distinct FWHM values, illustrating the model's accuracy across different pulse durations. The strong alignment between the actual and predicted spectra highlights the CfC model's effectiveness in capturing spectral changes with high precision, even as FWHM values vary. This analysis offers valuable insights into how pulse duration influences spectral behavior, providing researchers with a better understanding of the nonlinear dynamics in supercontinuum generation.



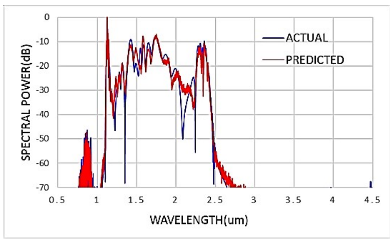
(a) Material I for Set I



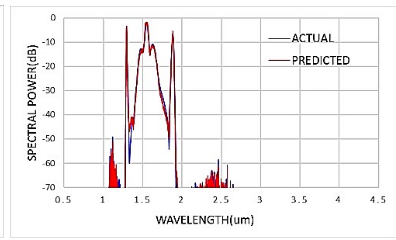
(b) Material II for Set I



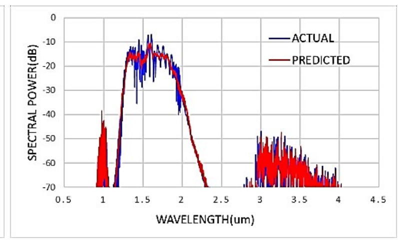
(c) Material III for Set I



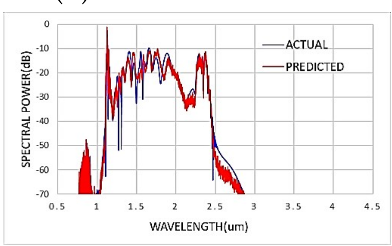
(d) Material I for Set II



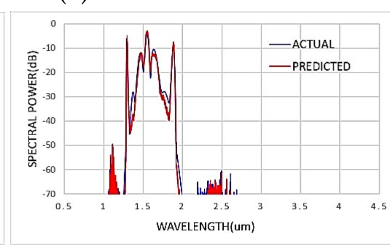
(e) Material II for Set II



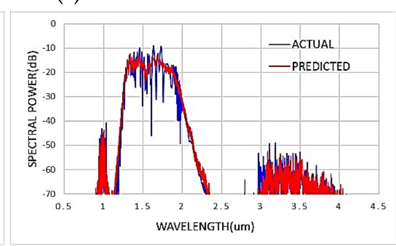
(f) Material III for Set II



(g) Material I for Set III



(h) Material II for Set III

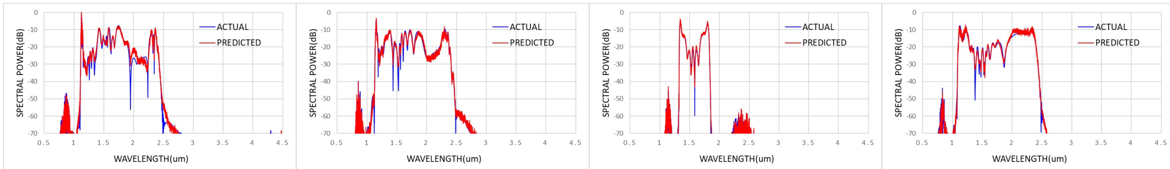


(i) Material III for Set III

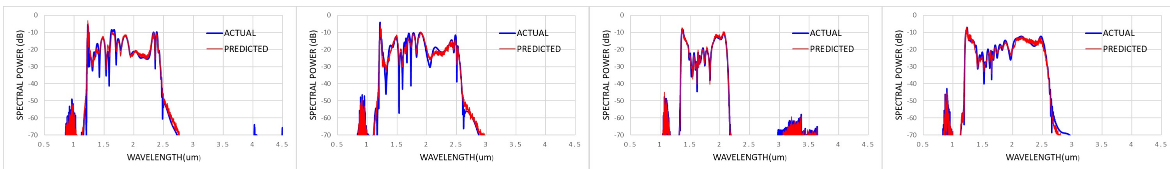
Figure 5.5: Comparison of Spectra for Materials 1, 2, and 3 with Changing FWHM values at PWL: 1550 nm and PPW: 4900 W. Parameters: Set I - FWHM: 50 fs; Set II - FWHM: 75 fs; Set III - FWHM: 100 fs, Section: 240.

5.2.3 Comparative Analysis of Actual and Predicted Spectra for Various Materials and Input variables in the Four-Material Dataset

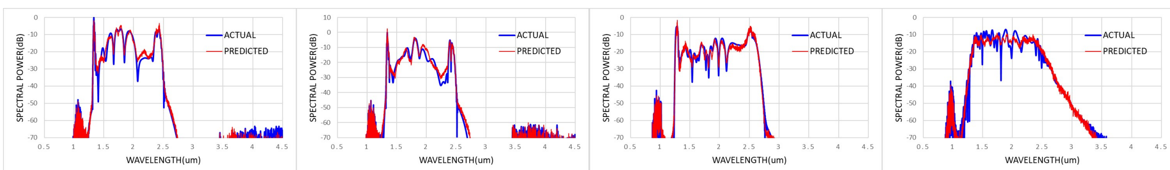
For the expanded dataset that includes four materials, the variations similar to those observed in the three-material dataset yield 36 distinct outcomes. Figures 5.6, 5.7, and 5.8 illustrate the comparison between actual and predicted spectra across various input parameters. These figures present spectral predictions for two different sets of input parameters, covering spectral data for four materials: Si_3N_4 (Material 1), $LiNbO_3$ (Material 2), SiC (Material 3), and Ta_2O_5 (Material 4). In total, nine unique sets of input parameters have been evaluated, consistent with the approach used in analyzing the three-material dataset.



(a) Material I for Set I (b) Material II for Set I (c) Material III for Set I (d) Material IV for Set I



(e) Material I for Set II (f) Material II for Set II (g) Material III for Set II (h) Material IV for Set II



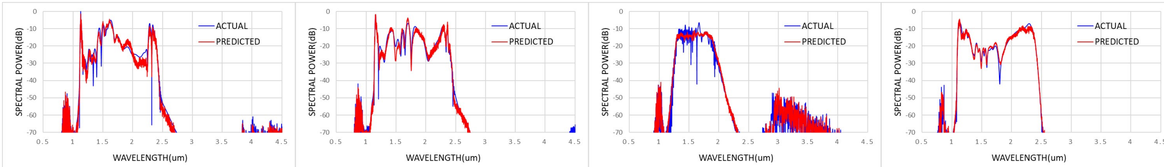
(i) Material I for Set III (j) Material II for Set III (k) Material III for Set III (l) Material IV for Set III

Figure 5.6: Comparison of Spectra with Varying PWL, PPW, and FWHM for Four Materials. Parameters: Set I - PPW: 4600 W, FWHM: 75 fs, PWL: 1550 nm; Set II - PPW: 4500 W, FWHM: 100 fs, PWL: 1650 nm; Set III - PPW: 4200 W, FWHM: 50 fs, PWL: 1750 nm.

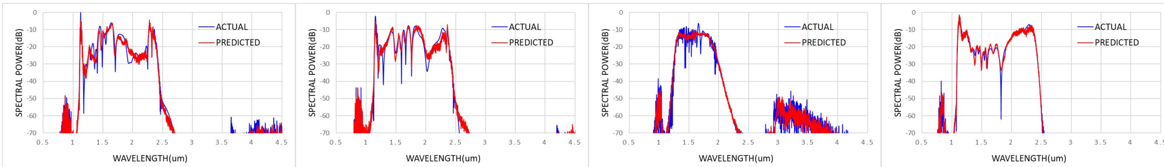
The inclusion of Ta_2O_5 in the dataset introduces additional variability, enabling a comprehensive evaluation of the model's generalization capabilities. Despite the expansion of the dataset to incorporate Ta_2O_5, the model continues to demonstrate a high level of accuracy, as evidenced by the close alignment between actual and predicted spectra. This indicates that the CfC model retains its robustness and reliability even when applied to new materials.

This detailed assessment highlights the model's effectiveness in predicting supercontinuum spectra across a wider range of materials and conditions. The ability to accurately generalize to additional materials underscores the model's strength and versatility, offering valuable insights into its performance and potential applications in the field of photonics.

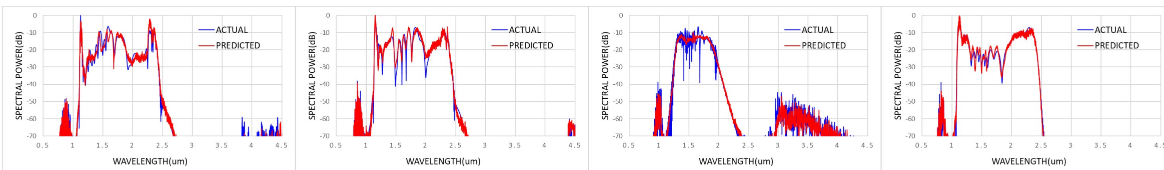
The continued strong performance of the CfC model, even with the addition of new materials, highlights its flexibility and reliability in predicting supercontinuum spectra across a diverse range of materials and operating conditions. The close cor-



(a) Material I for Set I (b) Material II for Set I (c) Material III for Set I (d) Material IV for Set I



(e) Material I for Set II (f) Material II for Set II (g) Material III for Set II (h) Material IV for Set II



(i) Material I for Set III (j) Material II for Set III (k) Material III for Set III (l) Material IV for Set III

Figure 5.7: Comparison of Spectra for Four Materials with Varying PWL and PPW at FWHM = 50 fs. Parameters: Set I - PPW: 4500 W, PWL: 1550 nm; Set II - PPW: 5000 W, PWL: 1550 nm; Set III - PPW: 5500 W, PWL: 1550 nm.

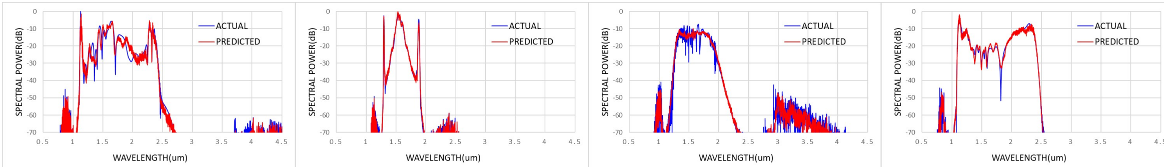
Table 5.2: Summary of Results on Test Set by the Neural Network

Epoch	No. of Materials	MSE Loss
100	3	7.4199×10^{-4}
100	4	6.9408×10^{-4}

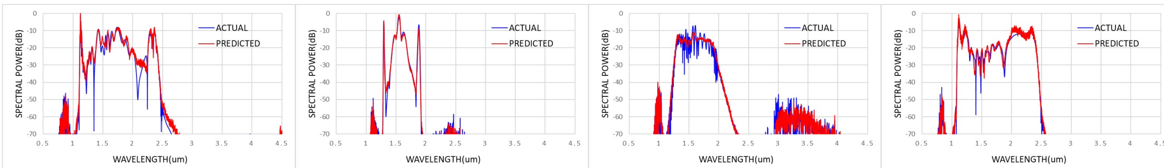
relation between the actual and predicted spectra underscores the model's ability to capture detailed spectral features and accurately forecast spectral behavior under various input conditions. The expansion of the dataset with additional materials further enhances the model's potential and accuracy.

This analysis provides valuable insights into the model's scalability and generalization strengths, paving the way for further advancements in predictive modeling and optimization of optical devices in photonics research. Figures 5.6, 5.7, and 5.8 offer a comprehensive understanding of the CfC model's capabilities and its potential applications across different domains of photonics and spectral analysis.

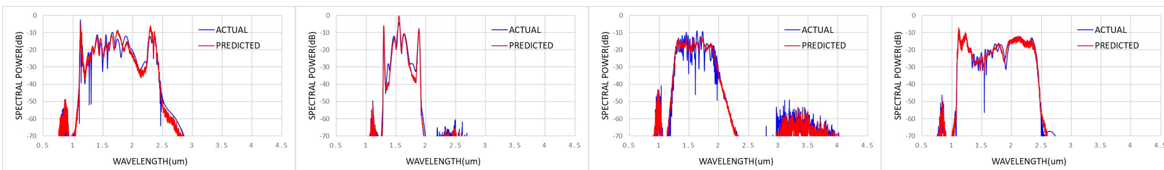
Furthermore, the Mean Squared Error (MSE) loss values presented in Table 5.2



(a) Material I for Set I (b) Material II for Set I (c) Material III for Set I (d) Material IV for Set I



(e) Material I for Set II (f) Material II for Set II (g) Material III for Set II (h) Material IV for Set II



(i) Material I for Set III (j) Material II for Set III (k) Material III for Set III (l) Material IV for Set III

Figure 5.8: Comparison of Spectra for Four Materials with Varying Input Parameters. Parameters: Set I - PWL: 1550 nm, PPW: 4900 W, FWHM: 50 fs; Set II - PWL: 1550 nm, PPW: 4900 W, FWHM: 75 fs; Set III - PWL: 1550 nm, PPW: 4900 W, FWHM: 100 fs.

provide quantitative measures of the model's performance across different materials. The MSE values for the models trained on both three and four materials are notably low, indicating that the CfC model is highly effective in learning and generalizing complex spectral patterns.

These results affirm the CfC model's strong predictive accuracy for supercontinuum spectra, demonstrating its potential for applications in photonics and related fields. The close alignment between the predicted and actual spectra, coupled with the low MSE values, highlights the model's robustness and reliability in accurately forecasting spectral behavior across diverse materials and operating conditions. Additionally, the model's ability to capture intricate spectral features further underscores its effectiveness in predicting and analyzing supercontinuum spectra.The CfC model's performance, as evidenced by the low MSE values and the close match between predicted and actual spectra, reinforces its potential for advancing predictive modeling in photonics. These findings suggest that the CfC model can be a valuable tool for optimizing and designing optical devices, offering precise and reliable predictions in various photonics applications.

The CfC model has demonstrated exceptional accuracy in predicting supercontinuum spectra, effectively capturing even the most subtle spectral details. Unlike conventional methods such as Artificial Neural Networks (ANNs), which may overlook minor spectral features and require substantial computational resources, the CfC model excels in identifying complex spectral patterns with greater efficiency. It surpasses traditional approaches, including Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, in both accuracy and computational efficiency.RNNs and LSTMs often face challenges related to vanishing gradients and difficulties in managing long-term dependencies, which can hinder their performance in accurately capturing spectral details. In contrast, the CfC model leverages its closed-form solutions to effectively address these challenges, providing accurate predictions without compromising performance.In the initial phase of this study, the CfC model was trained and evaluated using a dataset that included three different materials. This setup was crucial in demonstrating the model's ability to generalize across various material compositions. The subsequent incorporation of a fourth material into the dataset further showcased the model's scalability and robustness. This ability to expand the dataset by adding new materials underscores the CfC model's adaptability and its effectiveness in handling diverse scenarios. Overall, the CfC model's superior performance in capturing detailed spectral features and its scalability in incorporating additional materials highlight its potential as a powerful tool for predictive modeling in photonics. The model's efficiency and accuracy make it a valuable asset for various applications in optical device design and spectral analysis.

The findings here suggest that the CfC model not only achieves higher accuracy in predicting supercontinuum spectra but also more effectively captures detailed spectral characteristics than traditional methods like ANN, RNN, and LSTM. Its scalability and capacity for generalization position the CfC model as a valuable tool for a wide range of applications in photonics and spectral analysis.

5.3 Discussion

This study successfully deals with Closed-Form Continuous-Time Neural Networks (CfC) to predict supercontinuum spectra in planar waveguides for various materials. The CfC model exhibits remarkable accuracy in capturing spectral behavior under different operating conditions. Initially trained with data from three materials—silicon nitride (Si_3N_4), lithium niobate ($LiNbO_3$), and silicon carbide (SiC)—the model demonstrates its effectiveness through minimal Mean Squared Error (MSE) loss. Subsequently, the model was tested on an expanded dataset that includes a fourth material, Ta_2O_5, highlighting its ability to generalize and scale.

The CfC model's ability to accurately capture complex spectral patterns while maintaining computational efficiency is evident when compared to traditional methods such as Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) networks. The inclusion of Ta_2O_5 data further showcases the model's flexibility and scalability. Overall, these findings underscore the CfC model's effectiveness in predicting supercontinuum spectra, offering insights into nonlinear dynamics, and supporting the development of optical devices for diverse applications in photonics and related fields.

Despite the CfC model's impressive computational accuracy, handling the ever-increasing size of the dataset presents a challenge. As more materials are incorporated, the dataset grows larger, which may necessitate alternative approaches for managing the data. One potential solution is to refine the dataset by removing redundant information, focusing only on the spectral variations that are most critical and sensitive to changes in input parameters. This targeted approach could help in managing data size while maintaining the relevance of the information studied.Additionally, a limitation of the current study is the absence of a time-domain analysis of the input pulse. Investigating the temporal evolution of the pulse during propagation is essential for a comprehensive understanding of supercontinuum generation. Employing advanced neural network architectures, such as the CfC model, for time-domain studies could provide valuable insights into the dynamics of pulse propagation and further enhance the model's capabilities.Addressing these challenges and limitations will be crucial for future research. Developing methods to efficiently manage and analyze large datasets, along with incorporating time-domain analysis, will help in fully leveraging the CfC model's potential and improving its application in predictive modeling of supercontinuum spectra.

Moreover, this field may investigate expanding the dataset to encompass subsets of chalcogenide materials, hence facilitating the creation of predictions for a more extensive variety of materials with varied optical characteristics. The predictive power of the model could also be improved by adding a greater range of materials and altering the cladding material. By doing so, knowledge of the processes involved in supercontinuum creation and optimize optical devices for a wide range of photonics applications could be improved.

5.4 Conclusion

This study has demonstrated the effectiveness of Closed-Form Continuous-Time Neural Networks (CfC) in accurately predicting supercontinuum spectra across various core materials in planar waveguides, offering a significant improvement over traditional, computationally intensive methods. By training the CfC model on data from COMSOL Multiphysics simulations that incorporate key optical parameters, the research shows that the model can generalize across different materials, achieving low prediction errors close to 0.0006 for four materials. This approach not only reduces reliance on extensive numerical simulations but also accelerates the design process for optical devices, with broad implications for advancing photonic technologies in areas like optical communications, biomedical imaging, and spectroscopy.Besides this task has successfully investigated neural networks efficacy in dealing an extension of the existing dataset adding more complexity and variations.

Chapter 6

Applications of the Thesis

The research presented in this thesis has significant implications for various fields within photonics and optical communication. The integration of deep learning with photonic crystal fibers (PCFs) and waveguides opens up new possibilities for enhancing the design and optimization of these structures. Orbital Angular Momentum (OAM) mode transmission analysis using DL techniques, important part of this thesis, is increasingly vital in advancing modern optical communication systems. Some of the researches that explore innovative methods for generating and transmitting OAM modes can suggest the importance of OAM mode transmission. For instance, the partial arc transmitting scheme proposed by Xiong et al. [132] enables high-order OAM mode transmission and shows great potential for integration with photonic crystal fibers (PCFs), aligning well with this thesis's objectives. Moreover, Vigneswaran et al.'s [133]work on spiral photonic crystal fibers (S-PCFs) demonstrates the support of multiple OAM modes with low nonlinear effects and flat dispersion, crucial for space division multiplexing in optical fiber communication systems. Additionally, the design by Yang et al. [134] of PCFs that sustain 86 OAM modes with high mode quality and low nonlinear coefficients directly complements this thesis's focus on optimizing communication system performance through advanced OAM techniques. These studies underscore the significance of OAM mode transmission in achieving high-capacity and efficient communication systems, making this thesis a meaningful contribution to the ongoing advancements in OAM-based technologies. Another part of the thesis focuses on Non linear phenomenon, Supercontinuum generation prediction. Supercontinuum generation (SCG) plays a vital role in nonlinear optics, with applications that intersect closely with the research focus of this thesis on predicting supercontinuum spectra using deep learning techniques. Brès et al. [135]discuss how SCG in integrated photonics can bridge distant spectral regions, a capability that enhances the precision and efficiency of predicting spectral outputs—a challenge addressed through the deep learning methods developed in this thesis. Similarly, the work of Cao et al. [136] on SCG in ultra-silicon-rich nitride waveguides emphasizes the potential for creating CMOS-compatible light sources, which is directly applicable to the predictive modeling of SCG spectra for advanced photonic devices. De Sterke's [137] overview of SCG applications in nonlinear photonics highlights the real-world relevance of these phenomena, underscoring the importance of accurate and reliable spectral predictions, which this research aims to achieve through machine learning. Finally, Hary et al.'s [138]application of machine learning in controlling nonlinear fiber SCG for molecular spectroscopy illustrates the broader potential of integrating AI techniques with SCG, further validating the approach taken in this thesis. Collectively, these references support the integration of deep learning into the prediction of SCG spectra, enhancing the capabilities of next-generation photonic systems and aligning with the objectives of this research.

Chapter 7

Conclusion

This research has made significant contributions to the field of photonics by integrating advanced machine learning techniques with traditional numerical methods. Specifically, the study focused on enhancing the understanding and predictive capabilities related to the linear and non-linear properties of photonic crystal fibers (PCFs) and optical waveguides.

The first major outcome of this research involved the successful prediction of linear optical parameters such as effective refractive index, dispersion, mode purity, and effective area of PCFs using an Artificial Neural Network (ANN). The ANN model proivided MSE values close to 0.0218 with approximately 31 seconds of training time. This improvement in prediction accuracy, alongside the reduced computational time compared to traditional simulation methods, underscores the model's efficiency and effectiveness in practical applications.

The second key result pertained to the non-linear dynamics of supercontinuum generation in chalcogenide planar waveguides. By employing an ANN-based deep learning model, the research achieved precise modeling of pulse propagation dynamics. The predictive model captured the complex interaction between parameters such as pump power, wavelength, and waveguide dimensions, delivering insights with a significantly reduced computational burden.The computation accuracy yielded a low MSE value of 0.0032 with a computational training time of 472 seconds which proves to be efficient on both the metrics. The model's ability to replicate empirical results closely was evidenced by the accurate prediction of spectral evolution across varied pump powers, demonstrating its potential for real-time optimization of experimental setups.

Finally, the study introduced a novel approach using Closed-Form Continuous-Time Neural Networks (CfC) for predicting supercontinuum spectra in diverse material compositions. The CfC model outperformed traditional methods such as Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, particularly in terms of computational efficiency and scalability. The model maintained a low MSE while handling a dataset expanded to include four distinct materials, thereby proving its robustness and adaptability. The inclusion of additional materials like Ta_2O_5 further validated the model's generalization capabilities, ensuring consistent prediction accuracy under different operating conditions.The CfC model demonstrated high accuracy, with Mean Squared Error (MSE) values as low as 7.4199×10^{-4} when trained on a three-material dataset, and 6.9408×10^{-4} on a four-material dataset.The neural network was 15-16x faster compared to traditional methods.

The impact of this work extends beyond the immediate technical advancements in photonics and deep learning. By improving the accuracy and efficiency of predictive models for photonic devices, this research contributes to the broader societal goal of advancing optical technologies that are foundational to modern telecommunications, medical imaging, and environmental sensing. The reduction in computational time and resources, achieved through the integration of machine learning, also holds significant potential for accelerating the development of next-generation photonic devices, making advanced technologies more accessible and affordable. Furthermore, this work sets a precedent in the research community for the application of artificial intelligence in complex physical systems, encouraging interdisciplinary collaboration and inspiring future research that combines machine learning with other areas of physics and engineering. By demonstrating the feasibility and benefits of such an approach, this research opens new avenues for innovation, with potential ripple effects across various sectors that rely on photonics and optical technologies.

In summary, this research has demonstrated that integrating deep learning techniques with photonics not only enhances the accuracy of predictive models but also significantly reduces computational time. These advancements have practical implications for the design, optimization, and real-time control of photonic devices, paving the way for future innovations in the field.

Bibliography

[1] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, "Photonic crystals: Molding the flow of light," 10 2011. [Online]. Available: https://collaborate.princeton.edu/en/publications/photonic-crystals-molding-the-flow-of-light

[2] N. A. Kayed, M. R. Karim, N. Jahan, A. Hussain, and B. M. A. Rahman, "Mid-Infrared Supercontinuum Generation using Dispersion-Varying Silicon-Rich Silicon Nitride Waveguide," 8 2021. [Online]. Available: https://doi.org/10.1109/icsct53883.2021.9642601

[3] B. M. A. Rahman, F. Fernández, and J. Davies, "Review of finite element methods for microwave and optical waveguides," *Proceedings of the IEEE*, vol. 79, no. 10, pp. 1442–1448, 1 1991. [Online]. Available: https://doi.org/10.1109/5.104219

[4] M. Yetmez, *Finite element analysis*, 1 2016. [Online]. Available: https://doi.org/10.1007/978-3-319-20777-3_4

[5] R. Marklein, *The Finite Integration Technique as a General Tool to Compute Acoustic, Electromagnetic, Elastodynamic, and Coupled Wave Fields*, 10 2002, pp. 201–244.

[6] D. Burke and T. Smy, "Optical mode solving for complex waveguides using a finite cloud method," *Optics Express*, vol. 20, no. 16, p. 17783, 7 2012. [Online]. Available: https://doi.org/10.1364/oe.20.017783

[7] F. H. Alharbi, "Full-Vectorial meshfree spectral method for Optical-Waveguide analysis," *IEEE Photonics Journal*, vol. 5, no. 1, p. 6600315, 2 2013. [Online]. Available: https://doi.org/10.1109/jphot.2013.2244876

[8] Y. LeCun, Y. Bengio, and G. E. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 5 2015. [Online]. Available: https://doi.org/10.1038/nature14539

[9] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, A. Moschitti, B. Pang, and W. Daelemans, Eds. Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1724–1734. [Online]. Available: https://aclanthology.org/D14-1179

[10] G. E. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. W, Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, "Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 11 2012. [Online]. Available: https://doi.org/10.1109/msp.2012.2205597

[11] H. A. Pierson and M. S. Gashler, "Deep learning in robotics: a review of recent research," *Advanced Robotics*, vol. 31, no. 16, pp. 821–835, 8 2017. [Online]. Available: https://doi.org/10.1080/01691864.2017.1365009

[12] T. Zahavy, A. Dikopoltsev, D. Moss, G. I. Haham, O. Cohen, S. Mannor, and M. Segev, "Deep learning reconstruction of ultrashort pulses," *Optica*, vol. 5, no. 5, p. 666, 5 2018. [Online]. Available: https://doi.org/10.1364/optica.5.000666

[13] B. Sánchez-Lengeling and A. Aspuru-Guzik, "Inverse molecular design using machine learning: Generative models for matter engineering," *Science*, vol. 361, no. 6400, pp. 360–365, 7 2018. [Online]. Available: https://doi.org/10.1126/science.aat2663

[14] J. Carrasquilla and R. G. Melko, "Machine learning phases of matter," *Nature Physics*, vol. 13, no. 5, pp. 431–434, 2 2017. [Online]. Available: https://doi.org/10.1038/nphys4035

[15] Y. Rivenson, Z. Göröcs, H. Günaydın, H. Zhang, H. Wang, and A. Ozcan,

"Deep learning microscopy," *Optica*, vol. 4, no. 11, p. 1437, 11 2017. [Online]. Available: https://doi.org/10.1364/optica.4.001437

[16] M. Vijayan, S. S. Sridhar, and D. Vijayalakshmi, "A deep learning regression model for photonic crystal fiber sensor with xai feature selection and analysis," *IEEE Transactions on NanoBioscience*, vol. 22, no. 3, pp. 590–596, 2023.

[17] F. Ma, Y. Ma, P. Li, and W. Shi, "Inverse design of broadband dispersion compensation fiber based on deep learning and differential evolution algorithm," *IEEE Photonics Journal*, vol. 15, no. 3, pp. 1–7, 2023.

[18] M. I. Khalil and M. S. Islam, "Optical properties prediction of negative dispersion-compensating photonic crystal fiber using machine learning," in *2022 12th International Conference on Electrical and Computer Engineering (ICECE)*, 2022, pp. 32–35.

[19] M. A. Jabin and M. P. Fok, "Prediction of 12 photonic crystal fiber optical properties using mlp in deep learning," *IEEE Photonics Technology Letters*, vol. 34, no. 7, pp. 391–394, 2022.

[20] P. Verma, A. Kumar, and P. Jindal, "Machine learning approach for spr based photonic crystal fiber sensor for breast cancer cells detection," in *2022 IEEE 7th Forum on Research and Technologies for Society and Industry Innovation (RTSI)*, 2022, pp. 7–12.

[21] A. Zelaci, A. Yasli, C. Kalyoncu, and H. Ademgil, "Generative adversarial neural networks model of photonic crystal fiber based surface plasmon resonance sensor," *Journal of Lightwave Technology*, vol. 39, no. 5, pp. 1515–1522, 2021.

[22] A. Da Silva Ferreira, G. N. Malheiros-Silveira, and H. E. Hernandez-Figueroa, "Computing optical properties of photonic crystals by using multilayer perceptron and extreme learning machine," *Journal of Lightwave Technology*, vol. 36, no. 18, pp. 4066–4073, 9 2018. [Online]. Available: https://doi.org/10.1109/jlt.2018.2856364

[23] G. Alagappan and C. E. Png, "Modal classification in optical waveguides using deep learning," *Journal of Modern Optics*, vol. 66, no. 5, pp. 557–561, 12 2018. [Online]. Available: https://doi.org/10.1080/09500340.2018.1552331

[24] N. J. Anika and B. Mia, "Design and analysis of guided modes in photonic waveguides using optical neural network," *Optik*, vol. 228, p. 165785, 2 2021. [Online]. Available: https://doi.org/10.1016/j.ijleo.2020.165785

[25] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nature Photonics*, vol. 15, no. 2, pp. 77–90, 10 2020. [Online]. Available: https://doi.org/10.1038/s41566-020-0685-y

[26] S. Chugh, S. Ghosh, A. Gulistan, and B. M. A. Rahman, "Machine Learning Regression Approach to the Nanophotonic Waveguide analyses," *Journal of Lightwave Technology*, vol. 37, no. 24, pp. 6080–6089, 12 2019. [Online]. Available: https://doi.org/10.1109/jlt.2019.2946572

[27] G. Alagappan and C. E. Png, "Deep learning models for effective refractive indices in silicon nitride waveguides," *Journal of Optics*, vol. 21, no. 3, p. 035801, 2 2019. [Online]. Available: https://doi.org/10.1088/2040-8986/ab00d5

[28] M. I. H. Upal, D. Dutta, R. Rafi, M. R. Karim, M. R. Alam, and S. Ghosh, "Deep learning approach to determine the optical characteristics of photonic crystal fiber for orbital angular momentum transmission," in *2023 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 2023, pp. 1–6.

[29] G. Alagappan and C. E. Png, "Meshless optical mode solving using scalable deep deconvolutional neural network," *Scientific Reports*, vol. 13, no. 1, 1 2023. [Online]. Available: https://doi.org/10.1038/s41598-022-25613-4

[30] ——, "Group refractive index via auto-differentiation and neural networks," *Scientific Reports*, vol. 13, no. 1, 3 2023. [Online]. Available: https://doi.org/10.1038/s41598-023-29952-8

[31] Y. Kiarashinejad, S. Abdollahramezani, M. Zandehshahvar, O. Hemmatyar, and A. Adibi, "Deep learning reveals underlying physics of Light–Matter interactions in nanophotonic devices," *Advanced Theory and Simulations*, vol. 2, no. 9, 7 2019. [Online]. Available: https://doi.org/10.1002/adts.201900088

[32] B. Wetzel, M. Kues, P. Roztocki, C. Reimer, P.-L. Godin, M. Rowley, B. E. Little, S. T. Chu, E. A. Viktorov, D. Moss, A. Pasquazi, M. Peccianti, and R. Morandotti, "Customizing supercontinuum generation via on-chip adaptive temporal pulse-splitting," *Nature Communications*, vol. 9, no. 1, 11 2018. [Online]. Available: https://doi.org/10.1038/s41467-018-07141-w

[33] O. Tzang, A. M. Caravaca-Aguirre, K. Wagner, and R. Piestun, "Adaptive wavefront shaping for controlling nonlinear multimode interactions in optical fibres," *Nature Photonics*, vol. 12, no. 6, pp. 368–374, 5 2018. [Online]. Available: https://doi.org/10.1038/s41566-018-0167-7

[34] L. Salmela, N. Tsipinakis, A. Foi, C. Billet, J. M. Dudley, and G. Genty, "Predicting ultrafast nonlinear dynamics in fibre optics with a recurrent neural network," *Nature Machine Intelligence*, vol. 3, no. 4, pp. 344–354, 2 2021. [Online]. Available: https://doi.org/10.1038/s42256-021-00297-z

[35] L. Salmela, C. Lapre, J. M. Dudley, and G. Genty, "Machine learning analysis of rogue solitons in supercontinuum generation," *Scientific Reports*, vol. 10, no. 1, 6 2020. [Online]. Available: https://doi.org/10.1038/s41598-020-66308-y

[36] L. Salmela, M. Hary, M. Mabed, A. Foi, J. M. Dudley, and G. Genty, "Feed-forward neural network as nonlinear dynamics integrator for supercontinuum generation: erratum," *Optics Letters*, vol. 47, no. 7, p. 1741, 3 2022. [Online]. Available: https://doi.org/10.1364/ol.457941

[37] R. Paschotta, "effective refractive index," 2 2024. [Online]. Available: https://www.rp-photonics.com/effective_refractive_index.html

[38] S. K. Selvaraja and P. Sethi, *Review on Optical waveguides*, 8 2018. [Online]. Available: https://doi.org/10.5772/intechopen.77150

[39] M. Nitas, M.-T. Passia, and T. V. Yioultsis, "Fully planar slow-wave substrate integrated waveguide based on broadside-coupled complementary split ring resonators for mmWave and 5G components," *Iet Microwaves Antennas amp; Propagation*, vol. 14, no. 10, pp. 1096–1107, 6 2020. [Online]. Available: https://doi.org/10.1049/iet-map.2019.1014

[40] J. Alamán, M. López-Valdeolivas, R. Alicante, and C. Sánchez-Somolinos, "Optical Planar Waveguide Sensor with Integrated Digitally-Printed Light Coupling-in and Readout Elements," *Sensors*, vol. 19, no. 13, p. 2856, 6 2019. [Online]. Available: https://doi.org/10.3390/s19132856

[41] J. C. Knight, "Photonic crystal fibres," *Nature*, vol. 424, no. 6950, pp. 847–851, 8 2003. [Online]. Available: https://www.nature.com/articles/nature01940

[42] N. Mahnot, S. Maheshwary, and R. Mehra, "Photonic crystal fiber-an overview," 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:56450566

[43] J. M. Dudley, G. Genty, and S. Coen, "Supercontinuum generation in photonic crystal fiber," *Rev. Mod. Phys.*, vol. 78, pp. 1135–1184, Oct 2006. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.78.1135

[44] J. Laegsgaard, K. Hansen, M. Nielsen, T. Hansen, J. Riishede, K. Hougaard, T. Sorensen, T. Larsen, N. Mortensen, J. Broeng, J. Jensen, and A. Bjarklev, "Photonic crystal fibers," in *Proceedings of the 2003 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference - IMOC 2003. (Cat. No.03TH8678)*, vol. 1, 2003, pp. 259–264 vol.1.

[45] X. Qi and Y. Lian, "Photonic crystal fiber with double-layer rings for the transmission of orbital angular momentum," *Frontiers in Physics*, vol. 11, 7 2023. [Online]. Available: https://doi.org/10.3389/fphy.2023.1207182

[46] R. Tao, X. Wang, X. Hu, P. Zhang, and S. Liu, "Coherent beam combination of fiber lasers with a strongly confined tapered self-imaging waveguide: theoretical modeling and simulation," *Photonics Research*, vol. 1, no. 4, p. 186, 11 2013. [Online]. Available: https://m.researching.cn/articles/OJd24ee79ddec27820/figureandtable

[47] P. Zhao, Z. Ye, K. Vijayan, C. Naveau, J. Schröder, M. Karlsson, and P. A. Andrekson, "Waveguide tapering for improved parametric amplification in integrated nonlinear si3n4 waveguides," *Opt. Express*, vol. 28, no. 16, pp. 23 467–23 477, Aug 2020. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-28-16-23467

[48] M. R. Karim, N. A. Kayed, R. Rafi, and B. M. A. Rahman, "Design and analysis of inverse tapered silicon nitride waveguide for flat and highly coherent supercontinuum generation in the mid-infrared," *Optical and Quantum Electronics*, vol. 56, no. 1, 12 2023. [Online]. Available: https://doi.org/10.1007/s11082-023-05636-5

[49] M. A. Gaafar, H. Renner, M. Eich, and A. Y. Petrov, "Fourier optics with linearly tapered waveguides: light trapping and focusing," 2021.

[50] R. Jannesari, T. Grille, C. Consani, G. Stocker, A. Tortschanoff, and B. Jakoby, "Design of a curved shape photonic crystal taper for highly efficient mode coupling," *Sensors*, vol. 21, no. 2, p. 585, 1 2021. [Online]. Available: https://doi.org/10.3390/s21020585

[51] M. A. Butt, N. L. Kazanskiy, and , "Metal-Insulator-Metal Waveguide Plasmonic Sensor System for refractive Index sensing applications," *Advanced photonics research*, vol. 4, no. 7, 5 2023. [Online]. Available: https://doi.org/10.1002/adpr.202300079

[52] Y.-S. Liao, J. Wu, D. Thakur, J.-S. Hsu, R. P. Dwivedi, and S. H. Chang, "Power Loss Reduction of Angled Metallic Wedge Plasmonic Waveguides via the Interplay between Near-Field Optical Coupling and Modal Coupling," *Photonics*, vol. 9, no. 9, p. 663, 9 2022. [Online]. Available: https://doi.org/10.3390/photonics9090663

[53] B. Timotijević, F. Y. Gardes, W. C. Headley, G. T. Reed, M. Paniccia, O. Cohen, D. Hak, and G. Z. Masanovic, "Multi-stage racetrack resonator filters in silicon-on-insulator," *Journal of optics*, vol. 8, no. 7, pp. S473–S476, 6 2006. [Online]. Available: https://doi.org/10.1088/1464-4258/8/7/s25

[54] K. Prabhakar and R. M. Reano, "Fabrication of low loss lithium niobate RIB waveguides through photoresist reflow," *IEEE Photonics Journal*, vol. 14, no. 6, pp. 1–8, 12 2022. [Online]. Available: https://doi.org/10.1109/jphot.2022.3222184

[55] M. A. Butt, Ł. Kozłowski, and R. Piramidowicz, "Numerical scrutiny of a silica-titania-based reverse rib waveguide with vertical and rounded

sidewalls," *Appl. Opt.*, vol. 62, no. 5, pp. 1296–1302, Feb 2023. [Online]. Available: https://opg.optica.org/ao/abstract.cfm?URI=ao-62-5-1296

[56] S. Meister, B. A. Franke, H. J. Eichler, S. Kupijai, H. Rhee, A. Al-Saadi, and L. Zimmermann, "Photonic integrated circuits for optical communication," *Optik amp; Photonik*, vol. 7, no. 2, pp. 59–62, 5 2012. [Online]. Available: https://doi.org/10.1002/opph.201290052

[57] S. Stopiński, M. M. Smit, and X. X. Leijtens, "Photonic integrated circuits for data read-out systems," 2010. [Online]. Available: https://api.semanticscholar.org/CorpusID:59393309

[58] , S. V. Eliseeva, and , "Spectra of a Bragg Microresonator Filled with a Graphene-Containing Medium," *Photonics*, vol. 10, no. 4, p. 449, 4 2023. [Online]. Available: https://doi.org/10.3390/photonics10040449

[59] C. Guclu, S. Pan, P. Nazari, B.-K. Chun, L. Gilreath, P. Heydari, and F. Capolino, "Metamaterial waveguide with loss compensation," in *2011 XXXth URSI General Assembly and Scientific Symposium*, 2011, pp. 1–4.

[60] T. Amemiya, T. Kanazawa, S. Yamasaki, and S. Arai, "Metamaterial Waveguide devices for integrated optics," *Materials*, vol. 10, no. 9, p. 1037, 9 2017. [Online]. Available: https://doi.org/10.3390/ma10091037

[61] C. Deng, E. Igarashi, and Y. Honda, "Near-zero-index waveguide for beam steering," *arXiv (Cornell University)*, 7 2023. [Online]. Available: https://arxiv.org/abs/2307.01989

[62] M. Lyu and C. F. Gmachl, "Correction to the effective refractive index and the confinement factor in waveguide modeling for quantum cascade lasers," *OSA continuum*, vol. 4, no. 8, p. 2275, 7 2021. [Online]. Available: https://doi.org/10.1364/osac.431383

[63] R. Mouthaan, P. J. Christopher, J. Pinnell, M. H. Frosz, G. S. D. Gordon, T. D. Wilkinson, and T. G. Euser, "Efficient excitation of High-Purity modes in arbitrary waveguide geometries," *Journal of Lightwave Technology*, vol. 40, no. 4, pp. 1150–1160, 2 2022. [Online]. Available: https://doi.org/10.1109/jlt.2021.3124469

[64] M. H. Weik, *waveguide dispersion*, 1 2000. [Online]. Available: https://doi.org/10.1007/1-4020-0613-6_21018

[65] P.-H. Wang, N.-L. Hou, and K.-L. Ho, "Dispersion engineering of waveguide microresonators by the design of atomic layer deposition," *Photonics*, vol. 10, no. 4, p. 428, 4 2023. [Online]. Available: https://doi.org/10.3390/photonics10040428

[66] S.-P. Wang, T.-H. Lee, Y. Chen, and P.-H. Wang, "Dispersion engineering of silicon nitride microresonators via reconstructable SU-8 polymer cladding," *Micromachines*, vol. 13, no. 3, p. 454, 3 2022. [Online]. Available: https://doi.org/10.3390/mi13030454

[67] G. Li, Y. Li, F. Ye, Q. Li, S. H. Wang, B. Wetzel, R. Davidson, B. E. Little, and S. T. Chu, "Supercontinuum generation in dispersion engineered highly doped silica glass waveguides," *arXiv (Cornell University)*, 4 2023. [Online]. Available: https://arxiv.org/abs/2304.00800

[68] D. González-Andrade, A. Dias, J. M. Luque-González, J. Gonzalo Wangüemert-Pérez, A. Ortega-Moñux, R. Halir, Molina-Fernández, P. Cheben, and A. V. Velasco, "Dispersion-engineered nanophotonic devices based on subwavelength metamaterial waveguides," in *2020 IEEE Photonics Conference (IPC)*, 2020, pp. 1–2.

[69] B. Kibler, J. M. Dudley, and S. Coen, "Supercontinuum generation and nonlinear pulse propagation in photonic crystal fiber: influence of the frequency-dependent effective mode area," *Applied Physics B*, vol. 81, no. 2-3, pp. 337–342, 7 2005. [Online]. Available: https://doi.org/10.1007/s00340-005-1844-z

[70] Z. Luo, X. Yuan, D. Man, W. Ye, J. Ji, and M. Huang, "Mapping the modal effective area in submicrometer strip and slot waveguides," in *Proceedings of 2011 International Conference on Electronics and Optoelectronics*, vol. 2, 2011, pp. V2–325–V2–327.

[71] N. A. Mortensen, "Effective area of photonic crystal fibers," *Optics Express*, vol. 10, no. 7, p. 341, 4 2002. [Online]. Available: https://doi.org/10.1364/oe.10.000341

[72] L. Zhang, X. Zhang, X. Liu, J. Zhou, N. Yang, J. Du, and X. Ding, "Thermally tunable orbital angular momentum mode generator based on Dual-Core Photonic crystal fibers," *Nanomaterials*, vol. 11, no. 12, p. 3256, 11 2021. [Online]. Available: https://doi.org/10.3390/nano11123256

[73] H. Zhang, S. Fang, J. Wang, H. Feng, H. Li, W. Dong, X. Zhang, and L. Xi, "A Hybrid Cladding Ring-Core Photonic Crystal Fibers for OAM Transmission with Weak Spin–Orbit Coupling and Strong Bending Resistance," *Photonics*, vol. 10, no. 4, p. 352, 3 2023. [Online]. Available: https://doi.org/10.3390/photonics10040352

[74] A. Debnath and N. K. Viswanathan, "Measurement of orbital angular momentum of light using Stokes parameters and Barnett39;s formalism," in *Complex Light and Optical Forces XVII*, D. L. Andrews, E. J. Galvez, and H. Rubinsztein-Dunlop, Eds., vol. 12436, International Society for Optics and Photonics. SPIE, 2023, p. 124360S. [Online]. Available: https://doi.org/10.1117/12.2647985

[75] Y. Yue, C. Liu, A. Bogoni, and I. B. Djordjević, "Guest Editorial Orbital Angular Momentum Photonics," *Journal of Lightwave Technology*, vol. 41, no. 7, pp. 1915–1917, 4 2023. [Online]. Available: https://doi.org/10.1109/jlt.2023.3246879

[76] Q. Wang, C. Zhu, S. Yan, and T. Wu, "Analysis and design of refractive index sensor based on waveguide coupling," in *2023 5th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP)*, 2023, pp. 197–200.

[77] M. Neradovskiy, H. Tronche, D. S. Chezganov, E. A. Pashnina, E. O. Vlasov, P. Baldi, T. Lunghi, V. Y. Shur, F. Doutre, and M. P. D. Micheli, "Nonlinear characterization of waveguide index profile: Application to soft-proton-exchange in linbo$_3$," *Journal of Lightwave Technology*, vol. 39, no. 14, pp. 4695–4699, 2021.

[78] S. Yeoh, K. N. Mutter, M. Z. M. Jafri, and N. M. Hasan, "Graphene based nonlinear multimode interference waveguide for refractive index sensing," *AIP Conference Proceedings*, 1 2019. [Online]. Available: https://doi.org/10.1063/1.5136470

[79] G. P. Agrawal, *Nonlinear fiber optics*, 1 1989. [Online]. Available: https://booksite.elsevier.com/samplechapters/9780123695161/Sample_Chapters/01~FrontMatter.pdf

[80] Y. R. Shen and Y. Gao, *Theory of Self-Phase Modulation and Spectral Broadening*, 1 1989. [Online]. Available: https://doi.org/10.1007/978-1-4757-2070-9_1

[81] R. R. Alfano, *The Supercontinuum Laser Source*. Springer.

[82] C. Huang, F. Kou, K. Peng, T. Tu, S. Li, M. Guo, G. Yu, Y. Zhou, W. Bi, S. Zheng, C. Zhang, B. Zheng, and J. Wang, "Engineering the high-frequency components of coherent supercontinuum generation in hybrid optical fibers with yttrium aluminum garnet core," *Results in Physics*, vol. 34, p. 105233, 3 2022. [Online]. Available: https://doi.org/10.1016/j.rinp.2022.105233

[83] R. Paschotta, "Supercontinuum generation," RP Photonics Encyclopedia, Jul 2008, available online at https://www.rp-photonics.com/supercontinuum_generation.html. [Online]. Available: https://www.rp-photonics.com/supercontinuum_generation.html

[84] Y. R. Shen, *Stimulated raman scattering*, 1 1983. [Online]. Available: https://doi.org/10.1007/3-540-11913-2_7

[85] Y. Ozeki, K. Itoh, K. Nose, and C. Inc, "EP2681535A4 - Stimulated raman scattering detection apparatus - Google Patents," 3 2011. [Online]. Available: https://patents.google.com/patent/EP2681535A4/en

[86] M. Maier, "Applications of stimulated Raman scattering," *Applied physics*, vol. 11, no. 3, pp. 209–231, 11 1976. [Online]. Available: https://doi.org/10.1007/bf00897057

[87] I. Perez-Arjona, G. J. de Valcarcel, and E. Roldan, "Two-photon absorption," 2004.

[88] M. Kuna, *Finite element method*, 1 2013. [Online]. Available: https://doi.org/10.1007/978-94-007-6680-8_4

[89] O. Axelsson and V. A. Barker, *The finite element method*, 1 1984. [Online]. Available: https://doi.org/10.1016/b978-0-12-068780-0.50011-x

[90] M. A. Jusoh, Z. Abbas, M. A. A. Rahman, C. E. Meng, M. F. Zainuddin, and F. Esa, “Critical study of open-ended coaxial sensor by finite element method (fem),” *International Journal of Applied Science and Engineering*, vol. 11, pp. 343–360, December 2013. [Online]. Available: https://doi.org/10.6703/IJASE.2013.11(4).343

[91] B. Archambeault, O. M. Ramahi, and C. Brench, *The Finite-Difference Time-Domain method*, 1 1998. [Online]. Available: https://doi.org/10.1007/978-1-4757-5124-6_3

[92] M. A. Swillam, *Finite-Difference Time-Domain method in photonics and nanophotonics*, 12 2017. [Online]. Available: https://doi.org/10.1201/b16319-1

[93] J. W. Gooch, *Sellmeier Equation*, 1 2011. [Online]. Available: https://doi.org/10.1007/978-1-4419-6247-8_10447

[94] B. Li, Y. Li, and W. Zheng, “A new perfectly matched layer method for the Helmholtz equation in nonconvex domains,” *Siam Journal on Applied Mathematics*, vol. 83, no. 2, pp. 666–694, 4 2023. [Online]. Available: https://doi.org/10.1137/22m1482524

[95] S. Zhang, Z. Yang, B. Liu, J. Luo, and Z. H. Hang, “Optical perfectly matched layers based on the integration of photonic crystals and material loss,” *Optics Express*, vol. 31, no. 7, p. 11080, 3 2023. [Online]. Available: https://doi.org/10.1364/oe.486253

[96] “Meshing.” [Online]. Available: https://doc.comsol.com/5.5/doc/com.comsol.help.cfd/cfd_ug_quickstart.05.07.html

[97] C. Fan, L. Li, and F. Yu, “Soliton solution, breather solution and rational wave solution for a generalized nonlinear Schrödinger equation with Darboux transformation,” *Scientific Reports*, vol. 13, no. 1, 6 2023. [Online]. Available: https://doi.org/10.1038/s41598-023-36295-x

[98] F. Dong, B. Hu, and Q. Li, “Riemann-Hilbert problem associated with an integrable generalized mixed nonlinear Schrödinger equation,” *Complex Variables and Elliptic Equations*, pp. 1–20, 3 2023. [Online]. Available: https://doi.org/10.1080/17476933.2023.2185883

[99] U. Younas, M. Z. Baber, M. W. Yasin, T. A. Sulaıman, and J. Ren, "The generalized higher-order nonlinear Schrödinger equation: Optical solitons and other solutions in fiber optics," *International Journal of Modern Physics B*, vol. 37, no. 18, 12 2022. [Online]. Available: https://doi.org/10.1142/s0217979223501746

[100] "Split step fourier method based pulse propagation model for nonlinear fiber optics," 4 2007. [Online]. Available: https://ieeexplore.ieee.org/document/4287333

[101] A. Sharma and S. Deb, "lt;titlegt;Numerical method for solving the generalized equation for nonlinear pulse propagation through optical fiberslt;/titlegt;," *Proceedings of SPIE*, 12 1992. [Online]. Available: https://doi.org/10.1117/12.637005

[102] F. Bre, J. M. Giménez, and V. D. Fachinotti, "Prediction of wind pressure coefficients on building surfaces using artificial neural networks," *Energy and Buildings*, vol. 158, pp. 1429–1441, 1 2018. [Online]. Available: https://doi.org/10.1016/j.enbuild.2017.11.045

[103] A. Zhu, "Artificial neural networks," *International Encyclopedia of Geography*, pp. 1–6, 3 2017. [Online]. Available: https://doi.org/10.1002/9781118786352.wbieg0871

[104] J. Dancker, "A brief introduction to recurrent neural networks - towards data science," 1 2023. [Online]. Available: https://towardsdatascience.com/a-brief-introduction-to-recurrent-neural-networks-638f64a61ff4

[105] S. Ahlawat, *Recurrent neural networks*, 12 2022. [Online]. Available: https://doi.org/10.1007/978-1-4842-8835-1_4

[106] S. Arifin, A. Wijaya, R. Nariswari, I. G. A. A. Yudistira, S. Suwarno, F. Faisal, and D. Wihardini, "Long Short-Term Memory (LSTM): Trends and future research potential," *International journal emerging technology and advanced engineering*, vol. 13, no. 5, pp. 24–34, 5 2023. [Online]. Available: https://doi.org/10.46338/ijetae0523_04

[107] X. Chen, "The study for convolutional neural network and corresponding applications," *Theoretical and Natural Science*, vol. 5, no. 1, pp. 182–187, 5 2023. [Online]. Available: https://doi.org/10.54254/2753-8818/5/20230387

[108] N. Moustafa, "Decoding the CNN Architecture: Unveiling the power and precision of convolutional neural networks - Part I," 9 2023. [Online]. Available: https://www.linkedin.com/pulse/decoding-cnn-architecture-unveiling-power-precision-neural-moustafa

[109] R. Hasani, M. Lechner, A. Amini, D. Rus, and R. Grosu, "Liquid time-constant networks," *Proceedings of the ... AAAI Conference on Artificial Intelligence*, vol. 35, no. 9, pp. 7657–7666, 5 2021. [Online]. Available: https://doi.org/10.1609/aaai.v35i9.16936

[110] M. Bidollahkhani, F. Atasoy, and H. Abdellatef, "LTC-SE: Expanding the potential of Liquid Time-Constant Neural Networks for scalable AI and embedded systems," *arXiv (Cornell University)*, 4 2023. [Online]. Available: https://arxiv.org/abs/2304.08691

[111] R. Hasani, M. Lechner, A. Amini, L. Liebenwein, A. Ray, M. Tschaikowski, G. Teschl, and D. Rus, "Closed-form continuous-time neural networks," *Nature Machine Intelligence*, vol. 4, no. 11, pp. 992–1003, 11 2022. [Online]. Available: https://doi.org/10.1038/s42256-022-00556-7

[112] S. E. Whang, Y. Roh, H. Song, and J.-G. Lee, "Data collection and quality challenges in deep learning: a data-centric AI perspective," *The VLDB Journal*, vol. 32, no. 4, pp. 791–813, 1 2023. [Online]. Available: https://doi.org/10.1007/s00778-022-00775-9

[113] A. G. Ororbia, A. Mali, D. Kifer, and C. L. Giles, "Backpropagation-Free Deep Learning with Recursive Local Representation Alignment," *Proceedings of the ... AAAI Conference on Artificial Intelligence*, vol. 37, no. 8, pp. 9327–9335, 6 2023. [Online]. Available: https://doi.org/10.1609/aaai.v37i8.26118

[114] "What is Learning Rate in Machine Learning — Deepchecks," 12 2022. [Online]. Available: https://deepchecks.com/glossary/learning-rate-in-machine-learning/#:~:text=The%20learning%20rate%2C%20denoted%20by,network%20concerning%20the%20loss%20gradient%3E.

[115] V. Lyashenko, “Cross-Validation in Machine Learning: How to do it right,” 12 2023. [Online]. Available: https://neptune.ai/blog/cross-validation-in-machine-learning-how-to-do-it-right

[116] P. Baheti, “Activation Functions in Neural Networks [12 Types amp; Use Cases],” 4 2023. [Online]. Available: https://www.v7labs.com/blog/neural-networks-activation-functions#:~:text=drive%20V7’s%20tools.-,What%20is%20a%20Neural%20Network%20Activation%20Function%3F,prediction%20using%20simpler%20mathematical%20operations.

[117] A. Gupta, “A comprehensive guide on optimizers in deep learning,” 1 2024. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-deep-learning-optimizers/

[118] Rendyk, “Tuning the hyperparameters and layers of neural network deep learning,” 1 2024. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/05/tuning-the-hyperparameters-and-layers-of-neural-network-deep-learning/#:~:text=Frequently%20Asked%20Questions-,Neural%20Network%20Hyperparameters%20(Deep%20Learning),it%20adjusts%20its%20internal%20values.

[119] B. O. C. Science and B. O. C. Science, “What is a learning curve in machine learning? — Baeldung on Computer Science,” 5 2023. [Online]. Available: https://www.baeldung.com/cs/learning-curve-ml

[120] J. Brownlee, “How to use Learning Curves to Diagnose Machine Learning Model Performance,” 8 2019. [Online]. Available: https://machinelearningmastery.com/learning-curves-for-diagnosing-machine-learning-model-performance/

[121] “What is Overfitting? - Overfitting in Machine Learning Explained - AWS.” [Online]. Available: https://aws.amazon.com/what-is/overfitting/#:~:text=Overfitting%20occurs%20when%20the%20model,to%20several%20reasons%2C%20such%20as%3A&text=The%20training%20data%20size%20is,all%20possible%20input%20data%20values.

[122] "Mean Square Error (MSE) — Machine Learning Glossary — EnCorD — EnCoRD." [Online]. Available: https://encord.com/glossary/mean-square-error-mse/#:~:text=It%20measures%20the%20average%20squared,align%20with%20the%20ground%20truth.

[123] M. R. Karim, B. M. A. Rahman, and G. P. Agrawal, "Mid-infrared supercontinuum generation using dispersion-engineered ge11.5as24se64.5 chalcogenide channel waveguide," *Opt. Express*, vol. 23, no. 5, pp. 6903–6914, Mar 2015. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-23-5-6903

[124] C. Mei, J. Yuan, F. Li, B. Yan, X. Sang, Q. Wu, X. Zhou, K. Wang, C. Yu, and G. Farrell, "Efficient spectral compression of Wavelength-Shifting soliton and its application in integratable All-Optical quantization," *IEEE Photonics Journal*, vol. 11, no. 1, pp. 1–15, 2 2019. [Online]. Available: https://doi.org/10.1109/jphot.2018.2890424

[125] F. E. Seraji, A. Emami, D. R. Raf, and P. Sattari, "Dispersion characterization of photon crystal fiber using fully-vectorial effective index method," *Physics astronomy international journal*, vol. 4, no. 4, pp. 139–144, 7 2020. [Online]. Available: https://doi.org/10.15406/paij.2020.04.00212

[126] J. M. Dudley, G. Genty, and S. Coen, "Supercontinuum generation in photonic crystal fiber," *Reviews of Modern Physics*, vol. 78, no. 4, pp. 1135–1184, 10 2006. [Online]. Available: https://doi.org/10.1103/revmodphys.78.1135

[127] I. Rebolledo, Z. Ye, S. L. Christensen, F. Lei, K. Twayana, J. Schroeder, M. Zelán, and V. Torres-Company, "Coherent supercontinuum generation in all-normal dispersion Si3N4 waveguides," *Optics Express*, vol. 30, no. 6, p. 8641, 3 2022. [Online]. Available: https://doi.org/10.1364/oe.450987

[128] M. Yu, B. Desiatov, Y. Okawachi, A. L. Gaeta, and M. Lončar, "Coherent two-octave-spanning supercontinuum generation in lithium-niobate waveguides," *Optics Letters*, vol. 44, no. 5, p. 1222, 2 2019. [Online]. Available: https://doi.org/10.1364/ol.44.001222

[129] G. Pandraud, H. Pham, P. French, and P. Sarro, "PECVD SiC optical waveguide loss and mode characteristics," *Optics and Laser Technology*, vol. 39,

no. 3, pp. 532–536, 4 2007. [Online]. Available: https://doi.org/10.1016/j.optlastec.2005.10.014

[130] F. De Leonardis, R. A. Soref, and V. M. N. Passaro, “Dispersion of nonresonant third-order nonlinearities in Silicon Carbide,” *Scientific Reports*, vol. 7, no. 1, 1 2017. [Online]. Available: https://doi.org/10.1038/srep40924

[131] M. Belt, M. Davenport, J. E. Bowers, and D. Blumenthal, “Ultra-low-loss Ta2O5 core SiO2 clad planar waveguides on Si substrates,” *Optica*, vol. 4, no. 5, p. 532, 5 2017. [Online]. Available: https://doi.org/10.1364/optica.4.000532

[132] X. Xiong, S. Zheng, Z. Zhu, Y. Chen, Z. Wang, X. Yu, X. Jin, and X. Zhang, “(13) oam mode-group generation method: Partial arc transmitting scheme,” *arXiv: Signal Processing*, 2020.

[133] D. Vigneswaran, M. S. M. Rajan, B. Biswas, A. Grover, K. Ahmed, and B. K. Paul, “Numerical investigation of spiral photonic crystal fiber (S-PCF) with supporting high order OAM modes propagation for space division multiplexing applications,” *Optical and Quantum Electronics*, vol. 53, no. 2, 1 2021. [Online]. Available: https://doi.org/10.1007/s11082-021-02733-1

[134] Y. Yu, Y. Lian, Q. Hu, L. Xie, J. Ding, Y. Wang, and Z. Lu, “Design of PCF Supporting 86 OAM Modes with High Mode Quality and Low Nonlinear Coefficient,” *Photonics*, vol. 9, no. 4, p. 266, 4 2022. [Online]. Available: https://doi.org/10.3390/photonics9040266

[135] C.-S. Brès, A. D. Torre, D. Grassani, V. Brasch, C. Grillet, and C. Monat, “(3) supercontinuum in integrated photonics: generation, applications, challenges, and perspectives,” *Nanophotonics*, 2023.

[136] Y. Cao, B.-U. Sohn, H. Gao, P. Xing, G. F. R. Chen, D. K. T. Ng, and D. T. H. Tan, “(4) supercontinuum generation in a nonlinear ultra-silicon-rich nitride waveguide,” *Dental science reports*, 2022.

[137] C. M. de Sterke, *Supercontinuum generation*. Cambridge University Press, 1 2022. [Online]. Available: https://doi.org/10.1017/9781009067652.020

[138] M. Hary, L. Salmela, J. M. Dudley, and G. Genty, “Machine learning control of nonlinear fiber supercontinuum generation for application in molecular

spectroscopy," *Optica Advanced Photonics Congress 2022*, 1 2022. [Online]. Available: https://doi.org/10.1364/np.2022.nptu1g.2

MSC_THESIS_BOOK (2).pdf

ORIGINALITY REPORT

24%	12%	20%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Rakayet Rafi, M.R. Karim, Sampad Ghosh, B.M.A. Rahman. "Unified predictive modeling of supercontinuum spectra: Using multi-material data with Closed-Form Continuous Time Neural Networks", Optics Communications, 2024 Publication	5%
2	M.R. Karim, Rakayet Rafi, Dipta Dutta, Md. Iftekher Hossain Upal, Sampad Ghosh. "Employing deep learning for predicting supercontinuum generation in chalcogenide planar waveguide", Optik, 2024 Publication	3%
3	www.mdpi.com Internet Source	1%
4	Md. Iftekher Hossain Upal, Dipta Dutta, Rakayet Rafi, Mohammad Rezaul Karim, Mohammad Rafiqul Alam, Sampad Ghosh. "Deep Learning Approach to Determine the Optical Characteristics of Photonic Crystal Fiber for Orbital Angular Momentum	1%