INVERSE DESIGN OF METAMATERIALS FOR TAILORED LINEAR AND NONLINEAR OPTICAL RESPONSES USING DEEP LEARNING

A Dissertation
Presented to
The Academic Faculty

by

Lakshmi Raju

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
December 2022

INVERSE DESIGN OF METAMATERIALS FOR TAILORED LINEAR AND NONLINEAR OPTICAL RESPONSES USING DEEP LEARNING

Approved by:

Dr. Wenshan Cai, Advisor
School of Electrical and Computer Engineering, School of Material Science and Engineering
Georgia Institute of Technology

Dr. Andrew Peterson
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Azad Naeemi
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Ali Adibi
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Zhuomin Zhang
School of Mechanical Engineering
Georgia Institute of Technology

Date Approved:[September 22, 2022]

To my family and friends, whose support made this possible.

ACKNOWLEDGEMENTS

First, I would like to thank my advisor Professor Wenshan Cai. After joining his lab, I learned more than I ever imagined. His support and advice helped me reach my goal of completing this dissertation. I am also indebted to the kindness and help from my former lab mates, Dr. Sean Rodrigues, and Dr. Mohammad Taghinejad. Dr. Zhaocheng Liu and Dr. Kyu-tae Lee were instrumental in my success and were generous with their advice. To my lab mates who grew in the lab with me, Andrew Kim and Dayu Zhu, thank you for all your support! I would also like to wish good luck to our newest lab members, Anjan Ghoswani, and Chiyu Yang. My collaborators at the Air Force Research Lab – Dr. Augustine Urbas and Dr. Ekaterina Poutrina helped me learn so much about nonlinear optics. This research would not have been possible without my funding from the National Science Foundation Graduate Research Fellowship (Grant No. DGE-2039655). I would also like to thank my thesis committee for their support: Dr. Azad Naeemi, Dr. Zhuomin Zhang, Dr. Andrew Peterson, and Dr. Ali Adibi.

I was incredibly fortunate to also have mentors outside of my research. Thank you, Dr. Joy Harris, Anna Holcomb, Christina Bourgeois, Dr. Elliot Moore, Dr. Tammy McCoy – GT CETL, Dr. Rebecca Pope-Ruark – GT CETL, Tonia Valeck – GT WIE, Dr. Magnus Egerstedt, Dr. Phoebe Bronstein – GT LMC, Angela Elleby, Barbara Howard, and Dr. Daniela Staiculescu. You helped me grow my teaching and mentorship skills, and I will be forever grateful for your guidance. My work outside of research, mentoring students, giving tours, reshaping the student community in ECE, and teaching were all very special to me. I spent my entire university career at Georgia Tech and will be wrapping up my

third degree in December 2022, making it 9.5 years in total. I would like to thank the Institute for all that it has given me.

As an undergraduate student, I did not lack in research mentors. I would like to thank Dr. Tom Gaylord – GTECE, Dr. Chris Valenta - GTRI, Dr. Gisele Bennett - GTRI, Dr. Ben Klein - GTECE, Robert Drupp – Northrop Grumman, Aaron Cordes – Lockheed Martin, Dr. Brian Uhlhorn – Lockheed Martin, Dr. Greg Whaley – Lockheed Martin, Dr. Alan Hsu – MIT Lincoln Laboratory.

Science Fair was a huge part of my life before college, helped support my undergraduate education through scholarships, and furthered my passion for science. I would not have been able to as well as I did without the help of Dr. Virginia Vilardi, Dr. Jack Rogers – UAB, and Dr. Alexander Scheeline - UIUC.

My friends were a great support throughout this journey, but I would like to especially thank Emily Miller, who never doubted me.

Most importantly I would like to thank my family. My parents, Dr. Raghavan Raju and Sindhu Pillai, immigrated to this country with hopes of their children having great opportunities available to them. I believe that I have taken advantage of as many of them as I could and would not be here without them. I would be remiss if I didn't thank the feline addition to my family, J.J. Pillai, for his never-ending quest for comfort. I wish the best of luck to my brother, Srikrishnan Raju, my polar opposite, as he begins his journey towards grad school as well.

Thank you to all who have supported, helped, and guided me throughout this process. While it was difficult, I am proud to have written this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ALD	Atomic Layer Deposition
ANN	Artificial Neural Network
CCD	Charge-coupled device
CD	Circular Dichroism
CNN	Convolutional Neural Network
CPPN	Compositional Pattern-producing network
DOF	Degree of Freedom
ES	Evolution Strategy
GPN	Geometry Predicting Network
LSP	Localized Surface Plasmon
NIR	Near Infrared
PDF	Probability Density Function
PMMA	poly(methyl methacrylate)
SFG	Sum-frequency Generation
SHG	Second-harmonic Generation
SPN	Spectrum Predicting Network
SPP	Surface Plasmon Polariton
SRR	Split Ring Resonator

VAE Variational Auto-encoder

SUMMARY

The conventional process for developing an optimal design for nonlinear optical responses is based on a trial-and-error approach that is largely inefficient and does not necessarily lead to an ideal result. Deep learning can automate this process and widen the realm of nonlinear geometries and devices. This research illustrates a deep learning framework used to create an optimal plasmonic design for metamaterials with specific desired optical responses, both linear and nonlinear. The algorithm can produce plasmonic patterns that can maximize second-harmonic nonlinear effects of a nonlinear metamaterial. A nanolaminate metamaterial is used as a nonlinear material, and a plasmonic patterns are fabricated on the prepared nanolaminate to demonstrate the validity and efficacy of the deep learning algorithm for second-harmonic generation. Photonic upconversion from the infrared regime to the visible spectrum can occur through sum-frequency generation. The deep learning algorithm was improved to optimize a nonlinear plasmonic metamaterial for sum-frequency generation. The framework was then further expanded using transfer learning to lessen computation resources required to optimize metamaterials for new design parameters. The deep learning architecture applied in this research can be expanded to other optical responses and drive the innovation of novel optical applications.

CHAPTER 1. INTRODUCTION

The goal of this thesis research was to create trained deep learning algorithms to inversely design metamaterial structures for optimal linear and nonlinear optical responses. The research shows the capabilities for linear, second harmonic, and sum-frequency nonlinear optical responses. Expanding on the deep learning algorithms, the research shows the ability to boost the framework to be able to handle linear and nonlinear optical responses with a wider set of parameters.

1.1 Background

In general, inverse design in photonics is a method used to identify a set of parameters to define a photonic structure or device given the desired optical responses. This contrasts with traditional design processes of a back and forth, trial and error method starting with candidate patterns from empirical guesses, using parametric sweeps to slightly adjust parameters, and at times never reaching the desired goal due to limitations such as computing power, or geometric candidates.

Since the use of artificial neural networks (ANN) for inverse design can lead to fast, accurate, and sometimes counterintuitive results that cannot be achieved by traditional methods, it has been at the forefront of many research problems as large as wind turbine design, and even for smaller applications such as plasmonics [1-5].

The essence of deep learning is to learn from sizable datasets using algorithms loosely arising from models of biological nervous systems [6-10]. Deep learning has evolved quite remarkably in the last few years, outgrowing the designation as a subset of machine learning [11, 12]. The basic application of a deep learning model is that once

trained on a labeled set of data, can perform classification of data using ANNs with multiple layers of artificial neurons – hence the term 'deep' learning. The power of the ANN's increasing complexity allows for the achievement of accuracy that can equal and sometimes even exceed human performance [13-15].

Plasmonics is a field that studies and utilizes the optical properties of metal-dielectric interfaces on nanometer-scale structures. Plasmonic structures can serve as unit cells for metamaterials to find new phenomena, or to prove existing theories. Periodic unit cells, such as a plasmonic structure, lead to a metamaterial whose properties are drawn from the periodic structure, as well as the individual cells [16-18].

When designing plasmonic metamaterials and metasurfaces for the purposes of new phenomena, the linear optical regime, such as spectral control, dispersion engineering, and beam steering, has been well-researched, so can be used to validate a deep learning framework [19-21]. In the nonlinear regime, there is much to be explored.

Optical nonlinear processes include frequency mixing, such as second-harmonic generation (SHG), sum-frequency generation, optical rectification, as well as the Pockels effect and optical Kerr effect. Nonlinear optics is essential for the generation of new spectral components and active control of light.

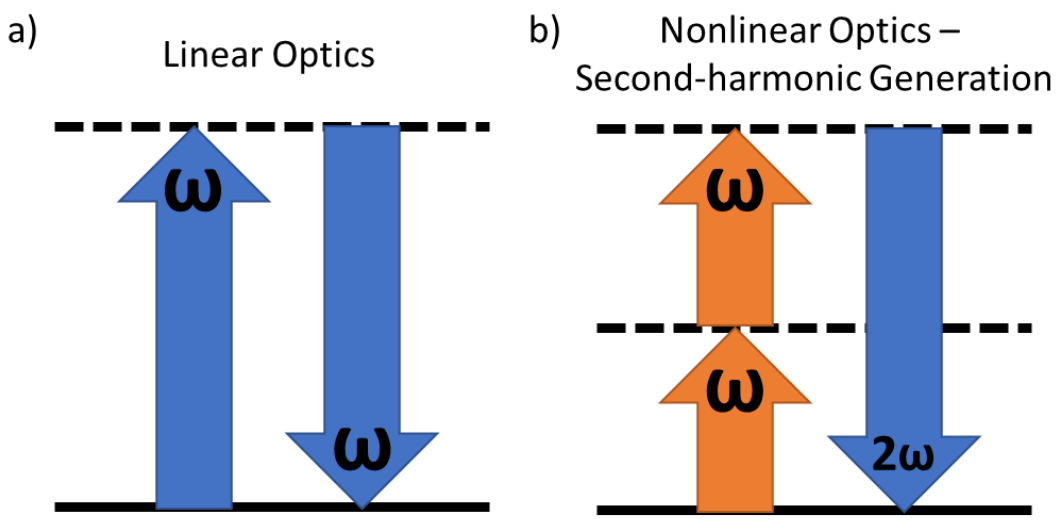

Figure 1.1 – Energy level description of a) linear optics such as reflection where the frequency, ω, of the light remains the same, and b) nonlinear optics such as second-harmonic generation, where the frequency of the input light is doubled to 2ω. The upwards pointing arrows indicate the incident light, and the downward pointing arrows are indicative of the output.

1.2 Tailored optical responses

The most important factor in linear optics is that monochromatic light entering a linear system will maintain its frequency, the light may be delayed, polarized, or otherwise acted upon but the frequency will be unchanged. This is illustrated in Fig 1.1a. In linear optics, the polarization of the material system, or the dipole moment per unit volume, describes the linear relationship between induced polarization and the electric field strength,

$$\boldsymbol{P} = \varepsilon_0 \chi^{(1)} \boldsymbol{E}. \tag{2.1}$$

In this polarization equation, the ε_0 represents the permittivity of free space, $\chi^{(1)}$ is the linear susceptibility, and $\boldsymbol{E}$ is the electric field strength [22].

Linear optical phenomena such as spectral control, dispersion engineering, and beam steering take advantage of the linear material properties like the refractive index of materials which is related to the linear susceptibility. To tailor linear optical responses, linear full wave simulations depend upon the linear material properties for phase control or amplitude modulation.

On the other hand, nonlinear optics is when the polarization, $\boldsymbol{P}$, non-linearly depends on the electric field strength, $\boldsymbol{E}$. For nonlinear optics, the polarization is expressed as a power series:

$$\boldsymbol{P} = \varepsilon_0\left[\chi^{(1)}\boldsymbol{E} + \chi^{(2)}\boldsymbol{E^2} + \chi^{(1)}\boldsymbol{E^3} + \cdots\right]$$

$$= \boldsymbol{P^1} + \boldsymbol{P^2} + \boldsymbol{P^3} + \cdots . \tag{2.2}$$

Here, the $\chi^{(n)}$ terms represent the n^{th} order nonlinear optical susceptibilities, so the second- and third-order, etc. All the $\chi^{(n)}$ terms are in fact tensors that depend on the applied field frequencies, as will be shown later. To refer to a specific polarization, for example, the second-order nonlinear polarization, we use:

$$\boldsymbol{P^2} = \varepsilon_0\chi^{(2)}\boldsymbol{E^2}. \tag{2.3}$$

General second-order nonlinear processes include second-harmonic generation, and sum-frequency generation. Both of these are frequency mixing processes that occur due to the applied electric field interacting with the nonlinear properties of the material allowing

for the generation of radiation at the second-harmonic frequency, as shown in Figure 1.1b, or sum-frequency [22].

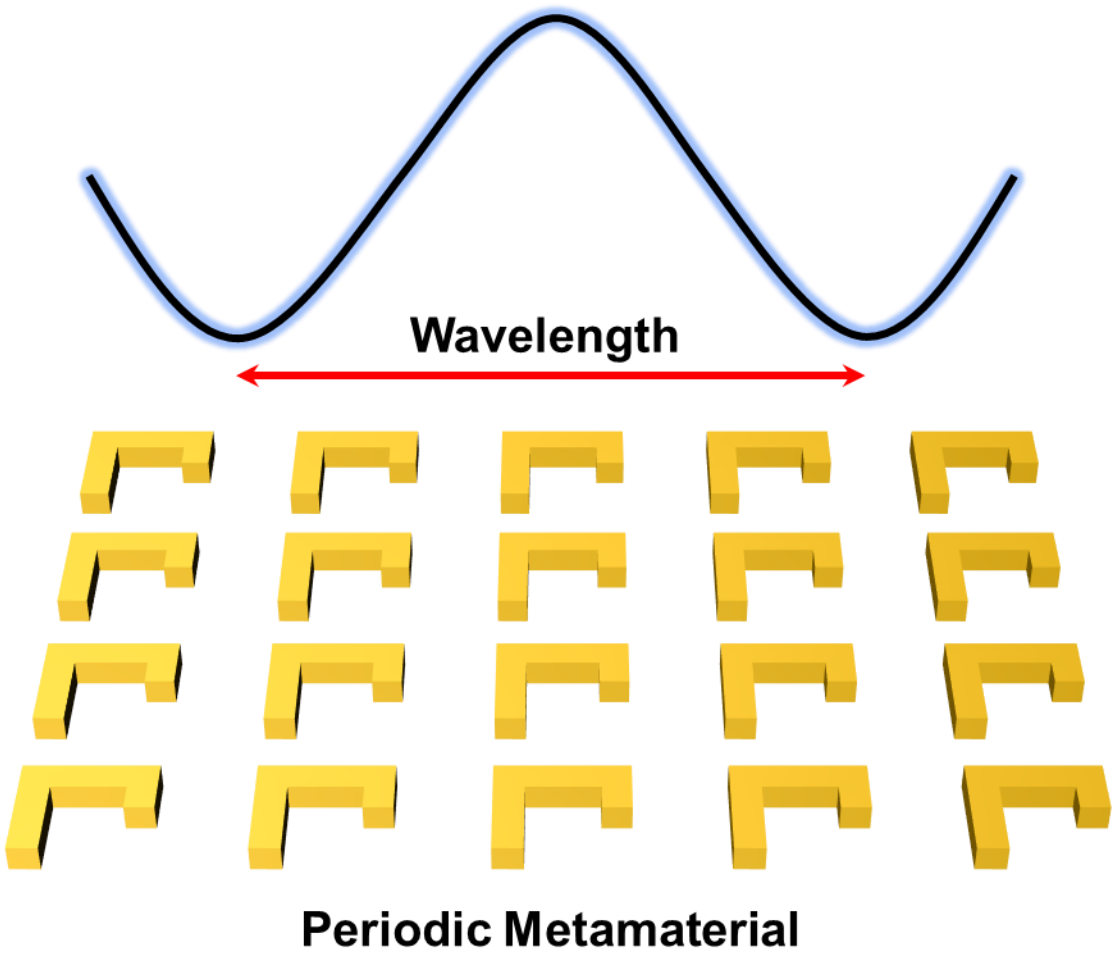

Figure 1.2 – Example of a metamaterial that is periodic at a sub-wavelength scale. The sub-wavelength scale is an important property of metamaterials.

1.3 Metamaterials

A metamaterial a material engineered to have electromagnetic functions usually not found in conventional materials, that derives its properties from the unit structure or meta-atoms and periodicity [23-24]. Another distinguishing factor is that these meta-

atoms, and periodicity are at a subwavelength scale, as shown in Figure 1.2. Metamaterials can be designed for both linear and nonlinear optical responses.

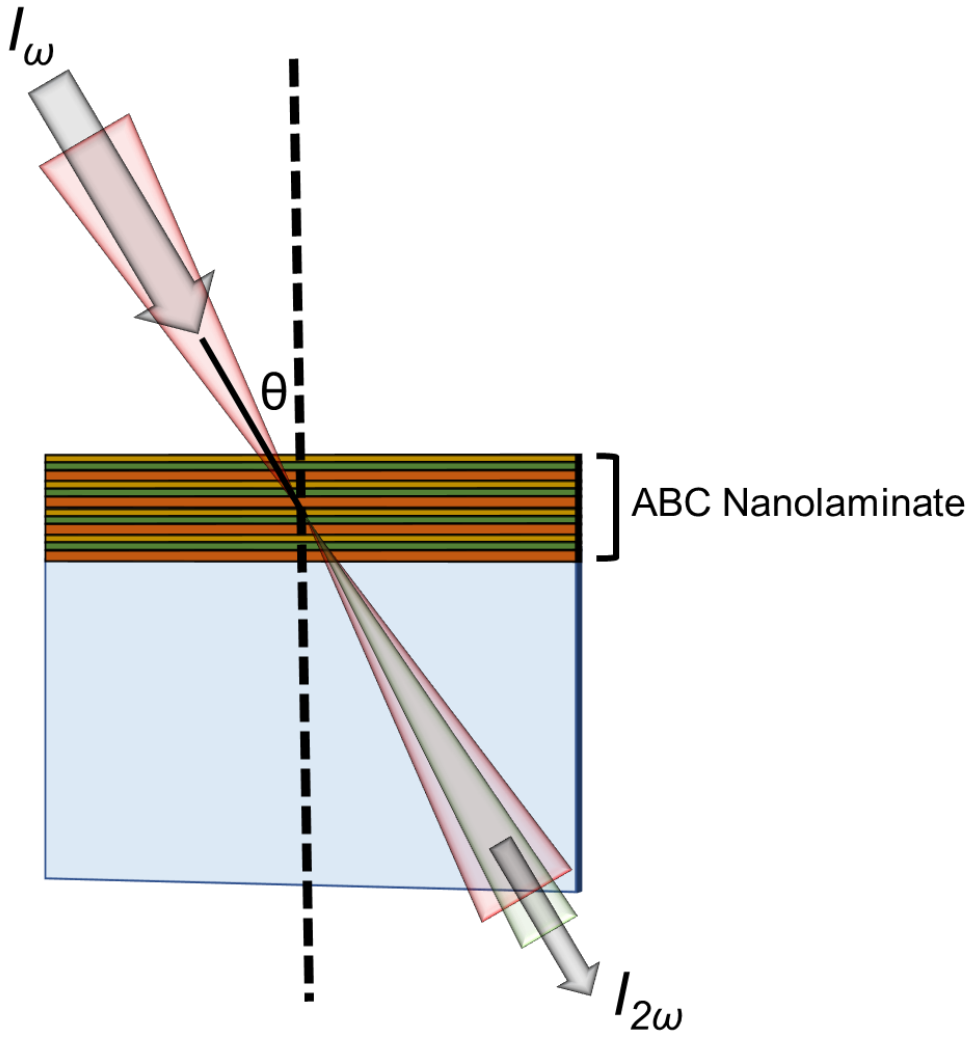

Figure 1.3 – Schematic of the ABC nanolaminate stack with second order nonlinear properties when light is incident at an angle.

While bulk thin-film nonlinear materials like $LiNbO_3$ are commercially available, an easy to fabricate and thinner film is possible using an ABC nanolaminate [25-28]. An ABC nanolaminate is a nonlinear metamaterial created using a repetition of thin films of

three different layers, an A, B, and C layer, repeated as ABCABC, as illustrated in Figure 1.3.

When using materials that have inversion symmetry, such as silicon, under the electron dipole approximation, there are no even order nonlinear effects, such as second-order effects, meaning there is no nonlinear susceptibility of the second type. At each of the layer interfaces, such as at A-B, or B-C, centrosymmetry is broken in the axis normal to the layers.

Once the centrosymmetry is broken, there can be second order nonlinear effects, or a nonlinear susceptibility in the bulk material along the perpendicular direction [25-28]. To utilize the break in centrosymmetry, the incident light is at an angle in order to possess a non-zero field component along the axial direction, inducing a second order nonlinear susceptibility.

With a TiO_2-Al_2O_3-HfO_2 ABC nanolaminate, following the material and geometric parameters from the successful demonstration by Alloatti et al., the manmade nonlinear oxide composite requires a simple deposition technique, such as atomic layer deposition (ALD), and can have an easily controllable thickness on the order of tens of nanometers, whereas thin-film bulk materials require complicated growth methods to get a crystalline structure, and typically the thickness is on the order of hundreds of nanometers [25].

The ABC nanolaminate is also a better choice, as it is compatible with silicon, which opens the door to numerous integrated optics applications on CMOS compatible platforms. Silicon is very important for integrated optics but is limited for second order nonlinear optics due to its centrosymmetry.

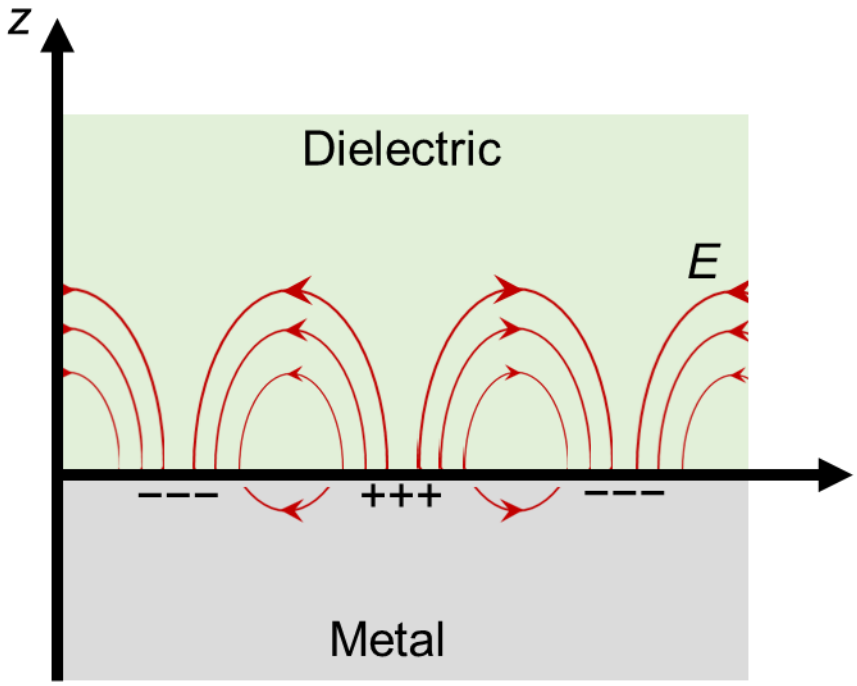

Figure 1.4 – Illustration of surface plasmons. At a metal-dielectric interface, the electromagnetic field at the interface excites and causes collective coherent oscillations (surface plasmons) of free electrons leading to a created charge density oscillations called surface plasmon polaritons.

1.4 Plasmonic structures

Plasmonic structures are often used in metamaterials as the surface plasmon effect can be taken advantage of. To explain this, we begin with surface plasmon polaritons (SPPs) [29]. The conduction band of metals, such as gold and silver, have free electrons. When a plasmonic structure creates a metal-dielectric interface, the electromagnetic field at the interface excites and causes collective coherent oscillations of free electrons, as shown in Figure 1.4. The oscillations lead to created charge density oscillations which we call SPPs [30]. The coherent oscillations of the free electrons are the surface plasmons.

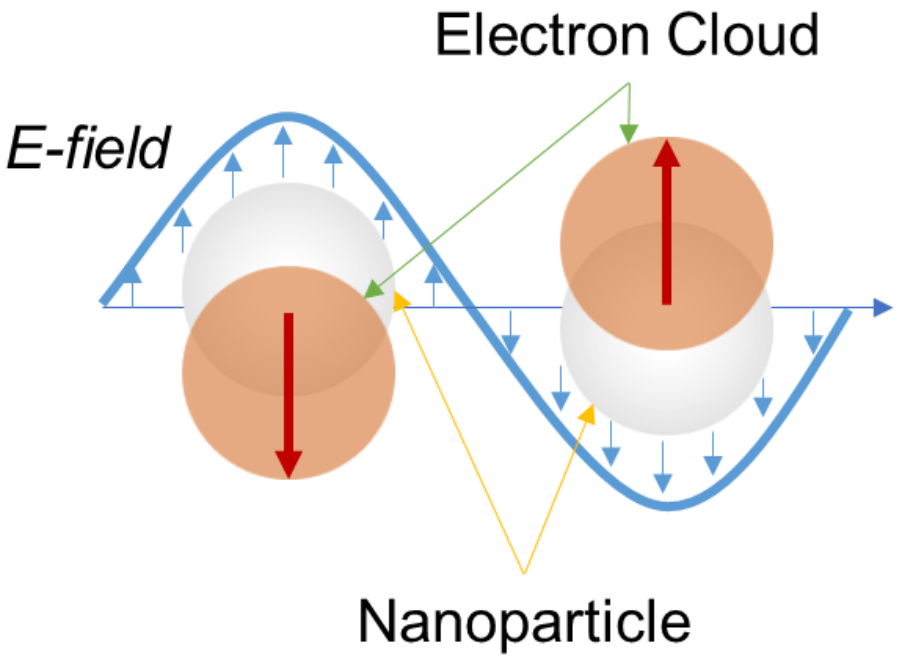

Figure 1.5 – Illustrated here is the confinement of surface plasmons in a nanoparticle leading to localized surface plasmons (LSPs).

The SPPs are electromagnetic waves that propagate along the metal-dielectric surface. As the wavelength of the SPPs are shorter than the incident wavelength, they can have large field enhancements, as described in Figure 1.5. The tight electric field confinement leads to a plasmon resonant frequency that can be tuned by the plasmonic geometry, refractive index of the plasmonic and surrounding material [31]. Characterizing the surface plasmons can be done through a few techniques including surface enhanced Raman scattering [32].

For metamaterials, periodic unit cells with plasmonic structures can lead to interesting optical phenomena, both in the linear and nonlinear regime. The plasmonic

structures have an effect even in the nonlinear regime, as the nonlinear polarization, ***P***, is still also dependent on the linear optical response.

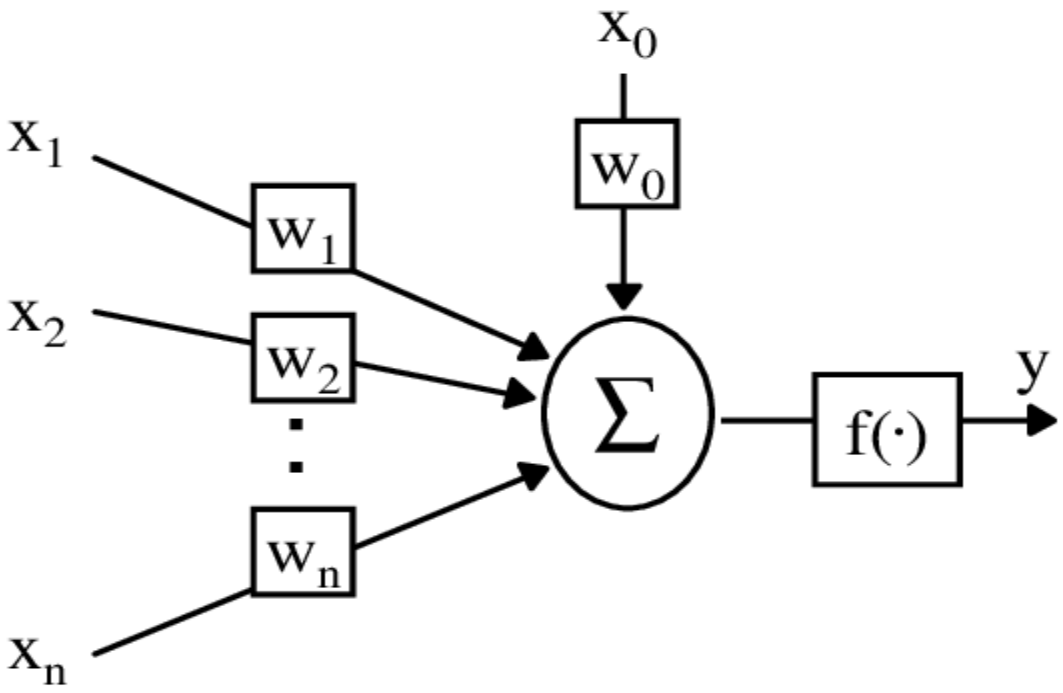

Figure 1.6 – Illustration of the MuCulloch-Pitts Neuron. McCulloch and Pitts were the first to put forth a mathematical mode for a neural network [33].

1.5 Deep Learning

While deep learning as a field is relatively new (~20 years), it's basis can be traced back to 1943 when Walter Pitts and Warren McCulloch created a mathematical neural network model, the McCullock-Pitts neuron, as illustrated in Figure 1.6 [34]. Even so, it wasn't until 1959 when machine learning was coined and created. Arthur Samuel is considered the father of machine learning, and it all began with computer programs

designed to play checkers [35]. From there, the field was born, and grew to include artificial neural networks, convolutional neural networks, gradient based learning, and much more.

For the inverse design of photonic structures, two separate deep learning methods are needed. One to generate patterns, beyond conventional geometries, and another to predict the optical response of patterns and optimize a pattern for the required needs. There are two general types of models – generative and discriminative. A generative model uses a probabilistic model to describe a data set. Using the probabilistic model, sampling can allow the generation of new data.

The other type is a discriminative model, this is more of a predictive model. The discriminative model, once trained on a dataset with classes, identifies the boundaries between the classes using conditional probability. For example, if a data set has a number of labelled animal images, “Dog”, “Cat”, “Cow”, etc., the model creates probability vectors for each class of animal. Given a new picture to classify, the model labels the image using the conditional probabilities.

For pattern generation, it seems natural to lean on a generative model, such as a variational auto-encoder. In order to predict the optical response of a photonic structure pattern as an image, a convolutional neural network(CNN) is perfect [13]. CNNs are designed to process arrays, like images, and find patterns in the data, through the use of multiple convolutional layers. The convolutional layers are made of convolutional kernels which identify patterns in the images in subsections. All of the kernels and layers working together yield data predictions for patterns.

1.6 References

[1] C. Forestiere *et al.*, "Genetically Engineered Plasmonic Nanoarrays," *Nano Letters*, vol. 12, no. 4, pp. 2037-2044, 2012/04/11 2012.

[2] D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training Deep Neural Networks for the Inverse Design of Nanophotonic Structures," *ACS Photonics*, vol. 5, no. 4, pp. 1365-1369, 2018/04/18 2018.

[3] I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via Deep Learning," *Light: Science & Applications*, vol. 7, no. 1, p. 60, 2018/09/05 2018.

[4] B. Sanchez-Lengeling and A. Aspuru-Guzik, "Inverse molecular design using machine learning: Generative models for matter engineering," *Science*, vol. 361, no. 6400, pp. 360-365, 2018.

[5] H. Zhao, H. Liu, W. Hu, and X. Yan, "Anomaly detection and fault analysis of wind turbine components based on deep learning network," *Renewable Energy*, vol. 127, pp. 825-834, 2018/11/01/ 2018.

[6] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[7] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[8] L. K. Hansen and P. Salamon, "Neural network ensembles," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 10, pp. 993-1001, 1990.

[9] A. V. Herz, T. Gollisch, C. K. Machens, and D. Jaeger, "Modeling single-neuron dynamics and computations: a balance of detail and abstraction," *science*, vol. 314, no. 5796, pp. 80-85, 2006.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.

[11] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep Learning for Computer Vision: A Brief Review," *Computational Intelligence and Neuroscience*, vol. 2018, p. 7068349, 2018/02/01 2018.

[12] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent Trends in Deep Learning Based Natural Language Processing [Review Article]," *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55-75, 2018.

[13] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84-90, 2017.

[14] J. Carrasquilla and R. G. Melko, "Machine learning phases of matter," *Nature Physics*, vol. 13, no. 5, pp. 431-434, 2017/05/01 2017.

[15] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep Learning-Based Classification of Hyperspectral Data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2094-2107, 2014.

[16] Z. Gong, C. Li, X. Hu, H. Yang, and Q. Gong, "Active control of highly efficient third-harmonic generation in ultrathin nonlinear metasurfaces," *Optical Materials*, vol. 60, pp. 552-558, 2016/10/01/ 2016.

[17] D. C. Hooper *et al.*, "Second Harmonic Spectroscopy of Surface Lattice Resonances," *Nano Letters*, vol. 19, no. 1, pp. 165-172, 2019/01/09 2019.

[18] D. Timbrell, J. W. You, Y. S. Kivshar, and N. C. Panoiu, "A comparative analysis of surface and bulk contributions to second-harmonic generation in centrosymmetric nanoparticles," *Scientific Reports*, vol. 8, no. 1, p. 3586, 2018/02/26 2018.

[19] K. Munechika *et al.*, "Spectral Control of Plasmonic Emission Enhancement from Quantum Dots near Single Silver Nanoprisms," *Nano Letters*, vol. 10, no. 7, pp. 2598-2603, 2010/07/14 2010.

[20] G. Rui, D. C. Abeysinghe, R. L. Nelson, and Q. Zhan, "Demonstration of beam steering via dipole-coupled plasmonic spiral antenna," *Scientific Reports*, vol. 3, no. 1, p. 2237, 2013/07/19 2013.

[21] X. Zhu *et al.*, "Dispersion Control in Plasmonic Open Nanocavities," *ACS Nano*, vol. 5, no. 8, pp. 6546-6552, 2011/08/23 2011.

[22] R. W. Boyd, *Nonlinear optics*, 3rd ed. Amsterdam ; Boston: Academic Press, 2008, pp. xix, 613 p.

[23] Y. Liu and X. Zhang, "Metamaterials: a new frontier of science and technology," *Chemical Society Reviews,* 10.1039/C0CS00184H vol. 40, no. 5, pp. 2494-2507, 2011.

[24] W. Cai, *Optical metamaterials fundamentals and applications*. New York: New York : Springer, 2010.

[25] L. Alloatti *et al.*, "Second-order nonlinear optical metamaterials: ABC-type nanolaminates," *Applied Physics Letters,* vol. 107, no. 12, p. 121903, 2015.

[26] S. Clemmen *et al.*, "Atomic layer deposited second-order nonlinear optical metamaterial for back-end integration with CMOS-compatible nanophotonic circuitry," *Optics Letters,* vol. 40, no. 22, pp. 5371-5374, 2015/11/15 2015.

[27] A. Hermans *et al.*, "On the determination of $\chi(2)$ in thin films: a comparison of one-beam second-harmonic generation measurement methodologies," *Scientific Reports,* vol. 7, no. 1, p. 44581, 2017/03/20 2017.

[28] H.-H. Hsiao *et al.*, "Enhancement of second-harmonic generation in nonlinear nanolaminate metamaterials by nanophotonic resonances," *Optics Express,* vol. 24, no. 9, pp. 9651-9659, 2016/05/02 2016.

[29] A. Artar, A. Yanik, and H. Altug, *Light tunneling in multi-layered plasmonic crystals* (SPIE NanoScience + Engineering). SPIE, 2010.

[30] S. Zeng, D. Baillargeat, H.-P. Ho, and K.-T. Yong, "Nanomaterials enhanced surface plasmon resonance for biological and chemical sensing applications," *Chemical Society Reviews,* 10.1039/C3CS60479A vol. 43, no. 10, pp. 3426-3452, 2014.

[31] K. L. Kelly, E. Coronado, L. L. Zhao, and G. C. Schatz, "The Optical Properties of Metal Nanoparticles: The Influence of Size, Shape, and Dielectric Environment," *The Journal of Physical Chemistry B,* vol. 107, no. 3, pp. 668-677, 2003/01/01 2003.

[32] M. Rycenga *et al.*, "Controlling the Synthesis and Assembly of Silver Nanostructures for Plasmonic Applications," *Chemical Reviews,* vol. 111, no. 6, pp. 3669-3712, 2011/06/08 2011.

[33] S. Doherty, "Control of pH in Chemical Processes Using Artificial Neural Networks," 01/01 1999.

[34] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics,* vol. 5, no. 4, pp. 115-133, 1943/12/01 1943.

[35] A. L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers," *IBM Journal of Research and Development,* vol. 3, no. 3, pp. 210-229, 1959

CHAPTER 2. NONLINEAR OPTICS AND DEEP LEARNING

This chapter is intended to give a deeper understanding into the nonlinear optics and deep learning used for the research conducted in this thesis. An overview of sum-frequency generation, nonlinear susceptibility, and deep learning examples will provide necessary background for the concepts put forth in later chapters.

2.1 Nonlinear Optics

It can be said that linear optics is noticed as a child – looking in a mirror and seeing a reflection, or when searching for the pot of gold at the end of a rainbow. Linear optics is displayed all around us, and thereby makes it much easier to grasp as it can be simply demonstrated for the most general cases of refraction, reflection, and diffraction.

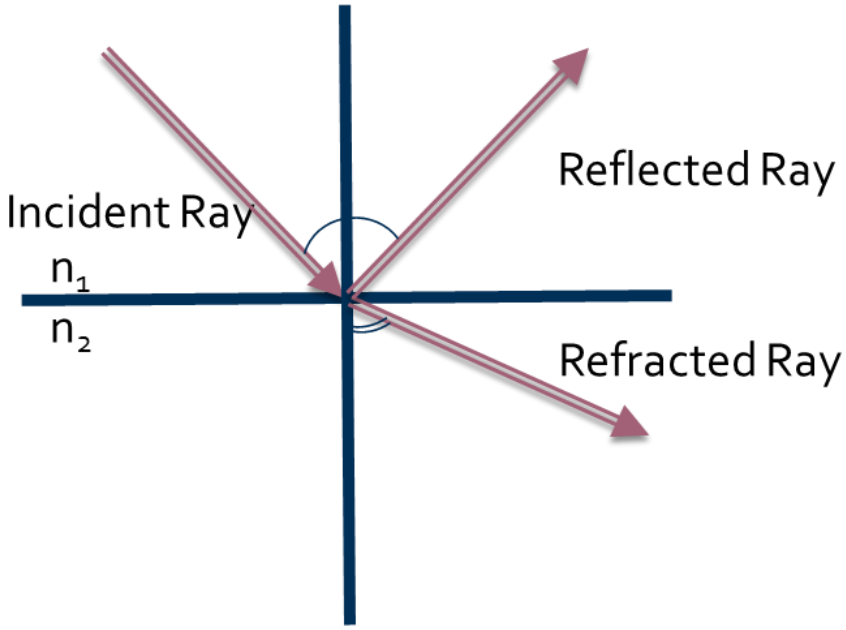

Figure 2.1 – Example of linear optics where an incident ray in a material with a refractive index of n_1 hits the surface with a refractive index of n_2. This results in a reflected ray and a refracted ray.

We can describe a linear optical response as one where monochromatic light entering a linear system maintains its frequency, an example of which is in Figure 2.1. While the light may be delayed, polarized, or otherwise acted upon, its frequency will remain unchanged. This is described through a polarization equation in Chapter 1. It is important to note that in a linear case, the refractive index is related to the linear susceptibility, or $\chi^{(1)}$ that is mentions in Equation 1.1. It can further be elucidated as such: $n = \sqrt{\chi^{(1)} + 1}$.

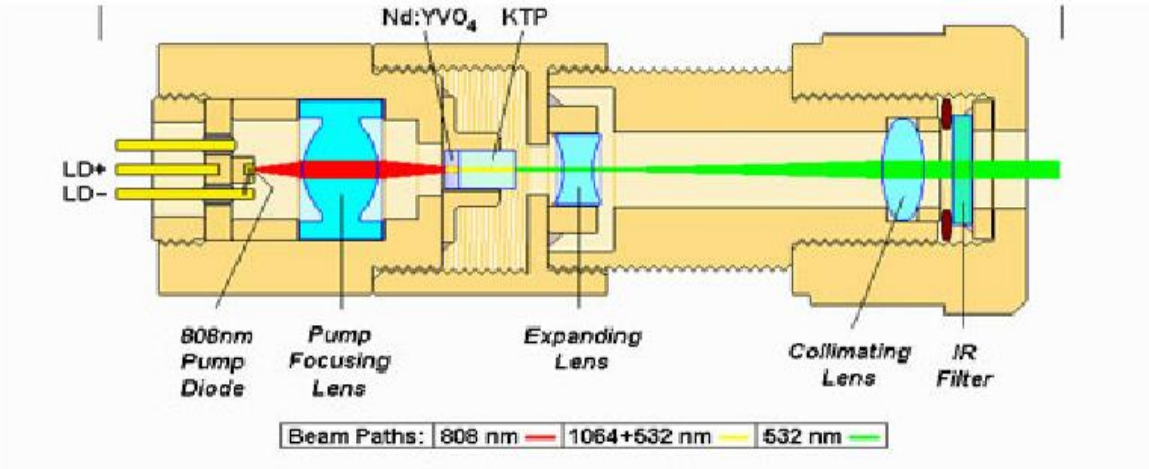

Figure 2.2 - Image illustrating how a green laser pointer functions [1].

On the other hand, nonlinear optics is much more complex, and in most scenarios, the processes can't be visualized without an expensive or complex setup. Even a simple example of how a green laser pointer works, relies on trust and imagination between the teacher and audience as shown in Figure 2.2. It is a much more hidden process.

Nonlinear optics can be described as frequency mixing processes. Frequency mixing processes that occur dur to the applied electric field interacting with the nonlinear properties of a material allow for the generation of radiation at different frequencies.

As described in Chapter 1, a nonlinear material can have a nonlinear susceptibility, for example the second order nonlinear susceptibility would be referred to as $\chi^{(2)}$. More specifically, these nonlinear susceptibilities are tensors, so we would actually refer to the 2nd order susceptibility as $\chi_{ijk}^{(2)}$, where *i*, *j*, and *k* are the cartesian components of the field[2].

It is possible to reduce the tensor components to a contracted notation for simplicity.

$$d_{ijk} = \frac{1}{2}\chi^{(2)}_{ijk} \tag{2.1}$$

This can further be simplified using Kleinman's symmetry to assume d_{ijk} is symmetric in the last two indices, leading to $d_{ijk} = d_{il}$[2]. Where:

$$\begin{matrix} jk: & 11 & 22 & 33 & 23,32 & 31,13 & 12,21 \\ l: & 1 & 2 & 3 & 4 & 5 & 6 \end{matrix}. \tag{2.2}$$

Leading to a 3 × 6 nonlinear susceptibility tensor:

$$d_{il} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix}. \tag{2.3}$$

Using the contracted notation, we can describe the nonlinear polarization of second-harmonic generation as such:

$$\begin{bmatrix} P_x(2\omega) \\ P_y(2\omega) \\ P_z(2\omega) \end{bmatrix} = 2\epsilon_0 \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix} \begin{bmatrix} E_x(\omega)^2 \\ E_y(\omega)^2 \\ E_z(\omega)^2 \\ 2E_y(\omega)E_z(\omega) \\ 2E_x(\omega)E_z(\omega) \\ 2E_x(\omega)E_y(\omega) \end{bmatrix}. \tag{2.4}$$

Generally, the way that the elements of nonlinear susceptibility tensors are calculated are by first identifying the nonvanishing elements and their symmetries based on the crystalline structure of the material. Then the nontrivial elements are experimentally measured.

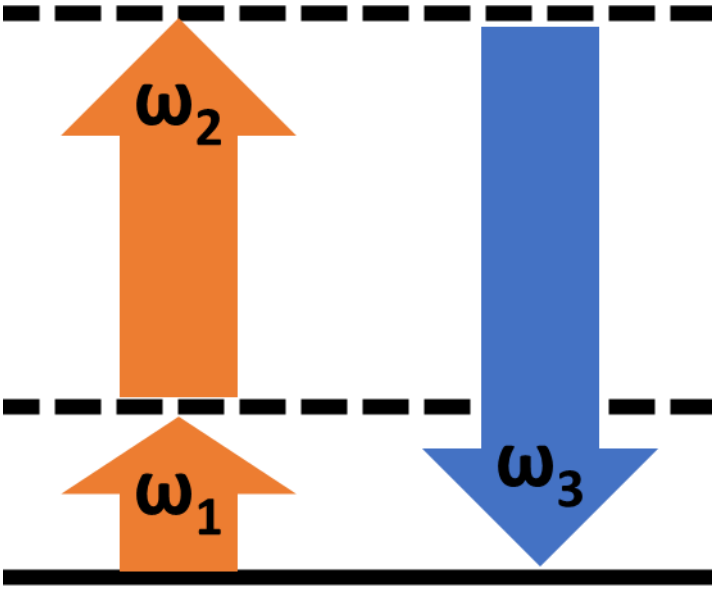

Figure 2.3 – Energy level description of sum-frequency generation. Two photons of frequencies ω_1 and ω_2 combine to form a photon of frequency $\omega_3 = \omega_1 + \omega_2$.

2.2 Sum-Frequency Generation

Second-harmonic generation is a specific case of sum-frequency generation (SFG). In second-harmonic generation, two photons of the same frequency, ω, are effectively combined to form a single photon of double the frequency, 2ω. We can say the second-harmonic generation is a process where $\omega + \omega \rightarrow 2\omega$, when interacting with a nonlinear material. In sum-frequency generation, two photons of different frequencies, ω_1 and ω_2, interact to result in the formation of a photon with a frequency of the sum of the two photons, $\omega_3 = \omega_1 + \omega_2$ as shown in Figure 2.3.

The nonlinear polarization, $P(\omega_1 + \omega_2) = 2\varepsilon_0\chi^{(2)}E(\omega_1)E(\omega_2)$, further shows the similarity to the second-harmonic generation.

The full nonlinear polarization for sum-frequency generation, in the contracted notation, is as follows,

$$\begin{pmatrix} P_x(\omega_1+\omega_2) \\ P_y(\omega_1+\omega_2) \\ P_z(\omega_1+\omega_2) \end{pmatrix} = \begin{pmatrix} P_x(\omega_3) \\ P_y(\omega_3) \\ P_z(\omega_3) \end{pmatrix} =$$

$$\epsilon_0 \begin{pmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{pmatrix} \begin{pmatrix} E_x(\omega_1)E_x(\omega_2) \\ E_y(\omega_1)E_y(\omega_2) \\ E_z(\omega_1)E_z(\omega_2) \\ E_y(\omega_1)E_z(\omega_2)+E_z(\omega_1)E_y(\omega_2) \\ E_x(\omega_1)E_z(\omega_2)+E_z(\omega_1)E_x(\omega_2) \\ E_x(\omega_1)E_y(\omega_2)+E_z(\omega_1)E_y(\omega_2) \end{pmatrix} \quad (2.5)$$

, and can be simplified for the ABC nanolaminate as follows,

$$\begin{aligned} \boldsymbol{P_x}(\omega_1+\omega_2) &= \boldsymbol{\epsilon_0 \chi_{xxz}}[E_x(\omega_1)E_z(\omega_2)+E_z(\omega_1)E_x(\omega_2)] \\ \boldsymbol{P_y}(\omega_1+\omega_2) &= \boldsymbol{\epsilon_0 \chi_{xxz}}[E_y(\omega_1)E_z(\omega_2)+E_z(\omega_1)E_y(\omega_2)] \\ \boldsymbol{P_z}(\omega_1+\omega_2) &= \boldsymbol{\epsilon_0 \chi_{zxx}}\big[E_x(\omega_1)E_x(\omega_2)+E_y(\omega_1)E_y(\omega_2)\big] + \boldsymbol{\epsilon_0 \chi_{zzz}}[E_z(\omega_1)E_z(\omega_2)] \end{aligned} \quad (2.6)$$

from these equivalencies, $\chi_{xxz} = d_{15} = d_{24}$, $\chi_{zxx} = d_{31} = d_{32}$,and $\chi_{zzz} = d_{33}$ [2, 3].

Utilizing the physics of sum-frequency generation, a process of photonic upconversion can take place. In nonlinear optics, upconversion is described as sum-frequency generation. Two, or more, photons are absorbed, and a photon of a shorter wavelength is emitted. This frequency mixing allows for the detection of light otherwise unable to be counted using available imaging processes. Sum-frequency generation is most utilized in spectroscopy, developed for the use of studying interfaces such as gas-liquid interfaces.

In this research, sum-frequency generation is used for photonic upconversion in order to convert infrared light to the visible regime. This research has many very important

applications. One important application for SFG upconversion of near infrared signals is for low-cost night vision/low light imaging systems. The uses for these imaging systems range from defense applications to autonomous vehicles and more.

SFG requires two incident beams of light, so the set up can be described as a pump-probe system. In this scenario, two infrared incident light sources – one of significance to be converted, and the other for the upconversion process, are focused on a nonlinear material such as an ABC nanolaminate. The resulting light in the visible regime would then be collected. For the ABC nanolaminate to be used as a nonlinear metamaterial, a plasmonic structure is required, in order to take advantage of the surface plasmon effect, as explained earlier.

Utilizing a deep learning framework, consisting of a pattern generator and a simulator, the geometry of the plasmonic device can be optimized, thereby optimizing the strength of the emitted visible light from the metamaterial allowing for more efficient SFG upconversion.

2.3 Deep Learning

Successful implementations of discriminative models will be introduced in this section for the design of photonic devices. Additionally, work accurately predicting optical responses of photonic structures with a deep learning model will be presented to show the growing and viable field. Such deep-learning-based methods enable the fast evaluation of the design performance of photonic structures during their computationally intensive optimization [4].

2.3.1 Design of Metallic Metasurfaces and Metamaterials

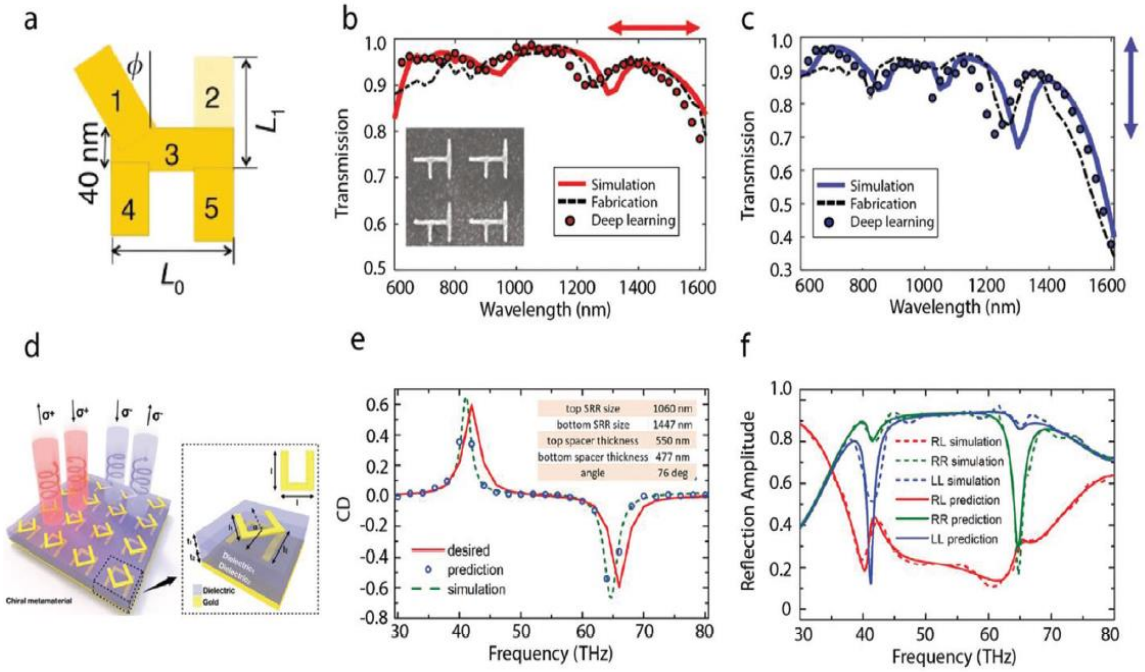

Figure 2.4 – Metasurfaces and metamaterials designed for amplitude and chirality manipulation are shown. a–c) "H" shaped metallic metasurface design. a) Schematic of the shape and parameters of the nanostructure. b-c) The simulated, measured, and deep learning retrieved spectra of the design when the incident light is horizontally and vertically polarized. The inset in b shows the SEM image of the fabricated. d–f) Chiral metamaterial design. d) Schematic of the designed chiral metamaterial. The inset is the zoomed-in structure of a single meta-atom unit. e) Desired, predicted, and simulated circular dichroism (CD) spectra. The insets list the retrieved geometric parameters. f) Predicted full reflection spectra along with the full-wave simulation results [4].

The suggestion of solving the inverse design of nanostructures in metasurfaces was first offered by Malkiel et al. using a few parameters through a bidirectional designed neural network [5]. Figure 2.4a shows the 'H' shaped particle. In order to solve the design problem, two networks are collectively built – a geometry predicting network (GPN) and a spectrum predicting network (SPN). Given a specific spectral response, the GPN predicts the geometric parameters, and then the spectrum of the input geometry is approximated by the SPN.

This is done as follows, during training, the GPN output is fed into the SPN. The idea is that if given a pair of geometries and their corresponding spectra as a training pair, minimizing the loss between the training pair and the GPN and SPN output lead to a usable architecture. Upon creating this bidirectional network and the subsequent training, it can be utilized to generate various desired responses for various designs by changing the objectives of the GPN.

An example, given two spectral behaviors with incident x and y polarized light, of a retrieved structure and the related SEM of a fabricated nanostructure is shown in Figure 2.4 b and c. The network-predicted, FEM simulated, and measured results of nanostructure are presented by solid, dashed, and circled lines, respectively. Analogous strategies leveraging two networks, one for simulating optical responses and one for calculating design geometries and parameters, have also been shown to be powerful in other design applications such as broadband highly reflective metasurfaces [6].

Figure 2.4d shows another bidirectional strategy that was developed and extended to the design of multilayered chiral metamaterials [7]. The chiral metamaterials unit cell is

composed of two stacked split ring resonators (SRRs) fabricated from gold. Circular dichroism (CD) is a measure of chirality and is used to characterize the chirality of the metamaterial. CD is the difference in absorption between the left and right circularly polarized incident light.

Again, here the overall network architecture utilizes the bidirectional framework, although two such networks are used. One is a primary, to learn the mapping between the optical spectra and design space, and the other is an auxiliary network, which tackles the relationship between the corresponding structure parameters and the CD. The addition of the auxiliary network further improves the accuracy of the prediction and design.

Figure 2.4e and f show an example of a predicted structure and its corresponding optical responses. The red line in Figure 2.4e represents the desired CD, and the simulated CD from the network shown by the dashed lines match the desired response. The simulated spectrum of the retrieved structure is given in Figure 2.4f.

2.3.2 Design of High DOP Photonic Device

Deep neural networks can be leveraged for the design of photonic systems with a high degree of freedom (DOF), with the capability of performing regression and classification of high dimensional data points.

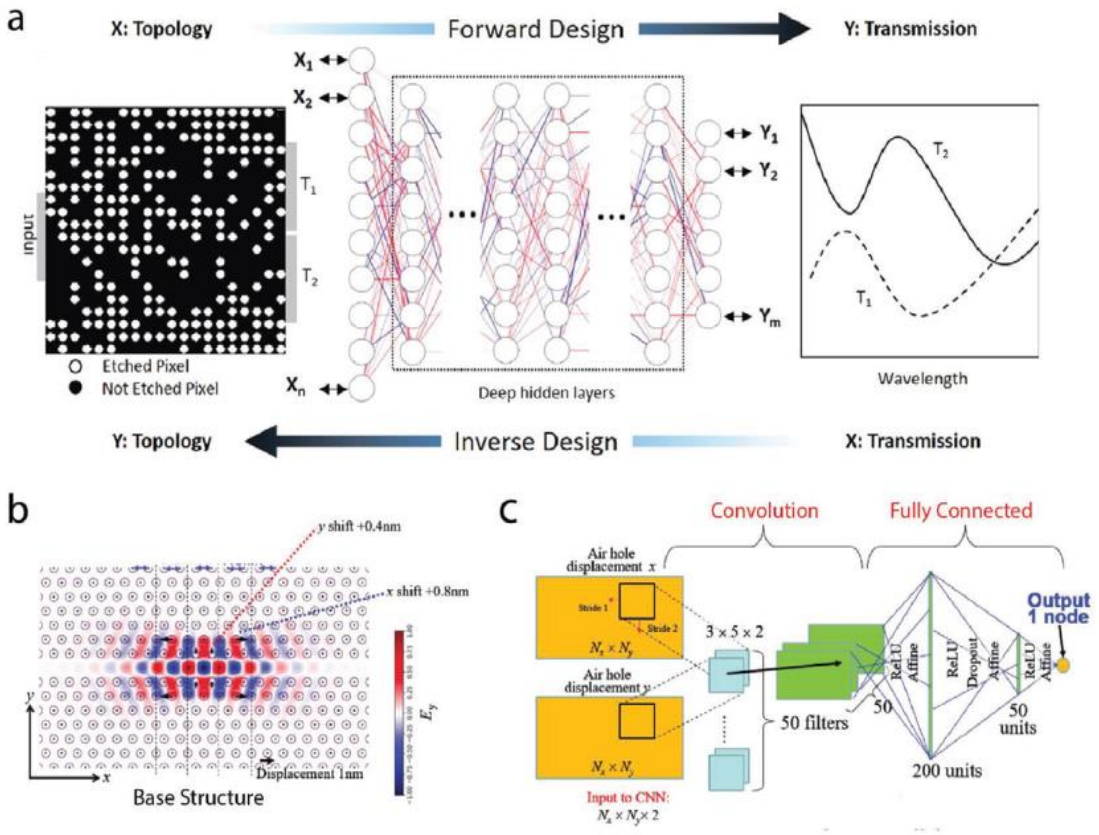

Figure 2.5 – Shown here are discriminative models for the design of high degree of freedom photonic devices. a) This is the design strategy and network architecture for 1 × 2 integrated photonic power splitters with various target splitting ratios. The etched pixels or holes represent the power splitters. The spectral response at the ports is simulated by the forward deep neural networks, and the choice of etching or not etching each pixel or hole is decided by the inverse network given a desired splitting ratio. b-c) CNNs are used to optimize a 2D photonic crystal nanocavity with CNNs. b) Illustration of the photonic crystal configuration. The circles indicate air holes formed in a silicon slab. c) Configuration of the neural network that can capture the relationship between the displacements of air holes and the Q factors of the photonic

crystal. The optimization is performed by treating the hole displacements as variables and iteratively optimize them through backpropagation [4].

A demonstration of a neural network enabled design of silicon-on-insulator based 1×2 integrated photonic power splitters with various target splitting ratios is presented in Figure 2.5a. The etched pixels represent a collection of holes or the power splitters in a binary image in Figure 2.5a. Another bidirectional network is implemented in this work as well. The forward simulation, or spectral response prediction at the output ports for a given structure encapsulate on of the network models. The other is for the identification of the optimal power splitter with the desired splitting ratio, or the backward design. As the only choices are if each hole of the splitter are to be etched or not etched, essentially a binary decision, a classification problem can be the inverse design process model [8].

CNNs with shared weights performing cross-correlation operations are adopted to efficiently process high dimensional data, because of the increasing DOF of the design problems. The local correlation of spatial information in images can be captured by a CNN. As such, given a photonic device, a CNN is an ideal candidate to process images representing photonic patterns and the spectral responses of the device. CNNs have been utilized in various optical problems, such as the inverse scattering problem, wave-front correction, digital coding metasurfaces, and the pre-diction of optical properties in complex photonic and materials systems [4, 9-15]. A few examples of photonic devices designed with CNNs will be discussed.

Asano and Noda reported work that utilizes a neural network consisting of CNNs to simulate the Q-factor (or quality factor) of photonic crystals and optimize that through

backpropagation [16]. An example of a heterostructure 2D photonic crystal nanocavity is shown in Figure 2.5b. The aim of the optimization is to identify the positions of these air cavities to maximize the Q-factors given a certain initial structure.

Figure 2.5c describes the network architecture, a CNN is concatenated by a fully connected network in order to predict the Q-factor of the input structure. It is quite important to note that by relying on their capability of processing large dimensional data, CNNs have become an indispensable architecture that deal with photonic devices with the high DOF [17, 18].

2.3.3 Generative Model - VAEs

Discriminative models can predict forward simulation with extraordinary accuracy to enable various optimizations of parameters defining photonic structures and their materials. However, when the dimension of the design space grows beyond thousands, it is unfeasible to generate sufficient data to train a surrogate model. Generative models, in this situation, are a great candidate to reduce the dimensionality of the design space in order to assist the universal optimization. Here, we will focus on the deep generative model of VAEs [19].

VAEs, as an essential member of deep generative models, have been implemented to reduce the dimensionality of photonic nanostructures and their related physical properties for efficient optical design. VAEs were utilized by Ma et al. to encode meta-atoms of metamaterials and their optical responses, enabling the investigation of the complex structure-performance relationship without extensive data collection [20].

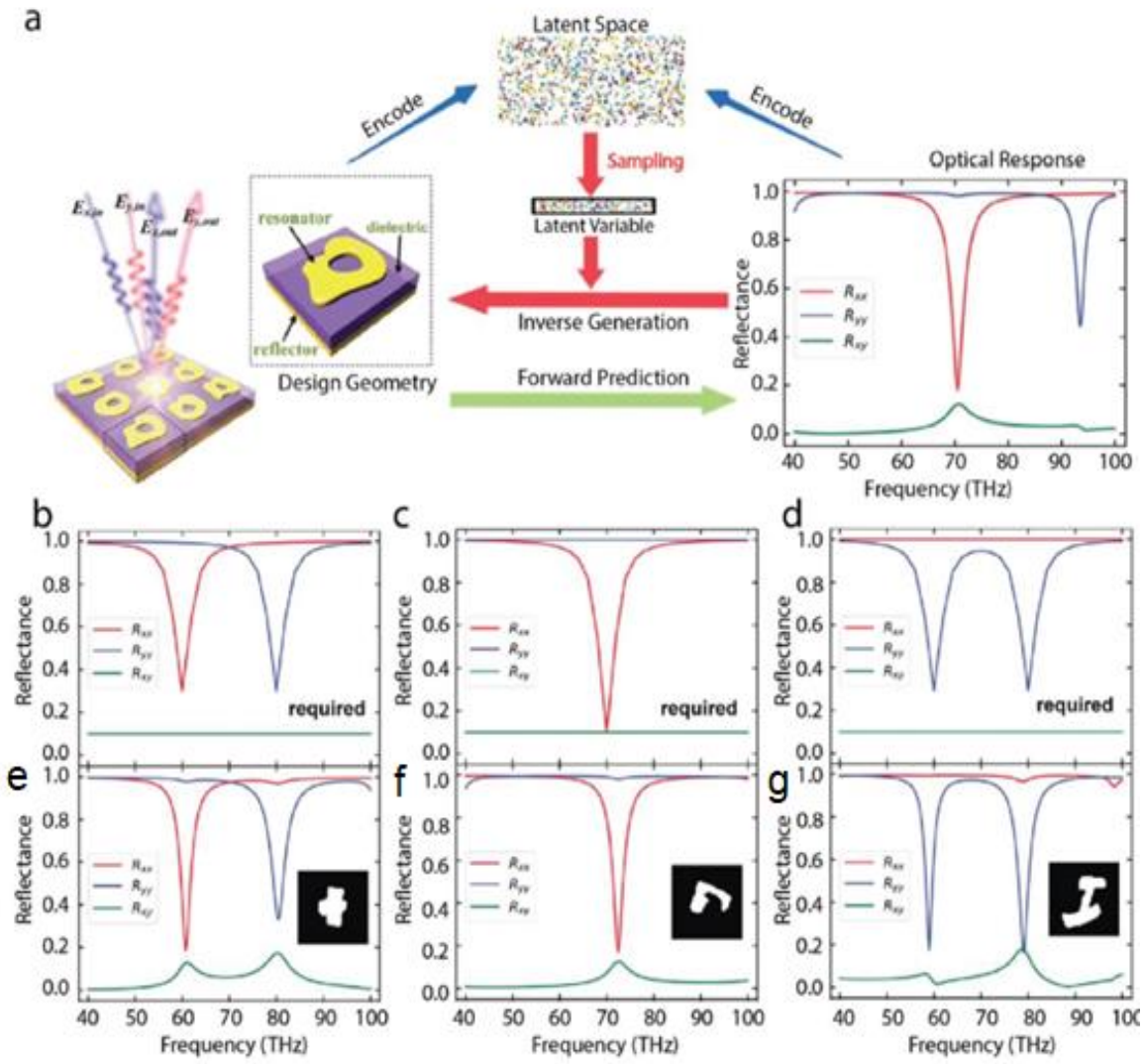

Figure 2.6 – Photonic structure design using a VAE-based strategy. a) The framework of the VAE network utilized for the design and characterization of reflective metamaterials. A latent space encoded with a set of meta-atoms and their optical responses leads to latent variables that are sampled for the inverse design in response to a specific objective. b–d) Specified on-demand reflection spectra input to the model to be designed for. e–g) Unit cells of the VAE-designed meta-atoms are shown in the insets and their optical responses for the corresponding inputs as shown in b–d [4].

A deep generative model of the design and characterization of a metamaterial is illustrated in Figure 2.7. To have similar designs and optical responses automatically amassed together, the metamaterial and its corresponding optical responses are encoded into the same latent space. Therefore, by sampling the latent space, candidate designs can be generated given strict requirements in the decoding process of the latent vectors. Three examples of generated metamaterial designs for on-demand spectra from the proposed generative model are shown in Figure 2.7b-g.

Figure 2.7b-d show the requested on-demand reflected spectra 40 THz – 100 THz, and the generated metamaterial designs together with their simulated spectra are displayed below in Figure 2.7e-g, respectively. Looking at the simulated and input spectra, it can be seen that they replicate one another. Similar techniques have also been applied to the design of multi-layered chiral metamaterials that satisfies various chiroptical response requirements [4].

2.3.4 Optimization

As a stand-alone technique, machine learning can analyze high dimensional complex datasets in order to capture the essential features for the simulation of the physical responses and the design of photonic structures. Even though traditional optimization algorithms, such as adjoint methods, genetic algorithms, and particle swarm optimization, have significant successes optimizing of photonic structures and devices, some drawbacks of these traditional optimization algorithms prevent their use as effective approaches for the design of high DOF devices in a universal manner.

Black-box optimization algorithms, such as simulated annealing, random search, and evolutionary algorithms, on the other hand, do not seek the gradients to iteratively update the parameters [4]. As an alternative, they stochastically update the structural parameters through probabilistic methods or by emulating physical and/or biological processes in order to identify a solution in a global manner.

However, without the gradient information from the problem's physical model, immense iterations of computation are necessary to explore the space to yield an optimal solution. These redundant computations limit black-box optimizations utilized in the design of photonic structures, materials, and devices that require computationally intensive simulations.

In order to ease the unnecessary and redundant computation, Hegde proposed a strategy that pairs a deep neural network with an evolutionary algorithm [21, 22]. The network model proposed, is trained as a surrogate model that partially replaces the expensive simulation for the preselection and optimization of optical thin-film systems [4]. This surrogate model enables massive, parallelized simulation thereby significantly reducing the time necessary for evaluating the fitness/cost function during optimization. Effectively, the network model is equivalent to a cache that stores the calculated results, but with a markedly compressed size, for the quick evaluation of the physical responses of photonic devices without the need for continual computations.

Generative models can be incorporated into optimizations, in this manner serving as a way to reduce dimensionality, bringing down the DOF. GANs, VAEs, and other machine learning algorithms can be used to construct a compact sparse representation of

photonic devices in a latent space with lessened dimensionality. It is also much more efficient to search the latent space of the generative models as compared to performing optimization on the original high-dimensional dataspace.

2.4 References

[1] E. Garmire, "Nonlinear optics in daily life," *Optics express,* vol. 21, pp. 30532-44, 12/16 2013.

[2] R. W. Boyd, *Nonlinear optics*, 3rd ed. Amsterdam ; Boston: Academic Press, 2008, pp. xix, 613 p.

[3] A. Hermans *et al.*, "On the determination of $\chi(2)$ in thin films: a comparison of one-beam second-harmonic generation measurement methodologies," *Scientific Reports,* vol. 7, no. 1, p. 44581, 2017/03/20 2017.

[4] Z. Liu, D. Zhu, L. Raju, and W. Cai, "Tackling Photonic Inverse Design with Machine Learning," *Advanced Science,* https://doi.org/10.1002/advs.202002923 vol. 8, no. 5, p. 2002923, 2021/03/01 2021.

[5] I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via Deep Learning," *Light: Science & Applications,* vol. 7, no. 1, p. 60, 2018/09/05 2018.

[6] E. S. Harper, E. J. Coyle, J. P. Vernon, and M. S. Mills, "Inverse design of broadband highly reflective metasurfaces using neural networks," *Physical Review B,* vol. 101, no. 19, p. 195104, 05/04/ 2020.

[7] W. Ma, F. Cheng, and Y. Liu, "Deep-Learning-Enabled On-Demand Design of Chiral Metamaterials," *ACS Nano,* vol. 12, no. 6, pp. 6326-6334, 2018/06/26 2018.

[8] M. H. Tahersima *et al.*, "Deep Neural Network Inverse Design of Integrated Photonic Power Splitters," *Scientific Reports,* vol. 9, no. 1, p. 1368, 2019/02/04 2019.

[9] B. Wu, K. Ding, C. T. Chan, and Y. Chen, "Machine Prediction of Topological Transitions in Photonic Crystals," *Physical Review Applied,* vol. 14, no. 4, p. 044032, 10/19/ 2020.

[10] T. Xie and J. C. Grossman, "Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties," *Physical Review Letters,* vol. 120, no. 14, p. 145301, 04/06/ 2018.

[11] A. Maksov *et al.*, "Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS2," *npj Computational Materials,* vol. 5, no. 1, p. 12, 2019/02/01 2019.

[12] Q. Zhang *et al.*, "Machine-Learning Designs of Anisotropic Digital Coding Metasurfaces," *Advanced Theory and Simulations,* https://doi.org/10.1002/adts.201800132 vol. 2, no. 2, p. 1800132, 2019/02/01 2019.

[13] L. Li *et al.*, "Machine-learning reprogrammable metasurface imager," *Nature Communications,* vol. 10, no. 1, p. 1082, 2019/03/06 2019.

[14] A. Turpin, I. Vishniakou, and J. d. Seelig, "Light scattering control in transmission and reflection with neural networks," *Optics Express,* vol. 26, no. 23, pp. 30911-30929, 2018/11/12 2018.

[15] L. Li, L. G. Wang, F. L. Teixeira, C. Liu, A. Nehorai, and T. J. Cui, "DeepNIS: Deep Neural Network for Nonlinear Electromagnetic Inverse Scattering," *IEEE Transactions on Antennas and Propagation,* vol. 67, no. 3, pp. 1819-1825, 2019.

[16] T. Asano and S. Noda, "Optimization of photonic crystal nanocavities based on deep learning," *Optics Express,* vol. 26, no. 25, pp. 32704-32717, 2018/12/10 2018.

[17] R. Lin, Y. Zhai, C. Xiong, and X. Li, "Inverse design of plasmonic metasurfaces by convolutional neural network," *Optics Letters,* vol. 45, no. 6, pp. 1362-1365, 2020/03/15 2020.

[18] Y. Li *et al.*, "Self-Learning Perfect Optical Chirality via a Deep Neural Network," *Physical Review Letters,* vol. 123, no. 21, p. 213902, 11/19/ 2019.

[19] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," *CoRR,* vol. abs/1312.6114, 2014.

[20] W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic Representation and Inverse Design of Metamaterials Based on a Deep Generative Model with Semi-Supervised Learning Strategy," *Advanced Materials*, https://doi.org/10.1002/adma.201901111 vol. 31, no. 35, p. 1901111, 2019/08/01 2019.

[21] S. P. Soumyashree and S. H. Ravi, "Transmission-mode all-dielectric metasurface color filter arrays designed by evolutionary search," *Journal of Nanophotonics*, vol. 14, no. 1, p. 016014, 3/1 2020.

[22] R. S. Hegde, "Photonics Inverse Design: Pairing Deep Neural Networks With Evolutionary Algorithms," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 1, pp. 1-8, 2020.

CHAPTER 3. LINEAR AND SECOND HARMONIC NONLINEAR OPTICAL RESPONSES

The research described in this chapter shows the capabilities for linear and second harmonic nonlinear optical responses. Deep learning algorithms are created and trained to inversely design metamaterial structures for optimal linear and nonlinear optical responses.

3.1 Framework design for deep learning algorithm

In order to optimize an arbitrary pattern to create an optimal photonic structure, a deep learning framework is used, as parameter-sweeping and standard geometries cannot yield the most enhanced optical response. This algorithm can be used to optimize a photonic structure for linear responses [1].

The framework can be broken into three parts, the generator, the simulator, and the optimizer. The generator is used to create 2D patterns that translate into the photonic structure. Once patterns can be generated, the simulator uses the generator to find the related optical response [2].

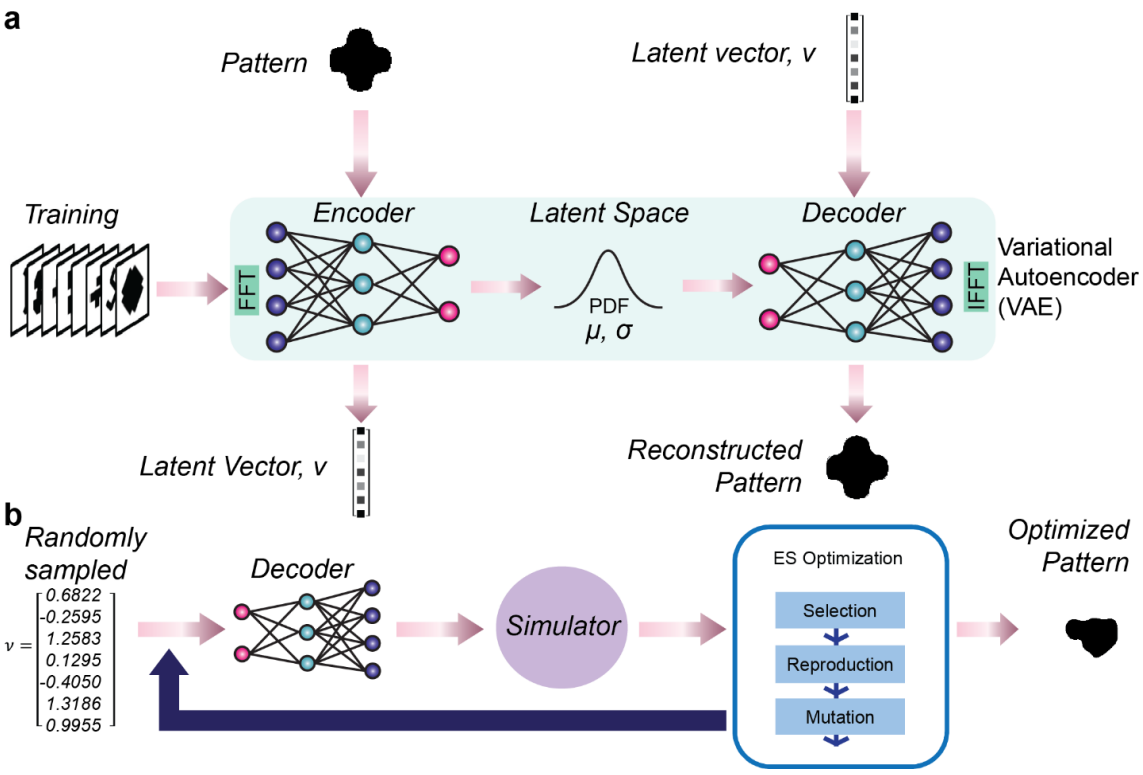

Figure 3.1 – General deep learning framework. (a) Overall flow of the generator, an algorithm used to represent the patterns in the training data using a probability density function (PDF). The generator is based on a Variational Autoencoder (VAE), and the PDF is used by the encoder and decoder to represent a pattern as a latent vector, *v*, through 'encoding', and then transformed v into a reconstructed pattern through 'decoding'. (b) By randomly sampling *v*, we can use the decoder from the generator to create patterns. These patterns are passed through the simulator, and a predicted optical response is output. In order to optimize a pattern to a specific optical response, an optimizer based on an Evolution Strategy (ES) is utilized. Passing through a system of selection, reproduction, and mutation, the optimizer circles back to try different *v*'s searching for better patterns until the closest fit is found. The final output is the best pattern for the desired parameters, and optical response [3].

To target a desired response, the optimizer is used to find the best pattern for the response requested. The generator algorithm is based on a variational auto-encoder (VAE), as shown in Figure 3.1a. First, 10,000 arbitrary created patterns (64×64 pixel images) form a training set for the generator.

The patterns are then passed through an encoder, which is a neural network. The purpose of the encoder is to 'encode' the data into a latent space, or probability density function, with a standard deviation, σ, and mean, μ, which represents all the possible patterns.

Then, a randomly sampled latent vector, v, can be fed to a decoder, which is another neural network, based on the same latent space, and be 'decoded' into a pattern. If a pattern is encoded, the output is a latent vector, and if that same v is input into the decoder, the output will be a reconstruction of the input image.

The generator allows for the creation of patterns that were not in the original data set, which can lead to a pattern with a response better than if inverse design was not utilized.

The simulator is based on the architecture of ResNet18 and is used to predict the linear optical response for the input image. To train the simulator, 10,000 patterns are created from the generator. Data collected from a full-wave simulation for linear responses is used to train the simulator. Once trained, the generator is used to produce patterns for the simulator to predict respective optical responses.

We adapted an Evolution Strategy as out optimizer in the framework, as shown in Figure 3.1b. Through a process of selection, reproduction, and mutation, while evaluating

a fitness score for the pattern, patterns that maximize the linear optimal response are created.

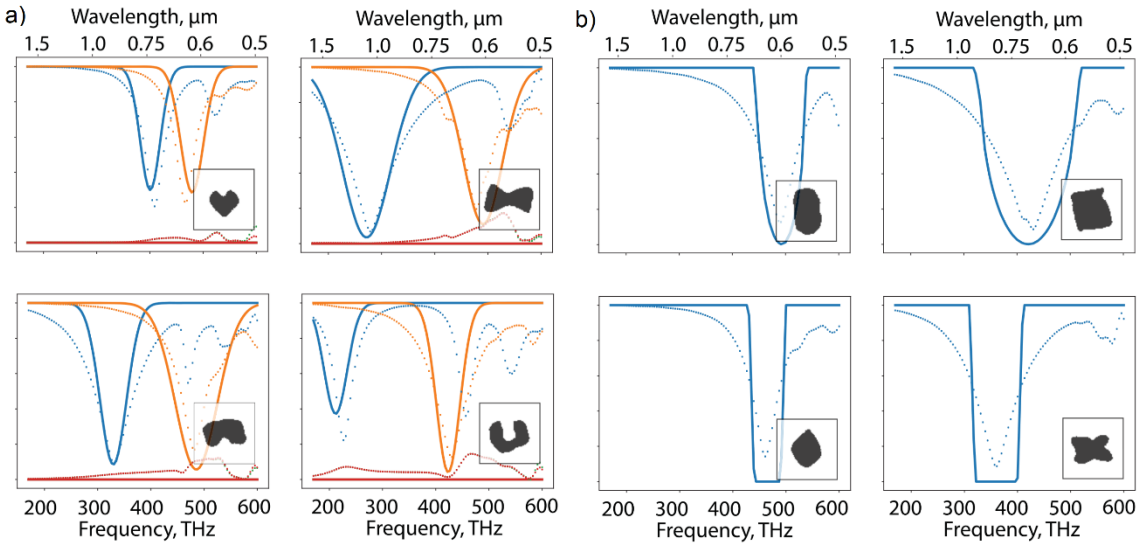

Figure 3.2 – Results of on-demand inverse design. a) Target Gaussian-like spectra. The desired spectra T_{xx} and T_{yy} at the input, shown as the solid lines, are two randomly generated Gaussian-like curves, and the T_{xy} and T_{yx} are zeros across the frequency range of interest. The generated patterns in the unit cell are depicted in the lower right corner of each panel, and the FEM-simulated spectra of the resultant nanostructures are represented in dashed lines. b) Target notch filters. he desired spectrum T_{xx} at the input has a band-stop transmission feature with specific central frequency and bandwidth. All these examples demonstrate the effectiveness of the framework, which can generate nanostructures that resemble the on-demand spectra fed at the input, and faithfully replicate major features in terms of the spectral location and the bandwidth.

To validate the data learning framework, two sets of experiments for on-demand designs are performed[1]. Figure 3.2a shows the results if the input spectra T_{xx} and T_{yy} are set to be two randomly chosen, Gaussian-shaped curves, and T_{xy} and T_{yx} are set to zero throughout the frequency range of interest. The output inversely designed patterns are shown in the insets of Figure 3.2a, the dashed lines identify the actual linear response from a full-wave simulation of the pattern, and the solid lines represent the desired input transmission spectra for the framework.

These results are distinctive as the independent manipulation of T_{xx}, T_{yy}, T_{xy} and T_{yx} is not readily achievable through conventional human design approaches, while our framework is able to accomplish simultaneous control of T_{xx}, T_{yy} and suppression of T_{xy}, T_{yx}.

Also, while independent manipulation of the transmission spectra is possible with the framework, the inversely designed patterns don't always exactly correspond to the desired transmission spectra. The reason being that not all desired input spectra can be achieved by the photonic structures in the given parameters.

In Figure 3.2b, we set the input T_{xx} as a notch filter without any specifications for the T_{yy}, T_{xy} and T_{yx} components. We note that within the family of single-layered metallic metasurfaces, it is not possible to identify a metasurface that perfectly replicates ultrasharp spectral features such as the steep cut-off slopes. However, the framework is able to generate a structure that accurately replicates the primary features of the desired filter, including the central frequency and bandwidth of the stop-band.

We also note that although the transmittance spectra are set as the design objective in these experiments, any photonic responses such as the diffraction behavior, optical

chirality, and field localization can be used as the intended design criteria without further adjustment of our framework.

These results show that the data learning framework created can be used for the inverse design of photonics structures to optimize linear optical responses.

3.2 Deep learning inverse design for Second-Harmonic Generation

There has been recent related research that discusses applications of inverse design and nonlinear optics using materials which have non-zero nonlinear optical susceptibilities [4-9]. The purpose of this research is to use inverse design to find an optimal plasmonic structure in order to induce the second harmonic generation and maximize the efficiency of the SHG of an ABC nanolaminate with normal incidence, where the second-order response is zero at normal incidence.

The deep learning algorithm finds the relationship between the parameters, the field profile, the $\chi^{(2)}$, and the SHG response to provide an optimal pattern for the metamaterial. The optimized plasmonic structure is designed with strong field confinement near the ABC nanolaminate and augmented symmetry engineering near the ABC nanolaminate, which will induce the second-harmonic generation in the ABC nanolaminate due to the electric field distribution in the few tens of nanometers range.

Using tailored algorithms to solve the challenging inverse design problem, previously unsolvable with conventional methods, of arbitrary topology with almost infinite degrees of freedom, the inverse design will allow the optimization of the metal pattern with user-defined constraints [1, 10]. This can be easily extended to output best

possible patterns for a number of scenarios given additional data relating to the new parameters, for example a different nonlinear material, or periodicity.

The goal of this research is to use a deep learning framework to find a metamaterial structure with an arbitrary plasmonic structure atop a nonlinear ABC composite material, that induces a maximal SHG response from the ABC composite. The pattern that will be fabricated as the plasmonic structure is based on a 64×64 canvas, meaning the inverse design problem essentially has over 2^{4096} degrees of freedom. Employing deep learning allows the solution of this problem, which cannot be solved using traditional parameter sweeping methods and geometries for true optimization.

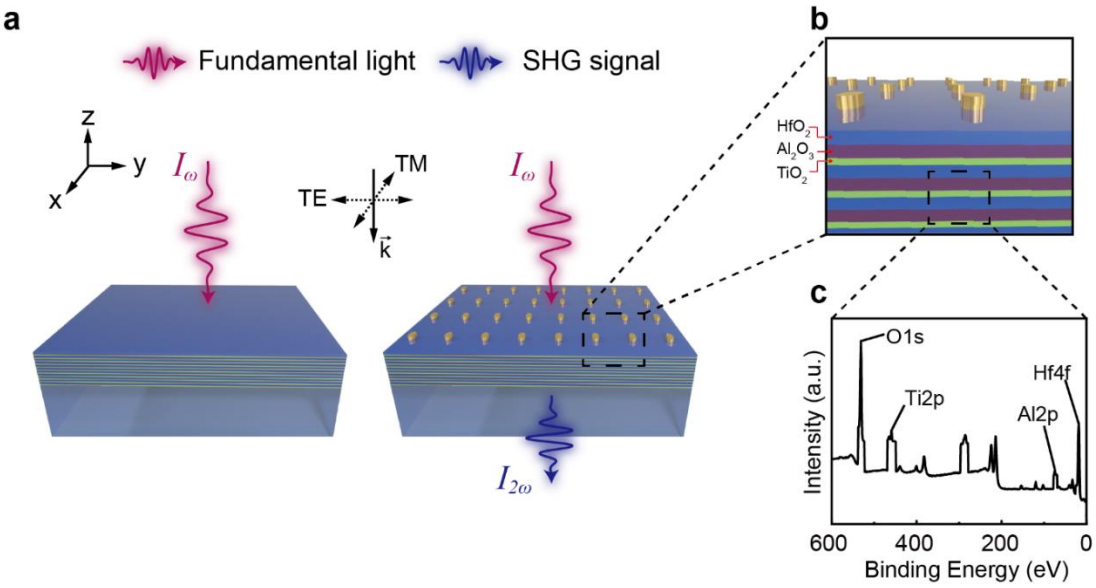

Figure 3.3 – Illustration of metamaterial device, as well as the material composition. (a) Schematic illustration of the ABC nanolaminate devices and definition of the TE and TM polarization of the incident fundamental light. Without plasmonic structures to generate an electric field with a z-component, the ABC nanolaminate will not emit a substantial second-harmonic response in the z-direction. (b) The nanolaminate is

comprised of three periodic layers, TiO_2, Al_2O_3, and HfO_2 and the plasmonic structure is patterned above the nanolaminate. (c) XPS survey spectra of the fabricated ABC nanolaminate consisting of TiO_2, Al_2O_3, and HfO_2 [3].

The general parameters for the metamaterial shown in Figure 3.3a are periodicity, p = 360 nm, thickness of the ABC composite, t_{ABC} = 75 nm, thickness of the gold plasmonic structure, t_{Au} = 45 nm, and normal incident wavelength, λ_ω = 850 nm. For the ABC composite, the individual layers also had specified thicknesses as seen in Figure 3.3b, the thickness of the Al_2O_3 layers, t_{Al2O3} = 0.9 nm, the thickness of the HfO_2 layers, t_{HfO2} = 0.9 nm, and the thickness of the TiO_2 layers, t_{TiO2} = 0.3 nm.

However, for the simulation portion of this research, the ABC nanolaminate was treated as a homogenized material with nonlinear susceptibilities. Referring back to Chapter 2, we can look at the nonlinear polarization to describe the second-harmonic response due to the nonlinear material. The nonlinear polarization of the layered structures with geometric features along the z-axis is calculated using this equation,

$$\begin{pmatrix} \boldsymbol{P_x(2\omega)} \\ \boldsymbol{P_y(2\omega)} \\ \boldsymbol{P_z(2\omega)} \end{pmatrix} = \epsilon_0 \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\chi_{xxz}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\chi_{xxz}} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\chi_{zxx}} & \boldsymbol{\chi_{zxx}} & \boldsymbol{\chi_{zzz}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{E_x(\omega)^2} \\ \boldsymbol{E_y(\omega)^2} \\ \boldsymbol{E_z(\omega)^2} \\ \boldsymbol{2E_y(\omega)E_z(\omega)} \\ \boldsymbol{2E_x(\omega)E_z(\omega)} \\ \boldsymbol{2E_x(\omega)E_y(\omega)} \end{pmatrix} \tag{3.1}$$

where,

$$\begin{aligned} \boldsymbol{P_x(2\omega)} &= \boldsymbol{2\epsilon_0\chi_{xxz}E_x(\omega)E_z(\omega)} \\ \boldsymbol{P_y(2\omega)} &= \boldsymbol{2\epsilon_0\chi_{xxz}E_y(\omega)E_z(\omega)} \\ \boldsymbol{P_z(2\omega)} &= \boldsymbol{\epsilon_0\chi_{zxx}(E_x(\omega)^2 + E_y(\omega)^2) + \epsilon_0\chi_{zzz}E_z(\omega)^2} \end{aligned} \tag{3.2}$$

and the $\chi^{(2)}$ susceptibilities used are $\chi_{zzz} = 0.667$ pm/V, $\chi_{xxz} = 0.254$ pm/V, and $\chi_{zxx} = 0.225$ pm/V [11, 12].

3.2.1 Deep Learning Framework

For this research, the deep learning framework outlined in section 3.1 was utilized. The framework is illustrated in Figure 3.1. These patterns are run through a full-wave simulation to find the actual SHG responses. The fundamental light electric field components are used to calculate the electric field components of the nonlinear response using the induced polarization of the ABC composite. Data collected from the full-wave simulation is used to train the simulator. Once trained, the generator is used to produce patterns for the simulator to predict respective SHG responses.

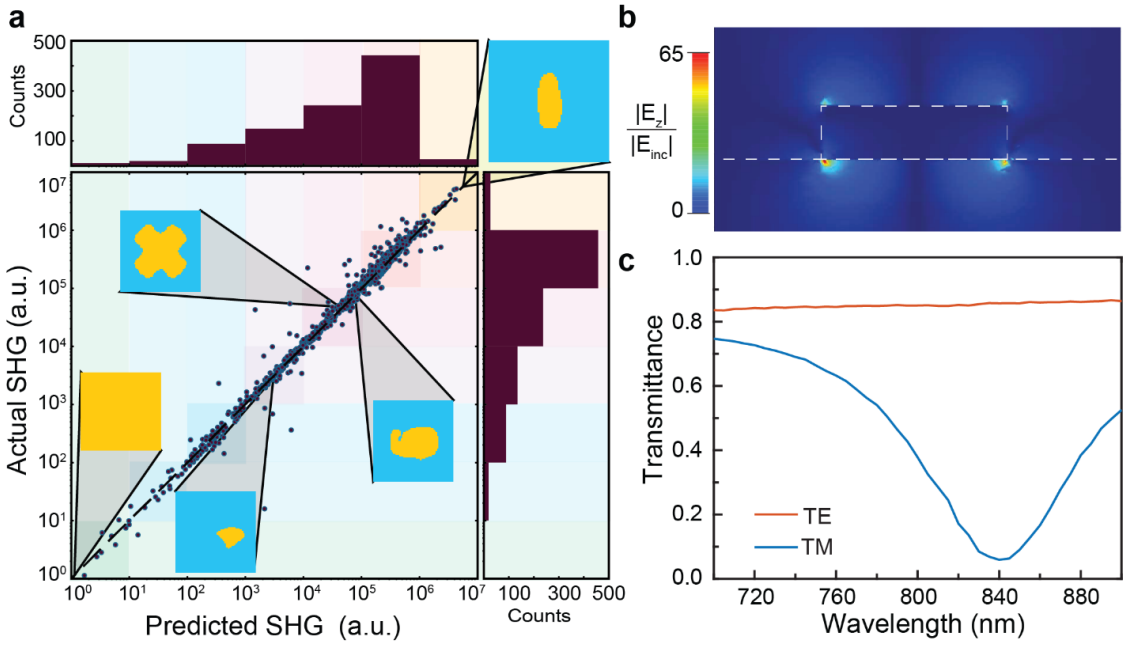

Figure 3.4 – Linear and nonlinear simulations based on deep learning algorithm. (a) The scatter plot is a comparison of predicted SHG enhancement values from the deep

learning simulator and actual SHG enhancement values from a full-wave simulation for the same set of patterns. The dotted line represents x=y, and the strong correlation between the line, and scattered points shows the simulator is very accurate for a wide range of SHG enhancement values. The histogram above the graph represents the spread of the predicted SHG enhancement values for the patterns from the simulator, and the histogram to the right represents the distribution of actual SHG enhancement values for the same patterns from a full-wave simulation. The SHG enhancement was calculated based on an unpatterned surface. The inset patterns in the graph are examples of patterns from the highlighted points in the scatter plot. The pattern in the top right is the pattern used for further simulation and fabrication in the paper. (b) Cross-sectional view, along the central x-direction or the length of the pattern, of the normalized z-component of the electric field with incident TM polarized light at a wavelength of 840 nm. The strong field enhancement due to the TM polarized light indicates that the SHG response induced in the ABC nanolaminate will be much higher when exposed to TM polarized light as compared to TE polarized light. (c) Calculated transmission spectra for TE and TM polarized light on the plasmonic ABC nanolaminate device [3].

The accuracy of the simulator, trained and modified for SHG responses, can be seen in Figure 3.4a, where 1000 patterns produced from the generator are run through both the full-wave simulation for the accurate SHG value, and the simulator, for the predicted value. The dotted line highlights the 1-1 correlation between the two. The histograms show that the spread of SHG values follows an expected curve for both the full-wave simulation data and simulator data. The whole image illustrates the reliability

and accuracy of the simulator algorithm. The optimizer is used to identify the pattern with the highest SHG response for the given parameters, and the resulting pattern is shown in the top right corner of Figure 3.4a.

The linear response, the field enhancement, and transmittance, of the identified optimal plasmonic structure was calculated using a full-wave simulation, as shown in Figure 3.4b and 3.4c, respectively. The plasmonic structure induces an electric field with a z-component, and the resonance in the linear response shows that a wavelength of 840 nm with TM-polarized light will lead to the most effective field enhancement. Utilizing the optimal field enhancement leads to a maximal SHG response from the ABC nanolaminate.

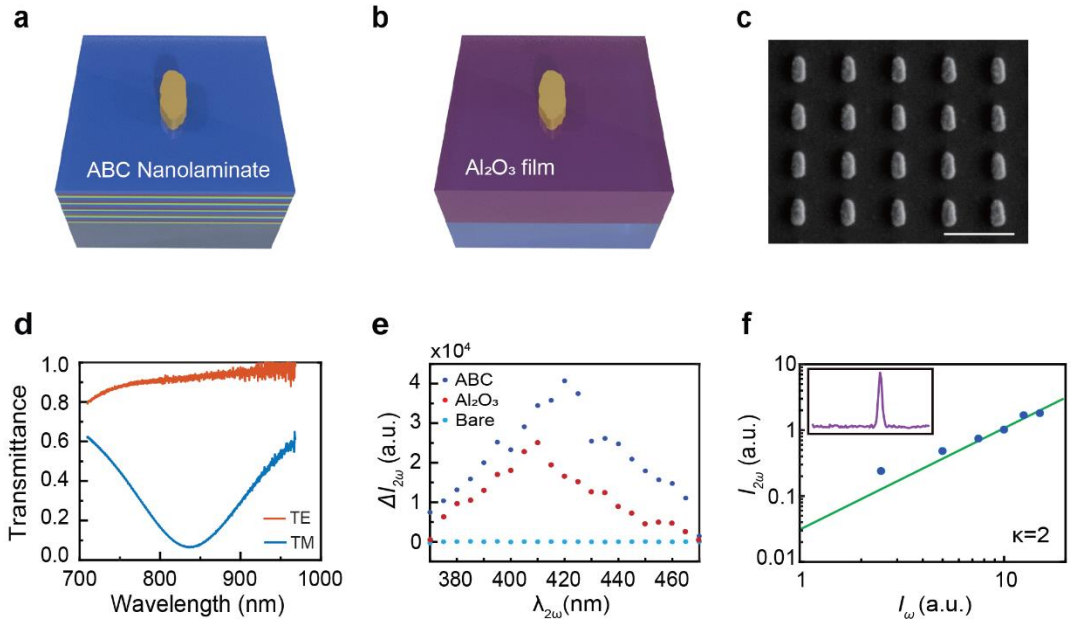

Figure 3.5 – Enhancement of the SHG from the ABC nanolaminate with the deep-learning designed plasmonic structure as well as linear and nonlinear measurements for the ABC nanolaminate device. (a, b) Visualization of a unit cell of ABC

nanolaminate device, and the control alumina device with an Al_2O_3 film, both with the same deep-learning optimized gold structure on the surface. (c) SEM image of the fabricated plasmonic structure on the ABC nanolaminate. The scale bar at the bottom represents 500 nm. (d) Experimental linear transmission spectra for TE and TM polarized light. The linear response for the ABC nanolaminate device is similar to the calculated linear response in Figure 3.4. The resonance wavelength of 840 nm will be used as the fundamental wavelength for subsequent nonlinear measurements. (e) Experimental SHG responses of the respective devices. The blue and red dots are from patterned ABC and alumina devices. The cyan dots represent the response from bare, unpatterned ABC nanolaminate. (f) Power dependence of the ABC patterned nanolaminate. The green line in the double logarithmic plot represents a slope of 2, confirming the second-order nature of the response. The inset shows the SHG spectrum for an incident wavelength of 850 nm measured from 370 nm-470 nm, to show the nature of the second harmonic response [3].

3.2.2 Experimental Results

For the experimental section of this research, two devices are fabricated – a device with the ABC nanolaminate and the gold pattern, as well as an alumina control device with an Al_2O_3 layer of the same thickness, replacing the ABC nanolaminate layer, below the same gold pattern, as shown in Figure 3.5a and 3.5b. The purpose of the alumina control device is to measure the nonlinear response due to the gold nanoparticle, to eventually differentiate the SHG response due to the ABC nanolaminate, and the SHG response from the gold structure. To ensure that the ABC nanolaminate was fabricated

correctly, an XPS analysis was done, the results are displayed in Figure 3.3c and demonstrates that all three oxides are present in corresponding ratios to the layer thickness. An SEM image, Figure 3.5c, was taken of the ABC nanolaminate device to validate that the patterned device closely matches the optimal plasmonic shape found through the deep learning algorithm. Figure 3.5d is the measured linear transmission response for the ABC device under both TM- and TE- polarized illumination at normal incidence. As the refractive index of the ABC nanolaminate (n = 1.714) and alumina (n = 1.7591) are similar, the linear responses are as well.

Comparing Figure 3.4c and Figure 3.5d, we can see that the simulated and experimental linear responses are in agreement, the slight differences are due to fabrication irregularities which are to be expected. The linear responses also indicate that the maximum SHG response for the ABC device should be expected at a wavelength 840 nm due to the resonance. The resonance in the linear response indicates that the plasmonic structure leads to the most extinction of light at the peak.

The optical nonlinear characterization in Figure 3.5e is the SHG response from the ABC nanolaminate. The device is excited by a fundamental wavelength, λ_ω, ranging from 740 to 940 nm, with a step size of 10 nm, from TM-polarized laser pulses at a constant intensity. The blue dots represent the SHG response from the patterned ABC device, the brown dots represent the frequency doubled output from the patterned control device, and the cyan dots represent the generated second-harmonic signal from the unpatterned ABC surface. The SHG values were evaluated by photon counting the intensity of the signal, and clearly show enhancement of the unpatterned ABC surface.

The last column of panels, Figure 3.5f, contains the power dependence plot for the ABC device. The plot shows the relationship between the fundamental light that is input, I_{ω}, and the intensity of the output SHG response, $I_{2\omega}$, on a log-log scale. The solid line represents $I_{2\omega} \propto I_{\omega}^{K}$, where K = 2, the quadratic dependency verifies the second-order nature of the signal. Referring to the SHG spectra, the maximum peaks show that the ABC device has clearly higher response, but to compare the two values is misleading. A proper comparison of the SHG response due to the plasmon induced second-harmonic response can be conducted by comparing the output polarization results.

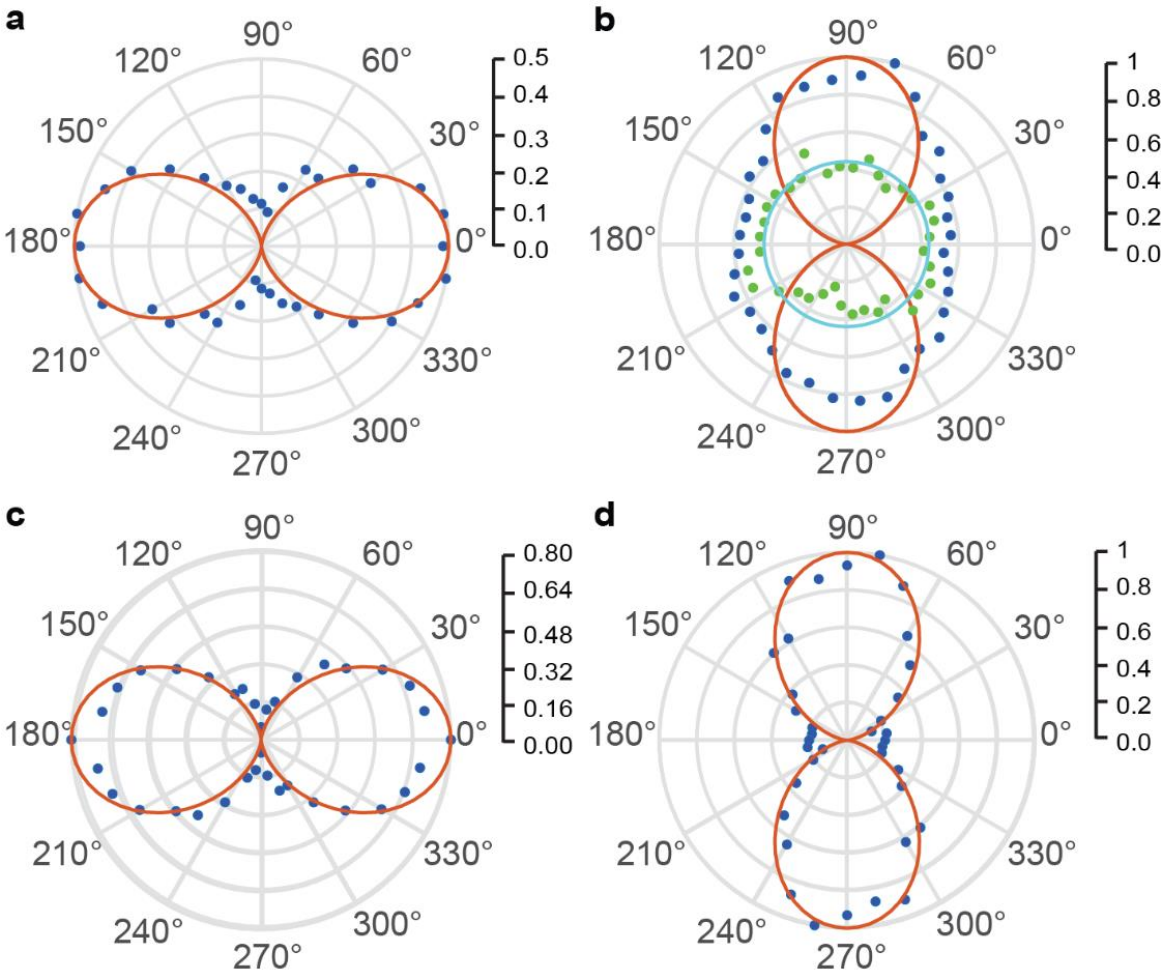

Figure 3.6 – Induced second harmonic response from ABC nanolaminate and SHG output polarization results for the devices. (a, b) For the patterned ABC device, the SHG output polarization in plotted in polar plots for TE and TM polarized incident fundamental light, respectively. The blue dots are the measured SHG response, and the red curve is the calculated SHG response. In (b), the green dots represent the SHG response induced from the ABC nanolaminate. The cyan circle best fits the SHG due to the ABC nanolaminate. (c, d) SHG output polarization plots for the patterned control device for TE and TM polarized incident fundamental light, respectively. The measured second-harmonic signal values are the blue dots, and the calculated second-

harmonic response is shown as the red curve. We expected the TE polarized incident light to lead to similar responses in the patterned ABC device, and the control device, and that is seen in (a) and (c). The loss of linear polarization observed in (a) and (b) indicate that the second harmonic response does not stem solely from the plasmonic structure and that is further confirmed by the degeneration of the peanut shape in (b) [3].

The SHG output polarization characterization reveals the presence of an SHG signal from the ABC nanolaminate under normal incidence. Figure 3.6a and 3.6b show the output polarization of the ABC device for TE- and TM-polarized incident fundamental light, respectively. The blue dots are the experimental data taken at every 5^{o}, and for the TE-polarization, the red peanut or two-lobe shape represents the data collected after a linear polarizer. The TM-polarization plot shows a wider shape than normally expected for an SHG signal. The peanut or two-lobe shape is expected for an SHG response, as the frequency doubled output is linked with the polarization angle.

Comparing this plot to the alumina control device for TE- and TM-polarized incident fundamental light (Figure 3.6c and 3.6d, respectively) elucidates the difference. It is essential to realize that for the control device, both the TM- and TE-polarized incident fundamental light lead to peanut shapes with a very tight ‘waist’, having zero waist indicates the light is linearly polarized, along the TM and TE polarization, respectively.

For the patterned ABC device with incident TE-polarization, it is easy to see that, while similar to the control device, the waist of the peanut is wider, or no longer perfectly

linearly polarized, with a nonzero ellipticity. Analyzing the incident TM-polarized fundamental light figure gives us a better insight into the phenomena.

In Figure 3.6b, the red curve represents the peanut best fit curve from the control device, we can refer to that as the polarization due to the gold pattern. The green dots in the figure were calculated by subtracting the measured SHG due to the gold pattern, from the measured SHG response of the patterned ABC device, to find the SHG response stemming solely from the ABC nanolaminate. The circle best fits the calculated data. Referring to the output polarization of the incident TE-polarized fundamental light, we can see there must be a very small circle relating to the ABC composite polarization present there too.

This shows that there is a large presence of the SHG response due to the gold pattern, which is why the earlier comparison of the $\Delta I_{2\omega}$ from the SHG spectra was misleading. To separate the SHG response from the plasmonic structure and the SHG response of the ABC composite, the incident TM-polarized fundamental light output polarization is crucial.

We have demonstrated an inversely designed plasmonic structure that maximizes the second-harmonic generation in an ABC nanolaminate nonlinear metamaterial. Further, we were able to differentiate the optical nonlinear response due to the plasmonic structure, and the response due to the $\chi^{(2)}$ of the ABC nanolaminate. The deep learning algorithm is a robust and flexible means for inverse design tasks that can lead to new and exciting metamaterial designs. Traditional guess-and-check methods involving parameter sweeping, and simple geometries cannot guarantee optimal results for complex nonlinear responses.

Occasionally, traditional methods also cannot reach a solution for certain desired optical responses. The deep learning framework utilized in this research is a capable instrument that can be applied to a variety of nonlinear optical responses, and more challenging applications. It is important to note that in this research, a single unit structure is designed and described. There have been many multi-unit structures identified and studied for plasmonic and nonlinear optical purposes. These include meta-lens and antenna structures [13-15].

Multi-unit patterns can increase the applications of the research and allow for more precise optimization. However, this research specifically highlighted single unit structures to simplify the algorithm in order to show the enhancement of nonlinear generation of light in the ABC nanolaminate is possible and successful.

Using the deep learning algorithm, our results represent the ability to optimize and demonstrate the nonlinear optical response of a thin-film nonlinear material using a plasmonic metamaterial device. The plasmonic ABC nanolaminate structure also allows for further investigation of optical phenomena that can be applied to areas such as integrated optics, as the ABC nanolaminate is compatible with inorganic materials and CMOS technology. The results of this research expand the scope of inverse design applications and shows the potential for new optical nonlinear phenomena.

3.2.3 Further Fabrication and Experimental Details

3.2.3.1 Device Fabrication

The fabrication process starts with depositing ABC nanolaminates on the glass substrate (Corning, C1737 glass) via atomic layer deposition technique (Cambridge Fiji

Plasma Atomic Layer Deposition System). The ABC layer is composed of 3.57 Å of TiO2, 9.56Å of Al2O3, and 9.11Å of HfO2 layer, the ABC layer deposition is repeated 34 times to get a 756 Å thickness of the ABC nanolaminate, following the material and geometric parameters of the ABC from Alloatti et al[16].

After the ABC nanolaminate is fabricated, the gold nanoparticles are formed on top of the nanolaminates in a three-step process: (i) standard electron beam lithography (E-beam litho., Elionix ELS-G100 EBL system) using poly(methyl methacrylate) (PMMA) as the positive tone electron resist, (ii) E-beam evaporation of 3 nm/45 nm Cr/Au metal, and (iii) an overnight lift-off process in acetone to resolve the plasmonic structures.

3.2.3.2 Linear Optical Characterization

A tungsten halogen lamp (B&W Tek BPS 120) is used as a broadband light source to characterize the linear response of the device. The polarization of the input light source is controlled by a set of linear polarizers and half waveplates. The transmittance spectra of the device at a normal incident angle is focused on the sample using 10× objective (NA: 0.25) and collected with a 20× objective (NA: 0.4).

The light collected by the objective is delivered to the spectroscopy system (Princeton Instrument Acton SP 2300i with PIXIS 400B camera). For the reflectance spectra of the device, a 20× objective (NA: 0.4) is used to both focus and collect the light.

3.2.3.3 Nonlinear Optical Characterization

The excitation source for the nonlinear optical characterization is a tunable Ti:Sapphire ultrafast oscillator (Spectra-Physics, Mai Tai HP, 690–1040 nm) with a 100

fs pulse duration and 80 MHz repetition rate. The power and polarization state are controlled by a set of halfwave plates and a Glan polarizer. The fundamental beam is delivered and focused on the device via a 10× objective (NA: 0.25), which results in a spot size of ~20 μm on the sample.

Both the fundamental light and the harmonic generated light from the sample are collected by a 20× objective (NA: 0.4), where the fundamental wave is eliminated as it passes through the bandpass filter. The harmonic signal is then characterized by the detector system which is composed of a monochromator (Princeton Instruments, IsoPlane) with a charge-coupled device (CCD) camera (Princeton Instruments, Pixis 400B).

The state of polarization of the harmonic generated signal is analyzed by using a rotating polarizer. The peanut or lob-shaped pattern shown in the polar diagram was formed by collecting the portion of the intensity of the harmonic generated light that passes through the rotating polarizer.

3.2.4 Additional Calculations and Details

3.2.4.1 SHG Conversion Efficiency

The conversion efficiency, $\eta = \frac{P_{avg}^{SH}}{P_{avg}^{FW}}$, for the ABC patterned device was calculated using the average fundamental wave power, $P_{avg}^{FW} = .00133W$, and the average second harmonic power, P_{avg}^{SH},which was calculated. To calculate P_{avg}^{SH}, first the integrated counts, N_c, were calculated from the intensity counts without post-processing multiplied

by the software control gain,

$N_c = 14181\ counts * 2\ \left(\frac{e^-}{cnt}\right) = 28362\ (integrated\ counts).$

The integrated counts were multiplied by the saturation counts and divided by the spectrometric well capacity, 300 ke$^-$, to get the photoexcited electrons, $N_c = \frac{N_e}{\zeta}, \zeta = \frac{300ke^-}{65535counts} = \frac{4.5777e^-}{cnts}, N_e = 129832.7274e^-$. The photoexcited electrons are then divided by the quantum efficiency of the sensor at λ_{SH},

$N_e = QE\ \times N_{Pixis}, QE = .6644, \qquad \lambda_{SH} = 420nm,\ N_{Pixis} = 195413.5\ photons.$

Using the total transmission and effective surface, this can be converted to the photons emitted at the device, $N_{PH} = \frac{N_{Pixis}}{T_{total} \times S_{Eff}}, T_{total} = 0.10088628, S_{eff} = 0.10262345, N_{PH} = 1.88745 \times 10^7 \approx 1.9 \times 10^7 photons$. The SH energy is calculated as follows,

$E_{SH} = N_{PH} \times \frac{hc}{\lambda_{SH}} = 1.9 \times 10^7 \times \frac{1.24}{0.420} \times 1.6022 \times 10^{-19} = 8.986 \times 10^{-12} W \cdot s.$

Which leads us to the average power of the SH emission and conversion efficiency,

$P_{avg}^{SH} = \Phi_e = \frac{E_{SH}}{t_e} = \frac{E_{SH}}{60s} = 1.4977 \times 10^{-13}\ W, \eta = 1.126 \times 10^{-10}$.

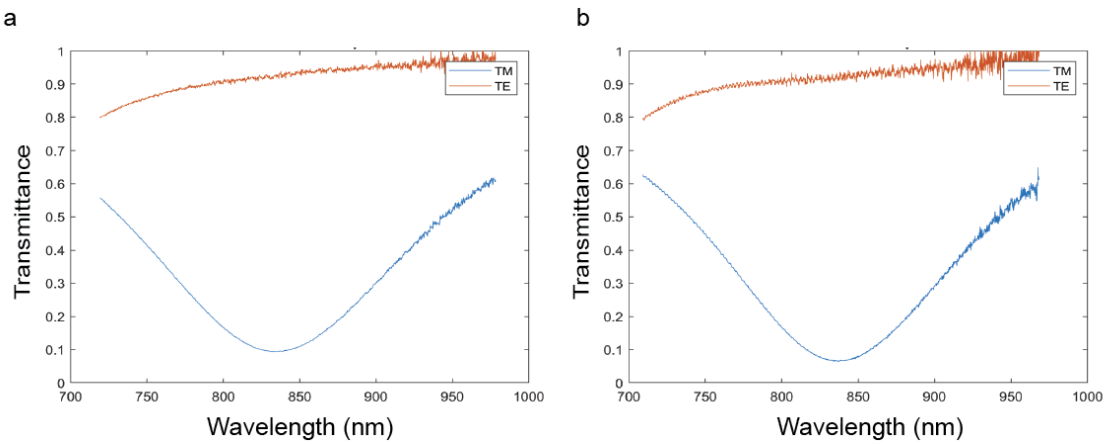

Figure 3.7 – Linear transmittance measured for the (a) alumina control device, and the (b) ABC nanolaminate device [3].

3.2.4.2 Experimental Linear Response

Experimental linear transmission spectra for TE and TM polarized light. The linear response for the alumina control device is similar to that of the ABC nanolaminate. The resonance of the alumina device in Figure 3.7 is slightly different due to the minor difference in the refractive indices of the two materials.

3.2.4.3 ABC Film Fabrication

The TiO2, Al2O3, and HfO2 (ABC) nanolaminates on the glass substrate were deposited via atomic layer deposition technique (Cambridge Fiji Plasma Atomic Layer Deposition System). The deposition temperature was 200 °C and the deposition rate of TiO2, Al2O3, and HfO2 was 0.51 Å/cycle, 1.195 Å/cycle, and 1.139 Å/cycle, respectively.

First, the refractive index of each film was characterized by ellipsometry (J.A. Woollam, M-2000), where the films were deposited up to 100 nm separately on the glass as shown in Figure 3.3a.

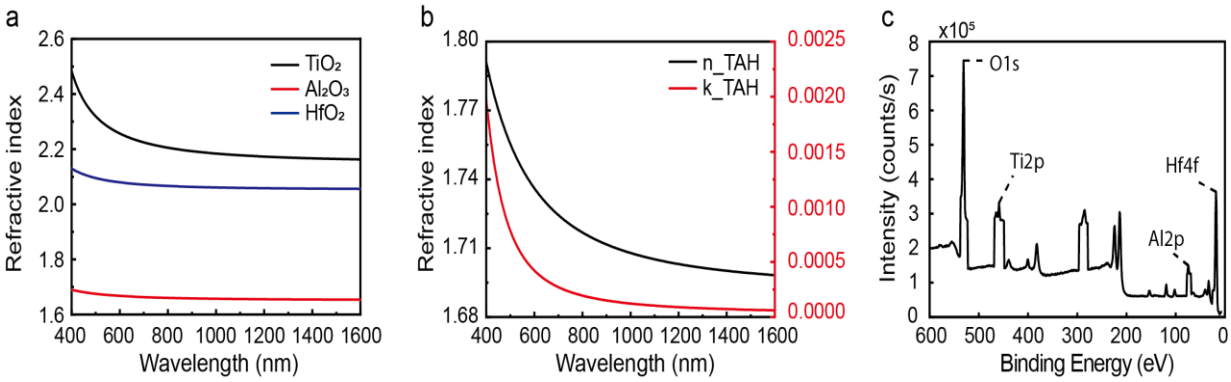

Figure 3.8 – Film characterizations. (a) Real part of the refractive index of TiO_2, Al_2O_3, and HfO_2. (b) Real and imaginary part of refractive index of ABC nanolaminate. (c) XPS results of prepared ABC nanolaminate on glass substrate [3].

Next, 756 Å thickness of the ABC nanolaminate was deposited on the glass substrate, and the refractive index and the composition of the film were characterized by ellipsometer and X-ray photoelectron spectroscopy (XPS, Thermo K-Alpha XPS, a monochromated small-spot XPS system), respectively. As shown in Figure 3.8c, both results clearly show that the ABC nanolaminate was well deposited on the glass substrate.

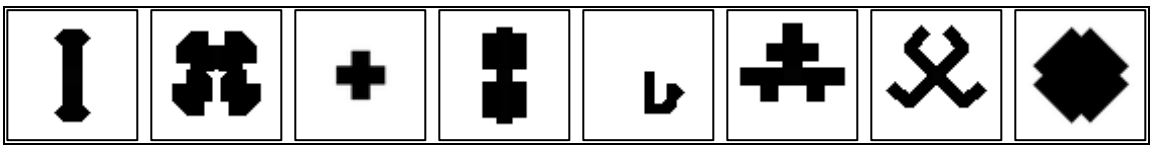

Figure 3.9 – Sample of 10,001 patterns created for training data to be run through Comsol to generate 10,001 SHG responses [3].

3.2.4.4 Pattern Optimization

After the structure was confirmed, 10,001 patterns were created in order to begin generating a set of training data. The patterns created were based on simple geometries,

which were then transformed to create a set of 10,001 patterns. It is important to note that these patterns are all single unit patterns, and not multi-unit patterns. Figure 3.9 shows a sample of patterns.

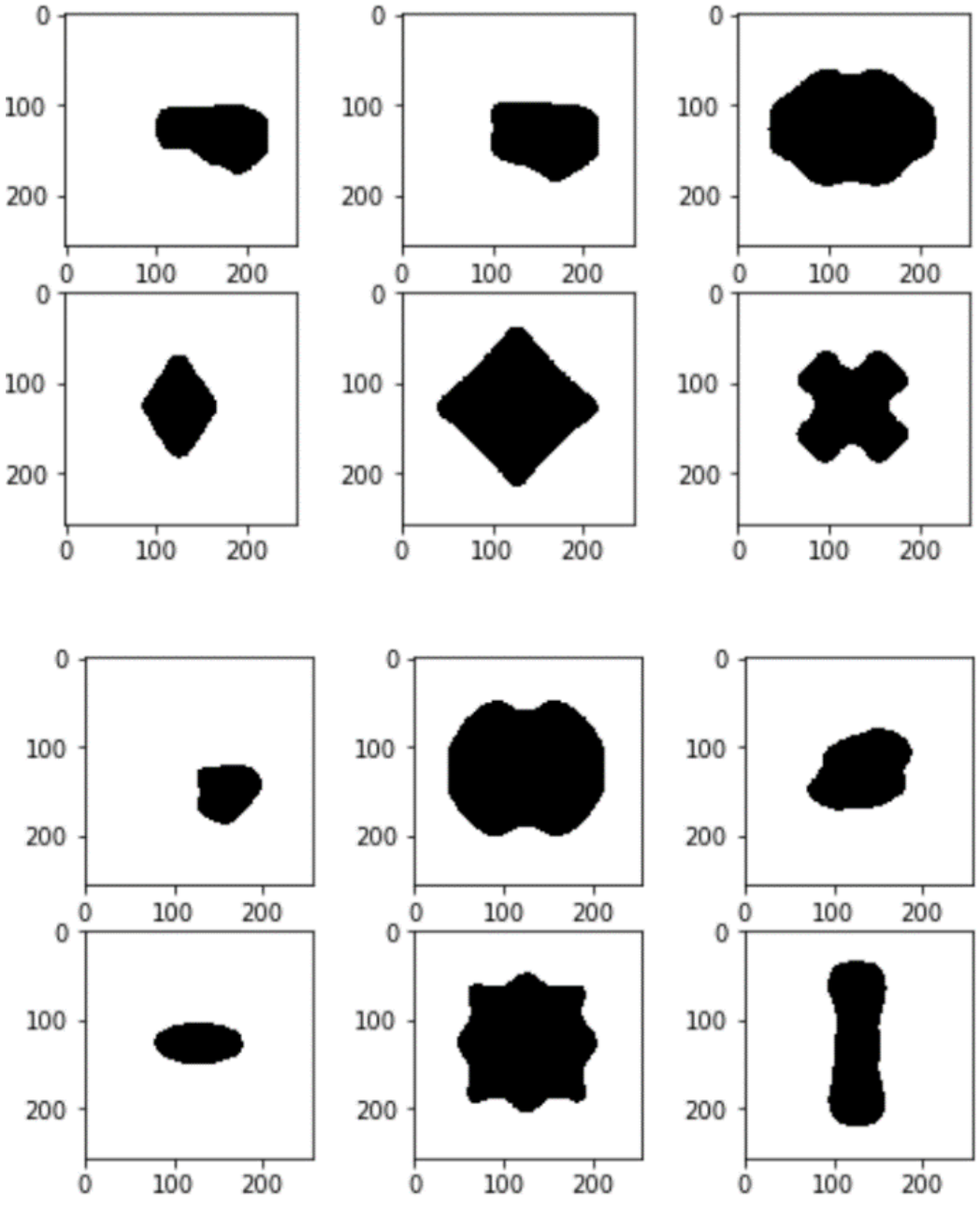

Figure 3.10 – 40 randomly generated patterns from the VAE [3].

These patterns were then used in conjunction with a Comsol model of the Au-LiNbO3 structure to generate 10,001 SHG responses. These responses can be considered the training data for the inverse design algorithms. Once the SHG responses were collected, a Variational Autoencoder (VAE) was used to be able to randomly generate patterns for the inverse design capabilities, sample images are in Figure 3.10.

The next step was to use the VAE to generate random patterns, and then to be able to predict the SHG value for each pattern. To do so, a simulator based on the ResNet19 architecture was used. To train the simulator, the 10,001 SHG responses and patterns were split into training and validation sets of sizes 8,201 and 1,800, respectively. Once the simulator model was trained, preliminary SHG responses, with no optimization, could be found.

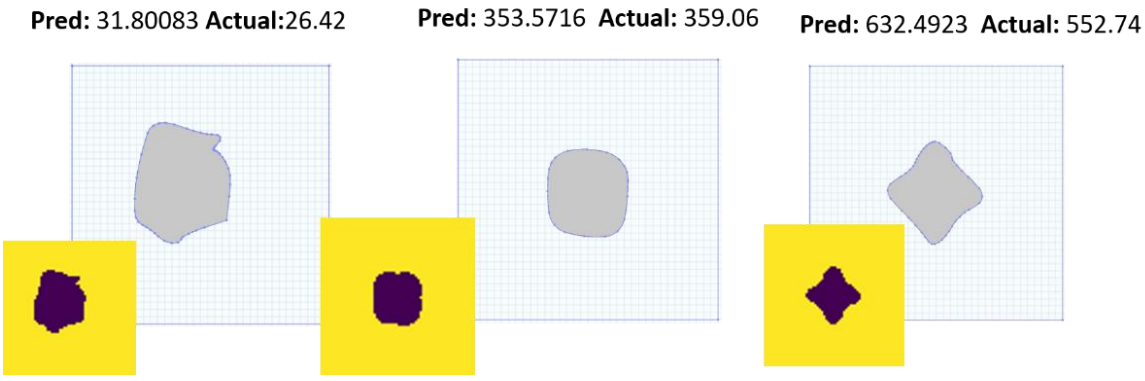

Figure 3.11 – Sample patterns with their actual SHG values from a full-wave simulation and the predicted from the algorithm. The values are SHG enhancement with arbitrary units for ease of comparison [3].

To see if the algorithm resulted in correct SHG responses, a few patterns were run through a full-wave simulation to compare the predicted (from the algorithm) and actual

values (from the full-wave simulation) as shown in Figure 3.11. Figure 3.11 also shows what the patterns look like when converted to .dxf files, the grey patterns shown are the .dxf files. This file conversion was necessary to be able to calculate the SHG values in the full-wave simulator. .dxf files are similar to the format used in cad software. The conversion also results in a loss of fine features.

Once the algorithm was found to be accurate, optimization was implemented utilizing an Evolution Strategy. Once applied, the algorithm could search for specified SHG values.

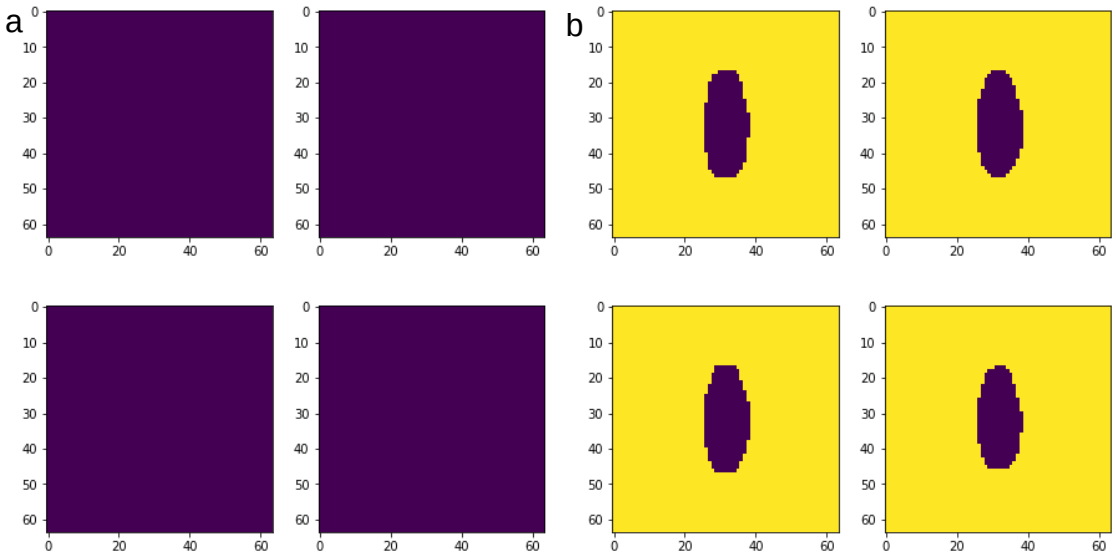

Figure 3.12 – Patterns returned from the algorithm optimized for (a) no SH generation, and (b) maximum SH generation [3].

Figure 3.12 shows that there is slight variation in the patterns that have maximum SH generation. This alludes to fabrication differences not largely affecting the SHG.

3.3 References

[1] Z. Liu, L. Raju, D. Zhu, and W. Cai, "A Hybrid Strategy for the Discovery and Design of Photonic Structures," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems,* vol. 10, no. 1, pp. 126-135, 2020.

[2] I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via Deep Learning," *Light: Science & Applications,* vol. 7, no. 1, p. 60, 2018/09/05 2018.

[3] L. Raju *et al.*, "Maximized Frequency Doubling through the Inverse Design of Nonlinear Metamaterials," *ACS Nano,* vol. 16, no. 3, pp. 3926-3933, 2022/03/22 2022.

[4] L. Carletti *et al.*, "Steering and Encoding the Polarization of the Second Harmonic in the Visible with a Monolithic LiNbO3 Metasurface," *ACS Photonics,* vol. 8, no. 3, pp. 731-737, 2021/03/17 2021.

[5] Y. Chen *et al.*, "Topology Optimization-Based Inverse Design of Plasmonic Nanodimer with Maximum Near-Field Enhancement," *Advanced Functional Materials,* vol. 30, no. 23, p. 2000642, 2020.

[6] C. Sitawarin, W. Jin, Z. Lin, and A. W. Rodriguez, "Inverse-designed photonic fibers and metasurfaces for nonlinear frequency conversion [Invited]: publisher’s note," *Photonics Research,* vol. 7, no. 4, pp. 493-493, 2019/04/01 2019.

[7] R. Wang, F. Liang, and Z. Lin, "Data-driven prediction of diamond-like infrared nonlinear optical crystals with targeting performances," *Scientific Reports,* vol. 10, no. 1, p. 3486, 2020/02/26 2020.

[8] L. Xu *et al.*, "Enhanced light–matter interactions in dielectric nanostructures via machine-learning approach," *Advanced Photonics,* vol. 2, no. 2, p. 026003, 2020.

[9] Y. Zhang, N. K. Grady, C. Ayala-Orozco, and N. J. Halas, "Three-Dimensional Nanostructures as Highly Efficient Generators of Second Harmonic Light," *Nano Letters,* vol. 11, no. 12, pp. 5519-5523, 2011/12/14 2011.

[10] A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vučković, "Inverse design and demonstration of a compact and broadband on-

chip wavelength demultiplexer," *Nature Photonics,* vol. 9, no. 6, pp. 374-377, 2015/06/01 2015.

[11] R. W. Boyd, *Nonlinear optics*, 3rd ed. Amsterdam ; Boston: Academic Press, 2008, pp. xix, 613 p.

[12] A. Hermans *et al.*, "On the determination of $\chi(2)$ in thin films: a comparison of one-beam second-harmonic generation measurement methodologies," *Scientific Reports,* vol. 7, no. 1, p. 44581, 2017/03/20 2017.

[13] X. Chen *et al.,* "Dual-polarity plasmonic metalens for visible light," *Nature Communications,* vol. 3, no. 1, p. 1198, 2012/11/13 2012.

[14] M. A. Nazirzadeh, F. B. Atar, B. B. Turgut, and A. K. Okyay, "Random sized plasmonic nanoantennas on Silicon for low-cost broad-band near-infrared photodetection," *Scientific Reports,* vol. 4, no. 1, p. 7103, 2014/11/19 2014.

[15] D. Zhu, Z. Liu, L. Raju, A. S. Kim, and W. Cai, "Building Multifunctional Metasystems via Algorithmic Construction," *ACS Nano,* vol. 15, no. 2, pp. 2318-2326, 2021/02/23 2021.

[16] L. Alloatti *et al.,* "Second-order nonlinear optical metamaterials: ABC-type nanolaminates," *Applied Physics Letters,* vol. 107, no. 12, p. 121903, 2015.

CHAPTER 4. SUM-FREQUENCY GENERATION

The objective of the sum-frequency generation research was to expand upon the demonstrated ability to optimize materials for linear optical responses and a nonlinear second harmonic optical response to other linear and nonlinear optical responses and advance the current data learning framework. Utilizing the framework from the previous research in Chapter 3, we can optimize a similar structure for sum-frequency generation.

In the previous framework for SHG, the geometries were single unit structures [1]. For this research, we will build upon the existing Compositional Pattern-producing Network (CPPN). A CPPN is essentially an ANN with genetic algorithms. The prior research used a VAE with no genetic evolution. Effectively, a CPPN can create infinite resolution encoded patterns.

The purpose of this device is to upconvert a near infrared (NIR) signal to the visible regime. This research has many very important applications. Using a lab fabricable nonlinear metamaterial such as the ABC nanolaminate decreases the cost and thickness of the device. Such as easy to fabricate device can also be easily manufactured.

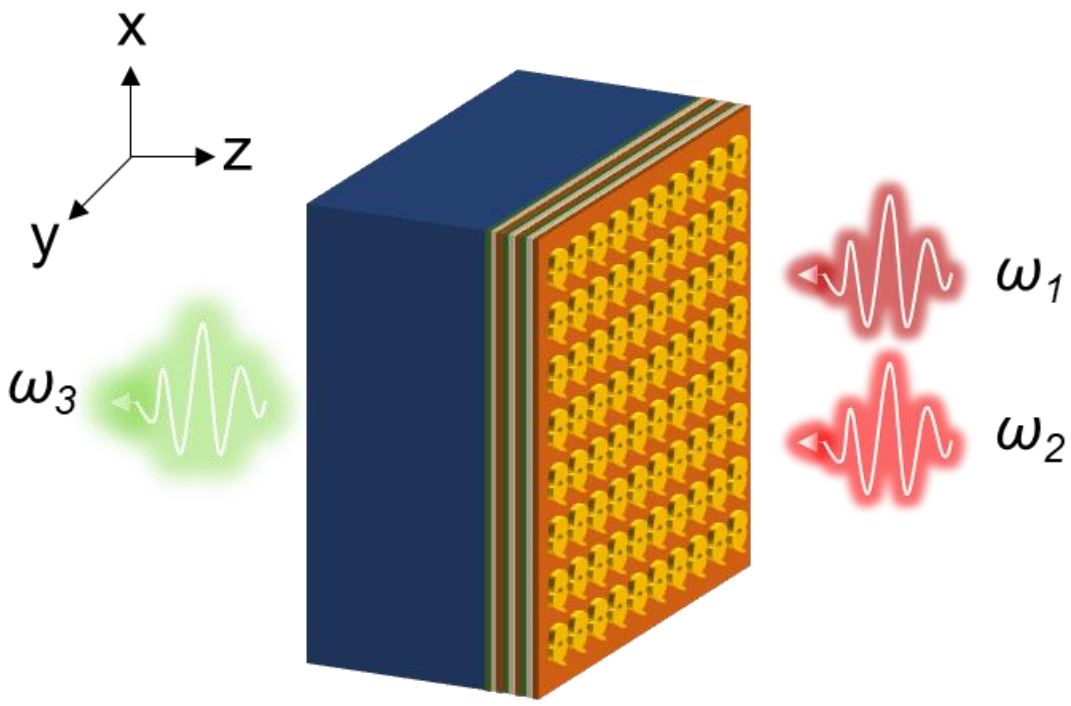

Figure 4.1 – Illustration of SFG nonlinear metamaterial device. The schematic shows the ABC nanolaminate with a periodic plasmonic structure atop. The incident infrared light, ω_1 and ω_2, is normal to the surface of the metamaterial device. The plasmonic structures here generate a z-component of electric field, allowing the ABC nanolaminate to emit visible light, ω_3 through sum-frequency generation. The nanolaminate is comprised of three periodic layers, TiO_2, Al_2O_3, and HfO_2 and the plasmonic structure is patterned above the nanolaminate.

4.1 SFG Nonlinear Metamaterial Parameters

The device parameters are as follows, periodicity, $p = 380\ nm$, ABC nanolaminate thickness, $t_{ABC} = 75\ nm$, gold plasmonic thickness, $t_{Au} = 45\ nm$, incident IR wavelengths,

$\lambda_1 = 800\ nm$ *and* $\lambda_2 = 1550\ nm$, output wavelength, $\lambda_3 = 528\ nm$. A schematic of the device is shown in Figure 4.1.

Once the device parameters are finalized, the full-wave simulation of sum-frequency generation in such a device was created. In this instance, Comsol was used as the full-wave simulation software. The full-wave simulation was validated in a number of ways, but most importantly by comparing the results of the sum-frequency generation simulator with the same parameters of the second-harmonic generation device.

As earlier described, SHG is a special case of SFG, where both incident frequencies are identical, and the output is double the incident frequency. Following this, computing the sum frequency generation where $\omega_1 = \omega_2$ led to the same results as shown in Chapter 3.

4.2 Data Learning Framework Update

The general framework for this work, focused on SFG, is similar to that from the SHG research, in that a pattern generation network and simulator network are required[1]. Given the SFG nonlinear metamaterial parameters, utilizing a pattern generator and simulator, we can find an optimal plasmonic geometry for the device to maximize the sum-frequency generation.

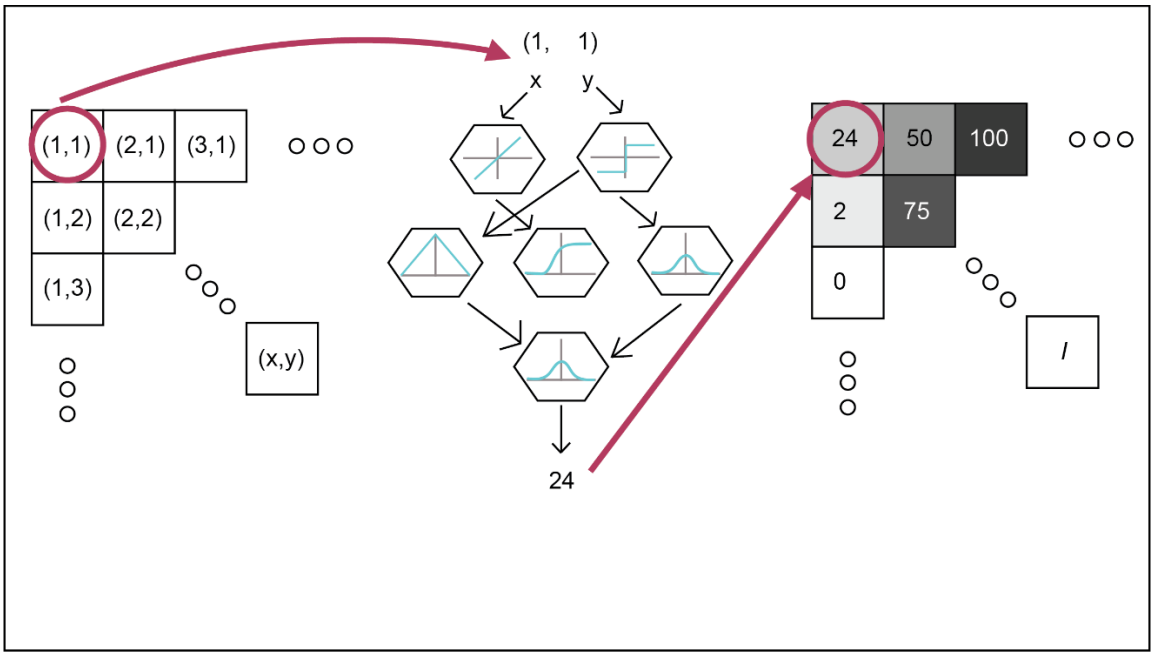

Figure 4.2 – Compositional Pattern Producing Network Architecture. The CPPN works by defining the image canvas into a grid of pixels with coordinates. Each pixel is then run through the CPPN and an intensity value is output. This is repeated for each pixel leading to a canvas full of generated intensities, in other words, an image.

As part of this data learning framework, here a CPPN is used for pattern generation. Previously a Variational Auto-Encoder was used and single unit geometries were generated. As SFG is a more complex process than SHG, a decision to utilize multi-unit geometries led to the need for a CPPN. Figure 4.2 shows the architecture of a CPPN.

The possibilities of complex CPPNs were highlighted in 2006 by K. Stanley, and it's use has only skyrocketed since then[2]. As described in Chapter 1, ANNs can be used to perform the classification of data with multiple layers of artificial neurons. CPPNs are a version of ANNs that also use genetic algorithms as an evolution method[3].

In order to truly understand the idea behind CPPNs, some background on activation functions is required. The way that ANNs work depends on artificial neurons. In ANNs, an activation function allows for the 'learning' of complicated patterns that might occur in the data. This is crucial for deep learning applications to be able to 'learn' and more importantly 'create' or 'identify' from patterns that humans can't easily find in large and complex data sets.

Activation functions add nonlinearities to networks, letting the networks essentially do more complex procedures. These functions effectively determine the neural network output using equations. These equations range from step and sigmoid functions to hyperbolic tangents and the identity function. Some examples of activation function plots are shown in Figure 4.2 in the CPPN architecture.

This collective use of varied activation functions leads to the naming of CPPNs – *compositional* pattern-producing networks. This also allows for the infinite number of possible geometries as the results can be sampled at any resolution. Practically speaking, in this research, CPPNs permit the generation of both single and multi-unit geometric structures for the plasmonic pattern.

The use of genetic algorithms comes to play in the choice of activation functions for each iteration or evolution. Not every pixel will go through the same relationship of activation functions, which is why CPPN's can create such complex patterns.

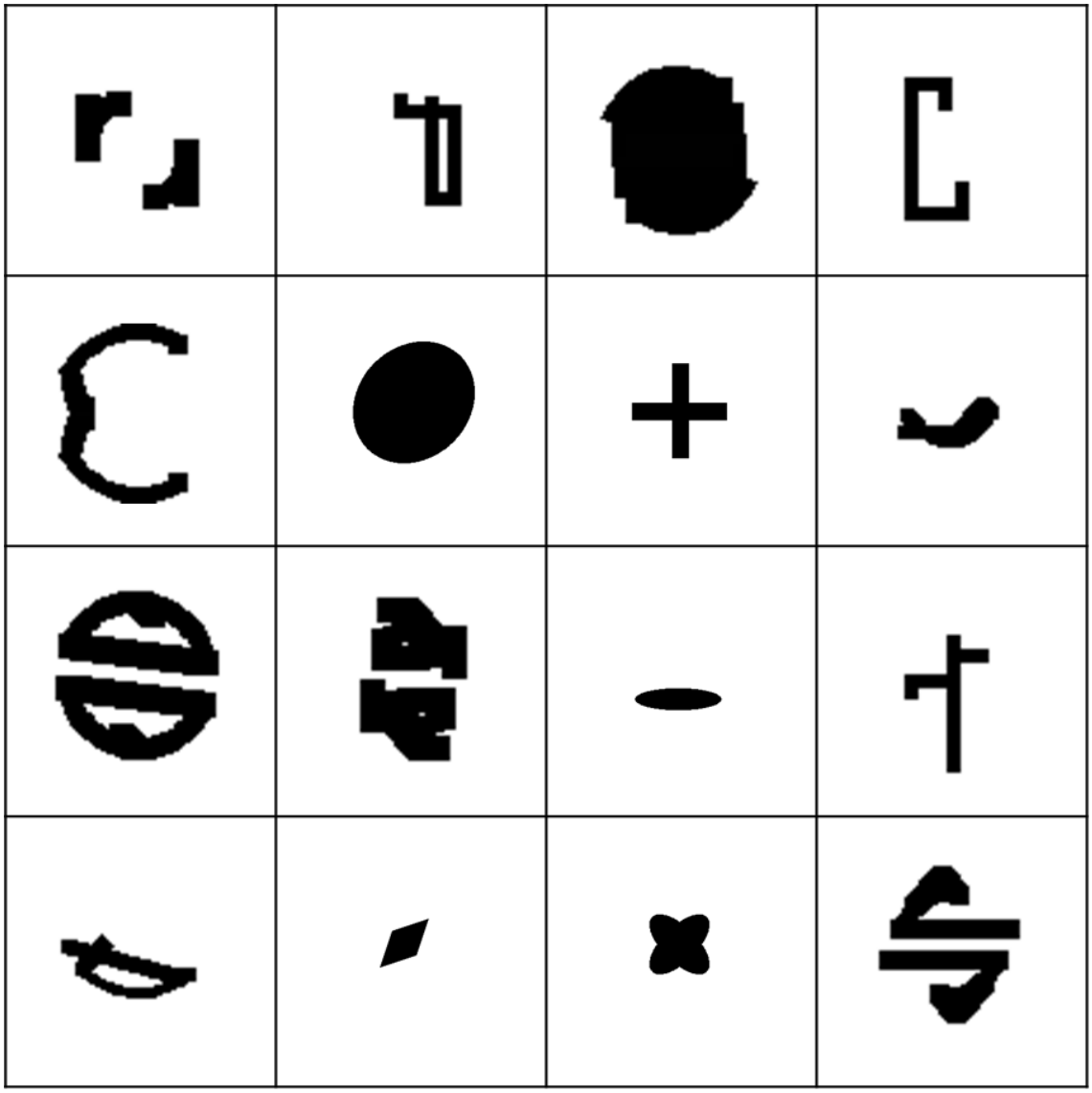

Figure 4.3 – Sample of patterns used as training data for CPPN. 10,000 patterns were used to train the CPPN or pattern generator.

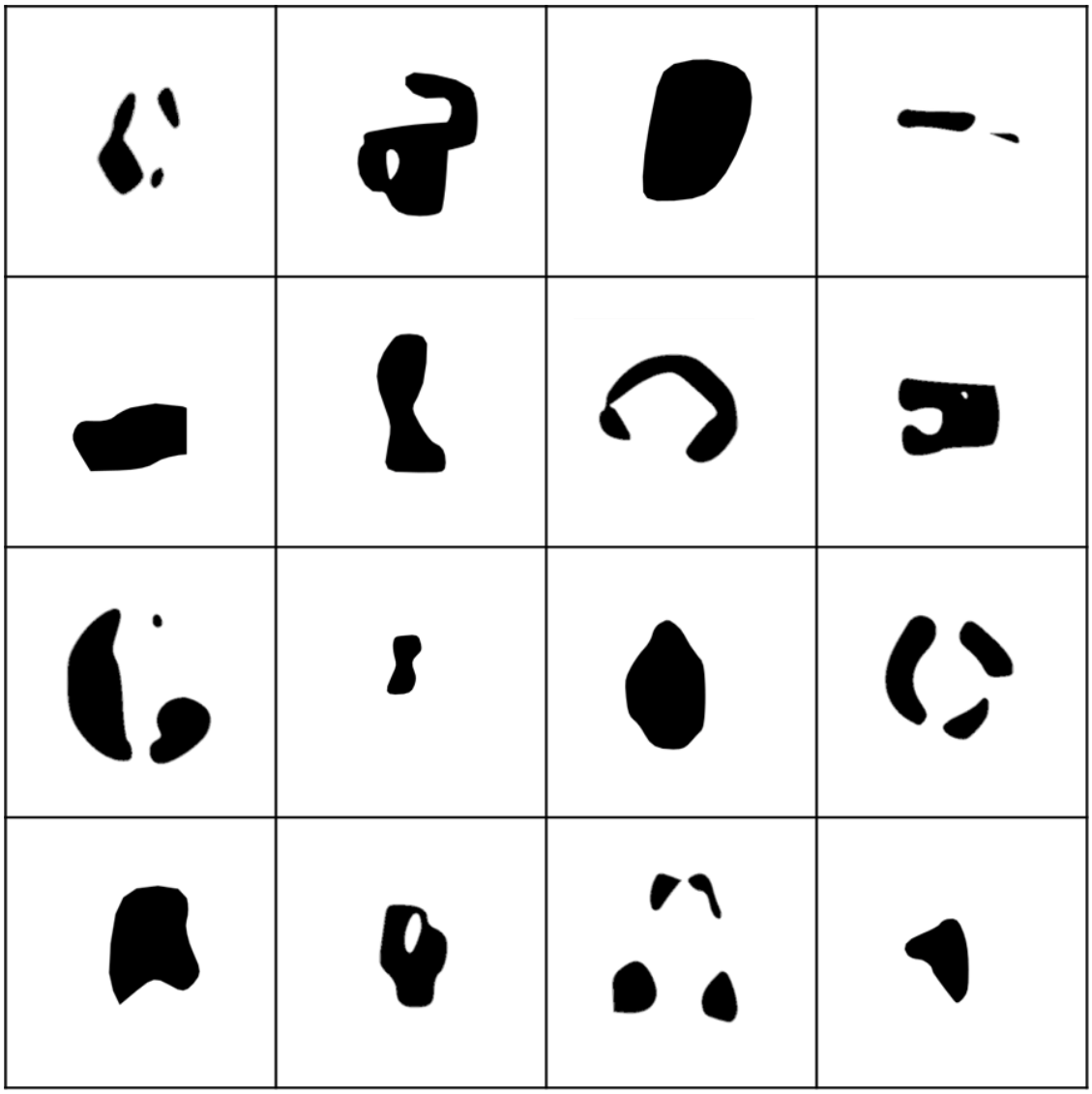

Figure 4.4 – Randomly generated patterns from CPPN. A variety of single and multi-unit geometries can be seen. There are also patterns that contain 'holes' that were not present in the patterns generated by the VAE. All of these show the increase in possible geometries when using the CPPN architecture for pattern generation.

For the maximization of SFG in our device, the CPPN pattern generator is crucial. Figure 4.3 shows a sample of patterns used to train the CPPN, and Figure 4.4 shows a sample of randomly generated patterns from the CPPN.

Once the pattern generator was created, then a deep learning simulator could be used to simulate the sum-frequency generation for any plasmonic geometry atop our nonlinear metamaterial. A similar simulator as described in Chapter 3, again based on a ResNet18 architecture, was created.

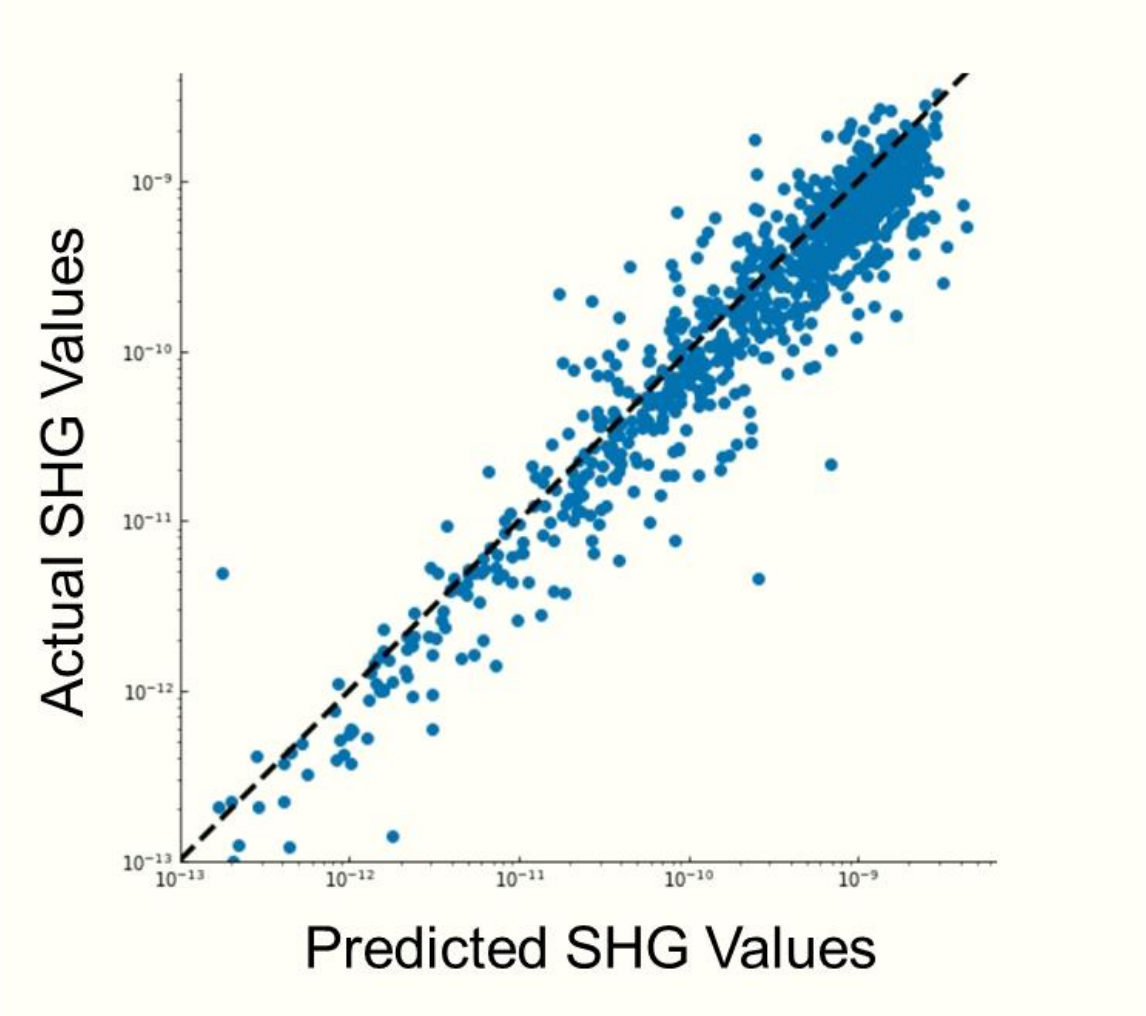

Figure 4.5 – Accuracy of simulator tested with 1,000 newly generated patterns run through both a full-wave simulator, and the simulator algorithm in arbitrary units. The dotted line represents a 1-1 correlation between the actual and predicted values.

The values near 10^{-9} represent a large SHG response, the values near the origin of the graph represent smaller SHG responses.

4.3 Sum-Frequency Generation Optimization

To train the simulator, 10,000 patterns were randomly generated from the CPPN. Then those patterns were input into the full-wave simulator to produce what will be called 'actual' SHG values. Both the patterns and corresponding SHG values are then used for training and validation of the simulator. Once the simulator was trained, the accuracy of the SHG 'prediction' needed to be tested.

The CPPN was used to generate 1,000 new patterns. These patterns were then run through the full-wave simulator for 'actual' SHG values, and then also run through the simulator for 'predicted' SHG values. The results are plotted in Figure 4.5. The values have arbitrary units, but the trend is significant, the following the 1-1 correlation of the dotted line, indicating good accuracy of the simulator.

Accuracy verification means that the results from the simulator algorithm can be trusted. Given a randomly generated pattern for the plasmonic structure, the simulator can predict the SHG. The last step is optimization. Utilizing an Evolution Strategy such as in Chapter 3, through a process of selection, reproduction, and mutation, the simulator can be modified to output a pattern for a particular SHG value [2].

Figure 4.6 – Results from Evolution Strategy for optimal plasmonic geometry to yield maximal SHG response.

For instance, searching for a pattern that does not lead to a second-harmonic response would lead to a completely covered canvas indicating a gold mirror would be the plasmonic device. Figure 4.6 shows the results for a maximal SHG response for the plasmonic ABC nanolaminate device. All four patterns look similar but do have minor variations. The SHG response for this pattern is an order of magnitude above anything in the accuracy test in Figure 4.5.

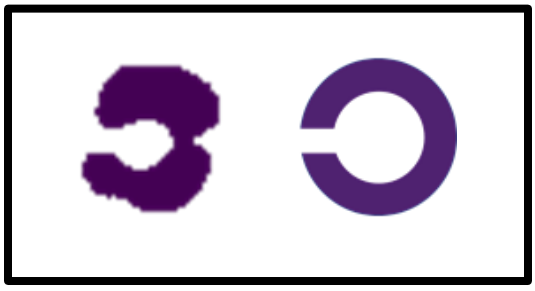

Figure 4.7 – Optimized structure and split ring resonator (SRR) with similar parameters. The optimized structure had a significantly higher second-harmonic nonlinear optical response as compared to the split ring resonator.

To show that the optimal result is indeed due to the unusual shape, and not the resemblance to a split ring resonator (SRR), a similarly shaped split ring resonator was simulated in a full-wave simulator to compare the results, as shown in Figure 4.7. The SRR had an SHG value two orders of magnitude less than the optimal shape.

4.4 References

[1] L. Raju *et al.*, "Maximized Frequency Doubling through the Inverse Design of Nonlinear Metamaterials," *ACS Nano*, vol. 16, no. 3, pp. 3926-3933, 2022/03/22 2022.

[2] K. O. Stanley, "Exploiting Regularity Without Development," in *AAAI Fall Symposium: Developmental Systems*, 2006.

[3] K. Ghaderi, F. Akhlghian, and P. Moradi, "A new digital image watermarking approach based on DWT-SVD and CPPN-NEAT," in *2012 2nd International eConference on Computer and Knowledge Engineering (ICCKE)*, 2012, pp. 12-17.

CHAPTER 5. DATA LEARNING FRAMEWORK ENHANCEMENT – TRANSFER LEARNING

When tackling a new problem, such as finding a metamaterial device with a specific optical response, each time the required parameters change, a new set of training patterns and corresponding optical response data will need to be created. Doing so, increases the data requirement to reuse the current data learning framework for a new optical response.

This is because, while new patterns are created, new full-wave simulation training data for these patterns will need to be run as well. In order to alleviate data requirements for the future, transfer learning can be implemented. Transfer learning uses the data sets already created, finds relationships between them, and requires a much smaller input of new training data to simulate a different set of parameters[1, 2].

This use of transfer learning would allow the easy and fast optimization of a much wider range of parameters for nonlinear optical responses. Once transfer learning has been implemented, we can quickly optimize new patterns for new parameters.

5.1 Transfer Learning Background

Transfer learning, in essence, is the ability to reuse an already trained neural network model for a new problem. This is assuming that the two models are trying to solve similar problems, otherwise the already ‘learned’ information would not lead to accurate results. The major selling point, of course, is the reduced amount of computation and data required to train the new model [1-8].

Transfer learning by itself of not an algorithm, but a method or technique that can be used to alleviate the resource concerns in network training. This is not only useful for new problems where the collection of training data is unwanted albeit possible, but for scenarios where the new similar problem does not have data that can be produced for training [6].

The technique dates back to 1976 when Bozinovski and Fulgosi explicitly addressed the method to train neural networks [9, 10]. Even so, it wasn't until the 1990's that the method took off, largely due to the work of Pratt and Thrun [8, 11, 12].

Transfer learning is currently used in numerous fields such as image processing, speech recognition, and natural language processing. The most daily use of transfer learning can be seen when speaking to Alexa or Siri and these systems being able to distinguish and understand multiple speakers and languages.

It can be a challenge balancing the performance of the new model with the usage of pre-trained results as there are positive relationships between parameters as well as non-useful information that must be distinguished.

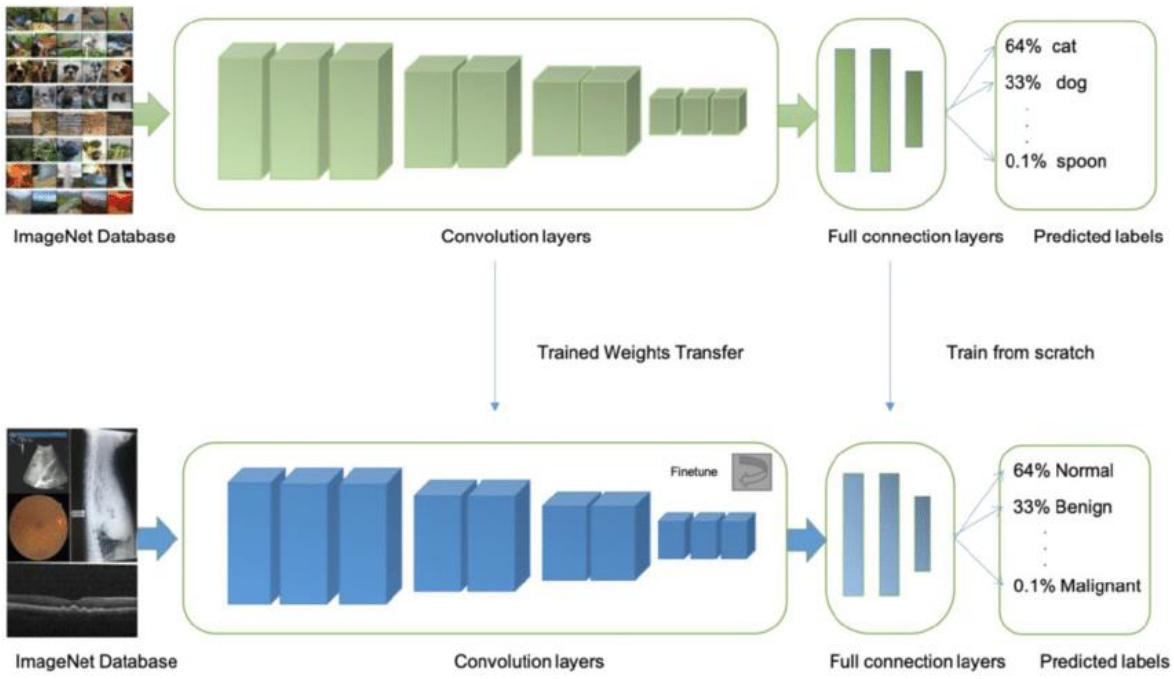

Figure 5.1 – Example to of the use of transfer learning to utilize a network trained to for image recognition a new problem of tumor classification. The convolution layers have trained weights transfer, along with fine-tuning, and the fully connected layers are trained with new data [2].

5.2 Transfer Learning Method

In transfer learning, there are two important processes, retraining and fine-tuning. Retraining refers to taking the pre-trained model that exists and contains generalized relationships between the input parameters and retraining the weights of the various layers for the new problem. Fine-tuning is where the model is trained further to increase the accuracy of the output. This is illustrated in an example in Figure 5.1 [2].

As described by Pan and Yang, transfer learning can be defined based on a domain and task [7]. The domain, $\mathcal{D} = \{\mathcal{X}, P(X)\}$, consists of the feature space, $\mathcal{X}$, and the

marginal probability distribution, $P(X)$, where $X = \{x_1, \dots, x_n\} \in \mathcal{X}$. What this means is that for a problem dealing with pattern classification, the feature space, $\mathcal{X}$, is the space containing the vectors x_i where the ith term refers to some pattern, and X is a specific training sample.

Given a domain, $\mathcal{D}$, we can then describe the task, $\mathcal{T} = \{\mathcal{Y}, P(y|x)\}$, as containing two parts, the label space or set of all labels, $\mathcal{Y}$, and the objective or target predictive function, $P(y|x)$, with a new instance of x. More specifically here, $x_i \in X$ and $y_i \in \mathcal{Y}$, creating training pairs of data $\{x_i, y_i\}$ [7].

Assuming we have only one source domain and one target domain, $\mathcal{D}_\mathcal{S}$ and $\mathcal{D}_\mathcal{T}$, respectively. Then we can say, $\mathcal{D}_\mathcal{S} = \{(x_{\mathcal{S}_1}, y_{\mathcal{S}_1}), \dots, (x_{\mathcal{S}_{n_\mathcal{S}}}, y_{\mathcal{S}_{n_\mathcal{S}}})\}$, where $x_{\mathcal{S}_i} \in \mathcal{X}_\mathcal{S}$ is the data instance, and $y_{\mathcal{S}_i} \in \mathcal{Y}_\mathcal{S}$ is the corresponding class label. Similarly, $\mathcal{D}_\mathcal{T} = \{(x_{\mathcal{T}_1}, y_{\mathcal{T}_1}), \dots, (x_{\mathcal{T}_{n_\mathcal{T}}}, y_{\mathcal{T}_{n_\mathcal{T}}})\}$, where $x_{\mathcal{T}_i} \in \mathcal{X}_\mathcal{T}$ is the input, and $y_{\mathcal{T}_i} \in \mathcal{Y}_\mathcal{T}$ is the corresponding output [7].

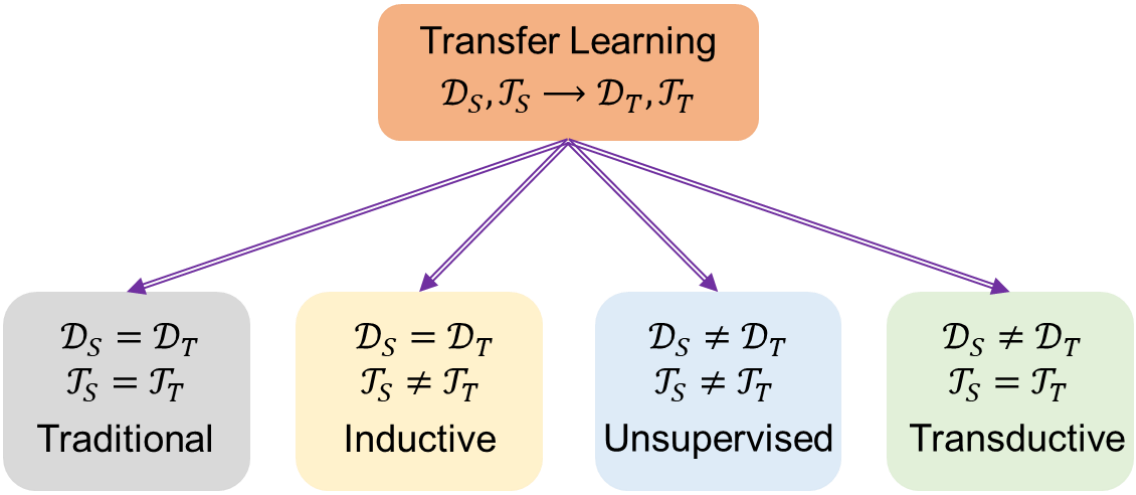

Figure 5.2 – Illustration of the types of transfer learning based on the definitions by Pan and Yang [7]. The classification is based on the source, S, and target, T, domain, $\mathcal{D}$, and task, $\mathcal{T}$. Based on the similarities of the domain and tasks of the source and target, the method of transfer learning can be classified into three main types, inductive, unsupervised, and transductive.

5.3 Types of Transfer Learning

The way that neural networks sometimes work, is that they focus on detecting the largest most general information and each subsequent layer focuses on more target-specific details or features [13-18]. Transfer learning mostly focuses on the early to mid-layers, as the deepest layers are saved for retraining. For example, if the original model was used to distinguish cars, and the goal of the new model is to distinguish trucks, the first layers would be focused on figuring out what a vehicle looks like, and the last layers would mostly

be focused on separating cars from other vehicles. Three main types of transfer learning techniques that exist, inductive, transductive, and unsupervised as shown in Figure 5.2.

Inductive transfer learning: For this type, the source and target tasks are different. Here $\mathcal{T}_S$ knowledge transfers in inductive learning an improved generalization for the $\mathcal{T}_\mathcal{T}$. The goal here is to improve the execution of the objective predictive function, $P(y|x)$.

Transductive transfer learning: In this type, the source and target tasks are the same, but the domains are distinct. Meaning that there is no labelled data in the target domain, but a lot of labelled data in the source domain exists.

Transductive transfer learning can be further categorized into two cases: 1) The feature spaces between the domains are different, $\mathcal{X}_S \neq \mathcal{X}_\mathcal{T}$; 2) Domain adaptation, where the feature spaces between the domains are the same, $\mathcal{X}_S = \mathcal{X}_\mathcal{T}$, but the marginal probability distributions of the input data are not, $P(X_S) \neq P(X_T)$.

Unsupervised transfer learning: Unsupervised transfer learning is similar to inductive transfer learning, as the target tasks of the source and domain are different. The distinction is that there is also no labelled data in the source and target domains. The goal here is to solve unsupervised learning tasks like clustering and dimensionality reduction in the target domain [7].

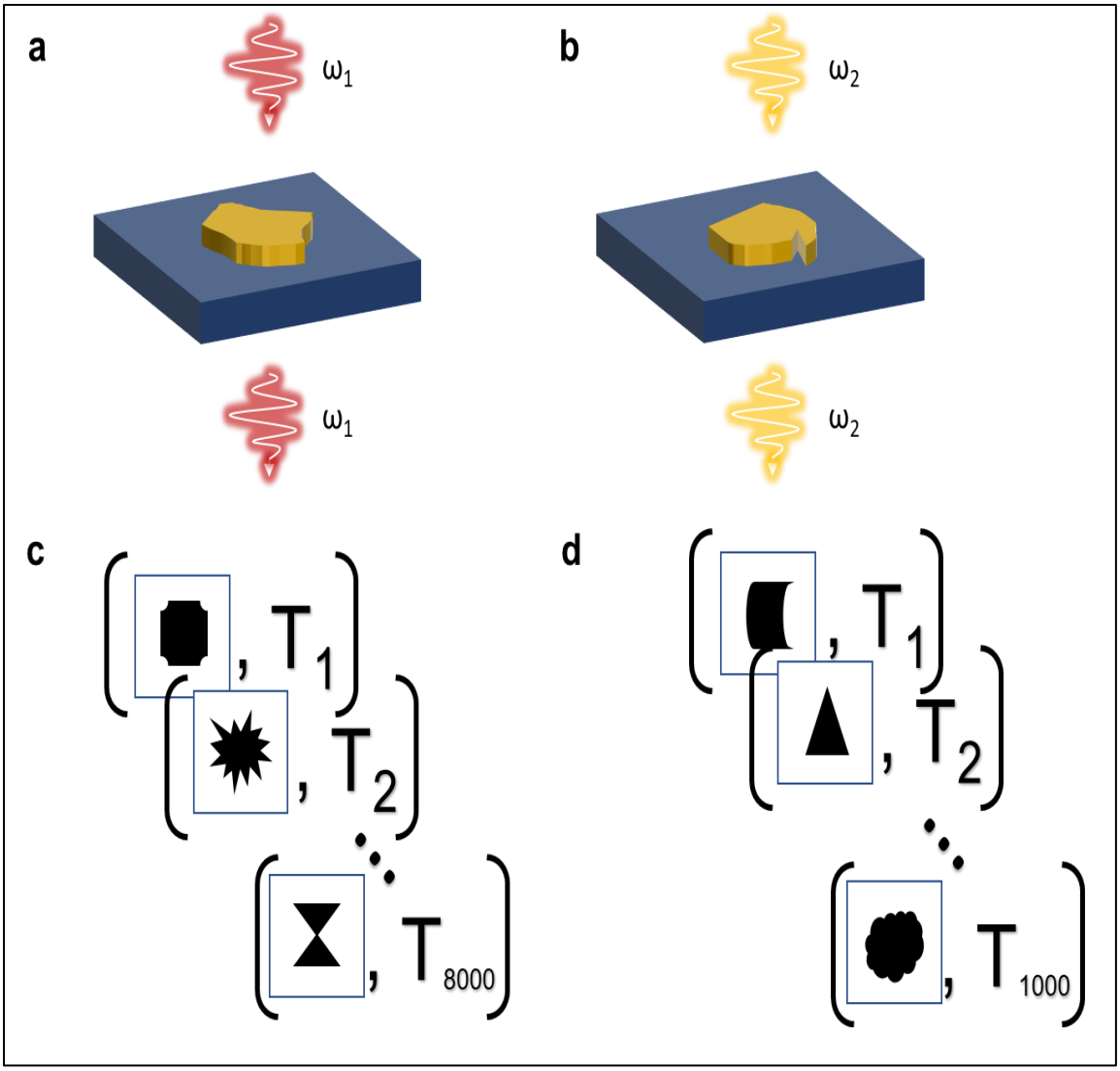

Figure 5.3 – Transfer learning applied to plasmonic metamaterial for linear optical response. a-b) Source design parameters for ω_1 and target design parameters for ω_2, respectively. c-d) Required training data pairs for source and target, the source (c) has a large dataset of 8000 pairs, the target (d) requires a much smaller training dataset of 1000 pairs.

5.4 Transfer Learning Applied to Linear Optical Problems

It is important to note that in transfer learning, the pattern generation network is not the network in question, but rather the simulation network. As transfer learning can benefit the training of networks for new problems, a sample case was implemented as a proof of concept. Figure 5.3 shows a generalized schematic for the work.

We begin with a simulation neural network that is trained for a specific problem – a periodic plasmonic structure with an incident normal light at 800 nm. The periodicity of the structure is $p = 380\ nm$, and the gold plasmonic thickness is $t_{Au} = 45\ nm$. Training data for this structure with various plasmonic geometries exist, and the network has been trained, so it can predict the linear optical transmission response for any pattern in the latent space of the pattern generation given an incident light of 800 nm.

If we wanted to do the same thing, simulate the linear optical response for a plasmonic pattern, given the same structure parameters, but a different wavelength, new training data would need to be collected. In this scenario, that would be 8,000-10,000

atterns and their corresponding linear optical responses. However, by implementing transfer learning, we can reduce that computation to a matter of a few hundred.

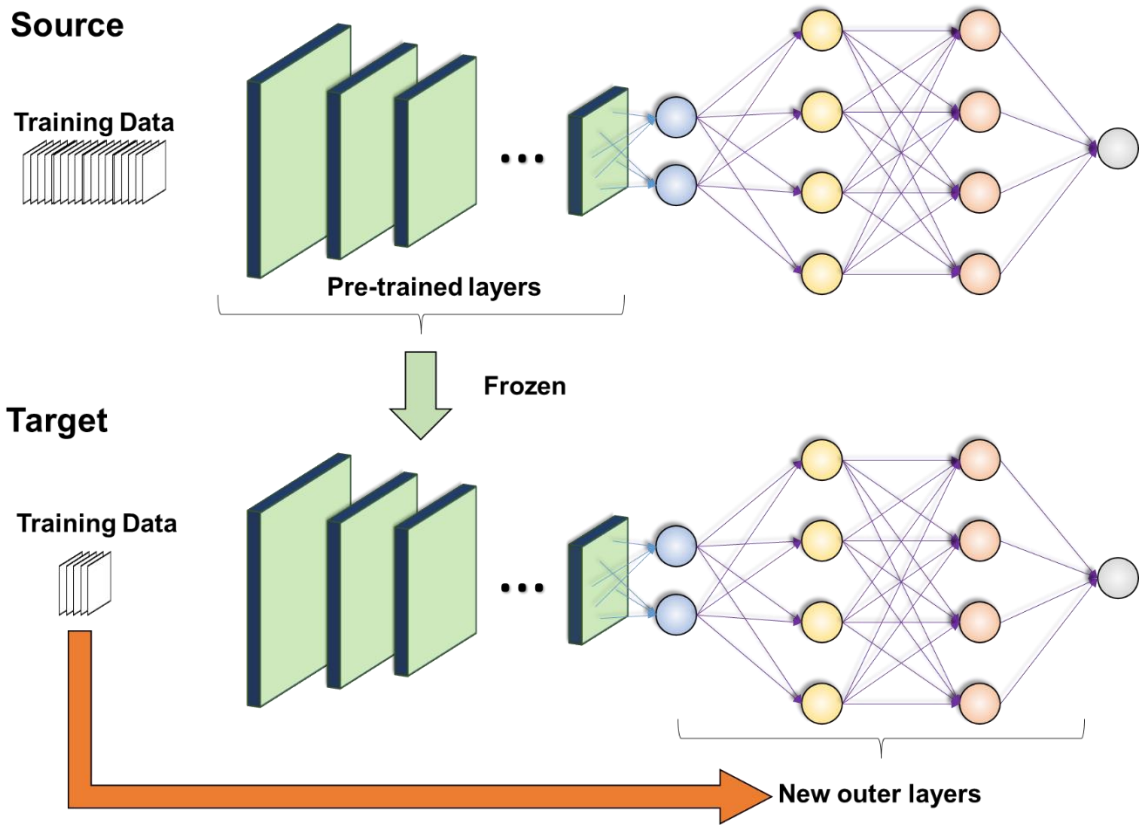

Figure 5.4 – Transfer learning process illustrated. Beginning with a source network that is already trained, the common inner layer is frozen and used for the target. Then new outer layers are created and trained with the new training data. An optional step of fine-tuning is possible by unfreezing the inner layers and training the whole network with the new data again.

The process is similar to what is illustrated in the example in Figure 5.1 and is shown in Figure 5.4. There are four basic steps to this method. First, a beginning model is needed.

In this case, the neural network model that is trained for incident 800 nm light. Second, layers that contain the generalized relationships, need to be 'frozen' [19-22]. Skipping this step would remove all the 'learned' data and would result in basically training the network from scratch. Freezing the data always the new network to take the important sharable information for itself.

On top of these frozen layers, we have the third step, which is creating new outer layers that will be trained with the new data. Finally, fine-tuning of the network is optional, but will make sure the results will be as accurate as possible. Fine-tuning involves training the new neural network, including unfreezing the frozen layers, with the target task's training data. This slightly adapts the pre-trained learning to the new scenario.

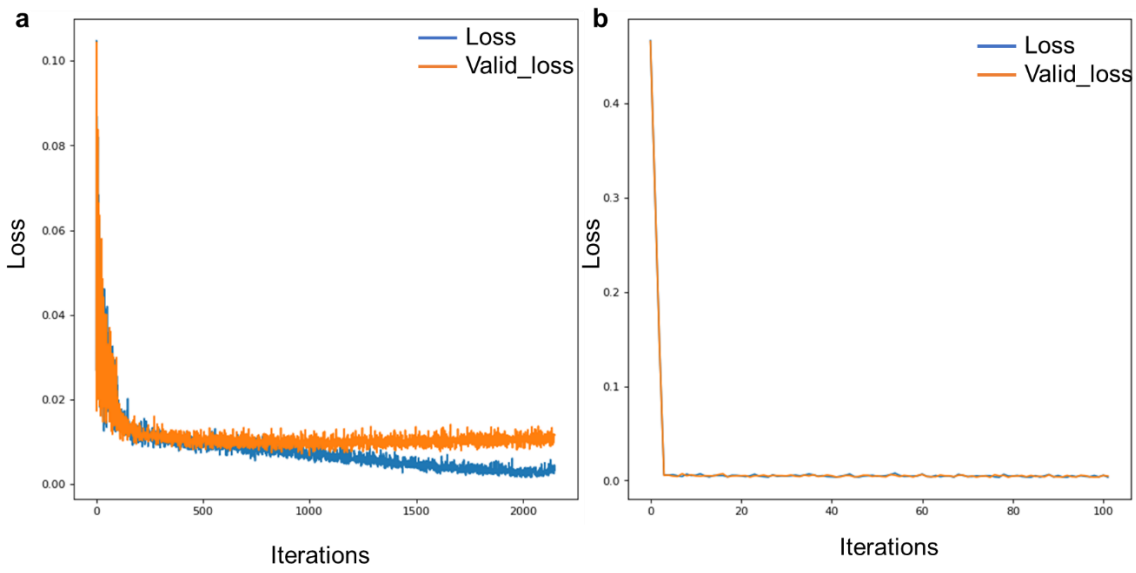

Figure 5.5 – Training and validation loss for the original neural network (source - a) and the new network (target - b).

The results of this research can be summarized by Figure 5.5. The loss while training and validating the neural network for the new parameter – wavelength – is acceptably low. This shows that the method of transfer learning can successfully be implemented to reduce computation resources.

5.5 References

[1] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1717-1724, 2014.

[2] J. Xu, K. Xue, and K. Zhang, "Current status and future trends of clinical diagnoses via image-based deep learning," *Theranostics*, vol. 9, pp. 7556-7565, 10/12 2019.

[3] Q. Yang and X. Wu, "10 challenging problems in data mining research," *International Journal of Information Technology & Decision Making*, vol. 5, no. 04, pp. 597-604, 2006.

[4] X. J. Zhu, "Semi-supervised learning literature survey," 2005.

[5] J. Ramon, K. Driessens, and T. Croonenborghs, "Transfer learning in reinforcement learning problems through partial policy recycling," in *European conference on machine learning*, 2007, pp. 699-707: Springer.

[6] R. Raina, A. Battle, H. Lee, B. Packer, and A. Ng, "Self-taught learning: transfer learning from unlabeled data," in *ICML '07*, 2007.

[7] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345-1359, 2010.

[8] S. Thrun and L. Pratt, *Learning to learn*. Springer Science & Business Media, 2012.

[9] S. Bozinovski and A. Fulgosi, "The influence of pattern similarity and transfer learning upon training of a base perceptron b2," in *Proceedings of Symposium Informatica*, 1976, vol. 3, pp. 121-126.

[10] S. Bozinovski, "Reminder of the first paper on transfer learning in neural networks, 1976," *Informatica*, vol. 44, no. 3, 2020.

[11] L. Pratt and S. Thrun, "Second special issue on inductive transfer-Guest editors' introduction," vol. 28, ed: KLUWER ACADEMIC PUBL SPUIBOULEVARD 50, PO BOX 17, 3300 AA DORDRECHT, NETHERLANDS, 1997, pp. 5-5.

[12] L. Y. Pratt, "Discriminability-Based Transfer between Neural Networks," pp. 204–211, 1992.

[13] B. Li, Q. Yang, and X. Xue, "Transfer learning for collaborative filtering via a rating-matrix generative model," in *Proceedings of the 26th annual international conference on machine learning*, 2009, pp. 617-624.

[14] Y.-Y. Tsai, P.-Y. Chen, and T.-Y. Ho, "Transfer learning without knowing: Reprogramming black-box machine learning models with scarce data and limited resources," in *International Conference on Machine Learning*, 2020, pp. 9614-9624: PMLR.

[15] S. Shao, S. McAleer, R. Yan, and P. Baldi, "Highly accurate machine fault diagnosis using deep transfer learning," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2446-2455, 2018.

[16] S. J. Pan, J. T. Kwok, and Q. Yang, "Transfer learning via dimensionality reduction," in *AAAI*, 2008, vol. 8, pp. 677-682.

[17] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big data*, vol. 3, no. 1, pp. 1-40, 2016.

[18] L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*: IGI global, 2010, pp. 242-264.

[19] J. Kunze, L. Kirsch, I. Kurenkov, A. Krug, J. Johannsmeier, and S. Stober, "Transfer learning for speech recognition on a budget," *arXiv preprint arXiv:1706.00290*, 2017.

[20] T. Q. Nguyen and D. Chiang, "Transfer learning across low-resource, related languages for neural machine translation," *arXiv preprint arXiv:1708.09803*, 2017.

[21] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A deep learning-based framework for automatic brain tumors classification using transfer learning," *Circuits, Systems, and Signal Processing*, vol. 39, no. 2, pp. 757-775, 2020.

[22] Y. Guo, H. Shi, A. Kumar, K. Grauman, T. Rosing, and R. Feris, "Spottune: transfer learning through adaptive fine-tuning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 4805-4814.

CHAPTER 6. CONCLUSION

The focus of the research presented in this thesis is the intersection between deep learning and plasmonic metamaterial devices. Specifically, creating a deep learning framework that has demonstrated the capability of optimizing plasmonic metamaterial structures for linear optical responses, and two nonlinear optical responses – second-harmonic generation, and sum-frequency generation[1, 2]. The secondary goal was to decrease the computational resources through a framework advancement by implementing transfer learning. All of this results in a framework that can be expanded to any nonlinear optical response, and multiple design parameters. The following sections will briefly highlight and conclude the work completed in the previous chapters.

6.1 Deep Learning Framework

As discussed earlier, traditional guess-and-check methods for solving plasmonic design problems with specific optical responses in mind is an intensive process that does not always lead to the best solution. Automating the process through the use of a deep learning framework not only makes the search more efficient, but also opens the door to new geometries not conventionally available. This is demonstrated in Chapter 3, first for linear optical responses, and then for nonlinear second-harmonic optical responses.

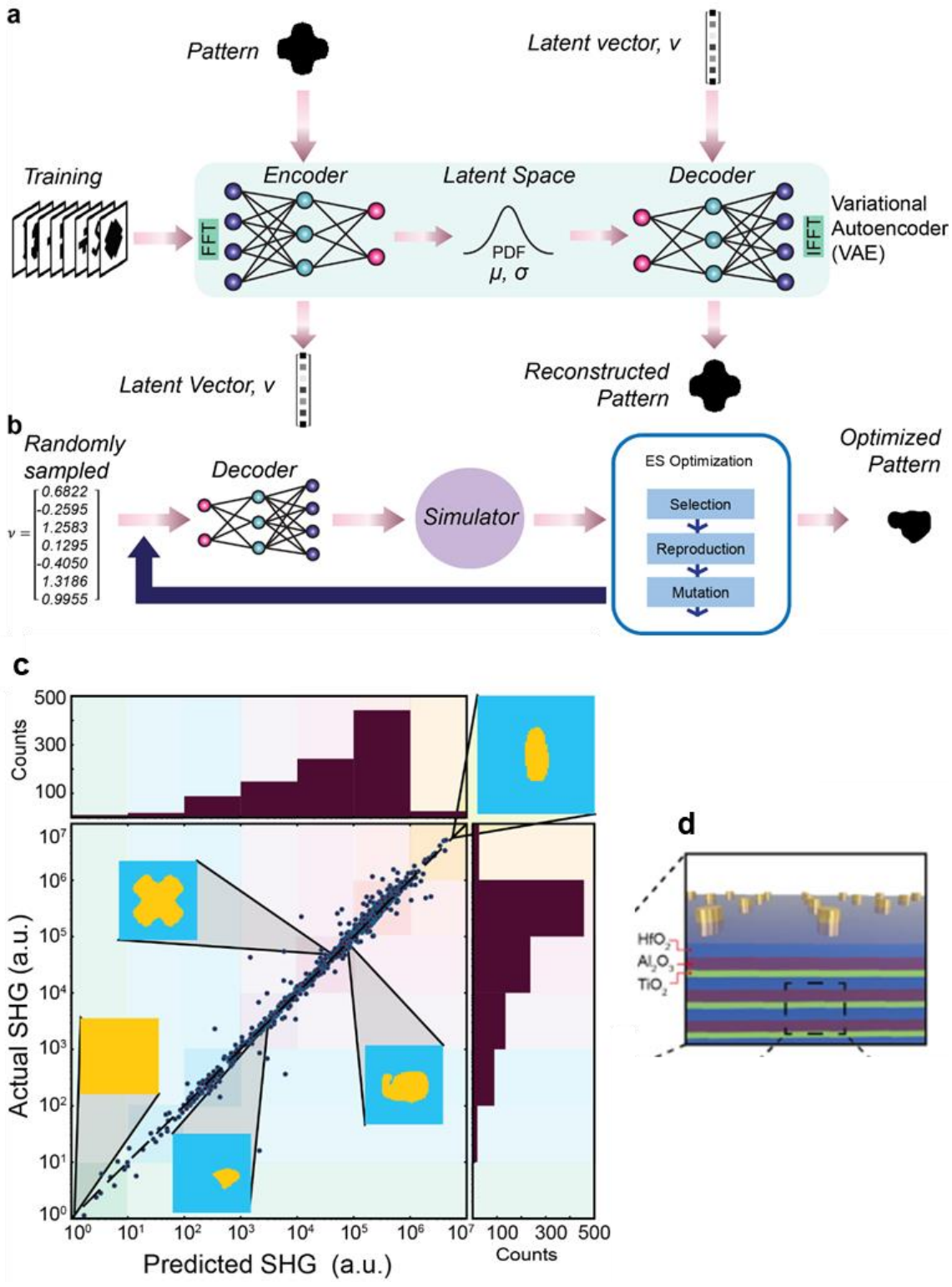

Figure 6.1 – Schematic of deep learning framework, accuracy results for SHG optical responses, and plasmonic nonlinear metamaterial device. a-b) Deep learning

framework including the pattern generation system in a and the simulation process in b. c) Accuracy of the simulator and spread of data represented shows that the deep learning framework results can be used to optimize the desired structure for SHG responses d) Optimized plasmonic nonlinear metamaterial structure schematic showing the ABC nanolaminate beneath the optimized plasmonic pattern [1].

In this deep learning framework described in Figure 6.1, there are two major parts – the pattern generator and the simulation process. Figure 6.1a explains how a Variational Autoencoder method is used to encode the possible patterns into a latent space. The latent space can then be sampled for latent vectors that are then decoded into 64×64 pixel canvas images. These images represent a plasmonic metamaterial device, where each canvas is a single unit of a periodic structure. Once we can generate possible geometries, or patterns, at will, it is then necessary to approximate the corresponding optical response for the device containing the sampled pattern.

Figure 6.1b illustrates the method by which such approximation takes place. Through training using data collected from a full-wave simulation, we are able to predict the optical response for each device in the latent space of the VAE. This is further expanded into an optimization process to find the best pattern for the corresponding optical response desired. Chapter 3 discusses the results for a linear optical response initially, then later discusses the results when second-harmonic generation is the optical response targeted.

The deep learning framework's accuracy for SHG responses is displayed in Figure 6.1c. The plot represents patterns whose SHG response was both predicted through the deep learning framework, and actually calculated using a full-wave simulation. The 1-1

correlation indicates good accuracy of the results. The pattern in the top right corner is the optimized structure for the desired optical response.

Shown in Figure 6.1d is a schematic of the optimized structure atop an ABC nanolaminate. The ABC nanolaminate is the nonlinear metamaterial utilized in this research [3]. It is unique, in that with normal incident light, there will be no second-harmonic generation. Light at a non-zero incident angle is necessary to induce a second-harmonic response in the material, as discussed in Chapters 1 and 3. The ABC nanolaminate is comprised of three repeating oxide layers as shown in Figure 6.1d.

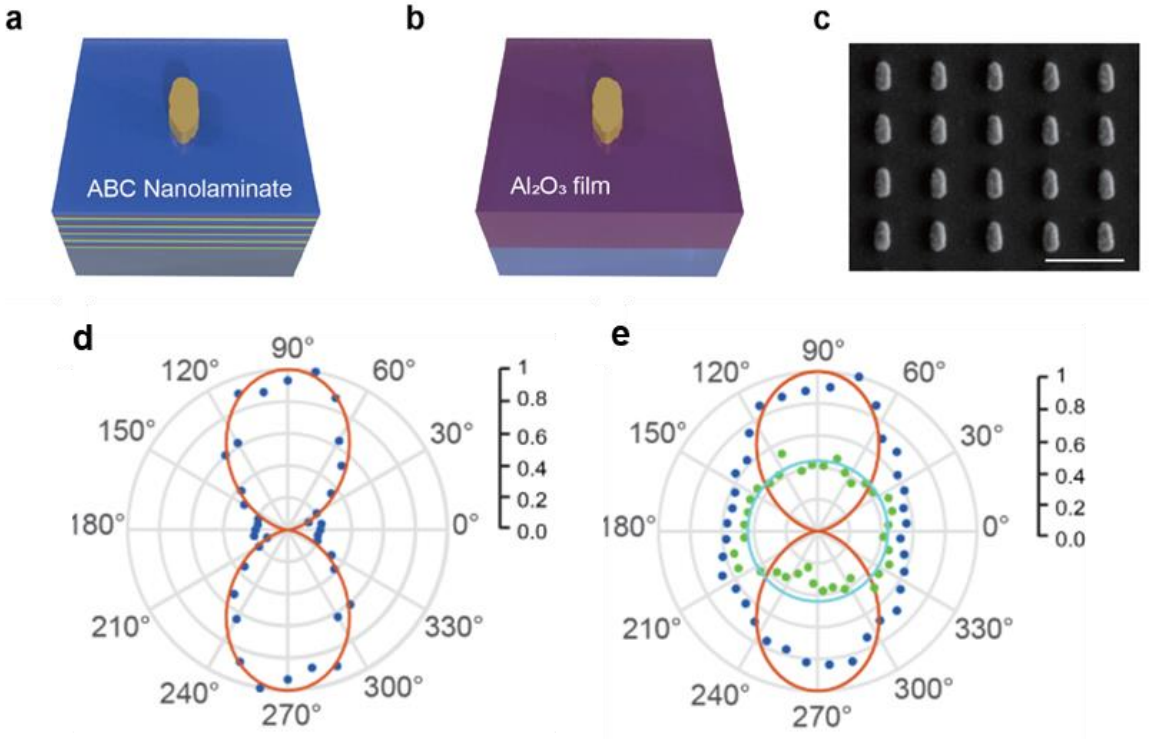

Figure 6.2 – Experimental results for the SHG optimized device. a-b) Schematics of the unit structures of the ABC and control device, respectively. c) SEM image of the

ABC device. d-e) Output polarization results for TM polarized incident light for the control and ABC device, respectively. The blue dots represent the output polarization experimentally measured. The red two-lobe shape represents second-harmonic response due to the plasmonic. The green dots represent the SHG response induced from the ABC nanolaminate, and the cyan circle is the best fit of the SHG due to the ABC nanolaminate.

For the optimized plasmonic nonlinear metamaterial designed for SHG, experimental measurements were performed in order to validate the deep learning framework. Figure 6.2 summarizes the results from Chapter 3. Essentially, not only was the deep learning framework validated, but it was also experimentally extracted that the ABC nanolaminate had a significant induced SHG response due to the plasmonic device above it.

The results of Chapter 3 indicate that the deep learning framework is a viable method to optimize structures for linear and nonlinear optical responses. In addition, the usage of the ABC nanolaminate allows for thin-film nonlinear metamaterial devices to be created. The resulting deep learning framework could be further investigated for other nonlinear optical responses, and that is what the research in Chapter 4 discusses.

6.2 Sum-frequency Generation Optimization

Chapter 4 builds upon the research conducted in Chapter 3 and expands the work for nonlinear sum-frequency generation. As explained in Chapter 2, second-harmonic generation is a special case of sum-frequency generation, so this research generalizes the work done in Chapter 3. Although the deep learning framework is similar, adjustments

were made to improve the framework and resulting patterns, the major adjustment being the replacement of the VAE with a Compositional-pattern producing Network.

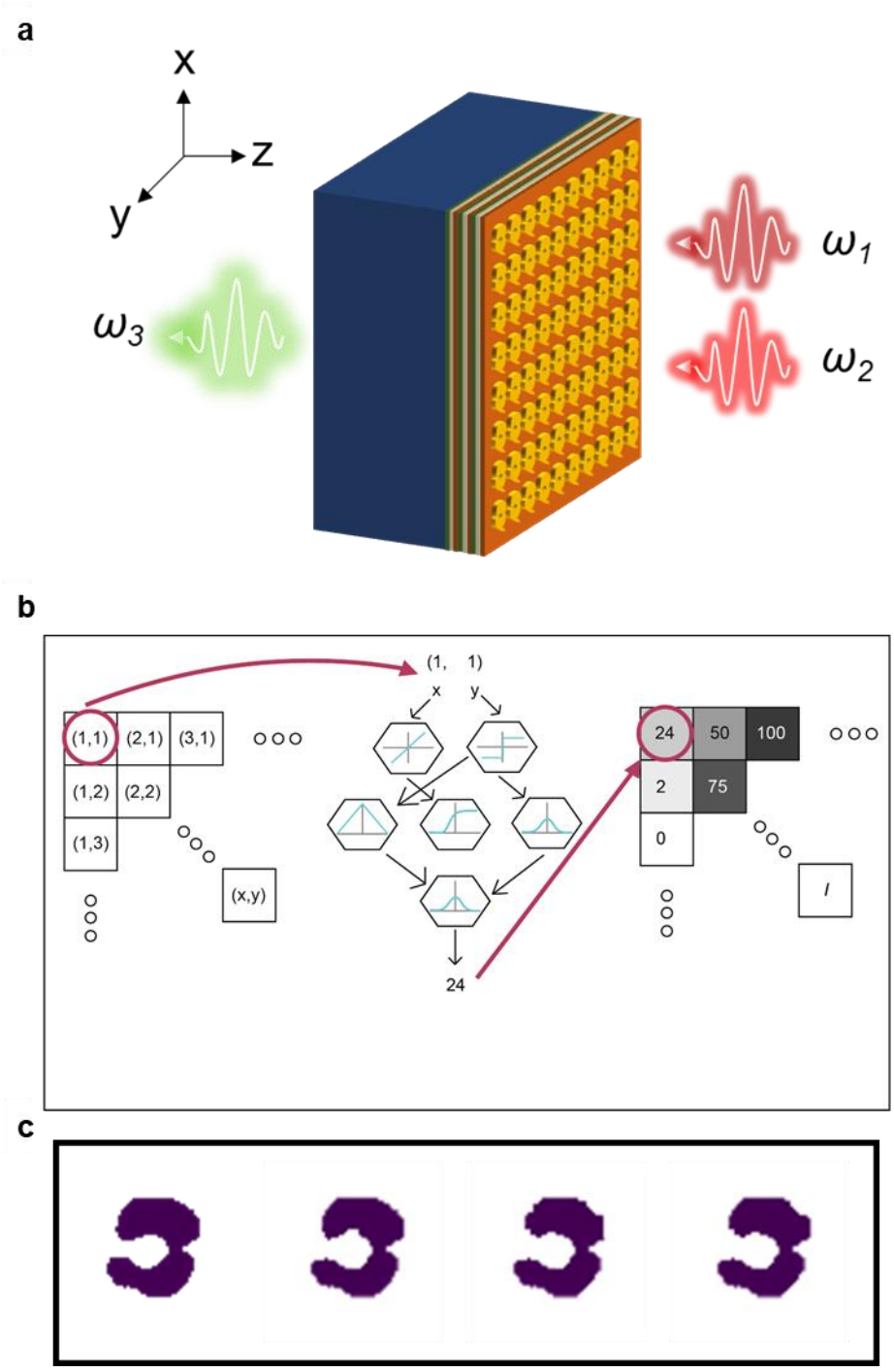

Figure 6.3 – Summary of Chapter 4 results. a) Schematic of nonlinear plasmonic metamaterial device for sum-frequency generation. b) CPPN methodology. c) Results of the evolution strategy optimization for the desired SFG response.

The metamaterial parameters set for the SFG optimization involved a plasmonic structure atop an ABC nanolaminate, similar to Chapter 3, as shown in Figure 6.3a. The biggest difference in terms of the physics between Chapter 3 and Chapter 4 is the involvement of two different incident wavelengths. As with sum-frequency generation, the output frequency is the sum of the input frequencies. This required a new full-wave simulation in order to create accurate training data.

To widen the possible geometries, a CPPN was implemented. The way that a CPPN works is illustrated in Figure 6.3b. As described in Chapter 4, a CPPN gives rise to an infinite number of possible patterns. The pattern generator in Chapter 3 was limited to single unit structures, here with the CPPN, multi-unit structures are possible. Various samples of patterns able to be produced by the CPPN are also shown in Chapter 4.

Finally, utilizing a similar simulation process as Chapter 3, an optimized structure through the use of an evolution strategy was found. The finalists, as seen in Figure 6.3c, represent a new geometry not conventionally used. Chapter 4 also discusses the possibility that a split ring resonator, the traditional geometry most closely related to the optimized structure, being of similar SFG response. The full-wave simulation comparison resulted in a difference of two orders of magnitude between the two, with the optimized pattern being superior.

The results of the research conducted in Chapter 4, reveals that the conclusions of Chapter 3 were indeed correct. That research could be furthered for other nonlinear optical responses. As other second-order nonlinear responses, physics-wise, are similar to sum-frequency generation, it seems that the deep learning framework can be used as-is simply

with new training, for those responses. The framework can also be expanded for other order nonlinear responses such as third order responses. As long as a full-wave simulation can be created for the nonlinear optical response, this framework can be used to optimize a corresponding plasmonic nonlinear metamaterial device.

6.3 Transfer Learning Framework Advancement

When discussing the expanded use of the deep learning framework, the issue of training and computation resources comes up. If the only difference between the original parameters and the new parameters, for example, is a slightly smaller periodicity, why is it necessary to create new training data pairs and retrain the neural networks with those? This problem can be solved with transfer learning, as described in Chapter 5.

The research completed in Chapter 5 posited the question of using a source network for a plasmonic metamaterial device that was trained for one wavelength and using a transfer learning method to train a target network using less computational resources.

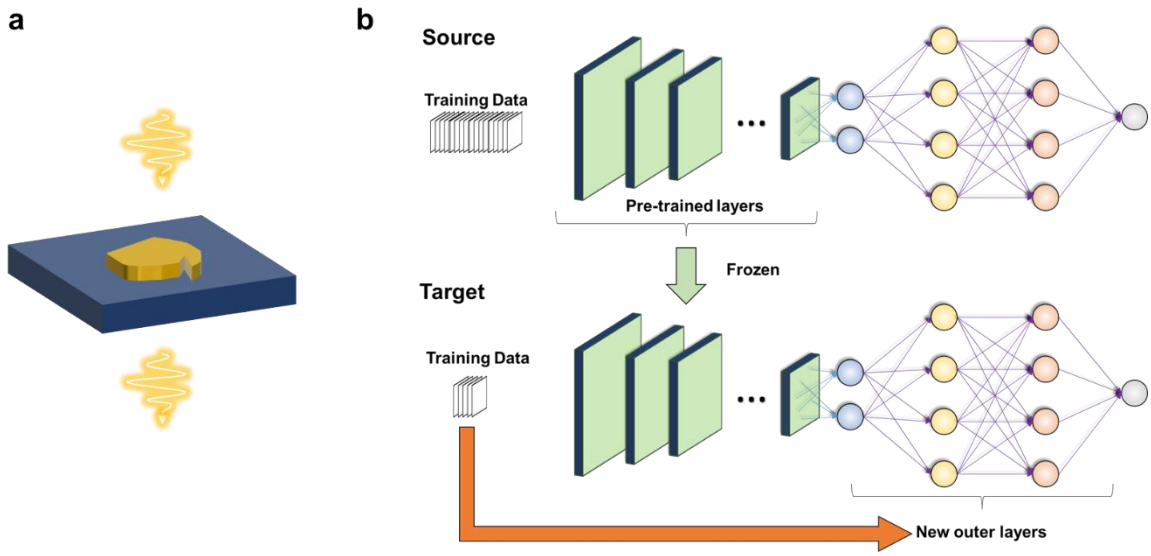

Figure 6.4 – Schematic of the plasmonic metamaterials for linear optical transmission response and transfer learning method from Chapter 5.

Figure 6.4b illustrates the method that was effectively used to employ transfer learning to the linear system, a schematic of which is illustrated in Figure 6.4a.

As discussed in Chapter 5, the use of transfer learning was successfully implemented and lead to a network whose loss was comparable to the source network. The loss calculated during training is also shown in Chapter 5. This indicates that the methodology used here can be extended for other similar uses.

Although the implementation for transfer learning in Chapter 5 is for a linear optical response, the method can be used for any optical response. This method can be used to decrease the computation resources for a new desired parameter, and this technique allows for a broader application of the deep learning framework for optical responses, both linear and nonlinear.

6.4 Outlook

This dissertation expands the scope of possible devices for optical phenomena. With the plasmonic ABC nanolaminate combined with the deep learning framework, applications include, but are not limited to, integrated optics, CMOS technology, low-cost lowlight/night vision for autonomous vehicles and defense application.

Not only does this enable the optimization of devices, but also generates patterns that one would not think of. The possible efficiency improvement of devices is something that could transform current optical devices.

With the future advancements in machine learning, the concepts and results of this research could develop into even more efficient frameworks for designing advanced devices for virtual reality, bio-photonics, integrated semiconductor devices and more. The potential for innovation is incredible, and possible as shown in this thesis.

6.5 References

[1] L. Raju *et al.*, "Maximized Frequency Doubling through the Inverse Design of Nonlinear Metamaterials," *ACS Nano,* vol. 16, no. 3, pp. 3926-3933, 2022/03/22 2022.

[2] Z. Liu, L. Raju, D. Zhu, and W. Cai, "A Hybrid Strategy for the Discovery and Design of Photonic Structures," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems,* vol. 10, no. 1, pp. 126-135, 2020.

[3] L. Alloatti *et al.*, "Second-order nonlinear optical metamaterials: ABC-type nanolaminates," *Applied Physics Letters,* vol. 107, no. 12, p. 121903, 2015.