Design and Optimisation of Optical Metasurfaces Using Deep Learning

GHASEMI, AMIR

Design and Optimisation of Optical Metasurfaces Using Deep Learning

AMIR GHASEMI

A Thesis presented for the degree of
Doctor of Philosophy

Department of Engineering
Durham University

March 2024

Dedicated to

My lovely wife, Narges, and my family without whom this work would likely not exist.

Abstract

This thesis centres on the design, processing, and fabrication of tunable optical metamaterials. It incorporates physics-based simulation, deep learning (DL), and thin film fabrication techniques to offer a comprehensive exploration of the field of optical metamaterials. Placing stiff resonators on a flexible substrate is a common type of mechanically tunable metasurface, whose optical responses are tuned by dynamically adjusting the spacing between resonators by applying mechanical force. However, the significant modulus mismatch between materials causes stress concentration at the interface, leading to crack propagation and delamination at lower strain levels (20-50%), and limiting the optical tunability of the structure. To address this challenge, we propose two designs to manipulate stress distribution. Under mechanical force, the structure enables localised deformation, redirecting stress from critical areas. This mechanism minimises the accumulation of stress in the interface, thereby diminishing the risk of material failure and improving stretchability up to 120% compared to traditional designs. This extreme stretchability leads to a 143 nm resonance shift, which is almost twice as large as that of conventional geometry. A universal machine learning (ML)-based approach was developed to optimise the metasurface design across three key aspects: geometric parameters, material development, and free-form shape configuration. In design parameters optimisation, a fully connected neural network (FCNN) was developed with a mean absolute error (MAE) of 0.0051, recommending a single geometry with a 10^4 order of magnitude decrease in computational time when compared to finite element method (FEM) simulations used for data generation. The suggested structure provides extensive coverage of the colour space, encompassing 27.65% of the standard RGB (sRGB) space. For the materials development part, an inverse design (ID) network was combined with effective medium approximation (EMA), navigating infinite materials composition space to identify new compositions for custom applications. The last network was tasked to explore boundless free-form shape space to propose the one for the on-demand optical properties with MAE of 0.21. The accuracy of all networks was experimentally validated.

Declaration

I declare that this thesis and the work presented in it are my own and have been generated by me as the result of my original research carried out at the Advanced Materials and Electronic Devices Research Node, the Department of Engineering of Durham University, Durham, England. I confirm that no part of this thesis has been submitted elsewhere for any other degree or qualification, and it is all my work unless referenced to the contrary in the text. Additionally, I acknowledge the use of OpenAI's language model to identify improvements in the writing style and check for grammatical errors. All experimental work is done by the author, except for the fabrication of samples and the microscopy analysis for the validation of the geometry optimisation network (as detailed in Section 6.6.), which were provided by my supervisor. The X-ray measurements, and microscopic analysis presented in Chapter 6 were performed at the GJ Russell Microscopy Facility at Durham University. I affirm that I was present for all stages, and the characterisation was completed solely by me.

Acknowledgements

I would like to thank my Supervisors Dr. Mehdi Keshavarz Hedayati and Prof. Dagou Zeze for their help and support. They have been outstanding mentors, making time and guiding the PhD.

I would like to acknowledge the invaluable support and insightful comments provided by Prof. Andrew Gallant and Dr. Stefan Szyniszewski throughout this research endeavour. I would like to express my gratitude to the staff and students I have met and worked with during my time at the Department of Engineering. In particular, thank you Rui Fang for her support and encouragement as I was setting out on my journey. Additionally, the Department of Engineering at Durham University has been a home away from home and provided a little microcosm which I could escape to. This PhD project was funded by the Durham University Doctoral Scholarship (DDS) program.

To my friends, your support has been invaluable. Our coffees, discussions, games and more have helped ground me, particularly during the odd times that were the pandemic.

Finally, I wish to thank my wife who means the world to me, and my parents and family whose encouragement has gotten me to both the start and the finish line of this PhD.

Contents

List of Figures

List of Tables

List of Publications

[1] **Ghasemi, A.**, O. Gheibi, Bowen, S. Leon, Giani, D. A. Zeze, and M. Keshavarz Hedayati, "AI-driven materials discovery: Custom materials development using neural network," *Submitted for Nature Materials*, (March 2024).

[2] **Ghasemi, A.**, O. Gheibi, D. A. Zeze, and M. Keshavarz Hedayati, "Inverse design of freeform nanophotonic structure by deep learning," *Paper is Ready to Submit*, 2024.

[3] R. Fang, **Ghasemi, A.**, D. A. Zeze, and M. Keshavarz Hedayati, "Machine learning-assisted inverse design of lateral hybrid metasurfaces," *Paper is Ready to Submit*, 2024.

[4] **Ghasemi, A.**, R. Fang, D. A. Zeze, and M. K. Hedayati, "Ultra-stretchable active metasurfaces for high-performance structural color," *AIP Advances*, vol. 13, no. 8, 2023.

[5] R. Fang, **Ghasemi, A.**, D. A. Zeze, and M. Keshavarz Hedayati, "Mechanically tunable metasurface with large gamut of color: Lateral hybrid system," *Journal of Applied Physics*, vol. 132, no. 13, 2022.

[6] **Ghasemi, A.**, R. Fang, D. A. Zeze, and M. Keshavarz Hedayati, "AI-based optical materials discovery," *SPIE Photonics Europe 2024 conference*, 2024 (Oral).

[7] R. Fang, **Ghasemi, A.**, D. A. Zeze, and M. Keshavarz Hedayati, "Inverse design of lateral hybrid metasurfaces with machine learning," *SPIE Photonics Europe 2024 conference*, 2024 (Oral).

[8] **Ghasemi, A.**, D. A. Zeze, and M. K. Hedayati, "AI-assisted design of ultra-stretchable metamaterial," *Photonics Electromagnetics Research Symposium (PIERS)*, 2023 (Oral).

[9] **Ghasemi, A.**, D. A. Zeze, and M. Keshavarz Hedayati, "100% deformable and reversible active metasurface for visible frequency," *16th International Congress on Artificial Materials for Novel Wave Phenomena – Metamaterials*, 2022.

Acronyms

2D	two-dimensional.
3D	three-dimensional.
AFM	atomic force microscopy.
AI	artificial intelligence.
ALD	atomic layer deposition.
ANN	artificial neural network.
BARC	bottom anti-reflective coating.
CAD	computer-aided design.
cGAN	conditional generative adversarial network.
CIE	Commission on Illumination.
CNN	convolutional neural network.
DC	direct current.
DL	deep learning.
DNN	deep neural network.
DOF	degree of freedom.
EBL	electron beam lithography.
ED	electric dipole.
EM	electromagnetic.
EMA	effective medium approximation.
FCNN	fully connected neural network.

FDTD finite-difference time-domain.
FEM finite element method.
FIB focused ion beam.
FNN forward neural network.
FOM* figure of merit.
FWHM full width at half maximum.

GA genetic algorithm.
GANs generative adversarial networks.
GST Germanium Antimony Telluride.

ID inverse design.
IR infrared radiation.
ITO indium tin oxide.

LMs liquid metals.
LSP localised surface plasmon.
LSPR localised surface plasmon resonance.

MAE mean absolute error.
MD magnetic dipole.
MEMS microelectromechanical systems.
MG Maxwell–Garnett.
ML machine learning.
MLP multilayer perceptron.
MSE mean squared error.

NIL nanoimprint lithography.
NN neural network.

PBCs periodic boundary conditions.
PBDL physics-based deep learning.
PCMs phase change materials.
PCs photonic crystals.
PDMS Polydimethylsiloxane.

PML perfectly matched layer.

ReLU rectified linear unit.

RF radio frequenc.

RI refractive index.

RL reinforcement learning.

RMS root mean square.

RMSE root mean square error.

RMSprop root mean squared propagation.

RNN recurrent neural network.

SCC Spearman's correlation coefficient.

SEM scanning electron microscopy.

SGD stochastic gradient descent.

SPP surface plasmon polariton.

SPR surface plasmon resonance.

sRGB standard RGB.

SRR split-ring resonator.

TEM transmission electron microscopy.

THz terahertz.

TMM transfer matrix method.

VAEs variational autoencoders.

XRR X-ray reflectometry.

Chapter 1

Introduction

1.1 Chapter Overview

This chapter provides a short introduction to metasurfaces by highlighting their recent applications. It also introduces the concept of "active" metasurfaces capable of altering their function post-production. Furthermore, it discusses the primary challenges associated with the limited stretchability of mechanically active metasurfaces and outlines the thesis's plan to address them. The chapter also outlines the objectives of the thesis and offers a brief overview of its structure.

1.2 Motivation and Objectives

Recent advances in nanophotonics have sparked a lot of interest in our ability to control how light interacts with matter on a very small scale, even smaller than the light's wavelength. Making nanophotonic elements smaller has given us a new category of compact and flat optical components known as "metasurfaces". These

are precisely designed, artificially created surfaces that act as alternatives to the old, bulky optical elements [1]. Metasurfaces, made of tiny blocks (called "meta-atoms"), can manipulate electromagnetic waves in terms of their colour, timing, or location with incredibly precise control, all at a scale much smaller than what was previously possible [1]. This level of control has enabled and accelerated critical developments in fields such as flat optics [2, 3, 4], quantum communications [5], and holography [6].

An active metasurface refers to a type of metasurface engineered to exhibit dynamic and controllable interaction with the incident electromagnetic wave [7]. Unlike passive counterparts, which have fixed properties determined during fabrication, active metasurfaces can change their behaviour or characteristics in response to external stimuli like magnetic field [8], thermal fluctuations [9], optical excitation [10], electrochemical tuning [11], mechanical force [12, 13], electrical actuation [14], and chemical reactions [15].

Active metasurfaces typically involve the ordered or random arrangement of stiff components, such as resonators or meta-atoms, on a flexible substrate. However, a significant challenge arises from the limited flexibility of this structure, typically falling within the range of 10-50% [16]. This narrow range is significantly lower than the maximum strain that the elastomer can endure, often exceeding 200%. Consequently, this limitation constrains the tunability range of active metasurfaces, particularly those dependent on mechanical forces for operation [17]. The primary objective of this thesis is to investigate the reasons for the limitation in the flexibility of mechanically active metasurfaces. Subsequently, the thesis aims to propose geometries, Kirigami-inspired and pillar-based, that can effectively enhance the flexibility of such metasurfaces. Although various forms of Kirigami-inspired patterns have been considered in various similar research fields like flexible electronics [18, 19], the pillar-based one is quite novel in the development of active metasurfaces.

Another equally important challenge of optical metasurfaces is the optimisation process. Conventional optimisation methods rely on intuition, trial and error, numerical simulations to solve Maxwell's equations, and extensive experimenta-

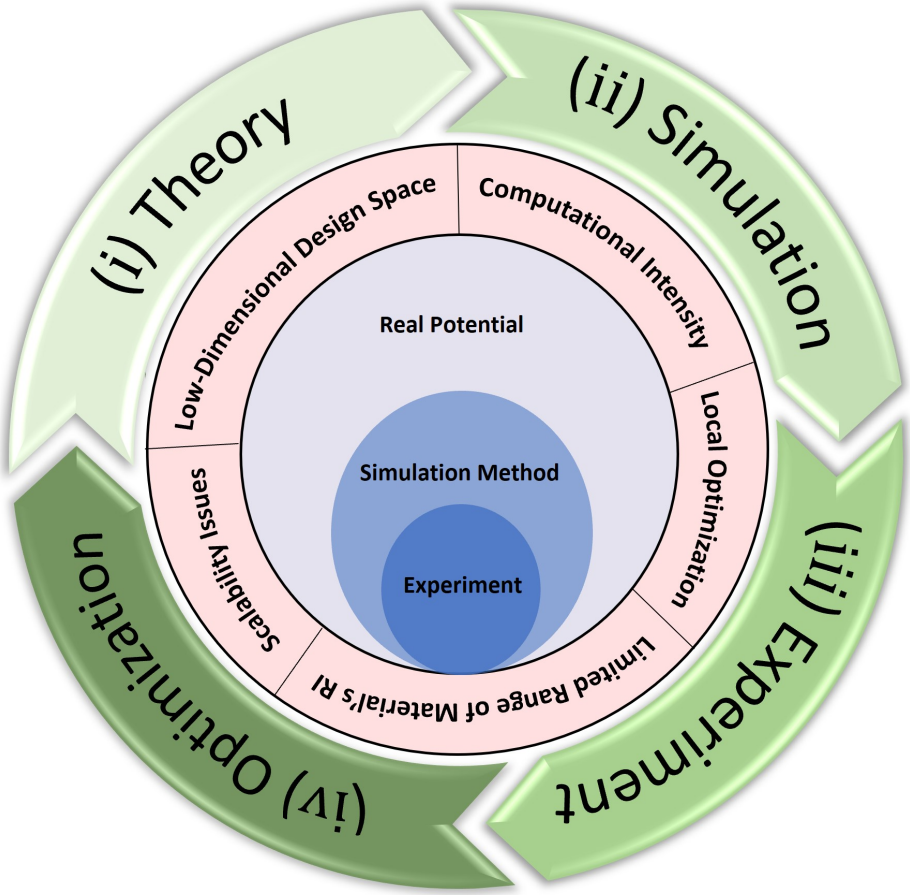


Figure 1.1: Schematic process of conventional materials design methods using iterative simulation and mass experiment optimisation. These methods are unable to find the design parameters and even the right pure elements or compounds fitting for a custom application directly.

tion [20]. There are two main challenges in the existing iterative simulation and mass experiment optimisation process which have left a substantial design space mostly unexplored (cf. Figure 1.1). Firstly, the current methods typically begin with theoretical concepts and computationally intensive simulations.

This initial phase is then followed by extensive experiments and model optimisation techniques. However, this process follows a "one-way" route (forward design) that often struggles to identify key design parameters for a custom photonic structure with predetermined optical responses.

Another challenge is the absence of a systematic approach to explore all possible material compositions and geometry space. The vast array of material combinations and ratios leads to an incredibly expansive compositional space, which

grows exponentially with the introduction of new materials. For example, recent studies on multi-component inorganic materials composed of more than one chemical element, have shown that even with constraints like charge neutrality, the design space for four-component materials exceeds 10^{10} combinations, while the five-component space surpasses 10^{13} [21]. Similarly, exploring the infinite geometry space to find the desired free-form geometry with a non-restricted shape in real space, allows for more diverse forms beyond traditional straight lines and polygons [22]. Navigating this extensive space is further complicated by the diverse processing parameters, where different synthesis methods produce varying material states, microstructures, and properties. As a result, it becomes impractical to systematically explore the practically infinite compositions solely through conventional numerical methods and extensive experimentation [23]. This thesis aims to address this limitation by developing a series of deep learning (DL)-based universal algorithms that enable the inverse design (ID) process. This method facilitates the design process by determining critical design parameters for custom structures with an order of 10^4 faster than the conventional methods. In addition, the optimised algorithm investigates the infinite space of accessible materials and geometries to find the ones which meet the requested optical properties.

1.3 Research Hypothesis

The hypothesis of this thesis is around the integration of rigid components into soft materials while keeping the overall materials' properties flexible via stress engineering and materials development. Two geometries including a high-relief pillar-based geometry and a kirigami-inspired geometry are considered to implement the hypothesis. Machine learning (ML)-based algorithm as a data-driven approach is used to optimise the developed structure in three main domains including geometric design parameters, material development, and freeform shape configurations.

1.4 Aims of the Thesis

The primary objective of this thesis is to comprehensively address the design, fabrication, and optimisation of a mechanically reconfigurable metasurface. This aim encompasses a series of specific objectives:

1. Geometrical Parameters and Stress Redistribution: This research begins by modelling various potential geometries to redistribute stress concentration in mechanically tunable metasurfaces effectively. A thorough examination of critical geometrical parameters is conducted to understand their impact.

2. Optimised Geometries and Optical Performance: The project proceeds by selecting the optimised parameters. These selected geometries are then studied to assess their optical performance. This evaluation involves a combination of mechanical and optical simulations to create metasurfaces that are not only mechanically tunable but also possess outstanding optical tunability.

3. Forward Machine Learning for Data Analysis: To enhance the project's data analysis capabilities, a forward ML algorithm is implemented. This advanced algorithm is used to analyse simulation data and generate additional data where needed, contributing to a more comprehensive understanding of the metasurface's behaviour.

4. Inverse Design Network for Structural Colours: The fourth objective focuses on training an ID network, particularly for a structural colour system. This demonstrates the advantages of optimised geometry by providing a single kirigami-inspired design that can produce a wide range of colours.

5. Inverse Design and Material Discovery Networks: To address the limitation of conventional methods in the design of nanophotonic structures, two sequential networks are trained. The first is an ID network, which specifies the thickness of each layer as well as the refractive indices of layers in the multilayer structure. The second one, the material discovery network, proposes amorphous metamaterials. This extends the possibilities for applications that require specific optical properties that cannot be achieved with pure materials.

6. Free-form Geometry Discovery Networks: An infinite design space is navigated by a developed algorithm to discover a free-form geometry that provides

custom optical behaviours.

7. Sample Fabrication and Network Validation: A series of experimental works are carried out to validate the performance and accuracy of the neural network (NN)s and their ability to translate theoretical designs into real-world outcomes.

1.5 Thesis Outline and Structure

This PhD thesis is organised into seven chapters as follows:

Chapter 1 provides a general introduction to the background of the thesis along with the research hypothesis and the motivation of this research.

Chapter 2 provides the definitions of metamaterials, metasurfaces, and active metasurfaces. Additionally, the chapter outlines the main tuning techniques and discusses the main challenges associated with the flexibility of mechanically tunable metasurfaces.

Chapter 3 introduces the concepts of ML and DL, providing a comprehensive overview of their general applications and classifications. It critically reviews the latest advancements of artificial intelligence (AI) in nanophotonics, focusing on three primary sections: design and optimisation, inverse design, and material selection. Additionally, it outlines the current limitations and gaps in harnessing the full potential of AI in each application area.

Chapter 4 presents two innovative geometries designed to achieve a high level of flexibility. This chapter discusses the details of the three-dimensional (3D) mechanical simulation for engineering stress redistribution at the interface between rigid resonators and flexible substrates. The optimised parameters are then used for the 3D optical simulation techniques to demonstrate the tuning potential of the ultra-stretchable designs. Furthermore, three applications of the device are discussed in the latter part of the chapter as proof of concept. Additionally, this chapter covers the details of data generation through simulation.

Chapter 5 introduces the developed DL networks for optimisation, inverse design, materials discovery, and free-form geometry discovery. It provides comprehensive details on each algorithm, including the data generation process, training procedures, challenges, and their accuracy. Additionally, a detailed comparison of the performance of the developed models with other reported works is presented, with a particular emphasis on the potential of the developed models.

Chapter 6 comprehensively reports the fabrication process and experimental procedures employed in this study to validate the achieved result in the previous chapters. Additionally, it discusses the main reasons for the observed mismatch between the properties achieved from the fabricated samples and those predicted by the DL algorithms.

Chapter 7 presents the thesis hypothesis and key conclusions drawn from the research undertaken. The chapter also outlines the main contributions made throughout the thesis, offering a comprehensive summary of the findings.

Bibliography

[1] A. Li, S. Singh, and D. Sievenpiper, "Metasurfaces and their applications," *Nanophotonics*, vol. 7, no. 6, pp. 989–1011, 2018.

[2] H.-D. Jeong, H. Kim, and S.-Y. Lee, "Review of metasurfaces with extraordinary flat optic functionalities," *Current Optics and Photonics*, vol. 8, no. 1, pp. 16–29, 2024.

[3] A. H. Dorrah and F. Capasso, "Tunable structured light with flat optics," *Science*, vol. 376, no. 6591, p. eabi6860, 2022.

[4] K. Shastri and F. Monticone, "Nonlocal flat optics," *Nature Photonics*, vol. 17, no. 1, pp. 36–47, 2023.

[5] A. S. Solntsev, G. S. Agarwal, and Y. S. Kivshar, "Metasurfaces for quantum photonics," *Nature Photonics*, vol. 15, no. 5, pp. 327–336, 2021.

[6] S. Mason, M. L. Meretska, C. Spagele, M. Ossiander, and F. Capasso, "Metasurface-controlled holographic microcavities," *ACS photonics*, vol. 11, no. 3, pp. 941–949, 2024.

[7] Y. Zhu, H. Niu, Y. Li, T. Lv, H. Li, X. Fan, and C. Bai, "Tunable metamaterial broadband perfect absorber based on double-layer graphene nanofilm," *Optical Materials*, vol. 149, p. 115085, 2024.

[8] S. Wu, Q. Ze, R. Zhang, N. Hu, Y. Cheng, F. Yang, and R. Zhao, "Symmetry-breaking actuation mechanism for soft robotics and active metamaterials," *ACS applied materials & interfaces*, vol. 11, no. 44, pp. 41 649–41 658, 2019.

[9] Y. Ke, I. Balin, N. Wang, Q. Lu, A. I. Y. Tok, T. J. White, S. Magdassi, I. Abdulhalim, and Y. Long, "Two-dimensional sio2/vo2 photonic crystals with statically visible and dynamically infrared modulated for smart window deployment," *ACS applied materials & interfaces*, vol. 8, no. 48, pp. 33 112–33 120, 2016.

[10] M. K. Hedayati, M. Javaheri, A. U. Zillohu, H. J. El-Khozondar, M. S. Bawa'aneh, A. Lavrinenko, F. Faupel, and M. Elbahri, "Photo-driven super absorber as an active metamaterial with a tunable molecular-plasmonic coupling," *Advanced Optical Materials*, vol. 2, no. 8, pp. 705–710, 2014.

[11] E. Kovalik, J. Eaves-Rathert, C. L. Pint, and J. G. Valentine, "Low-power electrochemical modulation of silicon-based metasurfaces," *ACS Photonics*, vol. 11, no. 2, pp. 445–451, 2024.

[12] S. Chen, J. Chen, X. Zhang, Z.-Y. Li, and J. Li, "Kirigami/origami: unfolding the new regime of advanced 3d microfabrication/nanofabrication with "folding"," *Light: Science & Applications*, vol. 9, no. 1, p. 75, 2020.

[13] C. Zhang, J. Jing, Y. Wu, Y. Fan, W. Yang, S. Wang, Q. Song, and S. Xiao, "Stretchable all-dielectric metasurfaces with polarization-insensitive and full-spectrum response," *ACS nano*, vol. 14, no. 2, pp. 1418–1426, 2019.

[14] N. Jeon, J. Noh, C. Jung, and J. Rho, "Electrically tunable metasurfaces: from direct to indirect mechanisms," *New Journal of Physics*, 2022.

[15] C.-W. Lee, H. J. Choi, and H. Jeong, "Tunable metasurfaces for visible and swir applications," *Nano Convergence*, vol. 7, pp. 1–11, 2020.

[16] P. Won, J. J. Park, T. Lee, I. Ha, S. Han, M. Choi, J. Lee, S. Hong, K.-J. Cho, and S. H. Ko, "Stretchable and transparent kirigami conductor of nanowire percolation network for electronic skin applications," *Nano letters*, vol. 19, no. 9, pp. 6087–6096, 2019.

[17] A. Ghasemi, R. Fang, D. A. Zeze, and M. K. Hedayati, "Ultra-stretchable active metasurfaces for high-performance structural color," *AIP Advances*, vol. 13, no. 8, 2023.

[18] M. Abbasipour, P. Kateb, F. Cicoira, and D. Pasini, "Stretchable kirigami-inspired conductive polymers for strain sensors applications," *Flexible and Printed Electronics*, vol. 8, no. 2, p. 024003, 2023.

[19] K. Yong, S. De, E. Y. Hsieh, J. Leem, N. R. Aluru, and S. Nam, "Kirigami-inspired strain-insensitive sensors based on atomically-thin materials," *Materials Today*, vol. 34, pp. 58–65, 2020.

[20] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nature Photonics*, vol. 12, no. 11, pp. 659–670, 2018.

[21] D. W. Davies, K. T. Butler, A. J. Jackson, A. Morris, J. M. Frost, J. M. Skelton, and A. Walsh, "Computational screening of all stoichiometric inorganic materials," *Chem*, vol. 1, no. 4, pp. 617–627, 2016.

[22] J. Park, S. Kim, D. W. Nam, H. Chung, C. Y. Park, and M. S. Jang, "Free-form optimization of nanophotonic devices: from classical methods to deep learning," *Nanophotonics*, vol. 11, no. 9, pp. 1809–1845, 2022.

[23] G. Lan, Y. Wang, and J.-Y. Ou, "Optimization of metamaterials and metamaterial-microcavity based on deep neural networks," *Nanoscale Advances*, vol. 4, no. 23, pp. 5137–5143, 2022.

Chapter 2

Theory

2.1 Chapter Overview

This chapter provides a comprehensive overview of general background of metamaterials. The discussion then transitions to metasurfaces, a subclass of metamaterials that have garnered significant attention in recent years. A brief history of metasurfaces is provided, tracing their evolution from the foundational concepts to their current state of the art. The main classifications and types of metasurfaces are outlined, highlighting their diverse functionalities and applications across different wavelengths and domains.

Moreover, the concept of active metasurfaces is discussed and the main tuning techniques are outlined in detail. A significant portion of the discussion centres on mechanically tunable metasurfaces, one of the most studied types of active metasurfaces. Despite their potential, mechanically tunable metasurfaces pose

certain limitations, which are explained in the last part of the chapter.

2.2 Background

The fundamental interaction between electromagnetic waves and different materials is the basis for almost all optical phenomena, playing a crucial role in how we perceive light [1]. Numerous natural materials have been studied throughout scientific history, leading to the development of innovative designs with a variety of applications, from energy to healthcare. However, sometimes the ever-growing demands of humanity in the 21st century are not met by conventional structures or materials. When nature is insufficient in such situations, scientists turn to the technique of molecularly altering materials to create whole new compounds with precisely specified properties [2]. For example, natural materials have a limited optical refractive index (RI), from 1 to 3, which has limited the advancement in the development of fast-responding computer chips [3].

An alternative approach for developing materials and structures with unique properties in nanophotonics is to purposefully design the components at the sub-wavelength scale. Even though these synthetic particles are larger than the atomic or molecular sizes of typical materials, their inhomogeneities are lower than the wavelength of interest, enabling to be defined in terms of **"effective"** parameters. These man-made materials are commonly known as **"metamaterials"** [4, 5].

The prefix "meta," which derives from the Greek word for "beyond", accurately describes these novel materials with properties beyond that of natural materials [6]. The word "metamaterial" was initially used in a study published by Smith et al. [7] in 2000. This study proposes a structured material having negative permittivity and permeability at microwave frequencies. Soon after this historic accomplishment, the subject of metamaterials experienced rapid expansion, attracting scientists from a wide range of research fields, including electrical engineering, physics, photonics, chemistry, and materials science.

The scientific community is still debating the exact definition of metamaterials. While some try to characterise their uniqueness using terms like "properties not

found in naturally developed components" or "absent in nature" [8], these definitions frequently fail to capture the wide range of properties that make metamaterials unique and set them apart from other artificially fabricated structures like photonic crystals. Cai's foundational work provides a more thorough, though not conclusive, definition: "Metamaterials are synthetic structures whose characteristics come from the structure of their units rather than the component materials. Their electromagnetic response is expressed through homogenised material properties, and they exhibit inhomogeneity on a scale much lower than the wavelength of interest" [9]. This definition notes the revolutionary potential of metamaterials which use engineered design to develop materials that go beyond what nature can provide.

Based on this definition, three levels of inhomogeneities are graphically represented in Figure 2.1, which exhibits distinct behaviours [10]. Starting with the smallest dimension, the subwavelength level is where the idea of metamaterials is first developed. Meta-atoms are the basic building components of metamaterials. Each meta-atom is smaller than the wavelength under consideration. The macroscopic inhomogeneity of the metamaterial is attributed to the subwavelength range of separation between these meta-atoms. This characteristic allows us to essentially view the metamaterial as a "material" rather than a specific device [11].

At a larger scale, photonic crystals exhibit inhomogeneity [12], influencing phenomena like diffraction and interference. Their periodic arrangements at or exceeding the wavelength scale provide optical bandgaps [13]. In contrast, conventional materials, with even larger inhomogeneities, follow the classical geometrical optics [11].

Metamaterials are often defined by a few constraints, primarily related to size [14]. In most cases, these materials consist of regularly arranged meta-atoms. However, there are examples where metamaterials are organised randomly, and they exhibit intriguing properties [10]. Artists have historically harnessed the distinctive qualities of metamaterials for creative purposes, even predating a comprehensive understanding of the associated physical principles. For example, the Lycurgus Cup is a remarkable ancient Roman glass chalice known for its unique optical proper-

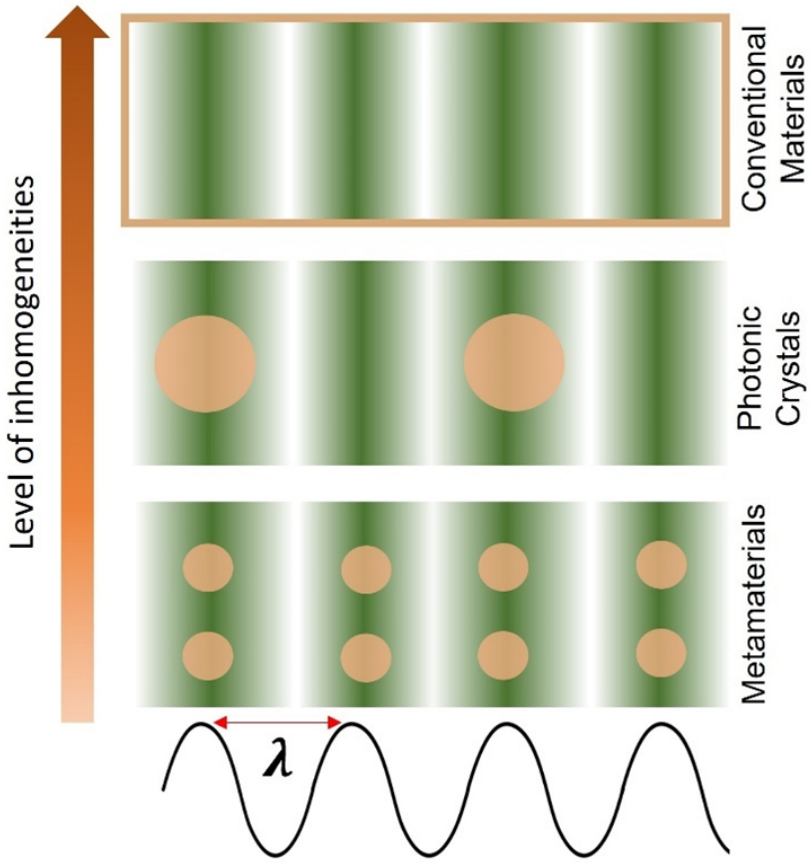


Figure 2.1: Optical material inhomogeneities at different levels. A light wave with two peaks separated by a wavelength is represented by red patterns.

ties and beautiful craftsmanship. This cup which dates to the fourth century AD and is currently kept in the British Museum, is made from a dichroic glass, which contains tiny amounts of random-sized gold and silver nanoparticles randomly dispersed in the structure [15]. What makes this cup particularly fascinating is its ability to change colour when illuminated from different angles. It appears green in reflected light, but when light passes through the glass, it turns reddish-brown, as Figure 2.2 illustrates [16].

While the term "metamaterial" was officially coined in 2000, microwave engineers had been studying the fundamental idea of using artificial dielectrics made of periodic configurations of subwavelength metallic wires, spheres, or plates for nearly seven decades. [18, 19]. Furthermore, several components, including split-ring resonators [20, 21], arrayed frequency filters [22], bi-anisotropic materials [23], and chiral materials [24], had been developed before the formal establish-

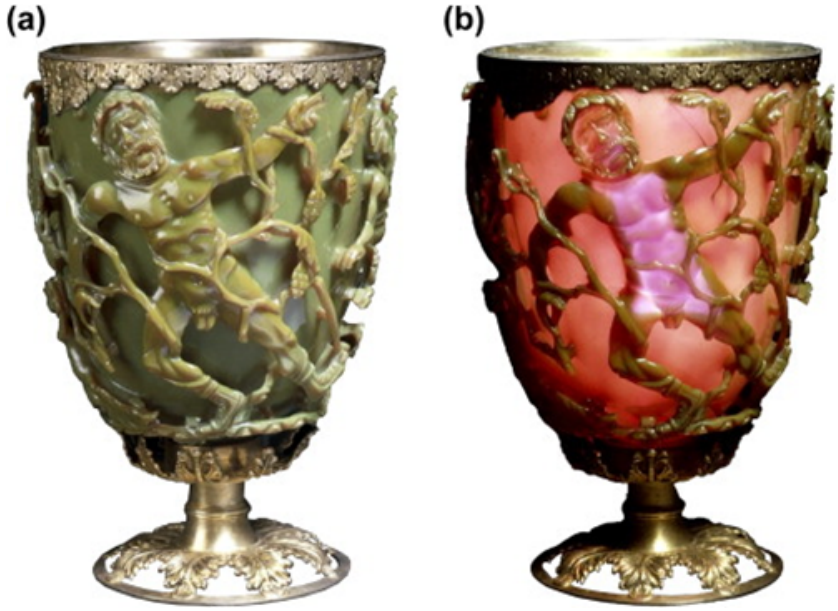


Figure 2.2: The colour of Lycurgus Cup in the a) reflected, and (b) transmitted mode. Reprinted by permission from Royal Society of Chemistry [17], Copyright (2012).

ment of the metamaterial field. This historical context highlights the longstanding exploration of metamaterial-like structures and components even before the field was officially recognised.

In the beginning, left-handed materials were the main focus of the majority of metamaterials research [25], also known as negative RI materials. Contemporary studies go well beyond materials with negative refraction. By adjusting the size, form, and composition of structural parts as well as their morphology, scientists are continuously pushing the boundaries to provide unique capabilities for a variety of applications. [26, 27, 28].

Metamaterials, initially developed for microwaves, have now been extended to the optical range of the electromagnetic (EM) spectrum, providing exceptional characteristics when exposed to visible light frequencies [29]. Progress in photonic metamaterial research has been made possible by modern nanofabrication techniques and enhanced capabilities for three-dimensional (3D) simulations [30]. Metamaterials bring a new approach to optical design. Instead of relying solely on the properties of materials, metamaterials are designed at a scale lower than the light wavelengths they interact with. This approach opens new possibilities

for controlling and manipulating electromagnetic waves, benefiting various areas of photonic engineering [30]. These applications include innovative waveguides [31], compact antennae [32], optical sensing [33], subwavelength imaging [34], and photonic circuits [35].

Several challenges have restricted the practical applications of optical metamaterials [36]. One of the primary obstacles lies in the 3D and large-scale fabrication process, particularly at the nanoscale. Achieving the precise and intricate structures required for metamaterials can be technically demanding and costly, often necessitating advanced nanofabrication techniques that are not always readily accessible. Additionally, these materials often exhibit properties that are highly sensitive to their structural parameters, adding further complexity to the design and manufacturing process [36]. Moreover, practical issues related to the loss of energy in metamaterials and their operational limitations in the optical spectrum also pose a significant challenge.

"Metasurfaces" are two-dimensional (2D) forms of metamaterials, possessing distinct properties and functionalities as schematically illustrated in Figure 2.3. By arranging subwavelength structures in a planar configuration, metasurfaces can precisely manipulate electromagnetic waves, offering novel properties and applications beyond conventional materials. Metasurfaces have relatively simple designs compared to 3D metamaterials, making them suitable for the fabrication of higher efficient miniaturised devices [36].

2.3 Metasurfaces: History, Classification, and Applications

Optical metasurfaces consist of engineered 2D arrays of subwavelength scatterers, known as meta-atoms (Figure 2.3). These meta-atoms are precisely designed to induce modifications in various characteristics of light, such as intensity, polarisation, phase, and wavelength. As a result, through the strategic choice of these meta-atoms, it becomes possible to finely control how light interacts with the structure. This interesting ability is widely used to introduce various planar func-

tional devices at a wide frequency range from microwave to terahertz (THz) [37] to visible light [38]. Furthermore, metasurfaces are often fabricated using such standard manufacturing methods as lithography and nanoimprinting, which are simpler than the fabrication of the counterpart metamaterials [39].

The concept behind metasurfaces has been observed by scientists long before the term itself was coined [40]. When Francis Hopkinson was staring through a thin silk handkerchief at a streetlamp in 1785, he spotted dark dots. Rittenhouse subsequently clarified this phenomenon by stating that white light is filtered into particular wavelength bands. This spectral filtering procedure was subsequently improved in the concept of diffraction gratings, which involves breaking down polychromatic light into its spectral components through a decomposition process [41].

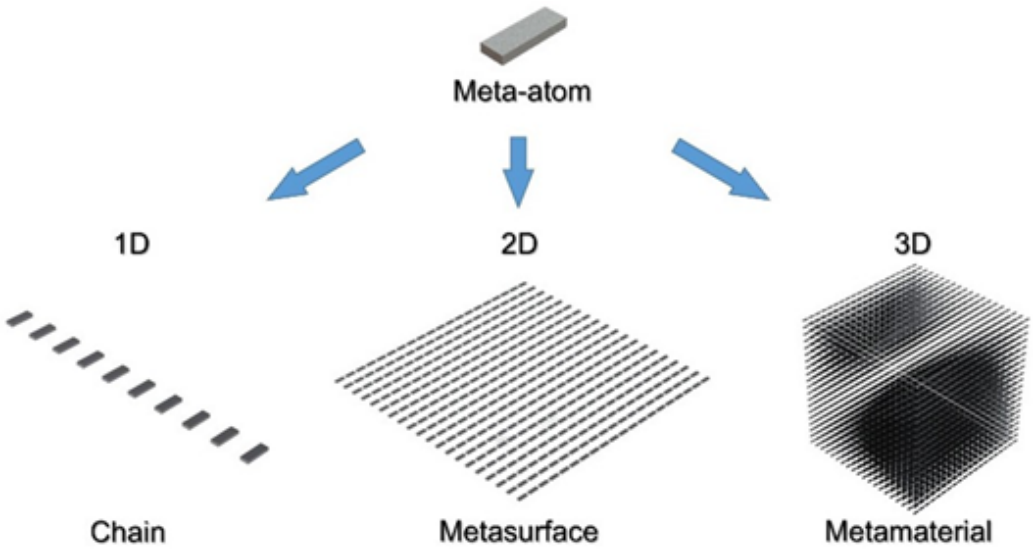


Figure 2.3: Schematic representation of meta-atom and its arrangement in a 1D chain, 2D metasurface and 3D metamaterials. Reprinted by permission from MDPI [42], Copyright (2016).

One of the earliest ideas in this field was to create metasurfaces with spatially varying optical responses. The Pancharatnam-Berry phase was first proposed by Hasman et al. [43] in the early 2000s to control light's polarisation state in space. They were able to manipulate light wavefronts by the fabrication of gratings made of metal-dielectric nanoparticles. This observation was completely overlooked by the optical scientific world for nearly ten years. In 2011, the generalised law of

refraction, which is based on optical surfaces that change spatially was introduced by Capasso's group [44]. This marked a significant turning point in the development and recognition of metasurfaces.

Like metamaterials, metasurfaces can achieve functionalities that are unattainable with standard optical elements. Conventional optical components, where phase changes accumulate as light propagates through various materials, follow Snell's classical law which explains the bending of light as it passes through different materials, such as air, glass, or water, leading to phenomena like refraction of light when it enters or exits a medium with a different RI (cf. Figure 2.4 (a)) [45]. The law is typically stated as:

$$\begin{cases} n_i \sin(\theta_i) = n_t \sin(\theta_t), \\ \theta_r = \theta_t \end{cases} \tag{2.3.1}$$

Where θ_i and n_i correspond to the angle of incidence of the incident light and RI of the first medium, respectively. However, θ_t and n_t are in the transmission (second) medium.

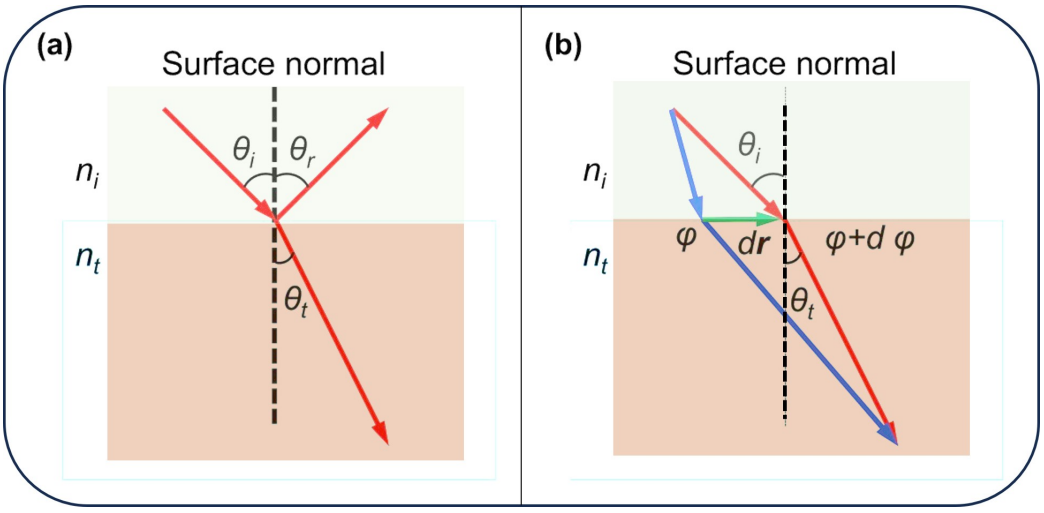


Figure 2.4: Illustration of Snell's Law: (a) traditional refraction in different media, and (b) extended law with phase gradient at media interface. θ_i, θ_t, and θ_r correspond to the incident angles, transmitted, and reflected beam, respectively. n_i and n_t represent the RI of the first and second media.

Conventional optical components follow traditional Snell's law, which explains how these variables relate to the angle of reflection, θ_r. But metasurfaces

provide a clear benefit by allowing sudden variations in the characteristics of incident light beams, which results in phase gradients at the two-media interface as shown in Figure 2.4 (b). In 2011, a generalised version of Snell's law was proposed to consider phase gradients and demonstrate the transformative capability of metasurfaces in altering optical interactions [44]. The generalised Snell's law can be expressed as:

$$\begin{cases} n_t sin\theta_t - n_i sin\theta_i = {k_0}^{-1} \nabla\theta \\ sin\theta_r - sin\theta_i = {n_i}^{-1} \, {\mathrm{k}_0}^{-1} \nabla\theta \end{cases} \tag{2.3.2}$$

where θ_i, θ_t, and θ_r stand for the incident angles, transmitted, and reflected beam, respectively. The phase gradient is represented by $\nabla\theta$, while the wavenumber in free space is represented by k_0.

The generalised Snell's law states that by adjusting the phase gradient each meta-atom introduces, it is possible to intentionally create unique far-field light patterns. Consequently, metasurfaces have the potential to control EM waves in novel ways that go beyond the limitations of traditional optical components [46, 35]. For example, the field of flat optics was developed to study how metasurfaces can be used to create small, low-profile optical components because of the great degree of control they provide over the characteristics of reflected or transmitted light [47].

Numerous perspectives exist for categorising optical metasurfaces. Each one focuses on a distinct facet of their operation. One crucial perspective to consider is the classification from a materials point of view, which considers the composition of the meta-atoms used in metasurface design into two main types: "plasmonic" and "dielectric" as shown in Figure 2.5. Plasmon metasurfaces are unit structures made of materials (mainly noble metals) that support plasmon, and the particles' optical behaviour determines the plasmon resonances of the structure. Conduction electrons in metallic particles depart from their balanced state with respect to the core ions when they interact with an electric field. This displacement results in the polarisation of the particle and the creation of a depolarising field [48]. This collective motion inside the metallic particle can be compared to a Lorentzian oscillator when it is exposed to a time-varying external field. This oscillator's dis-

placement amplitude peak, which is usually found at the resonance frequency, is its defining feature. Additionally, it is followed by a phase shift of π throughout the resonance's spectral width [48].

However, dielectric metasurfaces are normally made of materials with high RI like silicon and germanium. These dielectric particles provide both electric dipole (ED) and magnetic dipole (MD) responses, primarily based on Mie resonances [49]. These dielectric particles show MD (first Mie resonance) and ED (second Mie resonance) excitation when exposed to light waves with frequencies at or below the bandgap frequency of the material [42]. As a result, these Mie resonances enhance the MD and ED at the centre of the dielectric particle at optical frequencies, and this enhancement is closely tied to the inherent characteristics of the particles [50]. Therefore, the properties of their unit structures (e.g. arrangements, sizes, and materials) have a significant impact on how plasmonic and dielectric metasurfaces respond. By precisely selecting such parameters for each unit structure within the metasurfaces, it becomes possible to manipulate the EM field of the scattered light wave, thus enabling precise control over the metasurface's optical properties [51].

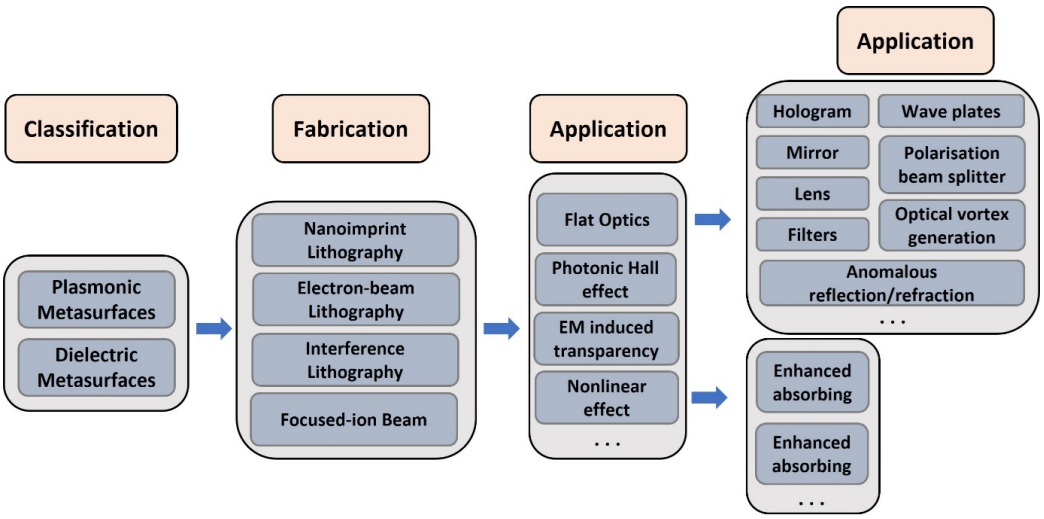


Figure 2.5: Various types, fabrication processes, and applications of optical metasurfaces. Reproduced by permission from MDPI [42], Copyright (2016).

Fabrication methods for optical metasurfaces vary in terms of precision, scalability, and cost-effectiveness [52]. The most effective strategy depends on the

particular application, including the required resolution and the resources that are available. Researchers are continually exploring new fabrication techniques and combinations of existing methods such as electron beam lithography (EBL) [53], atomic layer deposition (ALD) [54], focused ion beam (FIB) [55], interference lithography [56], and nanoimprint lithography (NIL) [57].

The unique properties of metasurfaces and the evolution of nanofabrication techniques have led to their application in diverse fields, including flat optics [47, 58], nonlinear effects [59, 60], photonic Hall effects [61], perfect absorber [62], electromagnetically induced transparency [63], cloaking [64], and more as shown in Figure 2.6. For example, in flat optics, metasurfaces have found applications in creating anomalous wave plates, controlling reflection and refraction, designing flat lenses, and developing mirrors, holograms, filters, optical vortex generation, and polarisation beam splitters, among other functions [65, 66]. However, the nonlinear effects of metasurfaces have been used in the second and third-harmonic generation and enhanced light absorption [67]. These diverse applications underscore the wide-ranging potential and significance of metasurfaces in advancing optical and photonic technologies [68].

As explained in the previous section, metasurfaces have shown impressive potential in the development of compact, high-performance, and cost-effective optical devices, sparking significant interest in the nanophotonics community. However, progress in this field has mostly focused on "passive metasurfaces". This class of metasurfaces have predefined properties that cannot be changed after they are fabricated. Essentially, they have been customised for specific applications, and designed with complex details to suit target purposes. The demand for flexibility and adaptability has led to a pioneering field of research centred on “active metasurface” or “tunable metasurfaces” [75]. The following parts explain the advancements, potential applications, and key challenges of active metasurfaces.

2.3.1 Tunable Metasurfaces

Traditional optical metasurfaces offer static and predefined optical functions that are primarily controlled by the shape, size, geometry, and arrangements of meta-

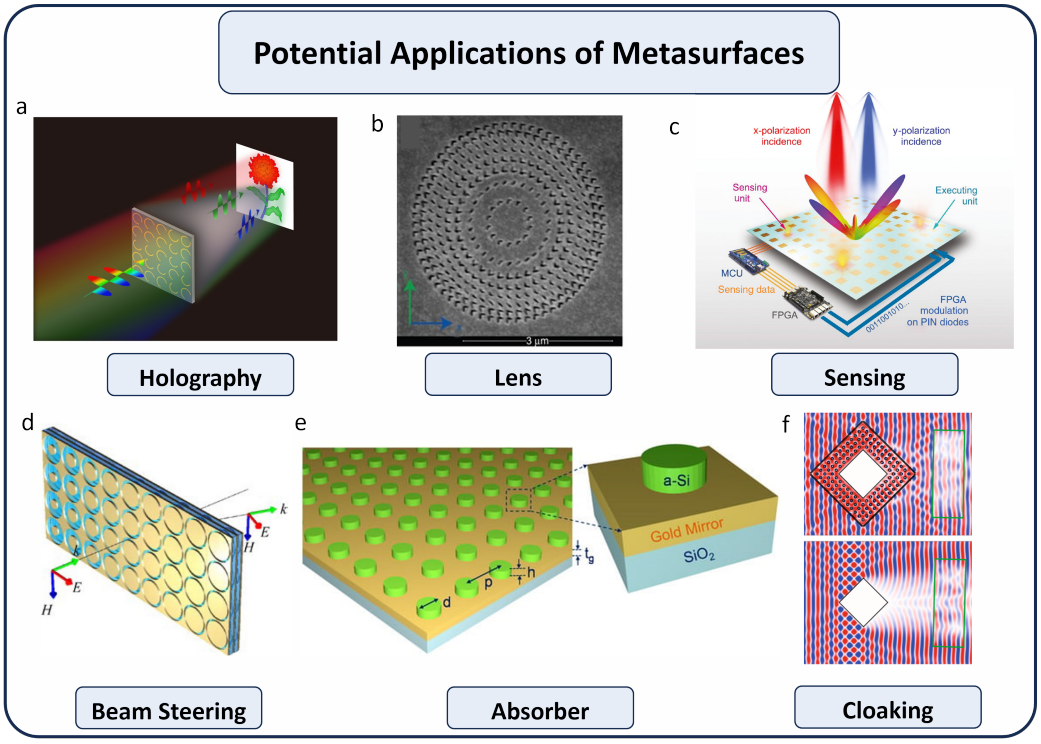


Figure 2.6: Some common applications of metasurfaces; (a) holography, Reprinted by permission from Nature [69], Copyright (2016), (b) lens, Reprinted by permission from Nature [70], Copyright (2013), (c) sensing, Reprinted by permission from Nanophotonics [71], Copyright (2020), (d) beam steering, Reprinted by permission from Optica [72], Copyright (2013), (e) absorber, Reprinted by permission from MDPI [73], Copyright (2018), (f) cloaking, Reprinted by permission from Nature [74], Copyright (2018).

atoms [76]. Nevertheless, achieving dynamic control over the interaction of electromagnetic waves and nanostructures has been a long-standing challenge in optics [77, 78]. Optical tuning can be achieved by relocating the resonators' spatial position since the optical response is dependent on their location as schematically illustrated in Figure 2.7 [79]. A tunable metasurface allows for changes to two or more pre-defined states, such as focal lengths and deflection angles whereas a passive metasurface keeps its function constant. A continuously tunable metasurface refers to the continuing adjustment of a certain property within a predetermined range, such as changing focal length or deflection angle. Lastly, a freely tunable metasurface allows for unrestricted and continuous adjustment of property, exemplified by the 3D modification of the structure [80].

Various external stimuli-responsive systems like magnetic field [81], thermal

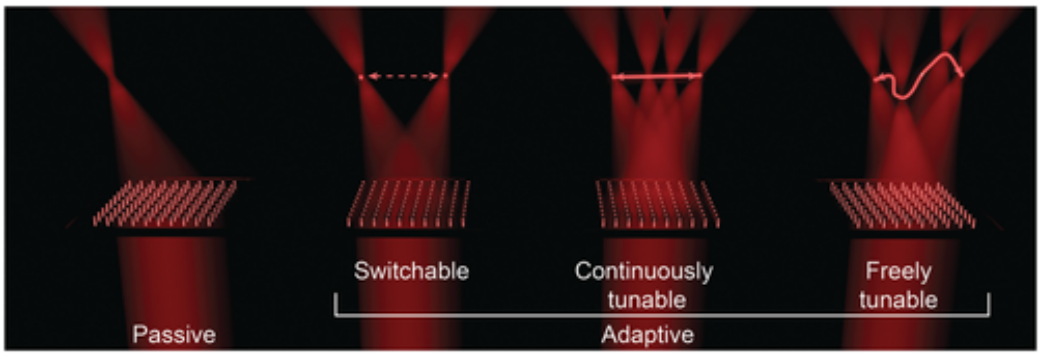


Figure 2.7: Metasurfaces go from passive to active (adaptive). An example of a continuously adjustable, freely tunable, passive, switchable metasurface is used to fabricate a metalens. Reprinted by permission from Wiley [80], Copyright (2019).

fluctuations [82], optical excitation [83], electrochemical tuning [84], mechanical force [85, 86], electrical actuation [87], and chemical reactions [88], etc., have been employed to make this tuning possible. The dielectric characteristics or physical dimensions of metasurface elements are altered by these external influences, which makes it possible to modulate the amplitude and phase responses of the structure. This ability to actively and dynamically vary metasurface properties opens the door to dynamic holograms and adaptive focusing lenses with adjustable focal lengths [89, 90, 91].

2.4 Tuning Techniques

Regardless of the type of stimuli used for parameter tuning, progress in active metasurface development includes diverse innovative techniques broadly classified into material-based techniques, geometry-based methods, and hybrid approaches.

2.4.1 Material-based Techniques

Materials-based techniques focus on exploiting the properties of specific materials such as phase change materials (PCMs) [92], ferroelectrics [93], 2D materials like graphene and MoS_2 [94], liquid metals (LMs) [95], and transparent conductive polymers [96] to achieve dynamic control over metasurface functionalities.

These materials possess a unique combination of properties that make them particularly attractive for metasurface engineering.

For example, PCMs, which consist of alloys of Germanium (Ge), Antimony (Sb), and Tellurium (Te), may change from a disordered-amorphous form to an ordered-crystalline one when heated or exposed to other stimuli [97]. To achieve this transition, the chalcogenide material is heated above its glass transition temperature (T_g) using a laser pulse or electronic excitation, which initiates a thermal transition process involving nucleation and crystallisation growth. Reverting to the amorphous state is accomplished by heating the material above its melting temperature (T_m), followed by rapid cooling. The significant contrast in optical and electronic properties between the two states is a result of the phase change that happens in the material [98]. Metasurfaces containing PCMs enable control over the reflected or transmitted phase of incident waves, leading to applications in tunable lenses, absorbers [92], holography, reflectors [99] and information encryption [97].

The PCMs have been used for the development of infrared radiation (IR) active metasurface, as demonstrated by Gholipour et al. [100]. They developed a non-volatile, all-optical modulation of resonance in the IR spectrum through a broadband switch (cf. Figure 2.8 (a-d)). They demonstrated that combining nanostructured plasmonic metamaterials with phase-change medium chalcogenide glass allows for this modulation. The hybrid metamaterial devices underwent a homogenous phase shift over sizable regions upon single-pulse laser activation, which profoundly altered their optical transmission and reflection properties.

Germanium Antimony Telluride (GST), as a common example of PCMs, was also used to fabricate an active metasurface wherein a non-volatile tuning of Au nanodisk resonance was accomplished across a wide spectral range spanning 500 nm by introducing a thin layer of GST as a RI-tunable medium (cf. Figure 2.8(e)). Figure 2.8(f) shows the transmission spectra of the hybrid plasmonic crystal at various intermediate phases. The phase-change thin film exhibited a red shift in the lattice resonance with increasing baking time, resulting in a more metal-like part. In particular, after 20 minutes of baking, the transmittance rises from 96% (black curve) to 72% (blue curve) at a wavelength of 1.89 μm [101].

While these examples highlight dynamic tuning of the intensity of physical response like scattered light, there are some active metasurfaces in which the modulation of the phase of scattered light through external control has been discussed. Various materials like vanadium dioxide (VO_2) and GST, which undergo phase transitions, allow for extensive index modulation [102]. However, they suffer from limitations in switching speed [103] and typically demand higher power consumption [104] when compared to field-effect modulation techniques. Additionally, when using external stimuli such as optical pumping, the range for altering the RI depends on the size of the stimuli. For example, a concentrated laser spot leads to a rather vast area that limits accurate control of individual metasurface components [105].

2.4.2 Geometry-based Techniques

Geometry-based methods rely on a local large but low-energy deformation that renders the collective transformations like Kirigami, Origami, and other shape-shifting strategies. Such strategies have been used to develop tunable metalenses [106], strain-multiplexed meta-holograms [107], and the active management of structural colouration [108]. This technique normally benefits from flexible substrates with higher levels of deformability. For instance, J. Reeves et al. [109] used a series of simulation and experimental techniques to propose geometrically controllable optical metasurfaces using a microelectromechanical systems (MEMS)-based stencil lithography technique. In this research, metallic meta-atoms were placed on a deformable microstructured polymer (Polydimethylsiloxane (PDMS)) scaffold, where the mechanical deformation of the substrate alters metasurface reflectivity in the mid-IR region. This geometry provided just a displacement in the direction normal to the applied force.

Origami, the art of paper folding, has inspired metasurface design by enabling the creation of complex 3D structures from 2D ones [85]. Researchers have used origami principles to fold metasurface sheets into various configurations, allowing dynamic tuning of optical properties. These folded structures can alter the metasurface's geometry to change its resonant frequency, polarisation, and dispersion

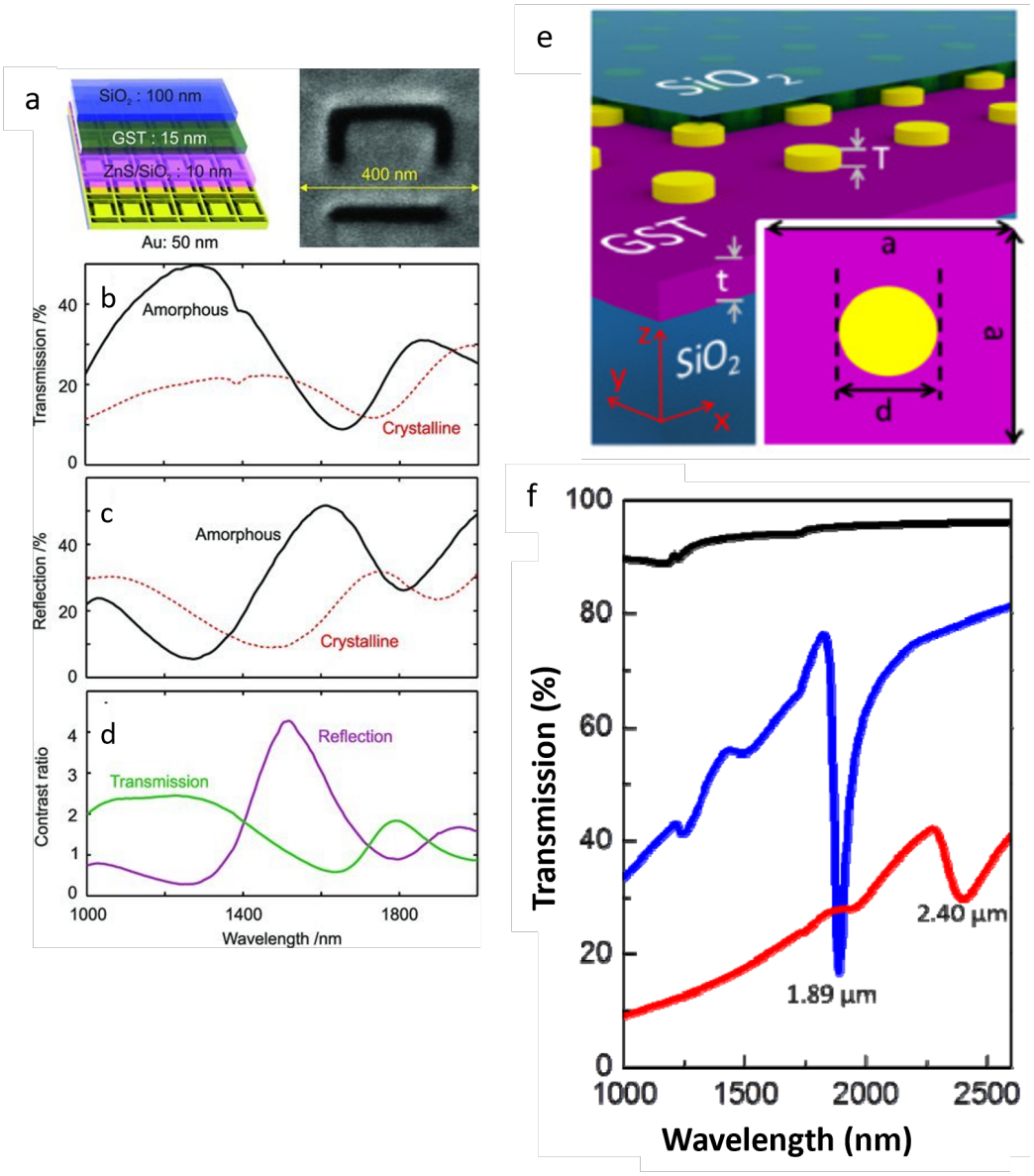


Figure 2.8: a) The structural features of the developed meta-switch layer thicknesses (left), and the scanning electron microscopy (SEM) image of the unit cell (right). Reprinted by permission from Wiley [100] Copyright (2013), b,c) the transmission and reflection curves of the hybrid device structure for the GST layer's crystalline and amorphous phases (as labelled). d) The transmission and reflection modulation contrast spectrum dispersion linked to the meta-device's chalcogenide phase switching. Reprinted by permission from Wiley [100], Copyright (2013). (e) a 3D schematic of an array of Au nanodisks on a thin layer of GST, (f) transmission spectra of continuous tuning at different baking time during the crystallisation process. Reprinted by permission from Optica Publishing Group [101], Copyright (2013).

characteristics. However, Kirigami is a technique that involves cutting patterns into materials. In metasurfaces, kirigami is used to introduce tailored cuts that

can be manipulated to control the propagation of electromagnetic waves. By selectively cutting metasurface elements, researchers have developed metasurfaces capable of reconfigurable wavefront manipulation, tunable absorption, and beam shaping [110].

For example, Wang et al. [111] located split-ring resonator (SRR) on a foldable Miura-ori unit as a common origami-based pattern to produce a programmable optical metasurface as shown in Figure 2.9 (a). Simulation results demonstrated that the 2D-to-3D transformation of flexible structure along the third dimension changes the spatial location of net ED and MD resonators parallel or anti-parallel to each other, providing a strong chiral response. This model is inherently unable to provide a wide range of displacement in two directions.

In another study, X. Liu et al. [112] fabricated a flexible optical metasurface comprising two-layered gold nanoribbon (340 nm) arrays on PDMS using EBL. In this approach, an out-of-plane design was employed to dynamically control the energy of surface plasmon polariton (SPP) in the visible to near-IR wavelength range. As shown in Figure 2.9 (b), the coupling characteristics between nanoribbons can be programmed by stretching the PDMS substrate. Moreover, they showed that resonance mode shifts to a lower frequency by increasing the incident angle.

Although various geometries have been developed to dynamically manipulate the inter-particle space in an active structure, designing a geometry without getting caught in a bifurcation which provides a more prominent frequency shift through resonator displacement in more than one direction, is a challenge. To address this challenge, Yan et al. [113] introduced the concept of self-folding. Figure 2.9 (c) shows the continuous and reversible 2D-to-3D shape transformation of a simple self-foldable structure guided by a unidirectional compressive buckling. In this study, they employed both theoretical modelling methods and experimental techniques to build an autonomic deployable origami-based 3D structure through an out-of-plane mechanical buckling.

Moreover, the implementation of nanoscale reconfigurable optical metasurfaces has rarely been considered due to the fabrication challenges associated. Another equally important remaining challenge in developing a polymer-supported con-

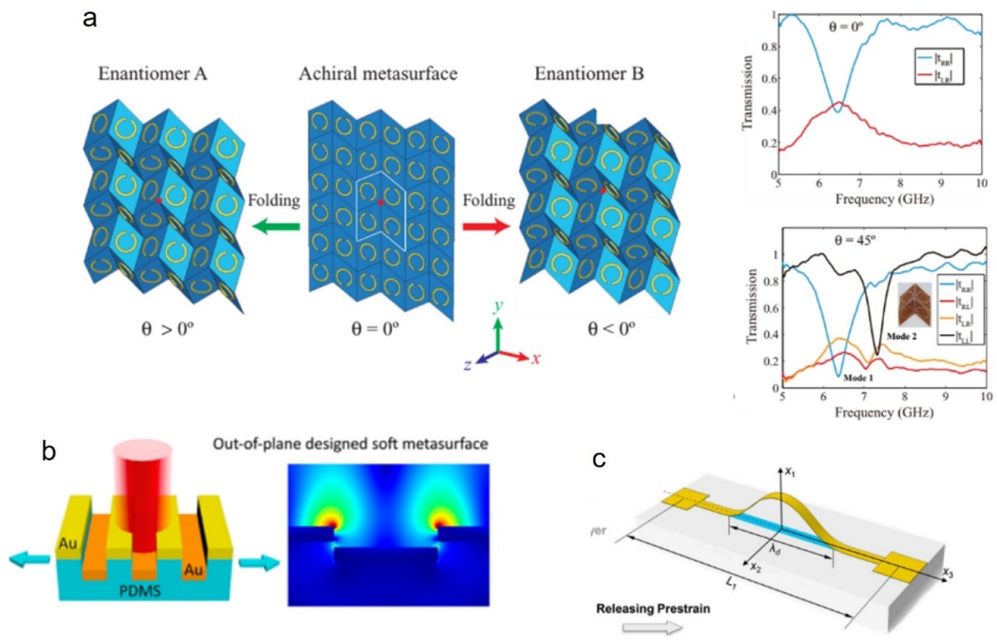


Figure 2.9: Reconfigurable optical metasurface guided by a 2D-to-3D shape transformation structure. a) a 2D origami-based pattern with two positive and negative folding(top), and transmission spectra at different folding angles. Reprinted by permission from Wiley [111], Copyright (2017). b) a schematic pattern (left) and simulated field intensity (right) of the out-of-plane soft metasurface. Reprinted by permission from ACS Publications [112], Copyright (2018). c) a schematic illustration of the continuous and reversible 2D-to-3D shape transformation of a simple self-foldable structure guided by a unidirectional compressive buckling. Reprinted by permission from Wiley [113], Copyright (2016).

tinuous metal film is controlling the rupture of structures when they are stretched, which is strongly dependent on the interfacial properties between substrate and film. It is reported that the ductility and flexibility of polymer-supported metal films improve as the interfacial strength increases, however, this increasing trend is generally affected by the interfacial stiffness [79].

2.4.3 Hybrid Techniques

Hybrid methods combine elements of both geometry-based and materials-based techniques to create metasurfaces with enhanced functionalities [114]. These approaches leverage the strengths of each method, resulting in metasurfaces that are highly versatile and adaptive.

2.5 Limitation of Mechanically Tunable Metasurfaces

The choice of stimulus depends on the requirements of the application, the level of tunability needed, and the compatibility with the surrounding environment. Among various systems, mechanically tunable ones have received substantial attention and are the subject of extensive research due to their high degree of tunability and versatility, offering low power consumption [115]. By applying controlled mechanical stress or strain, researchers can induce significant changes in the metasurface's geometry or structure. This versatility enables metasurfaces to adapt to a wide range of applications, making them a preferred choice for many research endeavours. The compatibility of mechanically tunable methods with existing technologies is another compelling reason for their popularity. Integrating mechanical tuning into established infrastructures and devices is relatively straightforward, making these metasurfaces more accessible for real-world applications [116].

One of the most well-known and extensively researched categories of mechanically tunable optical metasurfaces involves the placement of resonators on flexible substrates. In this configuration, the substrate's flexibility serves to dynamically alter the distances between resonators, consequently tuning the optical responses, including transmission and reflectance [108, 118, 119].
However, there exists a fundamental limitation with such bi-material structures. The significant disparity between the maximum range of elasticity of the elastomeric substrate and that of the resonators, such as metals and high RI dielectrics, severely constrains the stretchability of these devices. In ideal crystalline metals, the theoretical elastic strain limit can reach up to 10% [120]. However, most bulk metals exhibit elastic strain limits of less than 7% due to factors like inelastic relaxation involving defects such as dislocation, stress-induced phase transformations, and deformation twinning [121]. The elastic strain limits of common rigid components used in tunable metasurfaces are as follows: Copper (2.8%), Gold (1.9%), Silicon (2%), and Aluminium (1.27%) [122, 123, 124]. These values underline the practical limitations imposed by the materials themselves, which can hinder the achievable degree of stretchability in such resonator-substrate configu-

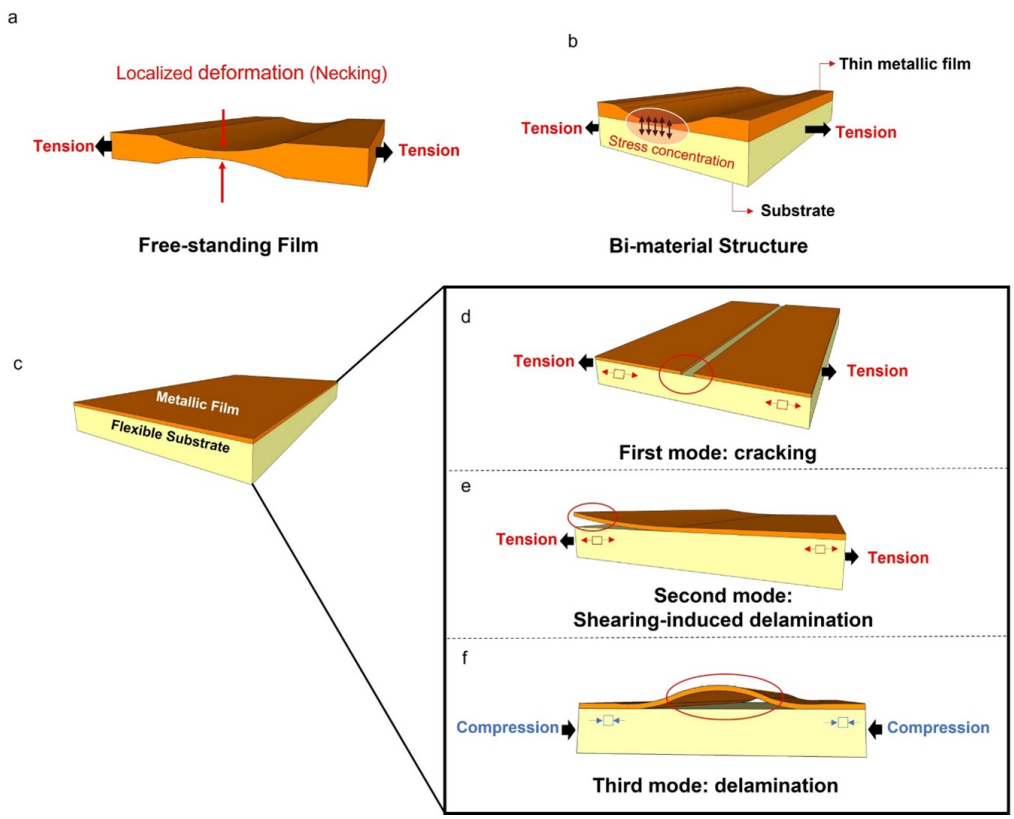


Figure 2.10: Schematic illustration of the general deformation under mechanical strains of a) freestanding, b) bi-material, and c) bi-layer structures with stiff metallic film and soft substrate. Typical failure modes are shown in d) cracking mode, e) shearing-induced delamination under tensile load, and f) delamination under compression load. Reprinted by permission from AIP Publishing [117], Copyright (2023).

rations.

Figure 2.10 (a) illustrates a comparison of the mechanical deformation in a metallic film, both in a free-standing structure and a bi-material one where the film adheres to a flexible substrate. In the case of the free-standing structure, when subjected to strain beyond its elastic limit, localised necking occurs, resulting in permanent deformation. However, when analysing the mechanical deformation and fracture behaviour of a bi-material structure under an external load, the situation becomes more complex. This complexity arises from the interplay of various phenomena, including cracking, necking, buckling, and delamination [125].

In a substrate-bonded film under tensile strain, the limited space for plastic deformation leads to a delocalised strain field in the metal film. This enables the film to undergo permanent deformation at a lower strain level, significantly ex-

tending the rupture point of the film compared to that of a freestanding film [126]. A bi-material structure can deform in various ways, as depicted in Figure 2.10 (d-f), and the specific mode of deformation depends on the shape, dimensions, and mechanical properties of both the rigid film and the soft substrate. In the first mode, when the applied uniaxial stress exceeds the point where further homogeneous deformation is not possible, strain-induced cracks form perpendicular to the applied stress (cf. Figure 2.10 (d)). These cracks propagate, ultimately causing the structure to fracture [127]. The second mode involves shearing stress at the interface due to a difference in the normal stress level. Delamination occurs when this shearing stress reaches the maximum shearing strength of the interface (See Figure 2.10(e)). In the third mode, a rigid/soft structure experiences compression, leading to the buckling of the rigid film. While high bonding strength can delay massive buckling, a wavelike buckling pattern on the surface is inevitable (See Figure 2.10 (f)) [128]. These deformation modes are prevalent in various rigid-soft materials, including metamaterials. Consequently, the deformability of metamaterials is intrinsically limited by the properties of their resonators. Resonators, being in direct contact with flexible substrates, are directly affected by external forces. Therefore, even in flexible metamaterials, elongations rarely exceed 30%, which falls well below the maximum elastic deformation of constituent elastomers as presented in Table 2.1 [108, 86, 86, 129, 130].

Despite achievements in the design of stretchable devices involving both rigid and soft components, such as serpentine or horseshoe geometries [131], nanocomposite-based structures [109], and ultrathin flexible substrates [129], devising a suitable structure to prevent the emergence of substantial cracks within a certain strain range, typically falling within 20%-50%, remains a significant challenge [132]. The performance of various geometries is provided in Table 2.1 The maximum reported strain is limited to 36%, which is by far below the maximum elasticity of the flexible substrates.

Table 2.1: Comparison of various types of stretchable optical MMs. Reprinted by permission from AIP Publishing [117], Copyright (2023)

Materials	Shape of resonator	Dimensions (nm)	Arrangment	Maximum Strain (%)	Ref.
Al/PDMS	rectangular	P[a]400-W[b]100-L[c]130-H[d]35	squared	32	[108]
Al/PDMS	disk	P320-D[e]200-H100	squared	31.6	[119]
Au/PDMS	split ring	L1=L2=2.4 μ m	-	5	[109]
TiO_2/PDMS	squared	P300-W230-H480	squared	30	[86]
Au/PDMS	rectangular	W100-L240-H70	Triangular	30	[129]
Si/PDMS	disk	P490-D305-H480	squared	36	[131]
Au-PS/PDMS	sphere	D1500	hexagonal	20	[132]
Au/PDMS	sphere	D11	close-packed	20	[133]
Silver/PDMS	grating	P665	-	23	[130]

[a] Periodicity; [b] Width; [c] Length; [d] Thickness; [e] Diameter of resonator.

2.6 Conclusion

Metasurfaces are a 2D version of metamaterials, comprising subwavelength components that can manipulate electromagnetic wave properties like phase, amplitude, polarisation, or direction of propagation. This chapter provided a comprehensive overview of the theory of metamaterials and metasurfaces, highlighting their importance and classifications. Unlike classical passive metasurfaces, whose features are set based on their design and composition, active metasurfaces may dynamically control their electromagnetic properties.

Tuning techniques employed for achieving tunability in active metasurfaces were generally classified into three main groups including geometry-based, material-based, and hybrid-based methods. The latest advancements in the methods and their limitations were discussed in this chapter.

Finally, despite the advancements of the mechanically active metasurfaces as the most straightforward tunable devices, their limitation in providing a higher level of stretchability in the commonly developed designs was discussed. The significant mismatch between the maximum elasticity of the elastomeric substrate and that of the resonators leads to a stress concentration in the interface of the structure. This localised stress causes plastic deformation and cracks formation in the metasurface. The proposed research hypothesis focuses on innovative geometries to redistribute the stress accumulation and enhance the stretchability of the device accordingly.

Bibliography

[1] J. Guan, J.-E. Park, S. Deng, M. J. Tan, J. Hu, and T. W. Odom, "Light–matter interactions in hybrid material metasurfaces," *Chemical Reviews*, vol. 122, no. 19, pp. 15 177–15 203, 2022.

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

[3] S. Kar, *Metamaterials and Metasurfaces: Basics and trends*. IOP Publishing, 2023.

[4] N. Suresh Kumar, K. C. B. Naidu, P. Banerjee, T. Anil Babu, and B. Venkata Shiva Reddy, "A review on metamaterials for device applications," *Crystals*, vol. 11, no. 5, p. 518, 2021.

[5] A. Komar, "Tunable all-dielectric metasurfaces: Fundamentals and applications," Ph.D. dissertation, The Australian National University (Australia), 2019.

[6] R. Kumar, M. Kumar, J. S. Chohan, and S. Kumar, "Overview on metamaterial: History, types and applications," *Materials Today: Proceedings*, vol. 56, pp. 3016–3024, 2022.

[7] D. R. Smith, W. J. Padilla, D. Vier, S. C. Nemat-Nasser, and S. Schultz, "Composite medium with simultaneously negative permeability and permittivity," *Physical review letters*, vol. 84, no. 18, p. 4184, 2000.

[8] J. B. Pendry and D. R. Smith, "The quest for the superlens," *Scientific American*, vol. 295, no. 1, pp. 60–67, 2006.

[9] W. Cai and V. M. Shalaev, *Optical metamaterials*. Springer, 2010, vol. 10, no. 6011.

[10] A. R. McGurn, *Introduction to nonlinear optics of photonic crystals and metamaterials*. IOP Publishing, 2021.

[11] M. Lapine and S. Tretyakov, "Contemporary notes on metamaterials," *IET microwaves, antennas & propagation*, vol. 1, no. 1, pp. 3–11, 2007.

[12] X.-J. Liu, Y. Liang, J. Ma, S.-Q. Zhang, H. Li, X.-Y. Wu, and Y.-H. Wu, "Two-dimensional function photonic crystals," *Physica E: Low-Dimensional Systems and Nanostructures*, vol. 85, pp. 227–237, 2017.

[13] A. Komar, R. Paniagua-Domínguez, A. Miroshnichenko, Y. F. Yu, Y. S. Kivshar, A. I. Kuznetsov, and D. Neshev, "Dynamic beam switching by liquid crystal tunable dielectric metasurfaces," *Acs Photonics*, vol. 5, no. 5, pp. 1742–1748, 2018.

[14] J. Bonache, G. Zamora, F. Paredes, S. Zuffanelli, P. Aguilà, and F. Martín, "Controlling the electromagnetic field confinement with metamaterials," *Scientific reports*, vol. 6, no. 1, p. 37739, 2016.

[15] D. J. Barber and I. C. Freestone, "An investigation of the origin of the colour of the lycurgus cup by analytical transmission electron microscopy," *Archaeometry*, vol. 32, no. 1, pp. 33–45, 1990.

[16] F. E. Wagner, S. Haslbeck, L. Stievano, S. Calogero, Q. Pankhurst, and K.-P. Martinek, "Before striking gold in gold-ruby glass," *Nature*, vol. 407, no. 6805, pp. 691–692, 2000.

[17] R. R. Arvizo, S. Bhattacharyya, R. A. Kudgus, K. Giri, R. Bhattacharya, and P. Mukherjee, "Intrinsic therapeutic applications of noble metal nanoparticles: past, present and future," *Chemical Society Reviews*, vol. 41, no. 7, pp. 2943–2970, 2012.

[18] W. E. Kock, "Metal-lens antennas," *Proceedings of the IRE*, vol. 34, no. 11, pp. 828–836, 1946.

[19] J. Brown, "Artificial dielectrics having refractive indices less than unity," *Proceedings of the IEE-Part IV: Institution Monographs*, vol. 100, no. 5, pp. 51–62, 1953.

[20] W. Hardy and L. Whitehead, "Split-ring resonator for use in magnetic resonance from 200–2000 mhz," *Review of Scientific Instruments*, vol. 52, no. 2, pp. 213–216, 1981.

[21] W. Froncisz and J. S. Hyde, "The loop-gap resonator: a new microwave lumped circuit esr sample structure," *Journal of Magnetic Resonance (1969)*, vol. 47, no. 3, pp. 515–521, 1982.

[22] J.-S. Hong and M. J. Lancaster, "Couplings of microstrip square open-loop resonators for cross-coupled planar microwave filters," *IEEE Transactions on Microwave theory and Techniques*, vol. 44, no. 11, pp. 2099–2109, 1996.

[23] F. Olyslager, "The behavior of electromagnetic fields at edges in bi-isotropic and bi-anisotropic materials," *IEEE transactions on antennas and propagation*, vol. 42, no. 10, pp. 1392–1397, 1994.

[24] M. M. Saadoun and N. Engheta, "A reciprocal phase shifter using novel pseudochiral or ω medium," *Microwave and optical technology letters*, vol. 5, no. 4, pp. 184–188, 1992.

[25] V. G. Veselago, "Electrodynamics of substances with simultaneously negative electrical and magnetic permeabilities," *Soviet Physics Uspekhi*, vol. 10, no. 4, pp. 504–509, 1968.

[26] J. B. Pendry, D. Schurig, and D. R. Smith, "Controlling electromagnetic fields," *science*, vol. 312, no. 5781, pp. 1780–1782, 2006.

[27] M. Decker, M. Klein, M. Wegener, and S. Linden, "Circular dichroism of planar chiral magnetic metamaterials," *Optics letters*, vol. 32, no. 7, pp. 856–858, 2007.

[28] E. Plum, V. Fedotov, A. Schwanecke, N. Zheludev, and Y. Chen, "Giant optical gyrotropy due to electromagnetic coupling," *Applied Physics Letters*, vol. 90, no. 22, 2007.

[29] R. S. Kshetrimayum, "A brief intro to metamaterials," *IEEE potentials*, vol. 23, no. 5, pp. 44–46, 2004.

[30] T. J. Cui, S. Zhang, A. Alu, M. Wegener, J. Pendry, J. Luo, Y. Lai, Z. Wang, X. Lin, H. Chen *et al.*, "Roadmap on electromagnetic metamaterials and metasurfaces," *Journal of Physics: Photonics*, 2024.

[31] D. C. Zografopoulos, J. F. Algorri, A. Ferraro, B. García-Cámara, J. M. Sánchez-Pena, and R. Beccherelli, "Toroidal metasurface resonances in microwave waveguides," *Scientific Reports*, vol. 9, no. 1, p. 7544, 2019.

[32] S. Painam and C. Bhuma, "Miniaturizing a microstrip antenna using metamaterials and metasurfaces [antenna applications corner]," *IEEE Antennas and Propagation Magazine*, vol. 61, no. 1, pp. 91–135, 2019.

[33] S. Tabassum, S. Nayemuzzaman, M. Kala, A. Kumar Mishra, and S. K. Mishra, "Metasurfaces for sensing applications: Gas, bio and chemical," *Sensors*, vol. 22, no. 18, p. 6896, 2022.

[34] W. J. Padilla and R. D. Averitt, "Imaging with metamaterials," *Nature Reviews Physics*, vol. 4, no. 2, pp. 85–100, 2022.

[35] A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Planar photonics with metasurfaces," *Science*, vol. 339, no. 6125, p. 1232009, 2013.

[36] G. Yoon, I. Kim, and J. Rho, "Challenges in fabrication towards realization of practical metamaterials," *Microelectronic Engineering*, vol. 163, pp. 7–20, 2016.

[37] X. Zang, B. Yao, L. Chen, J. Xie, X. Guo, A. V. Balakin, A. P. Shkurinov, and S. Zhuang, "Metasurfaces for manipulating terahertz waves," *Light: Advanced Manufacturing*, vol. 2, no. 2, pp. 148–172, 2021.

[38] M. Keshavarz Hedayati and M. Elbahri, "Review of metasurface plasmonic structural color," *Plasmonics*, vol. 12, no. 5, pp. 1463–1479, 2017.

[39] J. Kim, J. Seong, Y. Yang, S.-W. Moon, T. Badloe, and J. Rho, "Tunable metasurfaces towards versatile metalenses and metaholograms: a review," *Advanced Photonics*, vol. 4, no. 2, pp. 024 001–024 001, 2022.

[40] R. Mittra, C. H. Chan, and T. Cwik, "Techniques for analyzing frequency selective surfaces-a review," *Proceedings of the IEEE*, vol. 76, no. 12, pp. 1593–1615, 1988.

[41] A. E. Minovich, A. E. Miroshnichenko, A. Y. Bykov, T. V. Murzina, D. N. Neshev, and Y. S. Kivshar, "Functional and nonlinear optical metasurfaces," *Laser & Photonics Reviews*, vol. 9, no. 2, pp. 195–213, 2015.

[42] J. Wang and J. Du, "Plasmonic and dielectric metasurfaces: design, fabrication and applications," *Applied Sciences*, vol. 6, no. 9, p. 239, 2016.

[43] Z. Bomzon, V. Kleiner, and E. Hasman, "Formation of radially and azimuthally polarized light using space-variant subwavelength metal stripe gratings," *Applied Physics Letters*, vol. 79, no. 11, pp. 1587–1589, 2001.

[44] N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, "Light propagation with phase discontinuities: generalized laws of reflection and refraction," *science*, vol. 334, no. 6054, pp. 333–337, 2011.

[45] A. M. Shaltout, N. Kinsey, J. Kim, R. Chandrasekar, J. C. Ndukaife, A. Boltasseva, and V. M. Shalaev, "Development of optical metasurfaces: emerging concepts and new materials," *Proceedings of the IEEE*, vol. 104, no. 12, pp. 2270–2287, 2016.

[46] H.-H. Hsiao, C. H. Chu, and D. P. Tsai, "Fundamentals and applications of metasurfaces," *Small Methods*, vol. 1, no. 4, p. 1600064, 2017.

[47] N. Yu and F. Capasso, "Flat optics with designer metasurfaces," *Nature materials*, vol. 13, no. 2, pp. 139–150, 2014.

[48] N. Meinzer, W. L. Barnes, and I. R. Hooper, "Plasmonic meta-atoms and metasurfaces," *Nature photonics*, vol. 8, no. 12, pp. 889–898, 2014.

[49] S. Jahani and Z. Jacob, "All-dielectric metamaterials," *Nature nanotechnology*, vol. 11, no. 1, pp. 23–36, 2016.

[50] C. F. Bohren and D. R. Huffman, *Absorption and scattering of light by small particles*. John Wiley & Sons, 2008.

[51] M. Decker, I. Staude, M. Falkner, J. Dominguez, D. N. Neshev, I. Brener, T. Pertsch, and Y. S. Kivshar, "High-efficiency dielectric huygens' surfaces," *Advanced Optical Materials*, vol. 3, no. 6, pp. 813–820, 2015.

[52] G. Yoon, T. Tanaka, T. Zentgraf, and J. Rho, "Recent progress on metasurfaces: applications and fabrication," *Journal of Physics D: Applied Physics*, vol. 54, no. 38, p. 383002, 2021.

[53] Y. S. Tan, H. Wang, H. Wang, C. Pan, and J. K. Yang, "High-throughput fabrication of large-scale metasurfaces using electron-beam lithography with su-8 gratings for multilevel security printing," *Photonics Research*, vol. 11, no. 3, pp. B103–B110, 2023.

[54] R. C. Devlin, M. Khorasaninejad, W. T. Chen, J. Oh, and F. Capasso, "Broadband high-efficiency dielectric metasurfaces for the visible spectrum," *Proceedings of the National Academy of Sciences*, vol. 113, no. 38, pp. 10 473–10 478, 2016.

[55] M. V. Gorkunov, O. Y. Rogov, A. V. Kondratov, V. V. Artemov, R. V. Gainutdinov, and A. A. Ezhov, "Chiral visible light metasurface patterned in monocrystalline silicon by focused ion beam," *Scientific reports*, vol. 8, no. 1, p. 11623, 2018.

[56] C. Lu and R. Lipson, "Interference lithography: a powerful tool for fabricating periodic structures," *Laser & Photonics Reviews*, vol. 4, no. 4, pp. 568–580, 2010.

[57] D. K. Oh, T. Lee, B. Ko, T. Badloe, J. G. Ok, and J. Rho, "Nanoimprint lithography for high-throughput fabrication of metasurfaces," *Frontiers of Optoelectronics*, vol. 14, pp. 229–251, 2021.

[58] W. T. Chen, A. Y. Zhu, and F. Capasso, "Flat optics with dispersion-engineered metasurfaces," *Nature Reviews Materials*, vol. 5, no. 8, pp. 604–620, 2020.

[59] P. Vabishchevich and Y. Kivshar, "Nonlinear photonics with metasurfaces," *Photonics Research*, vol. 11, no. 2, pp. B50–B64, 2023.

[60] G. Li, S. Zhang, and T. Zentgraf, "Nonlinear photonic metasurfaces," *Nature Reviews Materials*, vol. 2, no. 5, pp. 1–14, 2017.

[61] Z. Zhang, H. Liang, T. He, Z. Wang, and X. Cheng, "Photonic spin hall effect based on broadband high-efficiency reflective metasurfaces," *Applied Optics*, vol. 59, no. 5, pp. A63–A68, 2020.

[62] M. K. Hedayati, M. Javaherirahim, B. Mozooni, R. Abdelaziz, A. Tavassolizadeh, V. S. K. Chakravadhanula, V. Zaporojtchenko, T. Strunkus, F. Faupel, and M. Elbahri, "Design of a perfect black absorber at visible frequencies using plasmonic metamaterials," *Advanced Materials*, vol. 23, no. 45, pp. 5410–5414, 2011.

[63] A. Bhattacharya, R. Sarkar, and G. Kumar, "Toroidal electromagnetically induced transparency based meta-surfaces and its applications," *Iscience*, vol. 25, no. 1, p. 103708, 2022.

[64] P.-Y. Chen, C. Argyropoulos, and A. Alù, "Broadening the cloaking bandwidth with non-foster metasurfaces," *Physical review letters*, vol. 111, no. 23, p. 233001, 2013.

[65] K. Shastri and F. Monticone, "Nonlocal flat optics," *Nature Photonics*, vol. 17, no. 1, pp. 36–47, 2023.

[66] Z.-L. Deng and G. Li, "Metasurface optical holography," *Materials Today Physics*, vol. 3, pp. 16–32, 2017.

[67] D. Smirnova, S. Kruk, D. Leykam, E. Melik-Gaykazyan, D.-Y. Choi, and Y. Kivshar, "Third-harmonic generation in photonic topological metasurfaces," *Physical review letters*, vol. 123, no. 10, p. 103901, 2019.

[68] J. Wang and J. Du, "Metasurfaces for spatial light manipulation," *Metamaterials: Devices and Applications*, p. 57, 2017.

[69] B. Wang, B. Quan, J. He, Z. Xie, X. Wang, J. Li, Q. Kan, and Y. Zhang, "Wavelength de-multiplexing metasurface hologram," *Scientific Reports*, vol. 6, no. 1, p. 35657, 2016.

[70] X. Ni, S. Ishii, A. V. Kildishev, and V. M. Shalaev, "Ultra-thin, planar, babinet-inverted plasmonic metalenses," *Light: Science & Applications*, vol. 2, no. 4, pp. e72–e72, 2013.

[71] Q. Ma, Q. R. Hong, X. X. Gao, H. B. Jing, C. Liu, G. D. Bai, Q. Cheng, and T. J. Cui, "Smart sensing metasurface with self-defined functions in dual polarizations," *Nanophotonics*, vol. 9, no. 10, pp. 3271–3278, 2020.

[72] Z. Wei, Y. Cao, X. Su, Z. Gong, Y. Long, and H. Li, "Highly efficient beam steering with a transparent metasurface," *Optics express*, vol. 21, no. 9, pp. 10 739–10 745, 2013.

[73] A. A. Rifat, M. Rahmani, L. Xu, and A. E. Miroshnichenko, "Hybrid metasurface based tunable near-perfect absorber and plasmonic sensor," *Materials*, vol. 11, no. 7, p. 1091, 2018.

[74] H. Chu, Q. Li, B. Liu, J. Luo, S. Sun, Z. H. Hang, L. Zhou, and Y. Lai, "A hybrid invisibility cloak based on integration of transparent metasurfaces and zero-index materials," *Light: Science & Applications*, vol. 7, no. 1, p. 50, 2018.

[75] J. Yang, S. Gurung, S. Bej, P. Ni, and H. W. H. Lee, "Active optical metasurfaces: comprehensive review on physics, mechanisms, and prospective applications," *Reports on Progress in Physics*, vol. 85, no. 3, p. 036101, 2022.

[76] J. Hu, S. Bandyopadhyay, Y.-h. Liu, and L.-y. Shao, "A review on metasurface: from principle to smart metadevices," *Frontiers in Physics*, vol. 8, p. 586087, 2021.

[77] B. O. Asamoah, "Light and matter interactions on plasmonic platforms," Ph.D. dissertation, Itä-Suomen yliopisto, 2021.

[78] M.-X. Ren, W. Wu, W. Cai, B. Pi, X.-Z. Zhang, and J.-J. Xu, "Reconfigurable metasurfaces that enable light polarization control by light," *Light: Science & Applications*, vol. 6, no. 6, pp. e16 254–e16 254, 2017.

[79] W. Xu, T. Lu, and F. Wang, "Effects of interfacial properties on the ductility of polymer-supported metal films for flexible electronics," *International Journal of Solids and Structures*, vol. 47, no. 14-15, pp. 1830–1837, 2010.

[80] C. U. Hail, A.-K. U. Michel, D. Poulikakos, and H. Eghlidi, "Optical metasurfaces: evolving from passive to adaptive," *Advanced Optical Materials*, vol. 7, no. 14, p. 1801786, 2019.

[81] S. Wu, Q. Ze, R. Zhang, N. Hu, Y. Cheng, F. Yang, and R. Zhao, "Symmetry-breaking actuation mechanism for soft robotics and active metamaterials," *ACS applied materials & interfaces*, vol. 11, no. 44, pp. 41 649–41 658, 2019.

[82] Y. Ke, I. Balin, N. Wang, Q. Lu, A. I. Y. Tok, T. J. White, S. Magdassi, I. Abdulhalim, and Y. Long, "Two-dimensional sio2/vo2 photonic crystals with statically visible and dynamically infrared modulated for smart window deployment," *ACS applied materials & interfaces*, vol. 8, no. 48, pp. 33 112–33 120, 2016.

[83] M. K. Hedayati, M. Javaheri, A. U. Zillohu, H. J. El-Khozondar, M. S. Bawa'aneh, A. Lavrinenko, F. Faupel, and M. Elbahri, "Photo-driven super absorber as an active metamaterial with a tunable molecular-plasmonic coupling," *Advanced Optical Materials*, vol. 2, no. 8, pp. 705–710, 2014.

[84] J. Eaves-Rathert, E. Kovalik, C. F. Ugwu, B. R. Rogers, C. L. Pint, and J. G. Valentine, "Dynamic color tuning with electrochemically actuated tio2 metasurfaces," *Nano Letters*, vol. 22, no. 4, pp. 1626–1632, 2022.

[85] S. Chen, J. Chen, X. Zhang, Z.-Y. Li, and J. Li, "Kirigami/origami: unfolding the new regime of advanced 3d microfabrication/nanofabrication with "folding"," *Light: Science & Applications*, vol. 9, no. 1, p. 75, 2020.

[86] C. Zhang, J. Jing, Y. Wu, Y. Fan, W. Yang, S. Wang, Q. Song, and S. Xiao, "Stretchable all-dielectric metasurfaces with polarization-insensitive and full-spectrum response," *ACS nano*, vol. 14, no. 2, pp. 1418–1426, 2019.

[87] N. Jeon, J. Noh, C. Jung, and J. Rho, "Electrically tunable metasurfaces: from direct to indirect mechanisms," *New Journal of Physics*, 2022.

[88] C.-W. Lee, H. J. Choi, and H. Jeong, "Tunable metasurfaces for visible and swir applications," *Nano Convergence*, vol. 7, pp. 1–11, 2020.

[89] L. Deng, J. Teng, H. Liu, Q. Y. Wu, J. Tang, X. Zhang, S. A. Maier, K. P. Lim, C. Y. Ngo, S. F. Yoon *et al.*, "Direct optical tuning of the terahertz plasmonic response of insb subwavelength gratings," 2013.

[90] J. Li, S. Kamin, G. Zheng, F. Neubrech, S. Zhang, and N. Liu, "Addressable metasurfaces for dynamic holography and optical information encryption," *Science advances*, vol. 4, no. 6, p. eaar6768, 2018.

[91] G. Kafaie Shirmanesh, R. Sokhoyan, R. A. Pala, and H. A. Atwater, "Dual-gated active metasurface at 1550 nm with wide (> 300) phase tunability," *Nano letters*, vol. 18, no. 5, pp. 2957–2963, 2018.

[92] Y. Chen, X. Li, X. Luo, S. A. Maier, and M. Hong, "Tunable near-infrared plasmonic perfect absorber based on phase-change materials," *Photonics Research*, vol. 3, no. 3, pp. 54–57, 2015.

[93] S. Yu, J. Zhang, L. Li, H. Zheng, and W. Zhang, "Low loss, high tunable bazr0. 2ti0. 8o3/basn0. 85ti0. 15o3 heterostructure thin films," *Applied Physics Letters*, vol. 109, no. 14, 2016.

[94] B. Lee, J. Park, G. H. Han, H.-S. Ee, C. H. Naylor, W. Liu, A. C. Johnson, and R. Agarwal, "Fano resonance and spectrally modified photolumines-

cence enhancement in monolayer mos2 integrated with plasmonic nanoantenna array," *Nano letters*, vol. 15, no. 5, pp. 3646–3653, 2015.

[95] M. D. Dickey, "Stretchable and soft electronics using liquid metals," *Advanced materials*, vol. 29, no. 27, p. 1606425, 2017.

[96] M. Vosgueritchian, D. J. Lipomi, and Z. Bao, "Highly conductive and transparent pedot: Pss films with a fluorosurfactant for stretchable and flexible transparent electrodes," *Advanced functional materials*, vol. 22, no. 2, pp. 421–428, 2012.

[97] A. Nemati, Q. Wang, M. Hong, and J. Teng, "Tunable and reconfigurable metasurfaces and metadevices," *Opto-Electronic Advances*, vol. 1, no. 5, p. 180009, 2018.

[98] N. Raeis-Hosseini and J. Rho, "Metasurfaces based on phase-change material as a reconfigurable platform for multifunctional devices," *Materials*, vol. 10, no. 9, p. 1046, 2017.

[99] L. Zou, M. Cryan, and M. Klemm, "Phase change material based tunable reflectarray for free-space optical inter/intra chip interconnects," *Optics express*, vol. 22, no. 20, pp. 24 142–24 148, 2014.

[100] B. Gholipour, J. Zhang, K. F. MacDonald, D. W. Hewak, and N. I. Zheludev, "An all-optical, non-volatile, bidirectional, phase-change meta-switch," *Advanced materials*, vol. 25, no. 22, pp. 3050–3054, 2013.

[101] Y. Chen, T.-S. Kao, B. Ng, X. Li, X. Luo, B. Luk'Yanchuk, S. Maier, and M. Hong, "Hybrid phase-change plasmonic crystals for active tuning of lattice resonances," *Optics express*, vol. 21, no. 11, pp. 13 691–13 698, 2013.

[102] Q. Wang, E. T. Rogers, B. Gholipour, C.-M. Wang, G. Yuan, J. Teng, and N. I. Zheludev, "Optically reconfigurable metasurfaces and photonic devices based on phase change materials," *Nature photonics*, vol. 10, no. 1, pp. 60–65, 2016.

[103] Z. Zhu, P. G. Evans, R. F. Haglund Jr, and J. G. Valentine, "Dynamically reconfigurable metadevice employing nanostructured phase-change materials," *Nano letters*, vol. 17, no. 8, pp. 4881–4885, 2017.

[104] A. Howes, W. Wang, I. Kravchenko, and J. Valentine, "Dynamic transmission control based on all-dielectric huygens metasurfaces," *Optica*, vol. 5, no. 7, pp. 787–792, 2018.

[105] R. Sokhoyan, P. Thureja, J. Sisler, M. Grajower, K. Shayegan, E. Feigenbaum, S. Elhadj, and H. A. Atwater, "Electrically tunable conducting oxide metasurfaces for high power applications," *Nanophotonics*, vol. 12, no. 2, pp. 239–253, 2023.

[106] A. She, S. Zhang, S. Shian, D. R. Clarke, and F. Capasso, "Adaptive metalenses with simultaneous electrical control of focal length, astigmatism, and shift," *Science advances*, vol. 4, no. 2, p. eaap9957, 2018.

[107] S. C. Malek, H.-S. Ee, and R. Agarwal, "Strain multiplexed metasurface holograms on a stretchable substrate," *Nano letters*, vol. 17, no. 6, pp. 3641–3645, 2017.

[108] M. L. Tseng, J. Yang, M. Semmlinger, C. Zhang, P. Nordlander, and N. J. Halas, "Two-dimensional active tuning of an aluminum plasmonic array for full-spectrum response," *Nano letters*, vol. 17, no. 10, pp. 6034–6039, 2017.

[109] J. B. Reeves, R. K. Jayne, T. J. Stark, L. K. Barrett, A. E. White, and D. J. Bishop, "Tunable infrared metasurface on a soft polymer scaffold," *Nano Letters*, vol. 18, no. 5, pp. 2802–2806, 2018.

[110] Z. Zhai, L. Wu, and H. Jiang, "Mechanical metamaterials based on origami and kirigami," *Applied Physics Reviews*, vol. 8, no. 4, 2021.

[111] Z. Wang, L. Jing, K. Yao, Y. Yang, B. Zheng, C. M. Soukoulis, H. Chen, and Y. Liu, "Origami-based reconfigurable metamaterials for tunable chirality," *Advanced materials*, vol. 29, no. 27, p. 1700412, 2017.

[112] X. Liu, Z. Huang, C. Zhu, L. Wang, and J. Zang, "Out-of-plane designed soft metasurface for tunable surface plasmon polariton," *Nano Letters*, vol. 18, no. 2, pp. 1435–1441, 2018.

[113] Z. Yan, F. Zhang, J. Wang, F. Liu, X. Guo, K. Nan, Q. Lin, M. Gao, D. Xiao, Y. Shi *et al.*, "Controlled mechanical buckling for origami-inspired construction of 3d microstructures in advanced materials," *Advanced functional materials*, vol. 26, no. 16, pp. 2629–2639, 2016.

[114] S. Abdollahramezani, O. Hemmatyar, M. Taghinejad, H. Taghinejad, Y. Kiarashinejad, M. Zandehshahvar, T. Fan, S. Deshmukh, A. A. Eftekhar, W. Cai *et al.*, "Dynamic hybrid metasurfaces," *Nano letters*, vol. 21, no. 3, pp. 1238–1245, 2021.

[115] D. Herle, A. Kiselev, L. G. Villanueva, O. J. Martin, and N. Quack, "Broadband mechanically tunable metasurface reflectivity modulator in the visible spectrum," *ACS photonics*, vol. 10, no. 6, pp. 1882–1889, 2023.

[116] M. Li, W. Chen, P. Liu, F. Xu, and Y. Chen, "Analysis of mechanically tunable metasurfaces for identifying multiple strains," *Optics Communications*, vol. 554, p. 130201, 2024.

[117] A. Ghasemi, R. Fang, D. A. Zeze, and M. K. Hedayati, "Ultra-stretchable active metasurfaces for high-performance structural color," *AIP Advances*, vol. 13, no. 8, 2023.

[118] X. Fan, Y. Li, S. Chen, Y. Xing, and T. Pan, "Mechanical terahertz modulation by skin-like ultrathin stretchable metasurface," *Small*, vol. 16, no. 37, p. 2002484, 2020.

[119] S. Song, X. Ma, M. Pu, X. Li, K. Liu, P. Gao, Z. Zhao, Y. Wang, C. Wang, and X. Luo, "Actively tunable structural color rendering with tensile substrate," *Advanced Optical Materials*, vol. 5, no. 9, p. 1600829, 2017.

[120] Z. Xu and Y.-S. Lin, "A stretchable terahertz parabolic-shaped metamaterial," *Advanced Optical Materials*, vol. 7, no. 19, p. 1900379, 2019.

[121] J. W. Morris Jr, C. Krenn, D. Roundy, and M. L. Cohen, "Elastic stability and the limit of strength," 2002.

[122] S. S. Brenner, "Tensile strength of whiskers," *Journal of applied physics*, vol. 27, no. 12, pp. 1484–1491, 1956.

[123] W. Robinson and S. Truman, "Stress-strain curve for aluminium from a continuous indentation test," *Journal of Materials Science*, vol. 12, pp. 1961–1965, 1977.

[124] S. Xu, T. Odaira, S. Sato, X. Xu, T. Omori, S. Harjo, T. Kawasaki, H. Seiner, K. Zoubková, Y. Murakami *et al.*, "Non-hookean large elastic deformation in bulk crystalline metals," *Nature communications*, vol. 13, no. 1, p. 5307, 2022.

[125] G. Pearson, W. Read Jr, and W. L. Feldmann, "Deformation and fracture of small silicon crystals," *Acta metallurgica*, vol. 5, no. 4, pp. 181–191, 1957.

[126] T. Li, Z. Huang, Z. Xi, S. P. Lacour, S. Wagner, and Z. Suo, "Delocalizing strain in a thin metal film on a polymer substrate," *Mechanics of Materials*, vol. 37, no. 2-3, pp. 261–273, 2005.

[127] T. Badloe, J. Lee, J. Seong, and J. Rho, "Tunable metasurfaces: the path to fully active nanophotonics," *Advanced Photonics Research*, vol. 2, no. 9, p. 2000205, 2021.

[128] T. Baëtens, E. Pallecchi, V. Thomy, and S. Arscott, "Cracking effects in squashable and stretchable thin metal films on pdms for flexible microsystems and electronics," *Scientific reports*, vol. 8, no. 1, p. 9492, 2018.

[129] H.-S. Ee and R. Agarwal, "Tunable metasurface and flat optical zoom lens on a stretchable substrate," *Nano letters*, vol. 16, no. 4, pp. 2818–2823, 2016.

[130] S. Olcum, A. Kocabas, G. Ertas, A. Atalar, and A. Aydinli, "Tunable surface plasmon resonance on an elastomeric substrate," *Optics express*, vol. 17, no. 10, pp. 8542–8547, 2009.

[131] X. Zhang, Y. Zhou, H. Zheng, A. E. Linares, F. C. Ugwu, D. Li, H.-B. Sun, B. Bai, and J. G. Valentine, "Reconfigurable metasurface for image processing," *Nano Letters*, vol. 21, no. 20, pp. 8715–8722, 2021.

[132] X. Zhu, S. Xiao, L. Shi, X. Liu, J. Zi, O. Hansen, and N. A. Mortensen, "A stretch-tunable plasmonic structure with a polarization-dependent response," *Optics Express*, vol. 20, no. 5, pp. 5237–5242, 2012.

[133] Y.-L. Chiang, C.-W. Chen, C.-H. Wang, C.-Y. Hsieh, Y.-T. Chen, H.-Y. Shih, and Y.-F. Chen, "Mechanically tunable surface plasmon resonance based on gold nanoparticles and elastic membrane polydimethylsiloxane composite," *Applied Physics Letters*, vol. 96, no. 4, 2010.

Chapter 3

Machine Learning in Nanophotonics

3.1 Chapter Overview

This chapter explores the effectiveness of artificial intelligence (AI) in the design of photonic materials, devices, and structures where conventional design methods prove impractical or ineffective. The chapter starts with a general introduction to the significance of AI in nanophotonics before outlining the details of the underpinning concept of machine learning (ML) models and advancements in deep learning (DL). Subsequently, model architectures commonly used in the design of metamaterials, plasmonic nanostructures, and tunable metasurfaces are discussed. These include the fundamental multilayer perceptron (MLP), sophisticated deep neural network (DNN), and hybrid models with additional optimisation techniques. Finally, the chapter discusses the current limitations of this mul-

tidisciplinary field, which has the potential to revolutionise research and practical applications in nanophotonics.

3.2 Conventional Design Approaches

The prediction of light-matter interactions is critically dependent on the design of nanophotonics devices. A precise design is essential whether they are simple like a two-layer thin film or complicated like three-dimensional (3D) metamaterials and photonic crystals made up of rows of metallic or dielectric building blocks. This process demands a comprehensive consideration of various parameters, including material properties, geometric configurations, and even fabrication constraints. Precise control over the composition of the material and the fabrication parameters is usually required to adjust its key properties, and this process can be time-consuming, expensive and difficult. However, optimising the vast topological design space of nanophotonic structures offers a variety of features without the need to directly alter the material composition. Moreover, one of the long-standing challenges in nanophotonics is understanding how best to configure both material properties and nano- or micro-scale structures to achieve custom functionalities like reflectance and transmittance [1].

The nanophotonics structures have been designed and optimised using two traditional methods. The first approach is **"physics-based methods"** which comprise scientific intuition, data from prior or related procedures, and generalised analytical models [2, 3]. For instance, the Mie theory accurately describes dielectric and metallic nanoparticles with basic shapes like core-shell structures or spherical nanoparticles. Multipolar resonances, both magnetic and electric, are used to calculate their scattering, absorption, and extinction responses [4]. The concept of split-ring resonators, proposed by John Pendry in the 1990s [5], was built upon the principles of electrical circuit theory and electromagnetics. To do this, changing the external magnetic fields creates a magnetic dipole (MD) and a current loop, which are then amplified around the resonance frequency set by the resonator's properties like inductance and internal capacitance [6]. Photonic crystals (PCs)

are another example of physics-based design methods motivated by comprehending the movement of electrons in the solid state periodic potential [7]. PCs allow the creation of 'photonic semiconductors' with complete bandgaps, preventing the propagation of light in the 'forbidden band' [7]. Despite providing essential guidelines, finding the appropriate structures for desired photonic properties becomes important, particularly as the geometry and spatial arrangements become more complex.

The second approach involves electromagnetic (EM) modelling through **"numerical simulation methods"**, using various commercial software tools like COMSOL Multiphysics, Lumerical solutions, ANSOFT, and CST Microwave Studio [8]. Typically, these computational techniques address the design problem by discretising Maxwell's equations spatially and temporally, starting from specific initial boundary conditions [8]. The optical characteristics of a given structure can be precisely calculated by arranging enough meshes and iteration steps. Using numerical algorithms like finite-difference time-domain (FDTD) and the finite element method (FEM) facilitates finding the optimised structure and predicts how light reacts to different conditions [9, 10]. However, achieving the desired responses often requires fine-tuning the geometry through iterative simulations. Only a restricted set of design parameters can be modified to discover the ideal structure thanks to simulation power and the time-consuming nature of the process [11]. For instance, in the early stage of a metasurface design, a unit cell model with periodic boundary conditions (PBCs) is simulated. The simulated data library is then used to help design the whole metasurface device by giving the phase or transmission responses for different structure parameters.

Parametric sweeps in the simulation process are often used to discover the importance of design parameters and optimise the nanophotonic devices accordingly. These are guided by physical models for classical functions. However, parametric sweeps may not be sufficient when dealing with extremely complex device functions or a large design space for random shapes of structures. Furthermore, the optimisation process with current numerical design techniques needs previous knowledge of the particular combination of materials and key design parameters

in the studied structure [12]. Therefore, a single, integrated solution that concurrently optimises across geometries and materials has not yet been achieved. Furthermore, the inverse design process is significantly more difficult since it includes directly retrieving the right structure and materials for the target optical performance, which involves exploring a far greater degree of freedom (DOF) in the design process [11].

The advent of AI techniques represents a paradigm shift in the design of nanophotonic devices. ML, a subset of AI, can learn and adapt from data, enabling a more efficient exploration of vast design spaces. ML techniques facilitate navigating complex patterns and nonlinear relationships within nanostructures and optical performance. This data-driven approach allows for the discovery of novel designs and optimisation strategies that may be challenging to achieve through discussed conventional methods [13].

3.3 Fundamentals of Machine Learning

ML is a subset of AI that empowers computers to learn from data, identify patterns, and make decisions with minimal human intervention. Its applications span across a multitude of domains such as industrial production, data security, education, medical diagnosis, transport, and economic analysis [14]. Unlike traditional rule-based programming, where explicit instructions dictate how a system should behave, ML systems learn directly from data. This data-driven approach allows machines to improve their performance over time without being explicitly programmed for every scenario. By recognising patterns and extracting insights from vast datasets, ML algorithms can uncover hidden relationships, predict future outcomes, and automate complex tasks with remarkable accuracy. Moreover, the model can make insightful predictions on data that has not been seen before [15]. During the model's training or use in ML, a deep understanding of the specific problem's physical or mathematical intricacies is often beneficial but not always required. This lack of requirement makes ML methods versatile and suitable for various tasks such as classification, regression, clustering, and structured predic-

tion. Generally, a greater quantity and higher quality of data lead to more precise predictions. Although collecting and training with a large dataset demands considerable time and effort, the subsequent use of the model proves to be convenient and fast [16].

The history of ML can be linked to the 1940s when Donald Hebb explored neural cell interaction in his book "The Organisation of Behavior" [17]. In the 1950s, IBM researcher Arthur Samuel pioneered ML in gaming by creating software for checkers [18]. He introduced a scoring function based on piece placements, employing a minimax approach, and termed his innovative method "rote learning". He proposed the term "machine learning" for the first time in 1952. The perceptron, combining Hebb's and Samuel's ideas, was created in 1957 but faced challenges in image recognition, leading to an "AI winter" [19]. The second wave of research on ML in the 1980s led to the introduction of artificial neural network (ANN). A significant milestone was the paper published by Rumelhart, Hinton, and Williams in 1986, modifying the back-propagation algorithm [20]. Geoffrey Hinton's work in 2006 turned DL into a modern era, demonstrating the effectiveness of training DNN [21]. The popularity of DL has significantly increased as researchers have trained networks with greater depth, highlighting the crucial role of model depth. Driven by the increasing volume of data, DL is currently the dominant focus of AI research, regularly surpassing prior models and even beating human performance.

3.3.1 Classification of Machine Learning Algorithms

ML algorithms can be classified based on the training approach into three main categories as described in the following parts.

Supervised Learning

In this approach, the algorithm is presented with input examples, each accompanied by an associated output label, forming a set of training data. Supervised learning algorithms aim to learn patterns, relationships, or mappings between the input features and the corresponding output labels [22].

During the training process, the algorithm iteratively adjusts its internal parameters, guided by the labelled examples. The adjustment is performed to minimise the difference, or error, between its predictions and the actual output labels. Through this iterative refinement, the algorithm endeavours to capture the underlying patterns within the data, enabling it to make accurate predictions on new, unseen instances [22].

Supervised learning comprises of various tasks, with two prominent categories being classification and regression. In terms of classification tasks, the algorithm learns to categorise input data into predefined classes or labels. For instance, it might distinguish between spam and non-spam emails or identify objects in images. On the other hand, regression tasks involve predicting numerical values, such as estimating house prices based on relevant features. The advantages of supervised learning lie in its ability to generalise from the labelled training data, allowing the model to make informed predictions on new, previously unseen data. This method has broad applications in a variety of fields, such as financial forecasting, picture recognition, medical diagnosis, and natural language processing [23].

In nanophotonics, supervised learning has been employed for tasks such as material selection and topology optimisation, where a model was trained on labelled data corresponding to the optical responses of various nanophotonic materials [24].

Unsupervised Learning

Unsupervised learning tasks involve extracting patterns, structures, or relationships within the data without explicit guidance in the form of labelled outputs [25]. In unsupervised learning, the algorithm explores the inherent characteristics of the data to uncover hidden structures or groupings. The absence of labelled outputs challenges the algorithm to identify intrinsic patterns independently. This approach is particularly valuable in scenarios where obtaining labelled data is impractical or costly [25].

Unsupervised learning, known for its versatility, finds applications in diverse re-

search areas. It excels in clustering similar data points, detecting anomalies, discovering associations between variables, and reducing dimensionality for enhanced data visualisation [26]. Unsupervised learning is also used in neural networks, particularly in learning efficient data representations through models like autoencoders. Additionally, K-Means Clustering is widely used for partitioning data points into distinct clusters based on similarities [27].

Unsupervised learning has been widely used in nanophotonics for inverse design problems, clustering meta-atom structures, and geometry optimisation [28, 29]. Unsupervised learning is also used to minimise a loss function in the implementation of generative neural network (NN)s like variational autoencoders (VAEs) and generative adversarial networks (GANs) in the nanophotonics optimisation process. These approaches discover new patterns from a massive dataset and use that information to generate new structures with better attributes [30, 31].

Semi-supervised learning is a technique which benefits from both supervised and unsupervised learning. This algorithm is valuable in situations where generating a sufficient amount of labelled data is expensive. In this approach, an algorithm combines both labelled and unlabeled datasets. The smaller labelled dataset is used to guide the classification process, providing explicit examples for the algorithm to learn from. Simultaneously, the algorithm is trained by the larger unlabeled dataset to extract features and identify patterns, enhancing its ability to generalise and make accurate predictions on new data [32].

Reinforcement Learning

Reinforcement learning (RL) is the process through which an agent gains decision-making skills via interaction with its surroundings. Reward or penalty points are given to the agent based on activities taken in the environment. The agent's objective is to learn a strategy or policy, that maximises the cumulative reward over time [33].

A classical example of RL is training an autonomous robot to navigate a maze. In this scenario, the maze environment is the setting, and the states are the various positions within the maze. The robot, acting as the agent, can take actions

like moving in different directions (left, right, forward, and backward). As the robot explores the maze, it receives feedback based on its actions. For instance, successfully reaching the exit might yield a positive reward, while hitting a wall could result in a negative penalty. The goal is for the robot to learn a strategy for deciding its actions in different states to efficiently navigate the maze and maximise its cumulative reward [34]. RL has been widely used in the optimisation of metasurfaces. For example, Seo et al. [35] proposed a global optimisation process for a Si-based beam deflector with a design space as large as $\approx 10^{17}$.

3.4 Deep Learning

ML has advanced by developing classical algorithms like support vector machines, decision trees, and Bayesian classifiers, and DL (cf. Figure 3.1). Among these, DL, particularly based on multilayer ANN, has found successful applications in various fields such as voice recognition, machine translation, and autonomous driving [36].

In biological NNs, neurons are the fundamental units connected to each other [37]. When a neuron's electrical potential surpasses a threshold, it activates and transmits chemicals to connected neurons, altering their electrical potential as shown in Figure3.2 (a). ANN is a mathematical model that mimics the behaviour of biological NN, carrying out information processing in a distributed parallel manner. NN models involve connected neurons sending input signals to neurons, which are then conveyed via weighted connections and added to the total input value [3]. This total value is then compared with the neuron's threshold and processed through an activation function to generate the neuron's output [37]. Most common DL models can be classified into three main types: ANN, convolutional neural network (CNN), and recurrent neural network (RNN).

3.4.1 Artificial Neural Networks

As shown in Figure 3.2 (b), an ANN typically consists of an input layer, an output layer, and multiple hidden layers [38]. The input layer neurons receive external

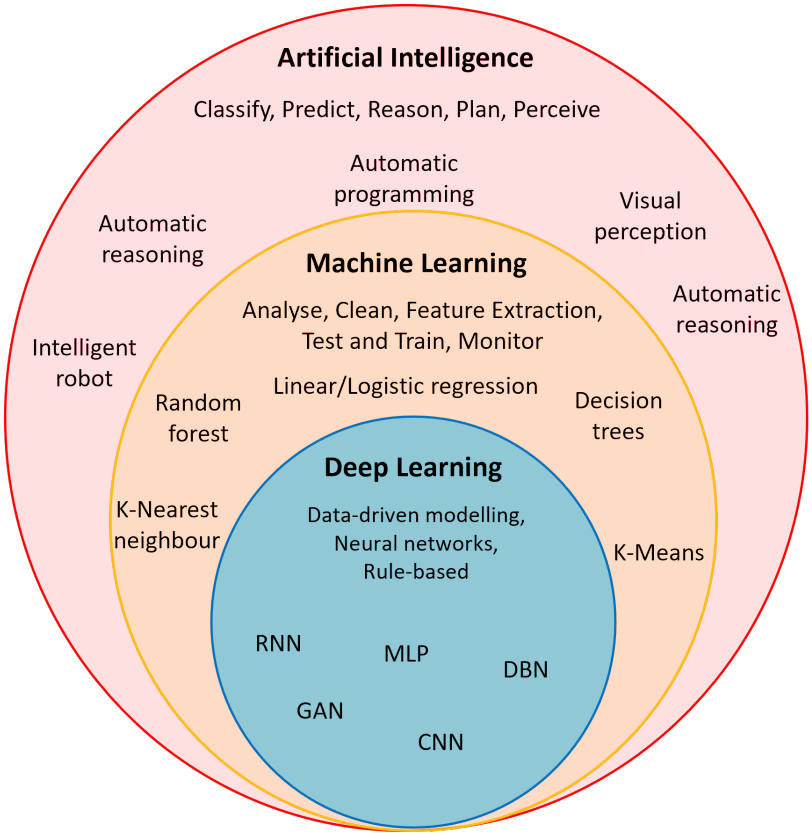


Figure 3.1: The relationship between AI, ML, and DL.

data, which undergoes processing through the hidden layers and then output ones. The final prediction value is produced by the output layer's neurons. The training process involves applying a substantial amount of data to iteratively adjust the weights of connections between neurons [3]. Several methods like the error back-propagation algorithm are used in this continuous adjustment until the neural network's output aligns with the target values. Following training, the ANN can accurately map inputs to the desired outputs. Figure 3.2 (b) illustrates the basic architecture of an ANN featuring two hidden layers.

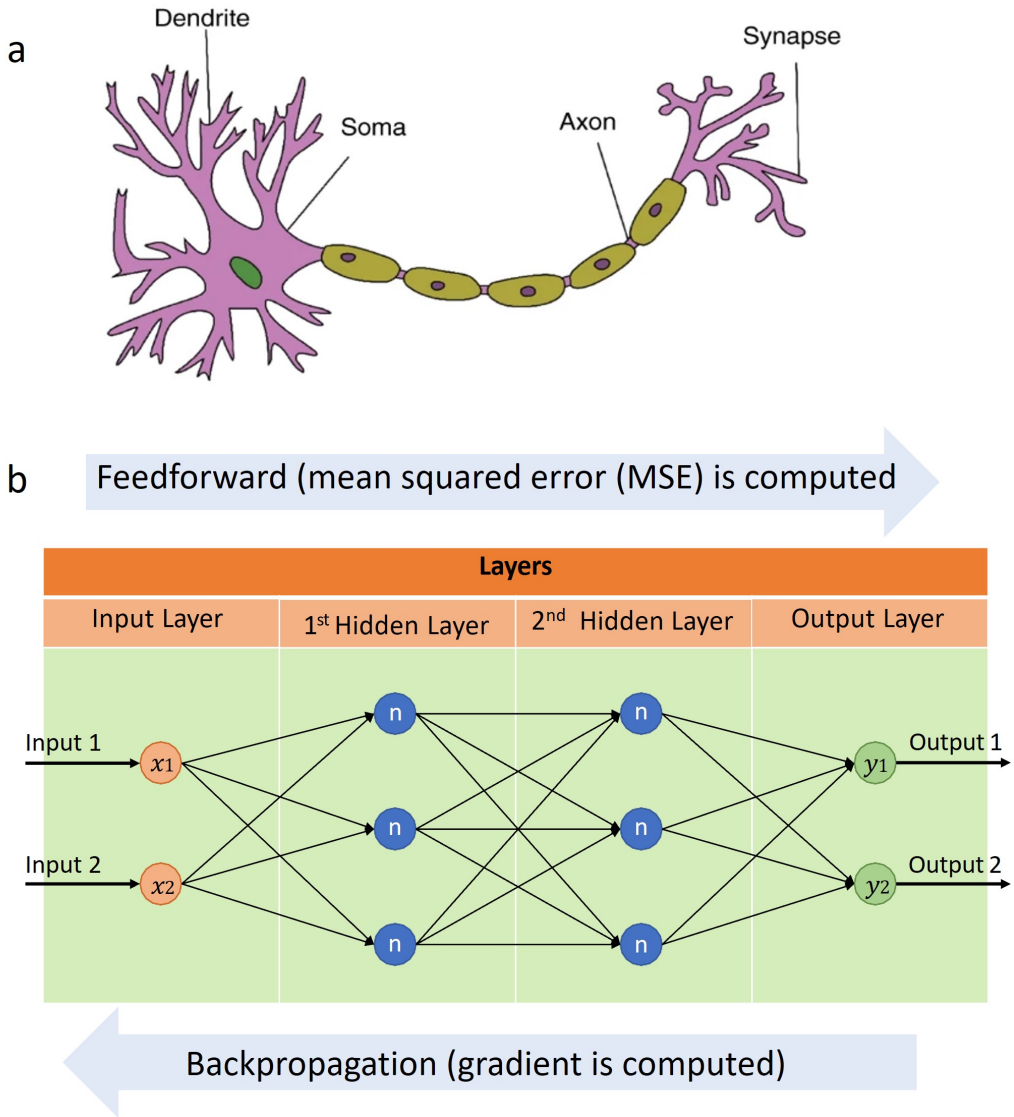


Figure 3.2: (a) A schematic of biological neurons, (b) a diagrammatic representation of a ANN consisting of two hidden layers.

Often referred to as feedforward neural networks, ANN processes inputs exclusively in the forward direction. The network's layers consist of nodes, each designed to emulate the behaviour of neurons in the brain (cf. Figure 3.2 (a)). The primary neurons of the brain are networked and communicate with one another through electrical impulses to support mental processes. Similarly, in ANN models, these neurons serve as nodes facilitating the flow of data and computations. These nodes receive input signals, originating from either the raw dataset or neurons in the preceding layer of the neural network. For each signal, the neuron calculates the sum after multiplying it by its corresponding weight and then

passes the result through an activation function. This output is subsequently transmitted to the next layer via a feedforward algorithm, as depicted in Figure 3.2 (b). Equation 3.4.1 elucidates the artificial neuron depicted in Figure 3.3, where θ represents a linear or nonlinear activation function.

$$x_j^{(l)} = \theta(s_j^{(l)}) = \theta(\omega_{0,j} + \sum_{i=1}^{n} \omega_{i,j}^{l} x_j^{(l-1)}) \tag{3.4.1}$$

In this equation, $x_i^{(l-1)}$ and $x_j^{(l)}$ denote the ith node at layer $(l-1)$ and the jth node at layer (l), respectively. The weight on the connecting node i at layer $(l-1)$ to node j at layer (l) is denoted as $\omega_{i,j}$. $\omega_{0,j}$ is the bias term for node j. Nonlinear activation functions, such as hyperbolic tangent, softmax, rectified linear unit (ReLU), and sigmoid are applied in this context. These functions enable the NN to generalise and adapt to diverse data, distinguishing between different output scenarios. ANN is recognised as a universal function approximator because it possesses the capability to learn any nonlinear function, employing a range of activation functions [38].

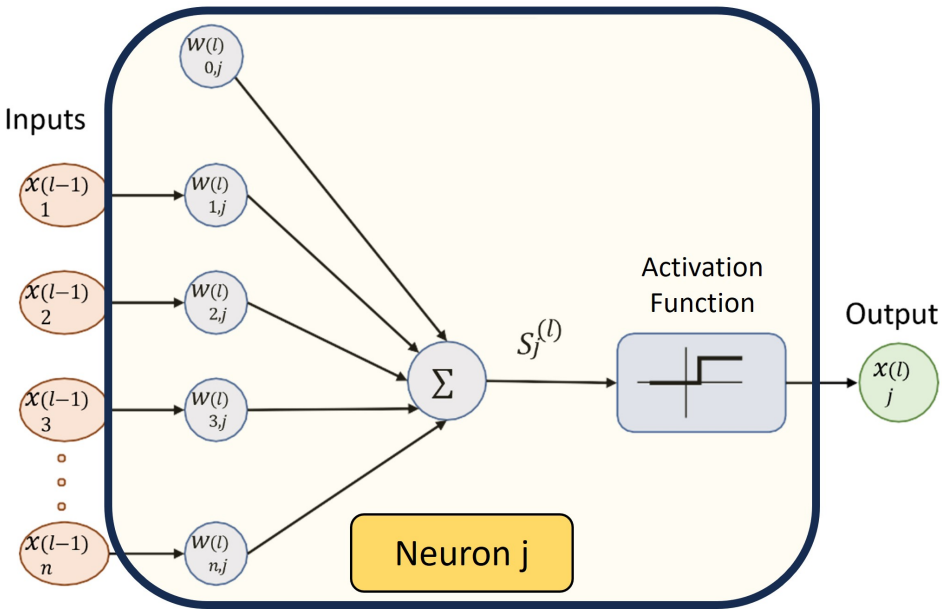


Figure 3.3: A visual depiction of an artificial neuron.

3.4.2 Convolutional Neural Networks

CNNs form a specific class of NN models designed for the analysis of two-dimensional (2D) image data. Nevertheless, they can be adjusted for applications involving one-dimensional and 3D data. The core element of CNN is represented by the convolutional layers, where a filter kernel systematically applies to input data using a simple convolution process. The mathematical expression for 2D convolution is given by equation 3.4.2, where $g(x,y)$ represents the filtered image, $f(x,y)$ is the original image, and ω denotes the filter kernel.

$$g(x,y) = \omega * f(x,y) = \sum_{d_x=-a}^{a} \sum_{d_y=-b}^{b} \omega(d_x,d_y) f(x-d_x, y-d_y) \tag{3.4.2}$$

Each filter kernel's componnet is taken into account within the ranges -a < d_x < a and -b < d_y < b. A feature map is produced by repeatedly applying the same filter—one that is smaller than the original image—to the input data. This process is visually depicted in Figure3.4. The resulting feature map indicates both the locations and significance of detected features within the input image. CNNs possess the distinctive capability to autonomously learn multiple filters tailored to the training dataset and the specific problem like image classification [38].

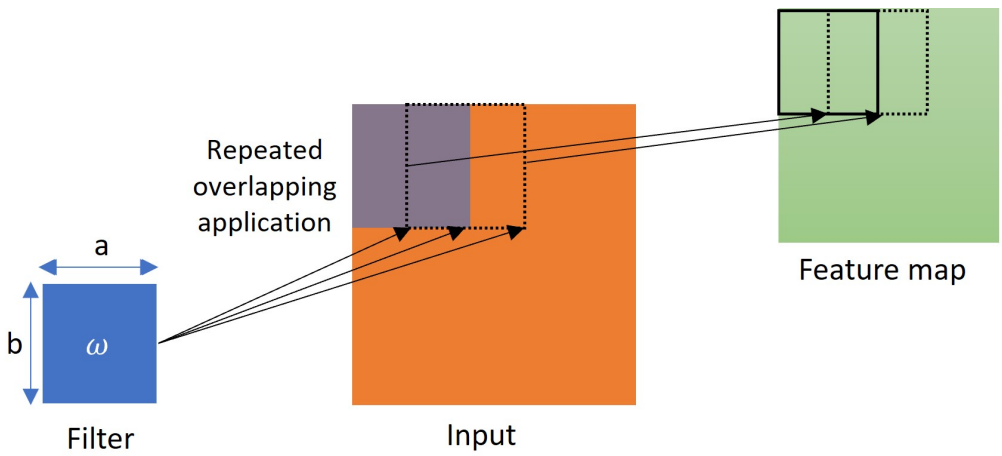


Figure 3.4: A schematic of a filter-based convolution process.

3.4.3 Recurrent Neural Networks

RNNs were proposed to address the limitations of ANNs and CNNs, which are only capable of processing current inputs and cannot discern sequential patterns in data or retain the memory of past inputs [39]. RNNs operate by maintaining internal memory, making them appropriate for applications such as natural language processing and time series analysis. There is a conventional forward path in which the desired output is connected to the original inputs. A feedback loop, on the other hand, also exists and returns each layer's output to its input. By doing this, it enhances the predictive outcomes for the subsequent probable situation [40].

3.4.4 Activation Function

Activation functions determine the activation status of a neuron by assessing the weighted sum and incorporating a bias (cf. Fig 3.3). These functions are differentiable operators, introducing non-linearity to the model and transforming input signals into corresponding outputs. The training convergence speed and total network performance in a given task are significantly influenced by the choice of activation function. Nonlinear, differentiable activation functions are commonly used due to their adaptability to complex datasets and their compatibility with backpropagation for NN training. Some common activation functions, including Sigmoid, tanh, ReLU [41], and hyperbolic tangent are briefly outlined for reference.

Sigmoid Function

The sigmoid function, usually referred to as the logistic function, is a commonly employed activation function in NN. It maps any real-valued number onto a range spanning from 0 to 1. The sigmoid function can be represented by the following mathematical expression:

$$\theta(z) = \frac{1}{(1+e^{-z})} \tag{3.4.3}$$

where e represents the base of the natural logarithm, and z is the input to the function. As shown in Figure 3.5 (a), the sigmoid function is S-shaped and has the characteristic of mapping large positive or negative inputs to values close to 1 or 0, respectively. It is useful in binary classification problems, where the output needs to be interpreted as a probability [42].

Rectified Linear Unit Function

ReLU creates non-linearity by directly outputting positive input; otherwise, it outputs zero as shown in Figure 3.5 (b). The mathematical expression for the ReLU function is defined as:

$$\theta(z) = max(0,z) = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{if } z \geq n \end{cases} \tag{3.4.4}$$

where z represents the input to the function. The ReLU functions help address the vanishing gradient problem associated with some other activation functions [42].

Leaky Rectified Linear Unit Function

An alternative to the ReLU function is the LeakyReLU activation function, denoted as $\theta(z)$ = max(α, z), where $\theta(z)$ represents the output. In LeakyReLU, a little subset of the input, defined by the parameter, is mapped to negative values 0 < α < 1 (e.g., α = 0.1), as illustrated in Figure 3.5 (c). Unlike ReLU, LeakyReLU allows neurons with zero input to produce a nonzero output, enabling them to contribute nonzero gradients during weight updates in training. Although the use of LeakyReLU may sacrifice convergence speed, it permits more neurons in the network to contribute, potentially enhancing overall performance.

Hyperbolic tangent (tanh) Function

The hyperbolic tangent function, often denoted as tanh, is another common activation function in the training process of a NN. Like the sigmoid function, tanh squashes input values to a specific range, but in this case, it ranges between -1 and

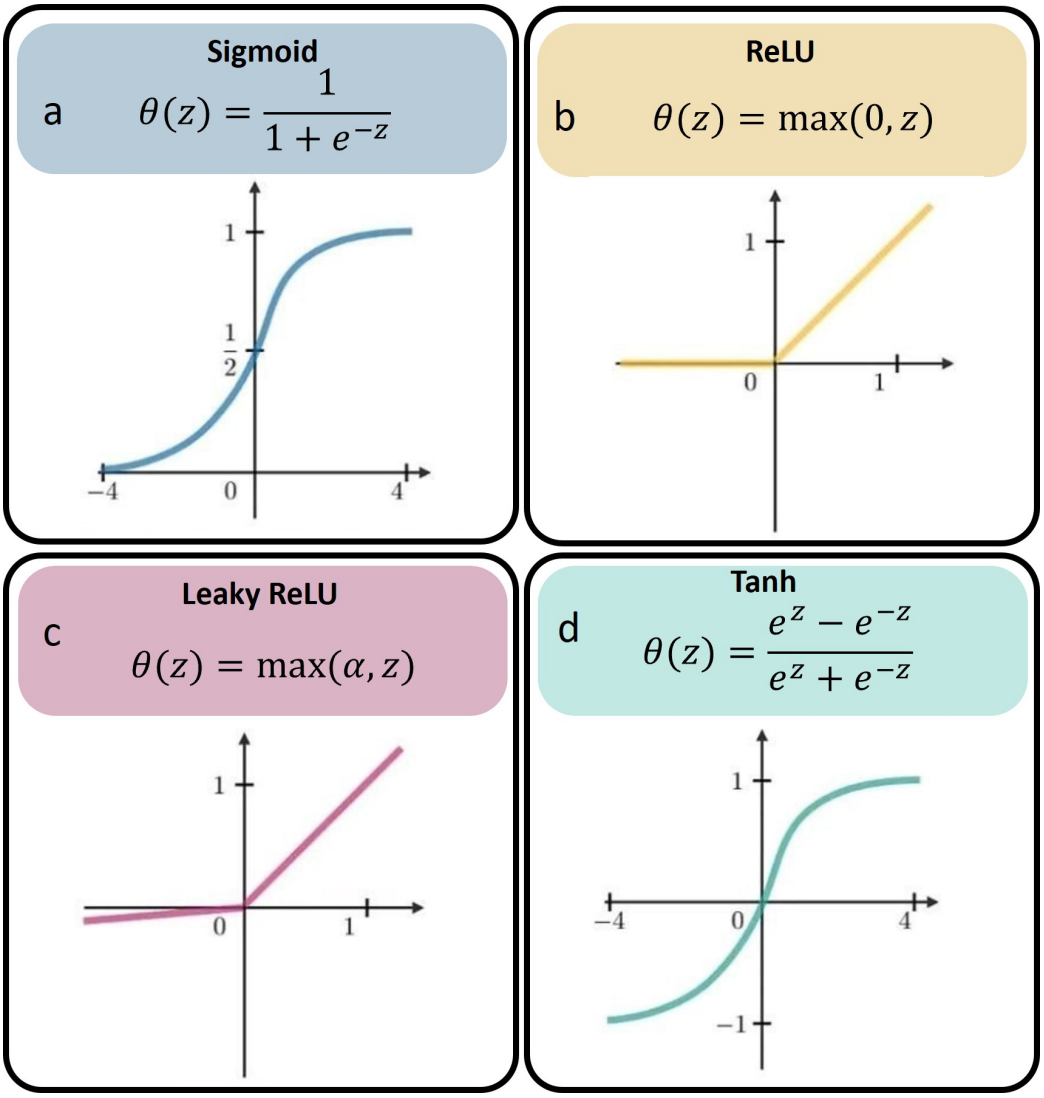


Figure 3.5: Common activation functions used in neural networks: (a) sigmoid, (b) ReLU, and (c) LeakyReLU, (d) tanh.

1 (cf. Figure 3.5 (d)). The mathematical expression for the hyperbolic tangent function is:

$$\theta(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \tag{3.4.5}$$

Where z is the input to the function. The tanh function is zero-centred, meaning that its output has a mean close to zero. This can be advantageous for certain optimisation algorithms during training [42].

Similar to the sigmoid function, the tanh functions suffer from the vanishing gradient problem for very large positive or negative inputs. However, it generally has the advantage of producing outputs in the range of -1 to 1, making it more conve-

nient for training neural networks that might benefit from values centred around zero [42].

3.4.5 Neural Network Training

Following the input of training data into the NN in the feedforward direction, the network generates an initial output prediction. This predicted output initially deviates significantly from the target labels in the training dataset due to the unadjusted weights of the network. To align the predicted result with the real values, the network receives feedback, leading to an iterative process for updating its weights [40]. A widely adopted strategy for training neural networks involves backpropagation. This method entails propagating the gradients of a defined loss function concerning the network's weights in the backward direction. As shown in Figure 3.2 (b), the **"gradient backward propagation"** process enables the network to systematically adjust its weights to minimise the overall loss. However, diverse methods exist for training NNs, offering alternative approaches to weight adjustment. Difference target propagation [20] involves modifying the network's weights based on the disparities between the predicted and target outputs. Decoupled neural interfaces using synthetic gradients take a unique route by independently training specific parts of the network, reducing reliance on direct gradient backpropagation [38].

Backpropagation

To understand the adjusting process, let's break down the backpropagation process step by step, using a simple NN with one hidden layer as an example. The network has an input layer (X), a hidden layer (H), and an output layer (Y). The error between the predicted output and the actual output is represented by a loss function (L).

- **Forward Pass:** The forward pass involves calculating the predicted output by propagating the input through the network.

$$H = \sigma(X.\omega_{in} + b_{in}) \tag{3.4.6}$$
$$Y = \sigma(H.\omega_{out} + b_{out})$$

Here,

- X is the input data.
- ω_{in} and b_{in} are the weights and biases of the input layer.
- σ is the activation function.
- ω_{out} and b_{out} are the weights and biases of the output layer.

- **Loss Calculation:** Calculate the loss (L) between the predicted output (Y) and the actual output (Y_{true}).

$$L = Loss(Y, Y_{true}) \tag{3.4.7}$$

- **Backward Pass (Gradient Calculation):** Calculate the gradients of the loss concerning the weights and biases. It computes the error term for each node in the output layer:

$$\delta_j^{(L)} = \frac{\partial C}{\partial z_j^{(L)}} \cdot \sigma'(z_j^{(L)}) \tag{3.4.8}$$

Here, $\sigma'(z_j^{(L)})$ represents the gradient function. It computes the error terms for each node in the hidden layers recursively using the chain rule:

$$\delta_j^{(l)} = \sum_k \delta_k^{(l+1)} \cdot \omega_{kj}^{(l+1)} \cdot \sigma'(z_j^{(l)}) \tag{3.4.9}$$

$$\Delta\omega_{ij}^{(l)} = -\alpha \cdot \delta_j^{(l)} \cdot x_i^{(l-1)} \tag{3.4.10}$$
$$\Delta b_j^{(l)} = -\alpha \cdot \delta_j^{(l)}$$

- **Weight Update:** Update the weights and biases using an optimisation algorithm, such as gradient descent.

$$\omega_{out} \leftarrow \omega_{out} - \alpha \nabla_{\omega_{out}} L \tag{3.4.11}$$

$$b_{out} \leftarrow b_{out} - \alpha \nabla_{b_{out}} L$$

$$\omega_{in} \leftarrow \omega_{in} - \alpha \nabla_{\omega_{in}} L$$

$$b_{in} \leftarrow b_{in} - \alpha \nabla_{b_{in}} L$$

where α corresponds to the learning rate. This is a hyperparameter that controls the weight updates' step size. This process is continued iteratively for several epochs until the model converges to a point where the loss is minimised, and the network accurately predicts unseen data.

Gradient Update Rules

Gradient update rules, also known as optimisation algorithms, dictate how the parameters (weights and biases) of a neural network are adjusted in the training process [38]. The primary goal is to find the optimum set of parameters that minimise the loss function. Various optimisation algorithms exist. The selection of the accurate algorithm often depends on such characteristics as the size of the dataset and computational resources [38]. Hyperparameters such as the learning rate, momentum terms, and decay rates need to be tuned for optimum performance. Here are some commonly used gradient update rules:

- **Adam Optimiser:** Adam combines ideas from momentum and root mean squared propagation (RMSprop). It adapts the learning rates for each parameter based on their past gradients and squared gradients. The update

rule is given by:

$$
\begin{aligned}
m_t &= \beta_1 \cdot m_{t-1} + (1-\beta_1) \cdot g_t \\
v_t &= \beta_2 \cdot v_{t-1} + (1-\beta_2) \cdot g_t^2 \\
\hat{m}_t &= \frac{m_t}{1-\beta_1^t} \\
\hat{v}_t &= \frac{v_t}{1-\beta_2^t} \\
\theta_{t+1} &= \theta_t - \frac{\eta}{\sqrt{\hat{v}_t}+\varepsilon} \cdot \hat{m}_t
\end{aligned}
\tag{3.4.12}
$$

where m_{t-1} and v_{t-1} are estimates of the first moment (the mean) and the second moment (the uncentered variance) of the gradients respectively, $\hat{m}_t$ and $\hat{v}_t$ are bias-corrected estimates of m_t and v_t, β_1 and β_2 are decay rates (typically close to 1). η represents the learning rate, and ε is a small constant to avoid division by zero.

- **Stochastic Gradient Descent:** Stochastic gradient descent (SGD) is a fundamental optimisation algorithm used in neural network training. It updates the parameters according to the gradient of the loss with respect to the parameters for each data point. The update rule for a parameter w is given by:

$$\theta_{t+1} = \theta_t - \eta \nabla J(\theta_t) \tag{3.4.13}$$

  Where $\nabla J(\theta_t)$ represents the loss function's gradient with respect to θ_t.

- **Mini-Batch Gradient Descent:** Mini-batch gradient descent is a compromise between SGD and batch gradient descent. Instead of updating parameters for each data point, it computes the gradient on a small random subset (mini-batch) of the training data. The update rule is similar to SGD but involves the average gradient over the mini-batch:

$$\theta_{t+1} = \theta_t - \eta \frac{1}{|B|} \sum_{i \in B} \nabla J(\theta_t; x^{(i)}, y^{(i)}) \tag{3.4.14}$$

  Where $|B|$ is the size of the mini-batch, and $(x^{(i)}, y^{(i)})$ are the input and target

of the i-th example in the mini-batch.

3.4.6 Hyperparameters

In NNs, hyperparameters are parameters that are not learned during the training process but are set before the training begins. They play a crucial role in determining the architecture and behaviour of the NN. For example, the number of layers and neurons per layer are the hyperparameters which change the architecture of the NN. They influence the model's capacity to learn complex patterns and its ability to generalise to unseen data. Activation functions, another set of hyperparameters, introduce non-linearity into the network and affect its ability to approximate complex functions [43].
Selecting optimal hyperparameters is crucial for achieving good performance and generalisation in NN. Several techniques are commonly employed for hyperparameter optimisation.

- **Grid Search:** Grid search involves exhaustively evaluating model performance for all possible combinations of predefined hyperparameters. While thorough, this approach can be computationally intensive, especially for models with numerous hyperparameters or large search spaces [44].

- **Random Search:** Random search is a more computationally efficient alternative to grid search, randomly sampling hyperparameter values from predefined ranges. Despite not exploring all combinations, it often yields satisfactory results.

- **Bayesian Optimisation:** Bayesian optimization is a method for efficiently optimising hyperparameters by building a probabilistic model of the objective function and selecting the next hyperparameter configuration to evaluate based on past observations. It balances exploration and exploitation to converge to the optimal configuration with fewer evaluations compared to brute-force methods. While efficient and flexible, it requires careful model selection and tuning and may converge to local optima [45].

- **Cross-Validation:** Cross-validation partitions the training data into subsets, enabling the estimation of model performance across different hyperparameter configurations. It aids in selecting hyperparameters that generalise well to unseen data [46].

3.5 Machine Learning in Nanophotonics

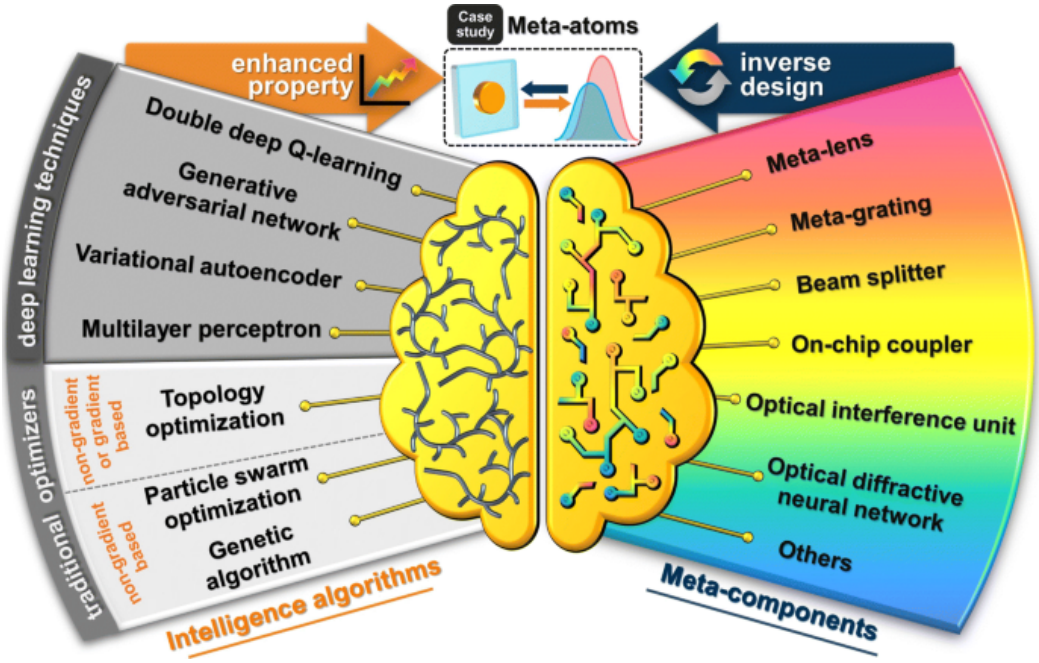
The advent of ML in nanophotonic research dates back to the 1990s, when the MLP, a basic form of a feedforward NN, used as a computer-aided design (CAD) instrument for rapidly prototyping microwave devices [47, 40].

The initiation of ML research in microwave domains has a couple of reasons. Firstly, most microwave device designs could be broken down into optimising specific parameters for a given target, conveniently expressed through a few variables. Given the well-developed physics, powerful simulation tools, and few challenges in data collecting, establishing relationships between these parameters was more straightforward for a MLP model. Furthermore, the scientific community had little understanding of more sophisticated photonic designs such as silicon photonic devices and metamaterials which were still in the early phases of development. Such devices with extensive design space and complex optical behaviour, posed challenges for the shallow NNs model prevalent at that time [40].

Following these early explorations into integrating data-driven methods in photonics research, deep MLP models with complex architectures and improved training strategies have been widely used in design of various devices like metamaterials [48, 49, 50], metagrating [51], beam splitters [52], structural colour systems [53, 54, 55], and plasmonic devices [56] (cf. Figure **??**).

3.5.1 Potential Applications

The main applications of ML in nanophotonics, especially metamaterials and metasurface, can be broadly categorised into three main sections: design and optimisation (forward design), inverse design, and materials discovery. The details of each category are discussed in the following sections.



Design and Optimisation of Structures

Conventional physics-based optimisation navigates the design space through predefined strategies for each case. However, DL seeks to comprehensively characterise the entire design space using training data as representative samples. This allows DL, with its generalisation capability within a specified design space, to generate rapid and precise designs without relying on case-specific and time-intensive numerical calculations. This concept is referred to as **"Forward Design"** in the literature [57, 11].

Forward design benefits from algorithms to predict and optimise the performance of nanophotonic devices before they are physically constructed. This process involves analysing data, selecting relevant features, training ML models, and then using these models to find the best design parameters for desired device properties. For example, core-shell nanoparticles provide fascinating phenomena, including multifrequency super scattering, directional scattering, and Fano-like resonance [58]. However, their increased DOF poses challenges in the design process. Peurifoy et al. [59] developed a MLP that predicts the scattering cross-section of a nanoparticle featuring silicon dioxide (SiO_2)/titanium dioxide (TiO_2) multilayered structures, as depicted in Figure3.6 (a, b). The network was trained

on 50,000 spectra obtained through the transfer matrix method (TMM). Results indicated that the developed network accurately computed spectra even when the input structure extended beyond the training data. This suggests that the DL goes beyond simple data fitting, uncovering underlying patterns and structures in the input and output data. However, this model architecture cannot achieve the inverse design of materials, and there are certain limitations on design freedom.

Figure3.6 (c) illustrates a typical forward design network for optimisation of a supercell all-dielectric metasurface [49]. Each unit-cell cylinder in the metasurface is defined by two parameters: radius and height, resulting in an 8-dimensional input vector. The effective simulation settings were established to calculate the scattering parameters as output. To prevent biasing the network towards specific geometries, the training set was constructed by randomly sampling the entire geometric hyperspace. The combination of all values provides approximately 13^8 million possible geometries. However, the mean squared error (MSE) was evaluated to reduce the training set to around 21,000 spectra. Consequently, a total dataset of this size was employed for the investigation, with 3,000 randomly selected pairs serving as a validation dataset. The training dataset comprised 18,000 pairs, representing only about 0.0022% of all possible geometries. The spacing between these values was carefully chosen to capture spectral variability adequately without introducing substantial redundancy among geometries with near structural features [49].

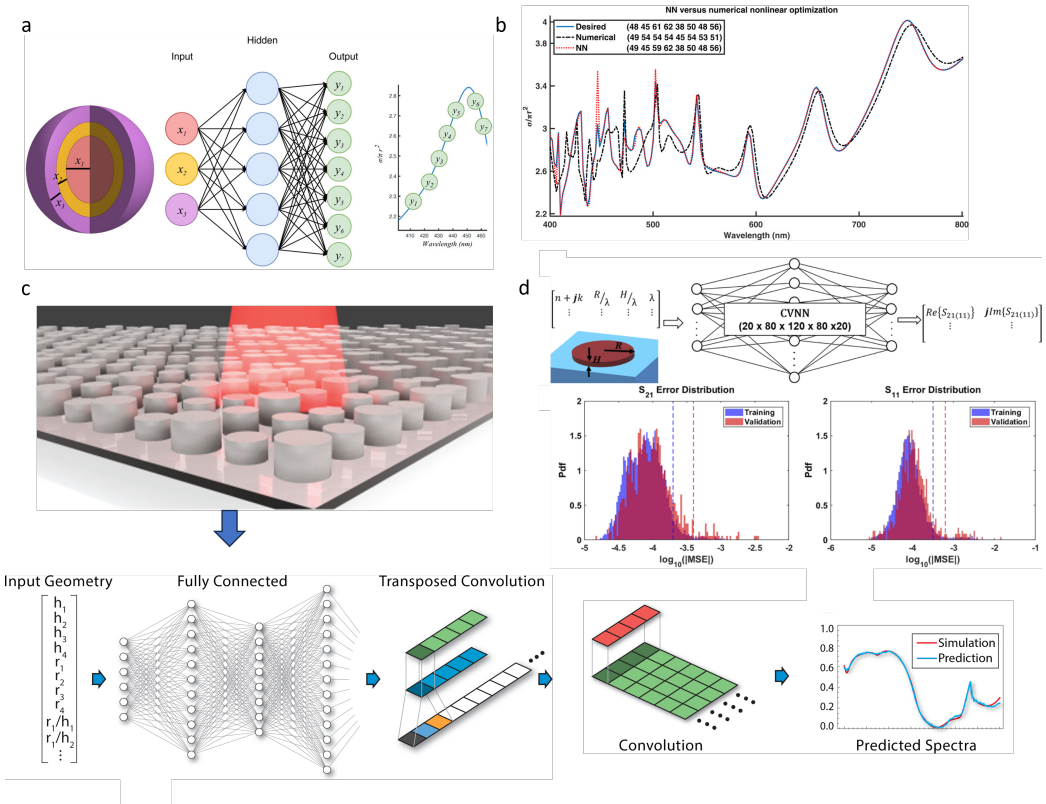


Figure 3.6: a) The neural network architecture takes the thickness of each shell of the nanoparticle as input and produces the scattering cross section at various wavelengths of the scattering spectrum as output, b) the accuracy of the inverse design NN for an eight-shell nanoparticle. Reprinted by permission from Optica Publishing Group [59], Copyright (2019). c) architecture of NN to predict the transmittance spectra of an all-dielectric metasurface with an array of cylindrical resonators with random height and radius. Reprinted by permission from Science [49], Copyright (2018). d) a complex-valued neural network with 5 fully connected hidden layers to learn the relationship between input (optical constants, dimension of cylindrical nanodisks, and wavelength of operation) and output (complex reflection or transmission coefficients). Reprinted by permission from Optica Publishing Group [60], Copyright (2021).

While most NN models for forward design were initially dedicated to passive metasurfaces, advancements have been made by incorporating the tunable optical properties of phase change materials (PCMs) as input in complex-valued NNs [60]. For example, as shown in Figure3.6 (d), a five-layer neural network was developed to learn the relationship between the input (optical constants, dimension of cylindrical nanodisks, and wavelength of operation) and output (complex reflection or transmission coefficients). This model predicts the output parameters with errors of less than 6.2×10^{-3} in 99% of the data. Given that optical prop-

erties involve complex numbers, complex-valued NNs are well-suited to address data error and accuracy challenges.

Inverse Design Problems

While the advancements of ML in photonics design and optimisation have been notable, current research is constrained to developing two-dimensional layouts for a single type of metasurface or nanophotonic structure. In addition, the analysis of these structures often maintains constant material characteristics and out-of-plane factors, such as layer thicknesses. The fundamental challenge that remains for optimisation processes using forward design models is a prior understanding of the specific capabilities offered by a particular category of design parameters [57]. However, relying on human intuition to select the optimised nanostructure, the starting point for numerical optimisation can be misleading when conflicting design goals arise.

Recently, the concept of **"Inverse Design"** has emerged in the literature to address these challenges. inverse design (ID) in nanophotonics refers to a methodology where ML techniques are employed to create novel and optimised nanophotonic structures or devices based on custom functionalities. Unlike forward design approaches, ID begins with defining the desired performance characteristics or functionalities and then uses algorithms to generate the corresponding structure and materials [61].

Malkiel et al. [62] introduced a fully connected bidirectional NN designed for predicting the geometries of metasurfaces based on the far-field responses and vice versa. As illustrated in Figure 3.7 (a), the focal structure is an H-shaped metallic structure characterised by eight geometrical parameters consisting of three continuous parameters corresponding to the rotation angle and length of each arm, and five discrete parameters corresponding to the arms. The main objectives consist of two reflection spectra when illuminated with light polarised along the horizontal or vertical direction. To reduce the number of input values, each reflection spectrum was discretised into 43 data points. In addition, the permittivity of the indium tin oxide (ITO) adhesion layer and hosting materials, were reduced to a

vector consisting of 25 parameters. The model was composed of two networks: a geometry-predicting network with eight group layers and a spectrum-predicting network with six layers. These networks were trained on a dataset consisting of 18,000 samples. This design enables the model to have dual functionality. Firstly, it serves as a rapid simulator, predicting the transmission of the structure for various geometries. Secondly, it acts as an ID tool, allowing the extraction of metasurface geometrical parameters based on the measured or desired transmission spectrum. The accuracy of the networks is shown in Figure 3.7 (b,c).

Upon further advancement of the design domain of core-shell nanoparticles, discussed in the previous section, an ID approach was developed to simultaneously determine the optical material and structural thickness based on the on-demand scattering spectra [24]. This was achieved by incorporating both classification and regression simultaneously, as illustrated in Figure 3.7 (d, e). The classification aspect determined the materials of each layer, while regression predicted the thickness of layers. The spectrum loss was computed through the MSE between target values of the spectra and the predicted ones generated by deep learning [24].

The "one-to-many" problem, where a single target optical response can be achieved by multiple solutions or designs, is a typical challenge faced in machine learning-aided ID. Novel training approaches, such as tandem networks [11, 54], and mixture density models [63], have been proposed to address this issue. For example, a metasurface configuration comprises arrays of metal disks suspended by polymeric pillars (Polydimethylsiloxane (PDMS)) with a metal (Al) back reflector, strategically designed for structural colour generation (cf. Figure 3.7 (f)) [11]. This geometry showed high interaction between the disk and hole arrays in the presence of hybridised reflectance modes. The parameter design space encompasses materials, pillar pitch (P), pillar diameter (d), pillar height (h), and metal thickness (t). The inverse model is trained using a tandem auto-encoder based on a pre-trained forward model. The generated dataset for self-learning was employed, resulting in an approximate 15% enhancement in the network's accuracy. The constrained inverse model enabled the instantaneous design of metasurfaces,

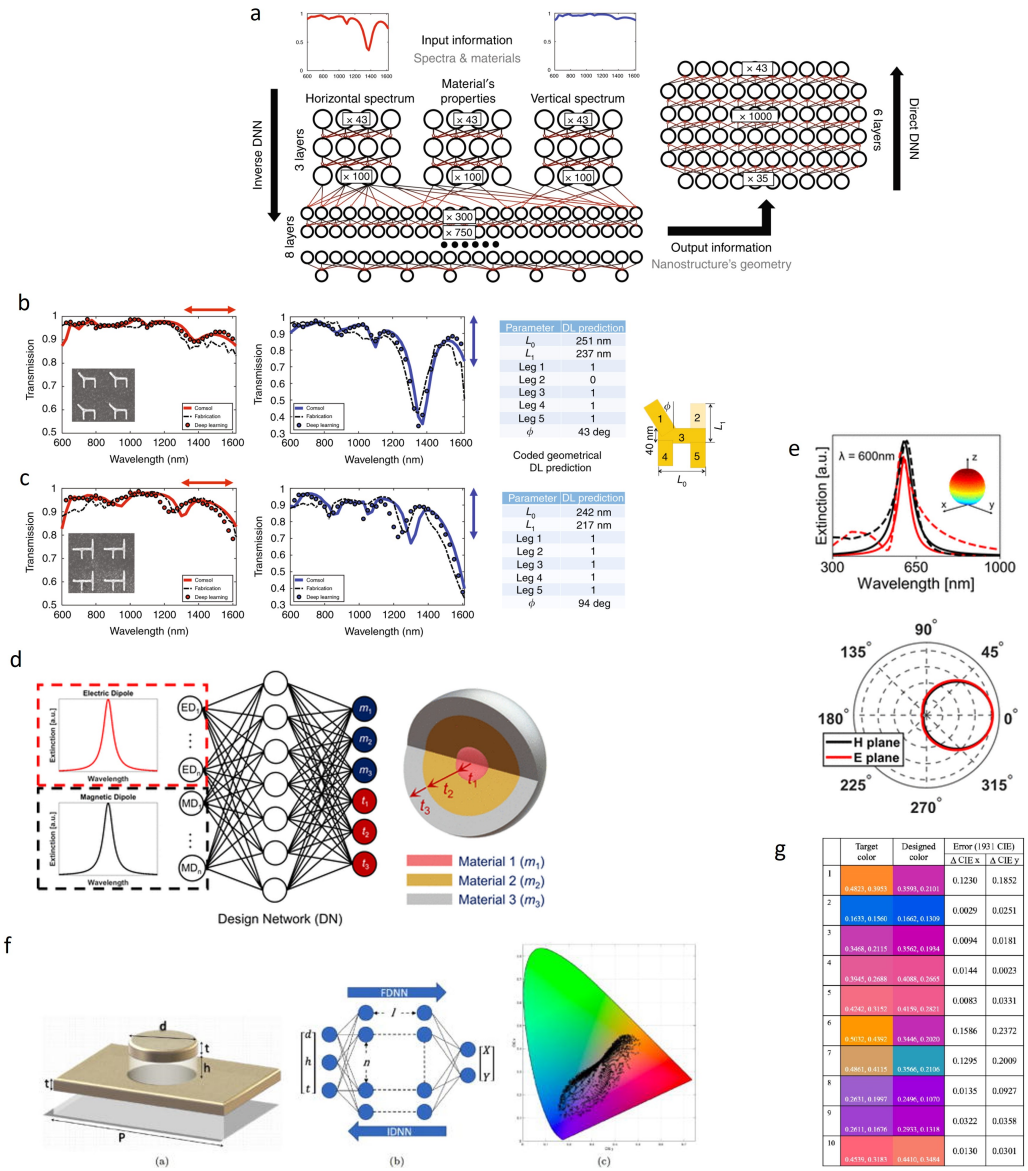


Figure 3.7: Examples of ID networks in nanophotonics. a) The bidirectional network's processing diagram encompasses two phases. b, c) A comparison of the transmission curves for samples derived from fabrication, simulation, and the DNN is presented, Reprinted by permission from Nature [62], Copyright (2018). d) the DL model to predict design parameters based on extinction spectra, e) the network's performance on the spectra of user-drawn values (solid lines) for the electric dipole (ED) (red) and MD (black) modes, and scattering radiation patterns at resonant wavelengths. Reprinted by permission from ACS Publications [24], Copyright (2019). f) a fully connected DNN with l hidden layers and n neurons in each hidden layer. The simulated metamaterials' range of spectral colours is represented on the Commission on Illumination (CIE) 1931 gamut, g) the accuracy of the model. Reprinted by permission from AIP Publishing [11], Copyright (2021).

achieving a precision of over 86% when targeting a desired colour as shown in Figure 3.7 (g) [11]. The suggested methodology is suitable for finding the most suitable parameters in the ID process. However, during the network training process, some solutions or solution spaces are either rejected or remain unreachable, which poses a serious challenge known as solution degeneracy [64].

To address this challenge, Dai et al. [64] developed a conditional generative adversarial network (cGAN). This approach enabled the generation of an average of 3.58 solution groups for a given ID of a transmissive colour filter. This network with various solutions empowers the selection of the best design for each colour, resulting in a record-high accuracy with an average index colour difference ΔE of 0.44. The effectiveness of the cGAN is experimentally verified through the inverse design of RGB colour filters [64].

In addition to the above-listed applications, ML has been used in the design of complex devices, including metagratings [65, 66], broadband highly reflective metasurfaces [67], multi-mode converters [68], multi-layer thin film [69], chiral metamaterials [70], and photonic crystals [71].

Material Selection

Materials selection is a critical aspect in the ID of metamaterials, particularly in the context of optical metamaterials and metasurfaces. The properties of these materials, such as refractive index (RI), play a critical role in determining the performance and functionality of the designed structures. Expanding the materials library for the NN-assisted design of nanostructures in metamaterials requires the consideration of more complex architectures. Moreover, a vast amount of data is necessary for training networks [72].

The materials library used for training these networks often involves indexed materials using integer or binary numbers represented by numerical values. These numerical indices simplify the training process, allowing for efficient handling of complex data. The fixed material in the library corresponds to a set of properties, including RI, that contribute to the overall performance of the metamaterial. The ID process typically involves selecting the most suitable material from the defined

library for a given set of specifications. The inverse problem is a combination of both classification and regression problems. It involves a classification problem to discover the appropriate material, and a regression one, seeking to determine the design parameters [24].

For example, an IDNN was trained with a dataset comprising 90,000 samples to predict the continuous and discrete design parameters of a typical multi-layer structure [73]. This structure involved the repetition of two different dielectrics (a random selection between Al_2O_3, Si_3N_4, and SiC) for N periods, as illustrated in Figure 3.8 (a). The combination of periods (N) with values of 4, 6, and 8, and the materials Al_2O_3, Si_3N_4, and SiC, leading to a total of 9 structure types. These structures were indexed with 0 and 1 for classification purposes. To explore the feasibility of ID with a reduced dataset and mitigate the time cost, three smaller datasets of 4,500, 9,000, and 18,000 samples were considered. The solid line in Figure 3.8 (b, c) represents the target spectrum, while the dashed lines show the proposed spectra of different models, each trained with a different dataset size.

Qiu et al. proposed an innovative multi-scenario training method to address imbalanced datasets in the inverse design of diverse multilayer nanoparticles and nanofilms [74] (cf. Figure 3.9 (a)). Optical scenarios with varying dataset sizes, some at the scale of 10^3 and others at the scale of 10^4, were considered. Compared to other methods, this approach used imbalanced datasets that were nearly four times smaller. The advantages of employing imbalanced datasets became increasingly pronounced with the expansion of the material library. Knowledge transfer learning was incorporated, involving the transfer of knowledge from a source scenario to a target scenario. The study employed a combination of four metals (Ag, Au, Cu, and Al) and four dielectric materials (TiO_2, SiC, Si_3N_4, and SiO_2) to create 16 optical scenarios. Distinct combinations of discrete 0s and 1s represented the materials in these scenarios, using a binary system to denote the two different materials (m_1 and m_2) in each case. For example, the scenario with Ag and TiO_2 was represented as "00," Au and SiC as "01," Cu and Si_3N_4 as "10," and Al and SiO_2 as "11." They implemented a multi-step process, training a base network (BaseNet) using the imbalanced multi-scenario method in the

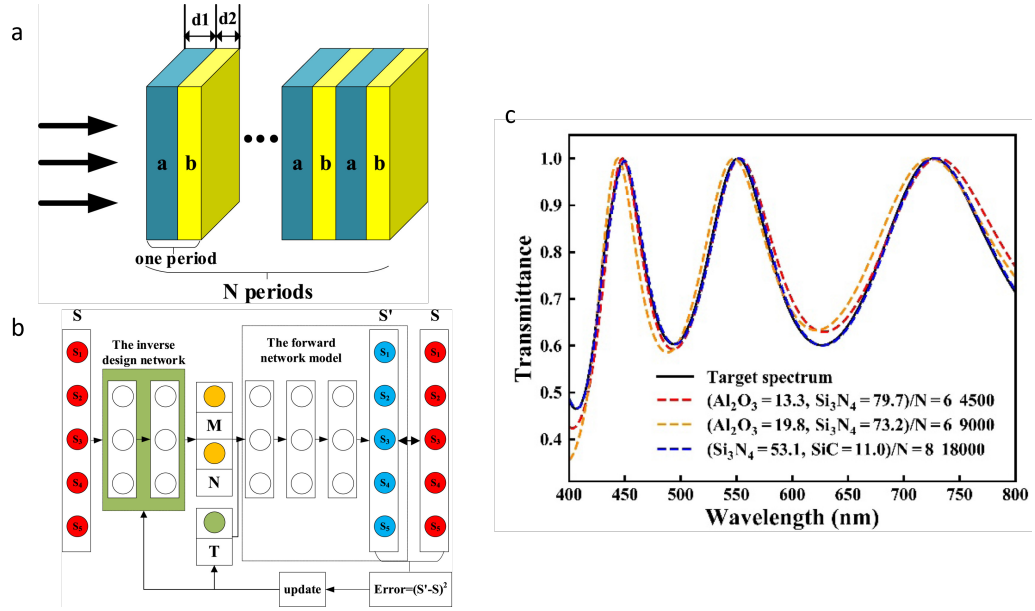


Figure 3.8: a) The materials a and b alternate placed to form a 1D photonic crystal structure. The first material thickness is d1 and the second material thickness is d2. b) The inverse design network updates M (materials) and N (periods) in the design process, c)Inverse design for few training samples. The solid line is the target spectrum, while the dashed lines are the three models' planned spectra. Reprinted by permission from Science Direct [73], Copyright (2021).

source scenario. Subsequently, part of the weights and biases from the hidden layers of BaseNet was transferred to a transfer network (TransferNet), which underwent fine-tuning using target scenario datasets (cf. Figure 3.9 (b)). Randomly selecting one, two, or three scenarios from the set of 169,000 samples were taken from these selected scenarios, while 900 samples were extracted from the remaining scenarios to create imbalanced datasets. The DNN model was then trained using the back-propagation algorithm with these imbalanced datasets. The accuracy of the ID network was validated through successful predictions and designs of nanophotonic structures as shown in Figure 3.9 (c) [74].

3.6 Limitations of Integrating AI in Nanophotonics

3.6.1 Data Collection

Due to the vast number of parameters in DNN, a large number of labelled data is typically required for effective training. However, generating such data often

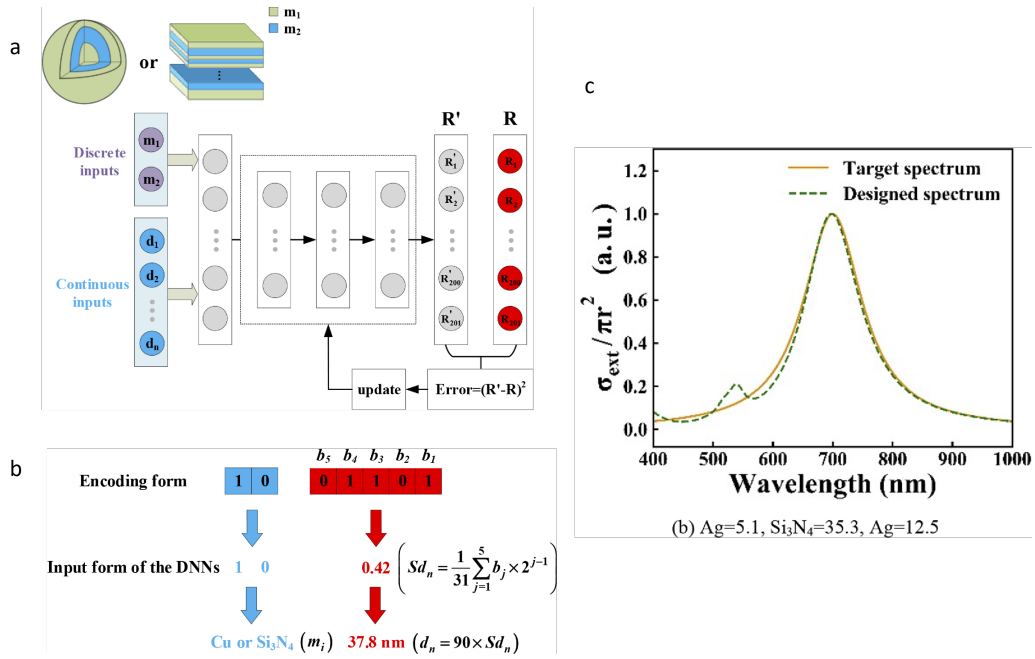


Figure 3.9: a) DNN for approximating multilayer nanoparticle transmittance spectra. The discrete inputs correspond to the material types, and the continuous values are nanostructure structural parameters. b) The genetic operation encoding form and DNN-recognised value of material and thickness parameters, c) the performance of the network in the design of a 3-layer nanoparticle structure. Reprinted by permission from Optica Publishing Group [74], Copyright (2021).

involves physical simulations or experimental measurements, making the collection of a massive dataset impractical in some situations [75]. Unsupervised and semi-supervised learning techniques can be employed to overcome this challenge by reducing the workload associated with data collecting. Unsupervised learning algorithms require only a small number of labels [40]. They provide efficient solutions to reduce the dimension and data clustering. These networks reveal the most significant design parameters, enabling the NN to recognise critical design data without unnecessary data collection [30, 31]. However, in the high-dimensional parameter spaces of the nanophotonic structures, the trained models with specific datasets may struggle to generalise to unseen data or different device geometries. Ensuring the transferability of trained models across various design scenarios with imbalanced datasets is a challenge.

Another approach to mitigate the challenge of massive data collection involves integrating physics-based deep learning (PBDL) models. The main idea behind PBDL is to integrate domain-specific knowledge of physical laws and equations

into the design and training of NN. By doing so, PBDL models can capture underlying physical principles, enabling them to make accurate predictions and generate meaningful insights [29]. Moreover, less reliance on data can result from understanding the underlying laws of physical processes. In this strategy, DL is often used to predict parameters from limited data for the regression of a few parameters that parametrise the models or equations of the system. PBDL models can also be used to generate solutions satisfying certain partial differential equations [76]. While these algorithms offer solutions for reducing reliance on extensive datasets, they encounter fundamental challenges when applied to multiphysics problems, such as optimising mechanically tunable metasurfaces. Multiphysics problems often involve nonlinear and coupled equations, leading to complex model architectures. Designing and training DL models that effectively capture these complexities while maintaining computational efficiency is a significant challenge [77].

3.6.2 Materials Discovery for Custom Structures

The discovery and advancement of new compositions play a significant role in driving technological progress, like the transformative impact of silicon on the microelectronics industry. Notably, advancements in critical technological domains, including sustainable energy, hinge on the capability to design and realise materials with optimised properties [78]. The process of discovering and designing materials requires synergistic cooperation between materials prediction, synthesis, and characterisation. The integration of computational tools, the generating of extensive materials databases, and advancements in experimental techniques have efficiently accelerated these endeavours [79].

In this research area, the concept of "materials by design" approaches has emerged to minimise the transition of materials from computer simulations to laboratory experimentation and practical consumer products [72]. The fundamental principle of the US Materials Genome Initiative revolves around the concept of the rational design of materials through the effective utilisation of data-driven methods [80, 81].

The initial incorporation of ML into materials science was highlighted by the generation of large databases primarily through computational methods and experiments. However, in recent years, an alternative approach has gained attention, leveraging less conventional ML tools like RL [82]) to expand small experimentally generated databases into larger sets and predict novel materials [78].

Up to now, optical technologies have primarily engaged in photonic optimisation within a restricted design space, mainly focused on structural topology (geometry) and shape. This optimisation approach frequently overlooks comprehensive feedback from other key design aspects, such as characterisation constraints and fabrication variables [83]. This challenge limits the application of current optimisation methods, preventing them from delivering efficient, globally optimised solutions. For example in multiphysics problems like phase-change platforms, the optical behaviour is closely connected to both the composition of materials and the fabrication procedures. The further development of advanced hybrid optimisation strategies is essential to address these optimisation challenges and investigate potential solutions for achieving the maximum performance of a photonic device [84].

By integrating pre-existing material knowledge into the training process, the design space can be enlarged to cover more complex compositions. In addition, by adding linear and nonlinear optical characteristics into the data generation process, it becomes feasible to optimise more features including the free-form geometry, or even the optimised deposition parameters of the target material [85].

3.7 Conclusion

Designing nanophotonic devices poses a significant challenge due to their complex and multifaceted nature. This chapter presented an overview of two traditional design processes, physics-based and numerical simulation methods, underlining their fundamental limitations such as high computational cost and reliance on trial-and-error procedures.

To address these challenges, the chapter introduced the concept of ML as a power-

ful tool capable of uncovering nonlinear relationships between design parameters and optical performance. It reviewed significant findings regarding the utilisation of ML-based algorithms in the design and optimisation of nanophotonic structures.

However, the current ML-based methods still face challenges, particularly in the training data generation, inverse design, and optimisation of mechanically tunable metasurfaces with infinite design space. Overcoming these challenges requires the development of innovative training approaches such as transfer learning, and the development of composite multifunctional models that are trained with a small dataset.

Bibliography

[1] D. Lee, W. Chen, L. Wang, Y.-C. Chan, and W. Chen, "Data-driven design for metamaterials and multiscale systems: A review," *Advanced Materials*, p. 2305254, 2023.

[2] D. Felbacq and G. Bouchitté, *Metamaterials Modelling and Design*. CRC Press, 2017.

[3] J. Chen, S. Hu, S. Zhu, and T. Li, "Metamaterials: from fundamental physics to intelligent design," *Interdisciplinary Materials*, vol. 2, no. 1, pp. 5–29, 2023.

[4] W. Hergert and T. Wriedt, *The Mie theory: basics and applications*. Springer, 2012, vol. 169.

[5] J. B. Pendry, A. J. Holden, D. J. Robbins, and W. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE transactions on microwave theory and techniques*, vol. 47, no. 11, pp. 2075–2084, 1999.

[6] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nature Photonics*, vol. 15, no. 2, pp. 77–90, 2021.

[7] L. Rayleigh, "Xvii. on the maintenance of vibrations by forces of double frequency, and on the propagation of waves through a medium endowed with a periodic structure," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 24, no. 147, pp. 145–159, 1887.

[8] A. R. McGurn, *Introduction to nonlinear optics of photonic crystals and metamaterials*. IOP Publishing, 2021.

[9] S. D. Campbell, D. Sell, R. P. Jenkins, E. B. Whiting, J. A. Fan, and D. H. Werner, "Review of numerical optimization techniques for meta-device design," *Optical Materials Express*, vol. 9, no. 4, pp. 1842–1863, 2019.

[10] K. Achouri and C. Caloz, "Design, concepts, and applications of electromagnetic metasurfaces," *Nanophotonics*, vol. 7, no. 6, pp. 1095–1116, 2018.

[11] N. B. Roberts and M. Keshavarz Hedayati, "A deep learning approach to the forward prediction and inverse design of plasmonic metasurface structural color," *Applied Physics Letters*, vol. 119, no. 6, 2021.

[12] M. M. Elsawy, S. Lanteri, R. Duvigneau, J. A. Fan, and P. Genevet, "Numerical optimization methods for metasurfaces," *Laser & Photonics Reviews*, vol. 14, no. 10, p. 1900445, 2020.

[13] K. Yao, R. Unni, and Y. Zheng, "Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale," *Nanophotonics*, vol. 8, no. 3, pp. 339–366, 2019.

[14] S. Angra and S. Ahuja, "Machine learning and its applications: A review," in *2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)*. IEEE, 2017, pp. 57–60.

[15] I. El Naqa and M. J. Murphy, *What is machine learning?* Springer, 2015.

[16] S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics*, vol. 9, no. 5, pp. 1041–1057, 2020.

[17] D. O. Hebb, *The organization of behavior: A neuropsychological theory*. Psychology press, 2005.

[18] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM Journal of research and development*, vol. 44, no. 1.2, pp. 206–226, 2000.

[19] F. Rosenblatt, "A bibliography of perceptron literature," *Scientific And Technical Information*, p. 189, 1963.

[20] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[21] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[22] S. B. Kotsiantis, I. Zaharakis, P. Pintelas *et al.*, "Supervised machine learning: A review of classification techniques," *Emerging artificial intelligence applications in computer engineering*, vol. 160, no. 1, pp. 3–24, 2007.

[23] A. A. Soofi and A. Awan, "Classification techniques in machine learning: applications and issues," *Journal of Basic & Applied Sciences*, vol. 13, no. 1, pp. 459–465, 2017.

[24] S. So, J. Mun, and J. Rho, "Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core–shell nanoparticles," *ACS applied materials & interfaces*, vol. 11, no. 27, pp. 24 264–24 268, 2019.

[25] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, R. Tibshirani, and J. Friedman, "Unsupervised learning," *The elements of statistical learning: Data mining, inference, and prediction*, pp. 485–585, 2009.

[26] M. Usama, J. Qadir, A. Raza, H. Arif, K.-L. A. Yau, Y. Elkhatib, A. Hussain, and A. Al-Fuqaha, "Unsupervised machine learning for networking: Techniques, applications and research challenges," *IEEE access*, vol. 7, pp. 65 579–65 615, 2019.

[27] G. Moon, J.-r. Choi, C. Lee, Y. Oh, K. H. Kim, and D. Kim, "Machine learning-based design of meta-plasmonic biosensors with negative index metamaterials," *Biosensors and Bioelectronics*, vol. 164, p. 112335, 2020.

[28] D. Sell, J. Yang, S. Doshay, R. Yang, and J. A. Fan, "Large-angle, multifunctional metagratings based on freeform multimode geometries," *Nano letters*, vol. 17, no. 6, pp. 3752–3757, 2017.

[29] A. Ghasemi, R. Fang, D. A. Zeze, and M. K. Hedayati, "Ultra-stretchable active metasurfaces for high-performance structural color," *AIP Advances*, vol. 13, no. 8, 2023.

[30] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[31] P. Pillai, B. Rai, and P. Pal, "Modified variational autoencoder for inversely predicting plasmonic nanofeatures for generating structural color," *Scientific Reports*, vol. 13, no. 1, p. 3536, 2023.

[32] S.-S. Learning, "Semi-supervised learning," *CSZ2006. html*, 2006.

[33] J. F. C. Mesa and J. C. C. Morales, "From artificial intelligence and bayesian statistics to neuroanatomy: Connections, analogies, and applications," *Migration Letters*, vol. 21, no. S1, pp. 162–182, 2024.

[34] J. Wang, S. Elfwing, and E. Uchibe, "Modular deep reinforcement learning from reward and punishment for robot navigation," *Neural Networks*, vol. 135, pp. 115–126, 2021.

[35] D. Seo, D. W. Nam, J. Park, C. Y. Park, and M. S. Jang, "Structural optimization of a one-dimensional freeform metagrating deflector via deep reinforcement learning," *ACS Photonics*, vol. 9, no. 2, pp. 452–458, 2021.

[36] L. Deng, D. Yu *et al.*, "Deep learning: methods and applications," *Foundations and trends® in signal processing*, vol. 7, no. 3–4, pp. 197–387, 2014.

[37] O. Eluyode and D. T. Akomolafe, "Comparative study of biological and artificial neural networks," *European Journal of Applied Engineering and Scientific Research*, vol. 2, no. 1, pp. 36–46, 2013.

[38] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[39] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, "Recent advances in recurrent neural networks," *arXiv preprint arXiv:1801.01078*, 2017.

[40] P. Naseri, "Inverse design of metasurfaces using machine learning techniques," Ph.D. dissertation, University of Toronto (Canada), 2023.

[41] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[42] T. Szandała, "Review and comparison of commonly used activation functions for deep neural networks," *Bio-inspired neurocomputing*, pp. 203–224, 2021.

[43] J.-Y. Kim and S.-B. Cho, "Evolutionary optimization of hyperparameters in deep learning models," in *2019 ieee congress on evolutionary computation (cec)*. IEEE, 2019, pp. 831–837.

[44] H. Alibrahim and S. A. Ludwig, "Hyperparameter optimization: Comparing genetic algorithm against grid search and bayesian optimization," in *2021 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2021, pp. 1551–1559.

[45] T. T. Joy, S. Rana, S. Gupta, and S. Venkatesh, "Hyperparameter tuning for big data using bayesian optimisation," in *2016 23rd International Conference on Pattern Recognition (ICPR)*. IEEE, 2016, pp. 2574–2579.

[46] K. Ito and R. Nakano, "Optimizing support vector regression hyperparameters based on cross-validation," in *Proceedings of the International Joint Conference on Neural Networks, 2003.*, vol. 3. IEEE, 2003, pp. 2077–2082.

[47] D. A. Cirovic, "Feed-forward artificial neural networks: applications to spectroscopy," *TrAC Trends in Analytical Chemistry*, vol. 16, no. 3, pp. 148–155, 1997.

[48] Y. Chen, J. Zhu, Y. Xie, N. Feng, and Q. H. Liu, "Smart inverse design of graphene-based photonic metamaterials by an adaptive artificial neural network," *Nanoscale*, vol. 11, no. 19, pp. 9749–9755, 2019.

[49] C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, "Deep learning for accelerated all-dielectric metasurface design," *Optics express*, vol. 27, no. 20, pp. 27 523–27 535, 2019.

[50] T. Qiu, X. Shi, J. Wang, Y. Li, S. Qu, Q. Cheng, T. Cui, and S. Sui, "Deep learning: a rapid and efficient route to automatic metasurface design," *Advanced Science*, vol. 6, no. 12, p. 1900128, 2019.

[51] L. Pilozzi, F. A. Farrelly, G. Marcucci, and C. Conti, "Machine learning inverse problem for topological photonics," *Communications Physics*, vol. 1, no. 1, p. 57, 2018.

[52] M. H. Tahersima, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, and K. Parsons, "Deep neural network inverse design of integrated photonic power splitters," *Scientific reports*, vol. 9, no. 1, p. 1368, 2019.

[53] O. Hemmatyar, S. Abdollahramezani, Y. Kiarashinejad, M. Zandehshahvar, and A. Adibi, "Full color generation with fano-type resonant hfo 2 nanopillars designed by a deep-learning approach," *Nanoscale*, vol. 11, no. 44, pp. 21 266–21 274, 2019.

[54] S. Head and M. Keshavarz Hedayati, "Inverse design of distributed bragg reflectors using deep learning," *Applied Sciences*, vol. 12, no. 10, p. 4877, 2022.

[55] I. Sajedian, T. Badloe, and J. Rho, "Optimisation of colour generation from dielectric nanostructures using reinforcement learning," *Optics express*, vol. 27, no. 4, pp. 5874–5883, 2019.

[56] T. Zhang, J. Wang, Q. Liu, J. Zhou, J. Dai, X. Han, Y. Zhou, and K. Xu, "Efficient spectrum prediction and inverse design for plasmonic waveguide systems based on artificial neural networks," *Photonics Research*, vol. 7, no. 3, pp. 368–380, 2019.

[57] X. Zheng, X. Zhang, T.-T. Chen, and I. Watanabe, "Deep learning in mechanical metamaterials: From prediction and generation to inverse design," *Advanced Materials*, p. 2302530, 2023.

[58] Y. Qin, J. Xu, Y. Liu, and H. Chen, "Multifrequency superscattering pattern shaping," *Chinese Optics Letters*, vol. 19, no. 12, p. 123601, 2021.

[59] J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Science advances*, vol. 4, no. 6, p. eaar4206, 2018.

[60] I. Tanriover, W. Hadibrata, J. Scheuer, and K. Aydin, "Neural networks enabled forward and inverse design of reconfigurable metasurfaces," *Optics Express*, vol. 29, no. 17, pp. 27 219–27 227, 2021.

[61] Z. Liu, D. Zhu, L. Raju, and W. Cai, "Tackling photonic inverse design with machine learning," *Advanced Science*, vol. 8, no. 5, p. 2002923, 2021.

[62] 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.

[63] H. Wang and L. J. Guo, "Neutron: Neural particle swarm optimization for material-aware inverse design of structural color," *Iscience*, vol. 25, no. 5, 2022.

[64] P. Dai, K. Sun, X. Yan, O. L. Muskens, C. de Groot, X. Zhu, Y. Hu, H. Duan, and R. Huang, "Inverse design of structural color: finding multiple solutions via conditional generative adversarial networks," *Nanophotonics*, vol. 11, no. 13, pp. 3057–3069, 2022.

[65] L. Xu, M. Rahmani, Y. Ma, D. A. Smirnova, K. Z. Kamali, F. Deng, Y. K. Chiang, L. Huang, H. Zhang, S. Gould *et al.*, "Enhanced light–matter interactions in dielectric nanostructures via machine-learning approach," *Advanced Photonics*, vol. 2, no. 2, pp. 026 003–026 003, 2020.

[66] H. Lin, J. Hou, Y. Wang, R. Tang, X. Shi, Y. Tian, W. Xu *et al.*, "Machine-learning-assisted inverse design of scattering enhanced metasurface," *Optics Express*, vol. 30, no. 2, pp. 3076–3088, 2022.

[67] 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, 2020.

[68] 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.

[69] R. Unni, K. Yao, X. Han, M. Zhou, and Y. Zheng, "A mixture-density-based tandem optimization network for on-demand inverse design of thin-film high reflectors," *Nanophotonics*, vol. 10, no. 16, pp. 4057–4065, 2021.

[70] 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.

[71] C.-X. Liu, G.-L. Yu, and G.-Y. Zhao, "Neural networks for inverse design of phononic crystals," *AIP Advances*, vol. 9, no. 8, 2019.

[72] A. Roitberg, "Machine learning for molecular and materials science," *Bulletin of the American Physical Society*, 2024.

[73] C. Qiu, X. Wu, Z. Luo, H. Yang, G. Wang, N. Liu, and B. Huang, "Simultaneous inverse design continuous and discrete parameters of nanophotonic structures via back-propagation inverse neural network," *Optics Communications*, vol. 483, p. 126641, 2021.

[74] C. Qiu, X. Wu, Z. Luo, H. Yang, G. He, and B. Huang, "Nanophotonic inverse design with deep neural networks based on knowledge transfer using imbalanced datasets," *Optics Express*, vol. 29, no. 18, pp. 28 406–28 415, 2021.

[75] P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, "Deep learning in nano-photonics: inverse design and beyond," *Photonics Research*, vol. 9, no. 5, pp. B182–B200, 2021.

[76] Y. Chen, L. Lu, G. E. Karniadakis, and L. Dal Negro, "Physics-informed neural networks for inverse problems in nano-optics and metamaterials," *Optics express*, vol. 28, no. 8, pp. 11 618–11 633, 2020.

[77] Z. Wang, R. Zhuang, W. Xian, J. Tian, Y. Li, S. Chen, and H. Xu, "Phononic metamaterial design via transfer learning-based topology optimization framework," in *International design engineering technical conferences and computers and information in engineering conference*, vol. 86229. American Society of Mechanical Engineers, 2022, p. V03AT03A048.

[78] D. Piccinotti, K. F. MacDonald, S. A. Gregory, I. Youngs, and N. I. Zheludev, "Artificial intelligence for photonics and photonic materials," *Reports on Progress in Physics*, vol. 84, no. 1, p. 012401, 2020.

[79] K. Alberi, M. B. Nardelli, A. Zakutayev, L. Mitas, S. Curtarolo, A. Jain, M. Fornari, N. Marzari, I. Takeuchi, M. L. Green *et al.*, "The 2019 materials by design roadmap," *Journal of Physics D: Applied Physics*, vol. 52, no. 1, p. 013001, 2018.

[80] 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.

[81] J. J. De Pablo, B. Jones, C. L. Kovacs, V. Ozolins, and A. P. Ramirez, "The materials genome initiative, the interplay of experiment, theory and computation," *Current Opinion in Solid State and Materials Science*, vol. 18, no. 2, pp. 99–117, 2014.

[82] A. Jiang and O. Yoshie, "A reinforcement learning method for optical thin-film design," *IEICE Transactions on Electronics*, vol. 105, no. 2, pp. 95–101, 2022.

[83] T. Gahlmann and P. Tassin, "Deep neural networks for the prediction of the optical properties and the free-form inverse design of metamaterials," *Physical Review B*, vol. 106, no. 8, p. 085408, 2022.

[84] D. Zhu, Z. Liu, L. Raju, A. S. Kim, and W. Cai, "Multifunctional metaoptic systems: Inversely designed with artificial intelligence," *arXiv preprint arXiv:2007.00130*, 2020.

[85] J. J. de Pablo, N. E. Jackson, M. A. Webb, L.-Q. Chen, J. E. Moore, D. Morgan, R. Jacobs, T. Pollock, D. G. Schlom, E. S. Toberer *et al.*, "New frontiers for the materials genome initiative," *npj Computational Materials*, vol. 5, no. 1, p. 41, 2019.

Chapter 4

Results Chapter: Mechanical and Optical Simulation

4.1 Chapter Overview

This chapter presents the simulation results carried out for this thesis, having two main goals: First, numerical simulation is used to find new structures that can provide higher levels of stretchability, which is important for designing mechanically tunable metasurfaces. In addition, the initial datasets required for training the machine learning (ML) algorithms were generated using simulations. The chapter starts with the mechanical simulation section that explores several geometries to determine which ones can achieve improved stretchability by redistributing stress concentration. The subsequent optical simulations evaluate reflectance levels as an optical characteristic, which informs the choice of effective designs. Two primary geometries, including pillar-based and kirigami-inspired patterns

that provide high stretchability and deformability potential, are introduced in this chapter. Then the effects of key geometrical parameters on mechanical stress distribution are presented with considerations of different resonators' arrangements. The optimised geometry achieved by mechanical simulations, is then combined with the optical simulations to determine its performance. The last part of the chapter presents the three-dimensional (3D) simulations used in dataset generation.

4.2 Mechanical Simulation

Despite the successful fabrication of various 3D structures with different geometrical and material configurations, their mechanical behaviour under applied loads remains largely unexplored. In this section, the mechanical responses of two distinct flexible 3D polymer-based structures are discussed through mechanically guided deformation.

4.2.1 Pillar-based Geometry

The primary challenge of mechanically tunable metasurfaces is their limited ability to undergo reversible stretching and relaxation through the application or removal of mechanical force. Typically, these structures adopt a bilayer system, integrating mechanically disparate materials—soft and rigid—within a single structure. The substrate commonly consists of elastic or viscoelastic polymers such as Polydimethylsiloxane (PDMS) or rubber, characterised by low elastic modulus (E $<$ 10 MPa), high elasticity (elongation at break $>$ 100%), Poisson's ratio (ν) close to 0.5, and thickness ranging from 10 μm to 1 mm [1]. In contrast, resonator materials, often used in thin films, exhibit stiffness (E in the GPa range), brittleness (fracture strain <5%), and thickness<100 nm. This large mismatch in material properties leads to a significant challenge [1].

Most mechanically tunable metasurfaces are designed by placing rigid resonators directly on flexible substrates as presented in Figure 4.1 (a, b). Here, the flexibility of the substrate is used to tune the optical responses (transmission and reflectance)

by dynamically adjusting the distances between the resonators. When force is applied to conventional rigid island geometries, stress accumulation occurs at the interface, leading to permanent deformation, cracking, or delamination, as discussed in Section 2.5.

In contrast to common "rigid island" geometries [2, 3, 4], the pillar-based geometry (Figure 4.1 (c)) is created with an engineered substrate's structure indicated by elevated areas referred to as "pillars", which possess adjustable geometrical parameters. For the majority of modes of deformation, the mechanical strain encountered by the pillars is significantly reduced in comparison to the strain applied to the entire substrate.

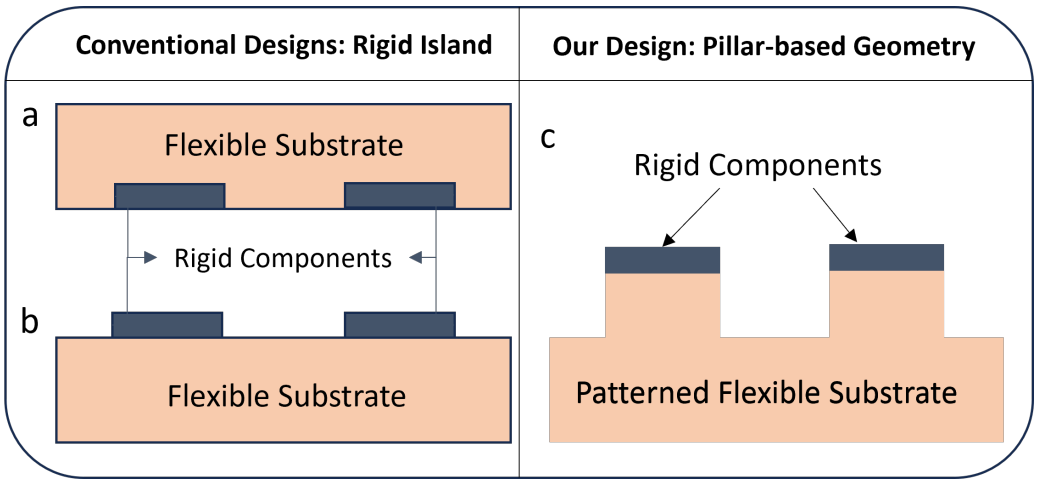


Figure 4.1: Schematic illustration of common bi-materials metasurfaces. The "rigid island approach" where the rigid materials characterised by higher Young's modulus are placed directly on the elastomeric substrate. These rigid regions can either be a) embedded within the substrate or b) deposited on it. c) our proposed geometry with an engineered substrate featuring high-pillar structures and rigid components on the top of the pillars.

Figure 4.2 illustrates the general pillar-based geometry, which includes a patterned substrate with a square arrangement of pillars (blue sections) and resonator discs (yellow sections) positioned on top of the pillars.

The finite element method (FEM) mechanical simulation was conducted using COMSOL Multiphysics. The primary structure, composed of PDMS, was assumed to exhibit linear elasticity. The material properties of PDMS, including Young's modulus, Poisson's ratio, and density, were set in accordance with

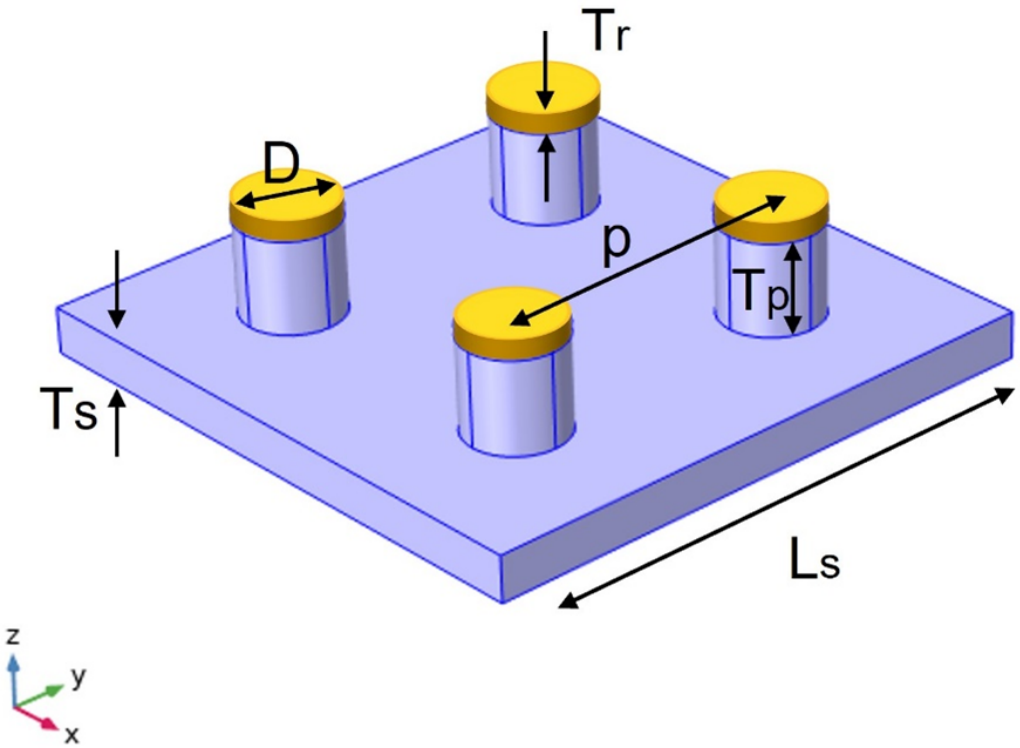


Figure 4.2: Schematic of the modelled geometry. T_s, T_r, T_p correspond to the thickness of the continuous substrate, the thickness of the resonator, and the pillars' height, respectively. D and p represent the diameter of the resonators and the pitch size.

literature references [5]. Additionally, the elastic modulus (E) and Poisson's ratio of aluminium nanodisks, considered as resonators, were specified as E_{Al} = 69 GPa and ν_{Al} = 0.34, respectively. To model the elastomeric substrate, a Mooney Rivlin hyper-elastic model was employed, with parameters C10 = 0.06757 MPa, C01 = 0.01689 MPa, and C11 = 0.48 MPa [6]. This model provides a good balance between accuracy and simplicity for describing the mechanical behaviour of elastomers and rubber-like materials. It calculates the nonlinear stress-strain relationships under large deformations [7]. The convergence of the mesh size was carefully ensured to maintain computational accuracy throughout the simulation process.

The initial findings illustrate the capabilities of pillar-based geometries in effectively redistributing stress concentrations within samples subjected to either uniaxial or biaxial tension. By strategically optimising various parameters, such as the pillar's height (cf. Figure 4.3 (a, b), it becomes feasible to minimise stress at

the interface between rigid and soft materials.

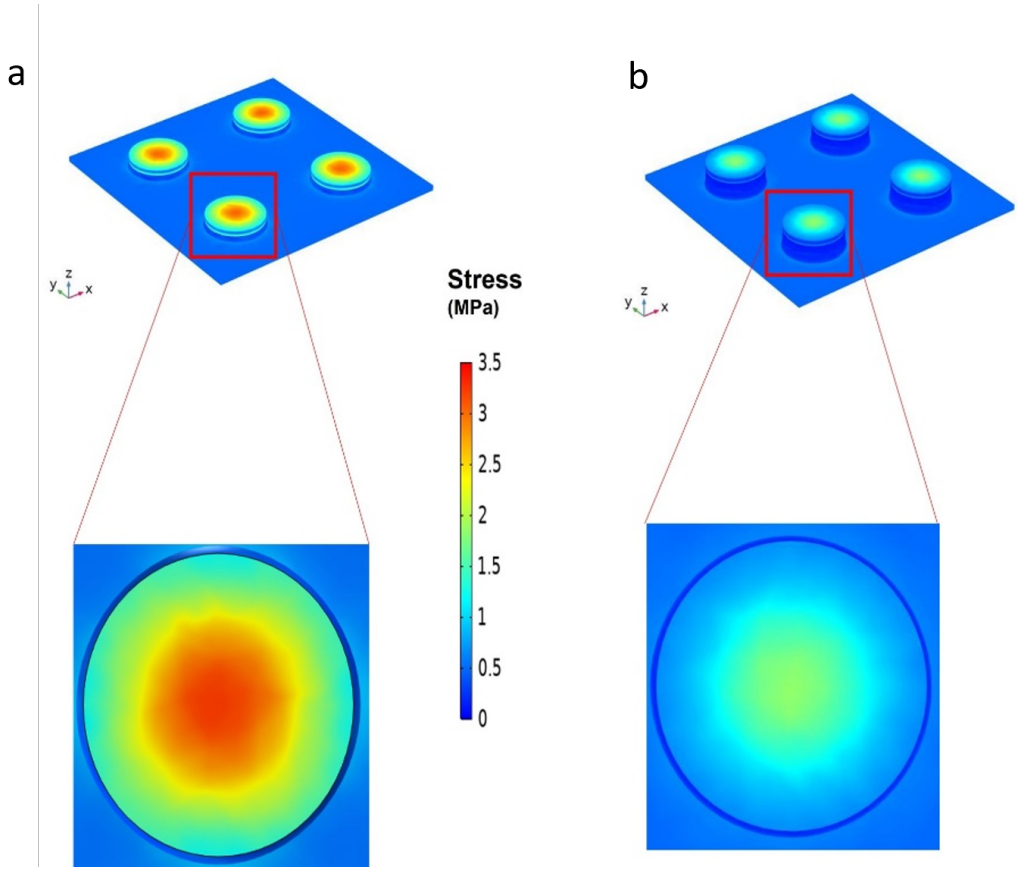


Figure 4.3: The comparison of the stress distribution in the simulation of a) rigid island and b) pillar-based geometries under bi-axial force. Increasing the height of the pillar reduces the stress on the stiff components, as observed from a top view of the pillar surface. Two stress bars are the same for both simulations.

To enhance the overall system's stretchability, a comprehensive understanding of how different parameters influence stress redistribution and minimise surface strain on the pillars is important. The modulation of geometrical parameters in various sections of the proposed geometry, including resonator, pillar, and substrate is systematically considered, as outlined in Table 4.1.

The structure is subjected to three levels of bi-axial strain (10%, 20%, and 30%) in the x-y plane. Two perpendicular sides of the substrate experience strain in the x-direction and y-direction, respectively. The remaining sides are fixed and cannot move or rotate. Additionally, a **"stress ratio"**, defined as the ratio of stress on the surface of the pillar to the stress at the bottom of the pillar, serves as a non-unit parameter for assessing the impacts of these parameters (cf. Figure 4.4

Table 4.1: List of different geometric design parameters for simulation with a range of values

Section	Parameter	Symbol	Values (nm)
Substrate	Length/Width	L_s	1000
	Thickness	T_s	100
Pillar	Diameter	D_p	50,100, 150, 200
	Thickness	T_p	0, 10, 100, 150, 250
	Periodicity	P	200, 300, 400, 600
Resonator	Diameter	D_r	50,100, 150, 200
	Thickness	T_r	25

(a)). A lower stress ratio indicates greater stretchability of the structure, thereby contributing to the flexible characteristics of the system.

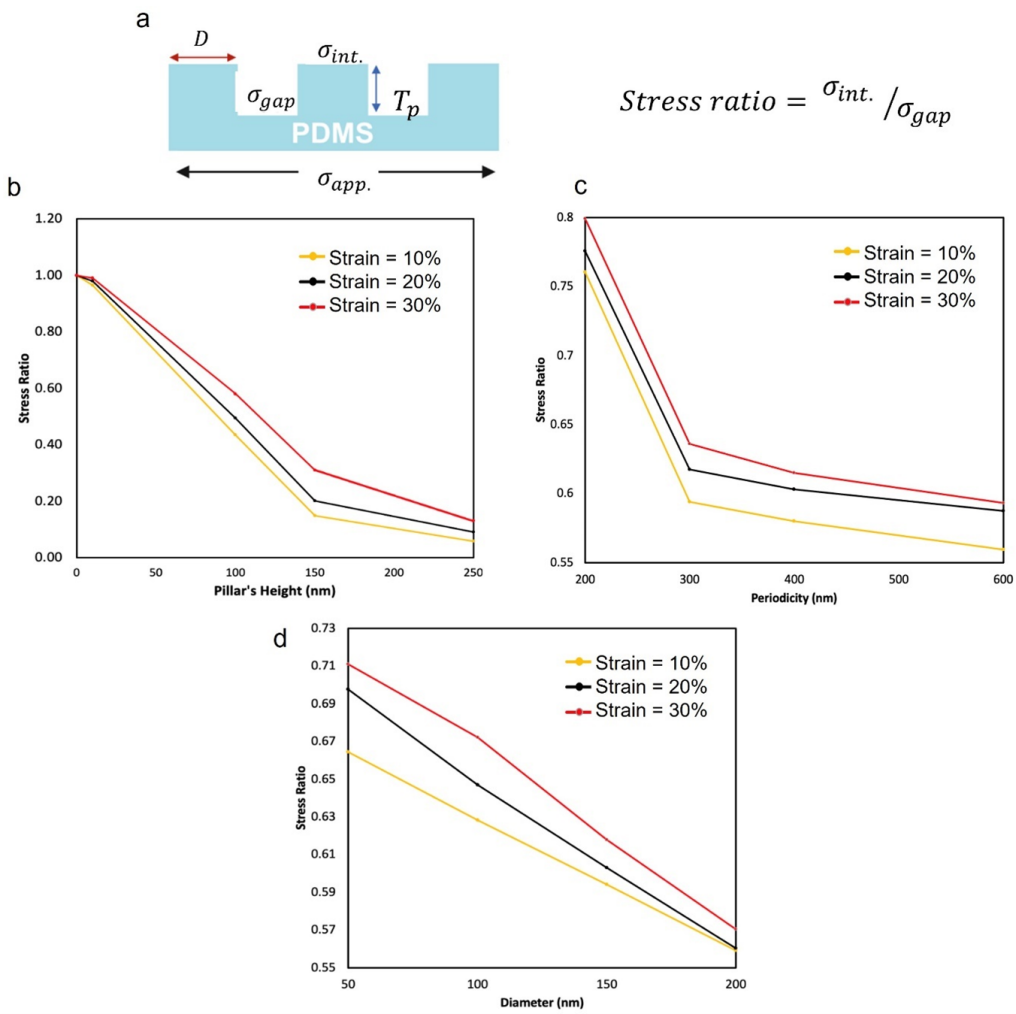


Figure 4.4: a) two-dimensional (2D) schematic of the of the proposed geometry. Stress ratio as a function of b) pillar's height, c) periodicity, and d) diameter of pillars in three different strain levels.

Figure 4.4 (b-d) show the effects of pillar height, periodicity, and diameter on the stress ratio. Notably, pillar height emerges as a key parameter to minimise the stress ratio. The results demonstrate that increasing pillar height, from 0 (conventional geometries) to 150 nm, substantially reduces the stress ratio by almost 80% (cf. Figure 4.4 (b)). As the pillar height continues to increase beyond 150 nm, the rate of decrease in the stress ratio diminishes and the minimum stress ratio of 0.15 was observed at a pillar height of 250 nm. This non-linear relationship between pillar height and stress ratio highlights a critical threshold in finding the optimised geometry. It was found that the h/d ratio (height-to-diameter ratio) higher than 1 leads to pillar collapse similar to previous reports. In Figure 4.4 (c), the significance of periodicity is demonstrated, indicating that minimum pillar-to-pillar distance results in an increased stress ratio due to stress accumulation in the gap, while an optimum periodicity, exceeding 300 nm, is identified. Excessive periodicity can adversely affect optical properties, which will be discussed in the optical simulation section. Lastly, Figure 4.4 (d) shows the stress ratio's dependency on diameter variations. While an increase in diameter generally diminishes the stress ratio, its influence is comparatively less than the other parameters.
Considering all investigations to design the optimised geometry, the maximum strain (120%) was achieved with a PDMS-based structure that has a total thickness of 200 nm. It consists of a 2D array of pillars, each with a diameter and height of 200 nm. The distance between each pillar is 400 nm. The resonators in the optimum geometry possess a thickness of 50nm and a diameter of 200nm.

In Figure 4.5 (a), the maximum stress level is depicted as a function of varying pillar height on the interface, illustrating the mechanical deformation of the sample under a uni-axial load. This configuration is particularly relevant for applications where asymmetric strain is applied. Figure 4.5 (a, b) further demonstrates the impact of loading in a single direction (the y-axis), resulting in a notably higher strain level in the load direction (approximately 140%) and a comparatively lower strain in the perpendicular direction (x-axis), reaching around 38%.

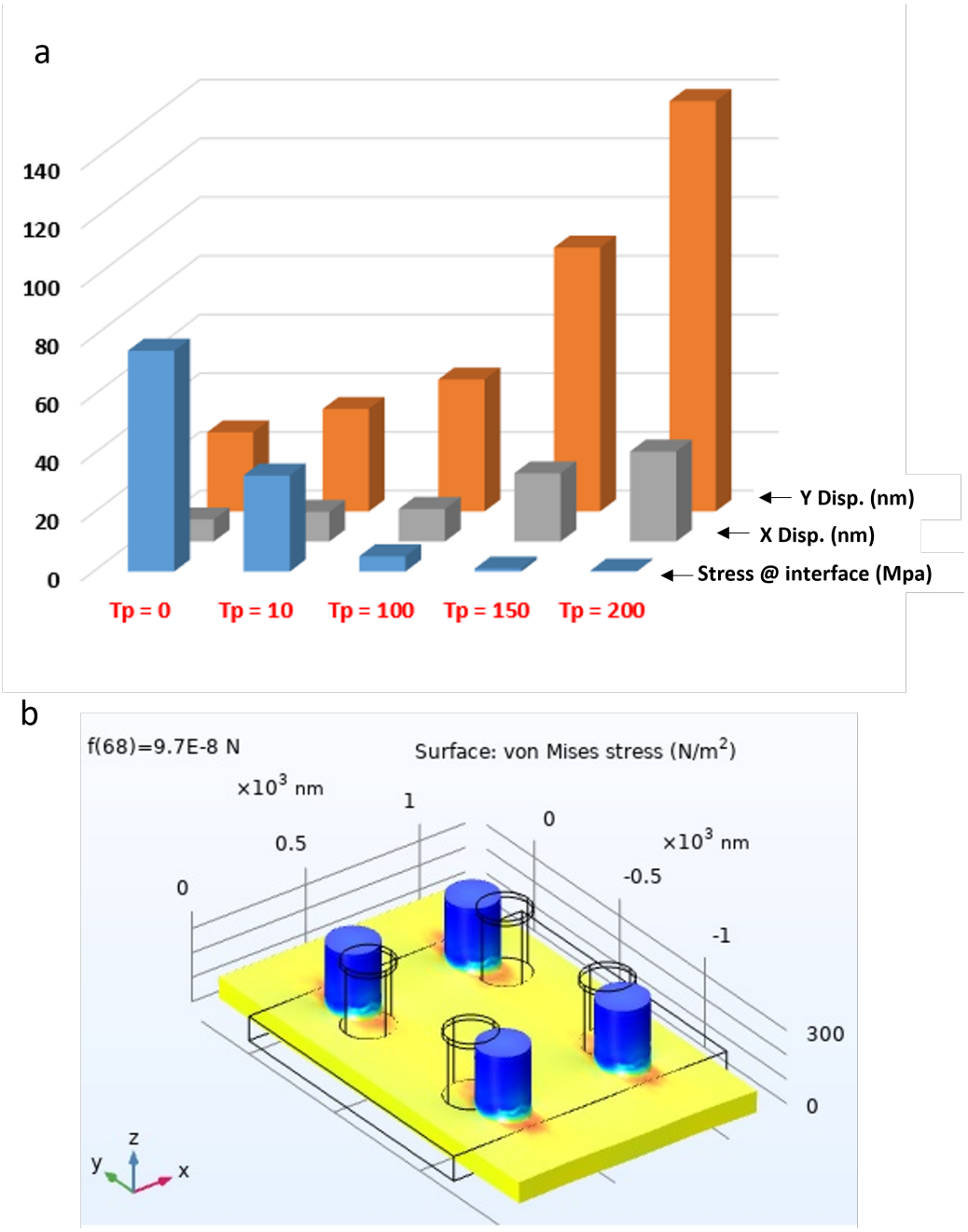


Figure 4.5: a) Effects of pillar's height on the maximum displacement in x-y directions, and stress concentration in the interface, b) stress distribution in the sample under uniaxial load.

4.2.2 Kirigami-inspired Geometry

Kirigami is a Japanese art similar to origami but involving both folding and cutting. It has emerged as a novel strategy to enhance the stretchability of various materials ranging from polymers to metals [8]. Unlike unpatterned materials, kirigami introduces localised stress concentrations at the cuts or notches in the

material. When the material is subjected to mechanical forces, these stress concentrations help distribute the strain more evenly throughout the structure. This redistribution of stress prevents the accumulation of localised strain in any one area, thus allowing the material to deform more easily without experiencing plastic deformation and failure [9].

Moreover, the cuts in kirigami structures enable certain regions of the material to relax their strain while others are being stretched. This strain relaxation mechanism further contributes to the material's ability to accommodate larger deformations without undergoing structural damage. Additionally, the presence of cuts or notches enhances the flexibility of the material, enabling it to bend and twist more easily in response to external forces. Furthermore, the geometric arrangement of cuts in kirigami patterns can be used to adjust the mechanical properties of the tunable metasurface [9].

Figure 4.6 (a) shows the kirigami architecture developed in this study. The configuration comprises an arrangement of aluminium nanodisks organised in a square lattice on a polymeric substrate based on kirigami principles. A flexible PDMS ribbon connects each nanodisk to its four adjacent counterparts. The entire assembly is placed on a pre-strained continuous elastomeric substrate that provides support. The nanodisks move in the x-y plane by relieving the substrate's strain as shown in Figure 4.6 (b). The substrate's elasticity, a crucial characteristic for material shaping and forming, relies on geometric factors such as width ratio and thickness ratio. To fabricate an "infinitely deformable optomechanical metamaterial", the width ratio (w/d, representing the width of interconnections to the diameter of disks) was adjusted to control local bending stiffness.

A 3D FEM was conducted to model the controlled mechanical buckling of the structure when subjected to compressive forces. The objective is to explore and find the optimised structure which provides the maximum stretchability. ABAQUS, a commercial software, was employed for simulating structures. Four-node shell elements were used to simulate the structures, while eight-node 3D stress elements were employed for the elastomer substrate. Mesh size convergence was ensured for computational accuracy. The elastomeric Kirigami structures were assumed

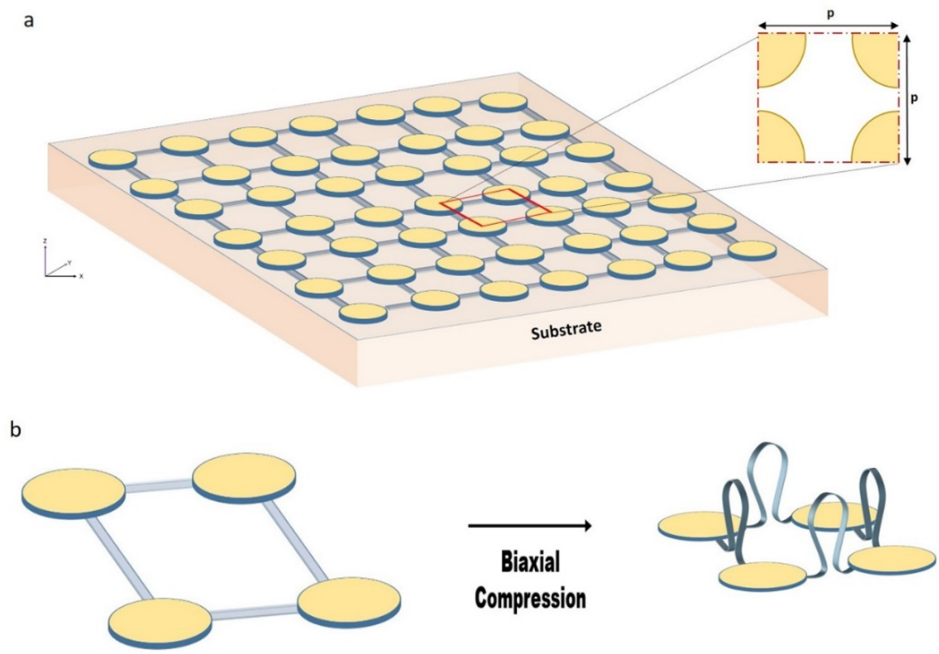


Figure 4.6: a) a schematic representation of the geometry, where thin and bendable ribbons made of PDMS connect a 2D array of Al nanodisks situated on a flexible site. Resonators fixed to the bottom substrate are indicated by the yellow sites, which move in both x and y directions. A continuous substrate is used to support the entire structure. b) The geometry's unit cell in the rest (left) and extreme (right) states. Reprinted by permission from AIP Publishing [10], Copyright (2023).

to exhibit linear elasticity. Similar to the previous geometry, the Mooney Rivlin hyper-elastic model was applied to represent the elastomeric substrate, with the same reported parameters.

As the bi-axial pre-strain is released, stress concentrates in the middle and both ends of the flexible ribbons. These regions have lower bending stiffness relative to the sections that hold the resonator (cf. Figure 4.7). The stress redistribution in the buckled structure leads to a 2D-to-3D shape transformation, changing the interparticle distance in a controllable manner. The buckling-induced strain must remain below the maximum strain limit for the elastomer connections as detailed in the reference [11].

The optimised design comprises nanodisks with a thickness of 100 nm and a diameter of 200 nm. The unit cell exhibits a period of p = 280 nm. For a more comprehensive comparison, two distinct modes were identified: the rest mode and the extreme mode. In the rest mode, characterised by a minimum gap of 80

nm (ε=0%), and in the extreme mode, identified by a maximum gap of 420 nm (ε=120%), these values respectively represent the states of maximum and minimum released strain.

Figure 4.7 (a, b) illustrate the Von-Mises stress and strain distribution of the optimised geometry subjected to a biaxial strain of extreme mode. In contrast to conventional rigid island geometries where the rigid resonators are directly placed on a continuous polymeric substrate, the stress concentration now shifts away from the interface to the ribbon-membrane connections. Figure 4.7 (c) shows the stress distribution in various strains from 0% to 120%. The significantly expanded range of stretchability is attributed to the minimised stress level at the rigid-soft interface, facilitating the spatial displacement of resonators over a greater distance. This enhances the optical tunability achieved by providing unprecedented degrees of freedom. Moreover, this design allows for highly adaptable resonators in both two and even three directions, enabling the use of plasmonic hybridisation and polarisation sensitivity to their fullest extent.

4.2.3 Alternative Arrangements

To generalise the concept, the arrangement of resonators was modified to demonstrate the geometry's ability in stress redistribution and strain-induced lattice deformation in other periodic metasurfaces. For example, Figure 4.8 (a, b) shows a diagonal arrangement of disks, where the central disk is free to move in all directions, while the other disks are fixed on PDMS-based pillars bonded to a continuous substrate. Applying strain to the main structure (by releasing the pre-strain of the continuous substrate) induces a biaxial compressive force that accumulates along the x and y directions. The kirigami-based pattern provides both in-plane and out-of-plane displacement as shown in Figure 4.8 (c).

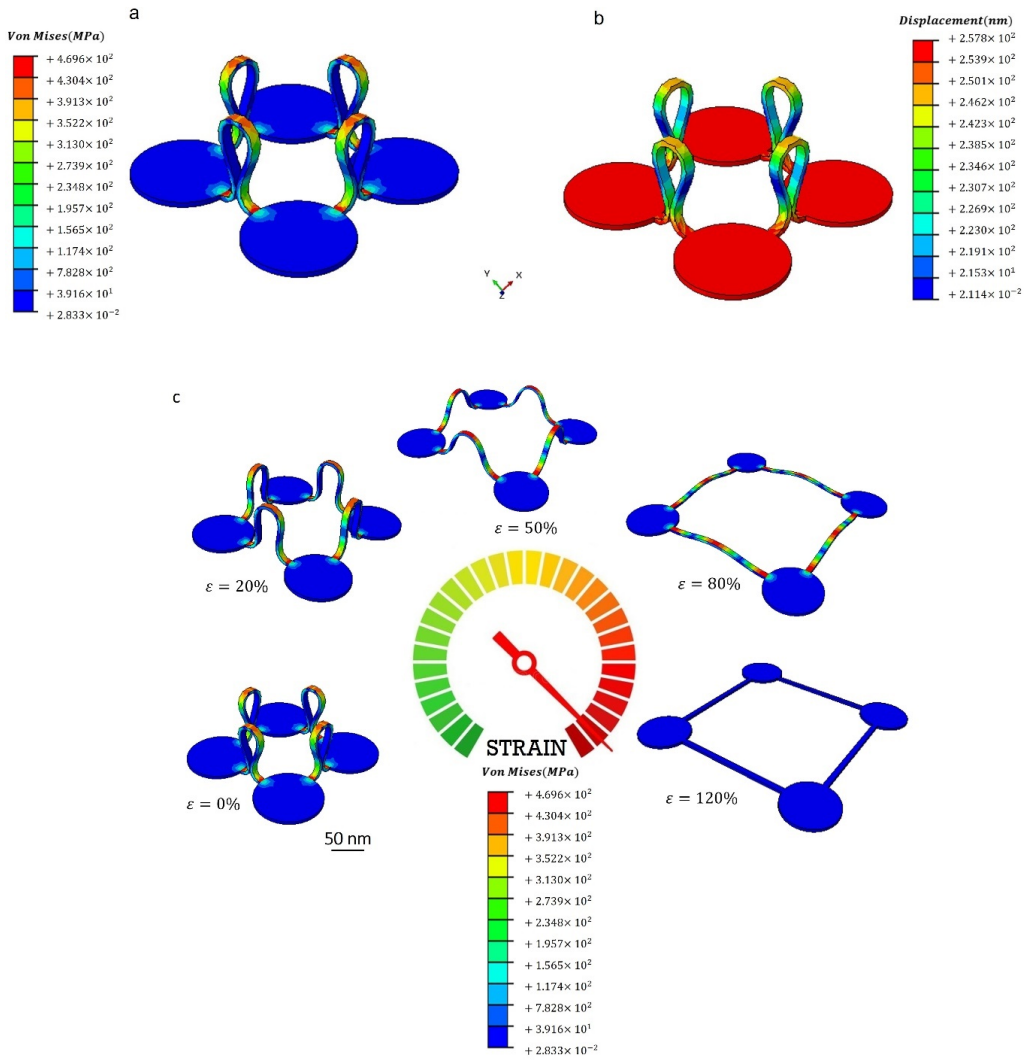
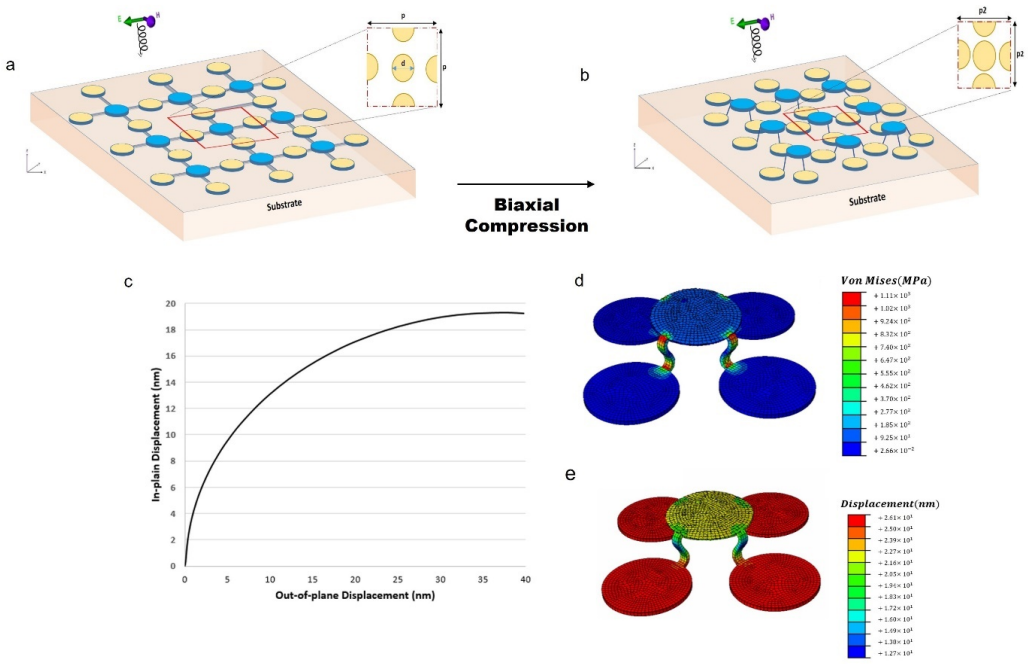


Figure 4.7: a) Von-Mises stress b) strain distribution of the original pattern in the extreme mode of the buckling process, and c) stress distribution in various strain levels. The minimum gap of 80 nm (ε=0%) and the maximum gap of 420 nm (ε=120%), respectively, represent the rest and extreme states. Reprinted by permission from AIP Publishing [10], Copyright (2023).



This compressive force acts as an upward trigger force on the flexible ribbons. Once the compressive load surpasses the elastic buckling strength, a threshold is reached, causing the interconnections to buckle. Simultaneously, the stress concentration shifts from the central parts to both ends. As depicted in 4.8 (d), in the extreme mode, stresses are predominantly distributed in the middle and both ends of the interconnections, with the disks experiencing negligible stress. This observation confirms the stability of this 3D structure.

According to our models, a minimum width ratio, defined as the width of the interconnection ribbon to the diameter of the resonator, (w/d $\geq$ 0.05) is necessary to induce an out-of-plane deformation in the interconnections; otherwise, the ribbons deform within the plane (2D). For instance, in Figure 4.9, the geometry with a width ratio of 0.04 is depicted. When the biaxial strain is released along a planar direction of the substrate, the localised stress at both ends of the flexible ribbons causes elastic bending of the connections, resulting in a reduction of the interparticle distance within the plane. The optimum geometry has a width ratio of 0.1.

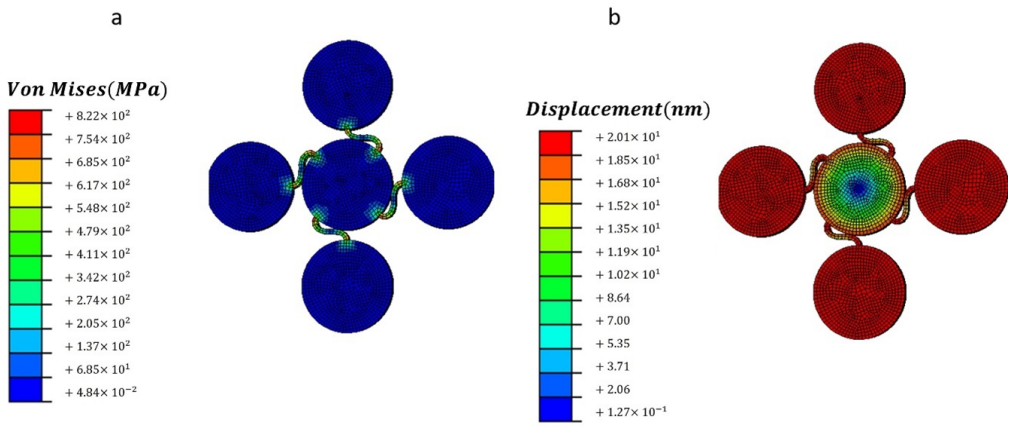


Figure 4.9: a) Von-Mises stress, b) strain distribution of a geometry with a width ratio of 0.04. Reprinted by permission from AIP Publishing [10], Copyright (2023).

4.3 Optical Simulation

Both discussed geometries in the previous sections exhibit a substantial stretchability range attributed to stress redistribution from the interface between rigid and

soft components to elastomer parts. This section focuses on reporting the optical behaviour of these geometries across different strain levels.

The optical simulation was conducted using COMSOL Multiphysics. The substrate (PDMS) was characterised by a thickness and refractive index of 100 nm and 1.45, respectively [12]. Al was employed as the metal, and its dielectric constant was determined from Palik's data [13].

Figure 4.10 (a) illustrates the schematic of a FEM model developed for the unit cell analysis. Floquet periodic boundary conditions (PBCs) is applied to the lateral faces to simulate an infinite array of elements. Perfectly matched layer (PML)s are positioned at the top and bottom boundaries to prevent reflections from the ports. Reflections were calculated for different periodicities (p) under normal incident light. The entire structure is meshed using triangular elements with a mesh size of 0.8 nm in the x and y directions. Elements from the x and y dimensions are projected to create the meshing in the z direction. The maximum mesh size is set in the air to reduce the simulation time. The nanodisks were specified with a diameter (d) of 200 nm and a thickness (h) of 100 nm.

The reflectance spectra of the geometry under different strains are shown in Figure 4.10 (b). The applied biaxial strain along a planar direction of the substrate changes the period of the unit cell in both the x and y directions. This adjustment consequently shifts the surface plasmon resonances in the kirigami-inspired pattern.

Resonator reflectance spectra are highly sensitive to gap size as shown in Figure 4.10 (b). When the gap size is less than 192 nm ($\varepsilon \leq 40\%$), two distinct reflectance peaks merge to form a single broad peak. By minimising the gap, The TM polarised incidence induces the bonding electric-electric dipole interaction by excitation of electric dipole (ED) along the disk dimer axis, which is similar to the plasmonic dimer situation [14].

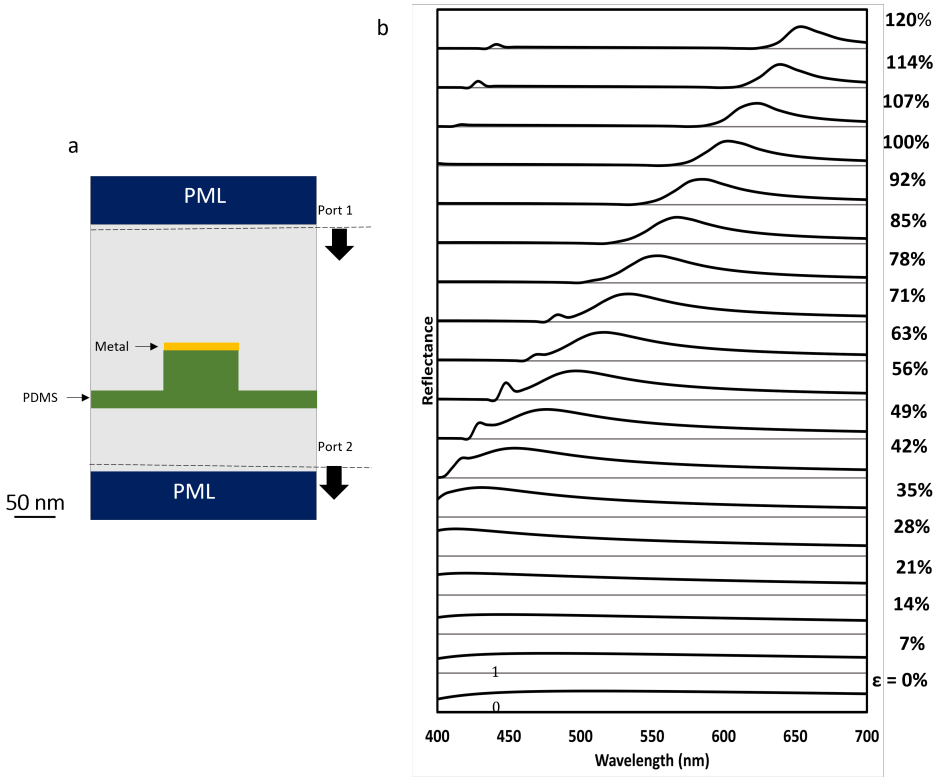


Figure 4.10: a) Schematic of a single simulation unit COMSOL. b) shows the reflectance curve as a function of strain level. Reprinted by permission from AIP Publishing [10], Copyright (2023).

The interaction has a significant impact on the energy of the exciting ED, causing a red-shift of the resonant wavelength of the electric mode (left peak), as depicted in Figure 4.10 (b). Simultaneously, the magnetic dipole (MD)s are excited in each particle but in a direction perpendicular to the dimer axis. The magnetic mode's energy level is raised by the repulsive force between the particle poles, which causes the resonant wavelength to move blue (right peak). These opposing peak shifts result in a single wide resonance centred at $\lambda = 508$ nm. This resonance is attributed to the plasmonic hybridisation of two modes with different local fields in the z-direction, as illustrated in Figure 4.10 (b) [15]. The single resonance closely aligns with the localised surface plasmon (LSP) resonance of a single disk, as observed in previous studies [16].

The reflectance spectra of the structure in the extreme mode, rest mode, and common island geometry's strain threshold reported in the literature are compared in Figure 4.11 (a). The two narrow resonance peaks at $\lambda = 440$ nm and $\lambda = 651$ nm

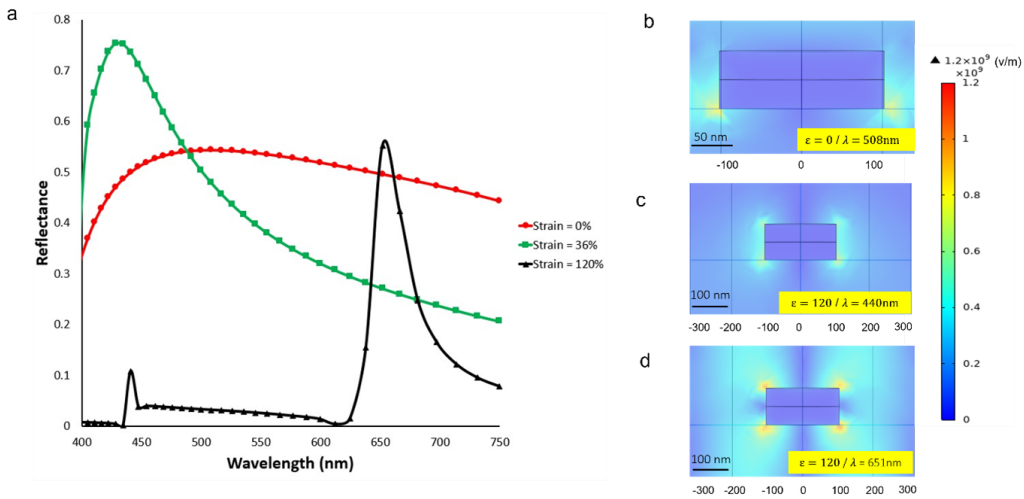


Figure 4.11: a) Reflection spectra of metasurface at (ε = 0, 36, 120%). The unit cell is comprised of Al nanodisks with a height of 100 and a diameter of 200 nm. b) The electric field distribution at the rest mode (ε = 0%) and in λ =508 nm resonant wavelength with the period p = 280 nm. c) The electric field distribution at the extreme mode (ε = 120%) and in λ =440 nm and (d) λ=651 nm resonant wavelengths with the period p = 620 nm. Reprinted by permission from AIP Publishing [10], Copyright (2023).

in the extreme mode, where the periodicity is highest, correspond to weakly hybridised LSP modes. It is worth highlighting that the strain threshold in literature is roughly 36%, leading to the 70-nm peak shift, however; the kirigami-inspired design may take a strain as high as 120%. In the extreme mode, this causes a peak shift of 143 nm (from 508 nm to 651 nm), which is nearly twice as great as the peak shift seen in the normal geometry.

According to Figure 4.11 (b), the electric field is located at the interface between the metal and dielectric and the corner of the metallic nanoparticles when the period (p) is 280 nm. Increasing the periodicity leads to isolating nanodisks, which provides stronger LSP resonances. As shown in Figure 4.11 (c), the electric field at the bottom corner of the first detected resonance (λ = 440 nm) gradually drops while increasing on both sides at p = 620 nm (ε = 120%). On the other hand, a greater electric field is focused in the disks' corner in the second detected resonance (λ = 651 nm), which is consistent with LSPR (Figure 4.11(d)).

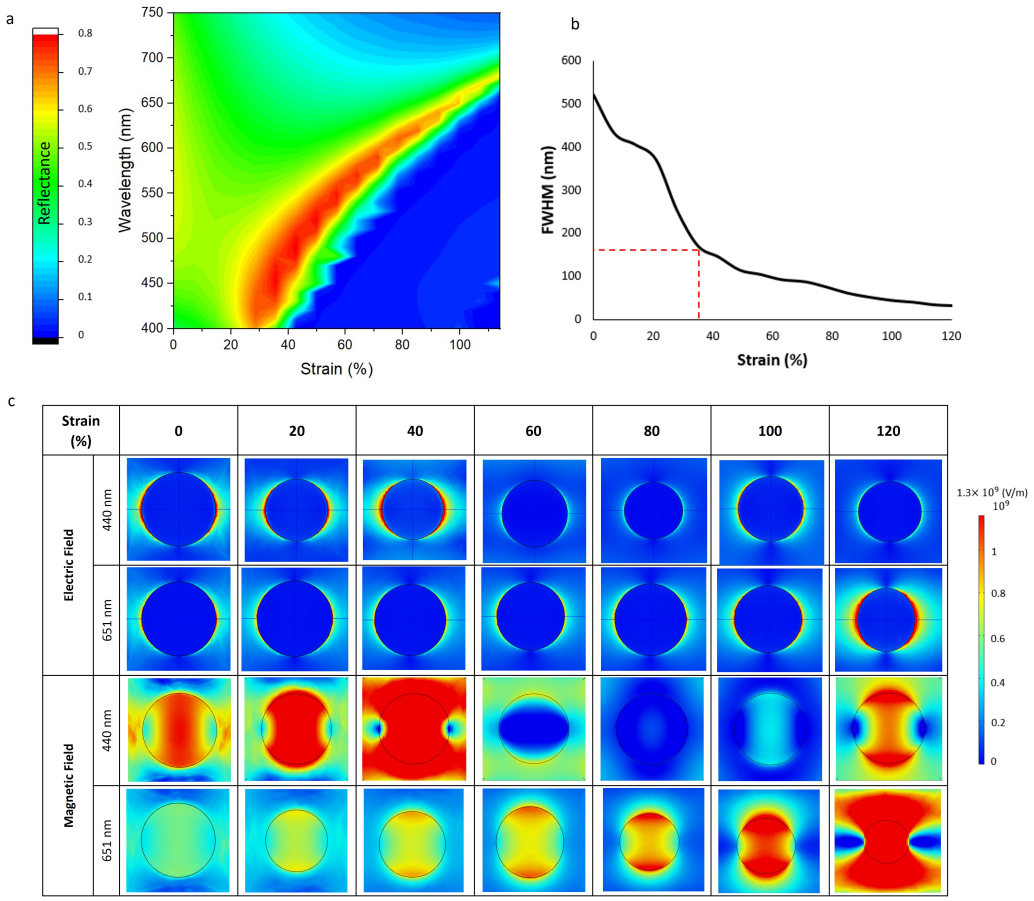


Figure 4.12: a) Contour map of the reflectance spectra as a function of strain, b) full width at half maximum (FWHM) values as a function of applied strain. The red dashed line represents the calculated FWHM (171 nm) at the maximum strain level (36%) in the conventional geometry, c) Electric field and magnetic field distribution in the xy plane for nanodisks at incident wavelengths of 440nm and 651 nm, corresponding to peaks revealed in simulation results. Reprinted by permission from AIP Publishing [10], Copyright (2023).

Figure 4.12(a) shows a contour map of the reflectance and a clear picture of the relative contributions from distinct phenomena, like localised surface plasmon resonance (LSPR) and surface plasmon resonance (SPR), which are related to different metal-substrate interfaces. When the smallest gap (strain) occurs, SPR is the predominant mechanism, producing a spectrum with a wide bandwidth [17]. The FWHM at different strain levels is shown in Figure 4.12(b). The FWHM was extracted from the reflection profiles that were fitted to a Gaussian function for this investigation. The results show that at greater strain levels, the suggested design can produce sharper peaks. FWHM, for example, drops to 33.1±0.1 nm at ε=120%, which is 5.1 times lower than the conventional counterparts at ε=36%.

Figure 4.12 (c) shows the distribution of the magnetic field. Within each nanodisk, the MD-induced electric field is primarily circulating. The magnetic field intensity within the Al nanodisks increases to ε = 40% at λ =440 nm, after which it decreases due to an increase in the strain level (gap size).

4.4 Exploring the Potential of Design

By redistribution of stress in the proposed geometries, their optical tunability increases significantly. This enhances their potential, giving them high-precision dynamic control, multifunctionality, and higher performance than conventional designs. Here, as a proof of concept, three applications of ultra-stretchable metasurface are discussed. Although these samples were not fabricated in this project, each example has great potential for further investigation.

Re-programmable Metadevice

Structural coloration has progressed significantly thanks to the extraordinary advancement in optoelectronic fields. However, the absence of a wide real-time tunability and rewritable colouration schemes has slowed down the progression in some relevant areas, especially the customised and reprogrammable optoelectronic devices [18]. So far, a series of optical devices have been created based on liquid crystals [19], micro/nanomechanical metamaterials [20, 21], phase-transition materials [22], and stimuli-responsive metasurface [23, 24] with critical drawbacks namely; a limited range of tunability, low efficiency, high fabrication /reconfiguration cost, and limited working temperature [25, 26].

The first application of the developed mechanically active metasurface is the design of a re-programmable and dynamic structural colour. Unlike conventional structural colour systems, the proposed one produces a wide range of colours from white to green and red to blue in a reversible manner. A single device that produces the colour from white to red and back to white is not easily possible with common designs, as they need to survive higher (>80%) strain levels as shown in Figure 4.13. Such optical devices may result in new kinds of gadgets with

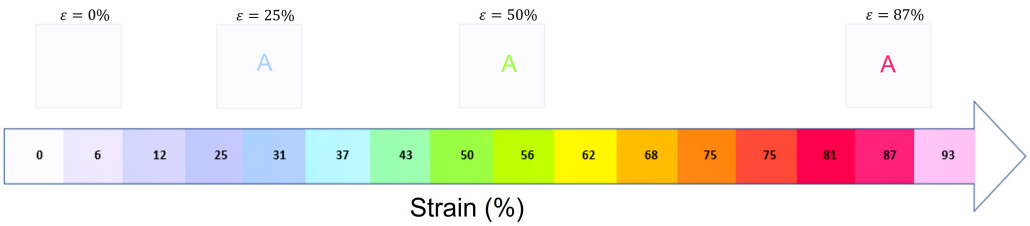


Figure 4.13: Deciphering embedded domain by strain. The presence of slight traces of the character on the left panels (ε= 0%) can be attributed to the interruption of the engineered periodic structure of the Al lattice, which can be deliberately designed to alleviate or eliminate the effect.

rewritable functions and optical multitasking. Moreover, high-resolution projectors, optical encryption, and real-time 3D displays might all employ them.

The variation in absorption and reflection is caused by the square lattices' increased periodicity upon stretching. Information is decoded when this strain-induced difference creates a sharp contrast between it and the background. For instance, we embedded a pattern of "A" character in a $50 \times 50\ \mu m^2$ region created by a square lattice of Al nanodisks with a height of 100nm, diameter of 200 nm, and periodicity of 320 nm on a white background. The sample was stretched, and the evolution of the square area was used to measure the colour change. A significant contrast in the reflection spectrum was observed once the strain reached approximately 25%. Notably, the letter was observed against a whitish background in reflection mode under bi-axial strain. In addition, in contrast to the previously reported devices which worked mainly in red to near-infrared regions with limited contrast and tunability, the proposed system presents a strong contrast form after stretching with a strain >25% under the reflected, offering a wide range of colour change in the presence of mechanical force (cf. Figure 4.13).

Plasmonic Sensors

Another application of mechanically tunable metasurface with high flexibility is plasmonic biosensing. Despite impressive achievements in designing tunable plasmonic sensors with high accuracy for refractive index (RI) sensing using phase change materials (PCMs), operating based on an unprecedented non-volatile change in their RI upon transition between amorphous and crystalline states, there

is still room for proposing a materials-independent solution to design similar sensors [27, 28, 29].

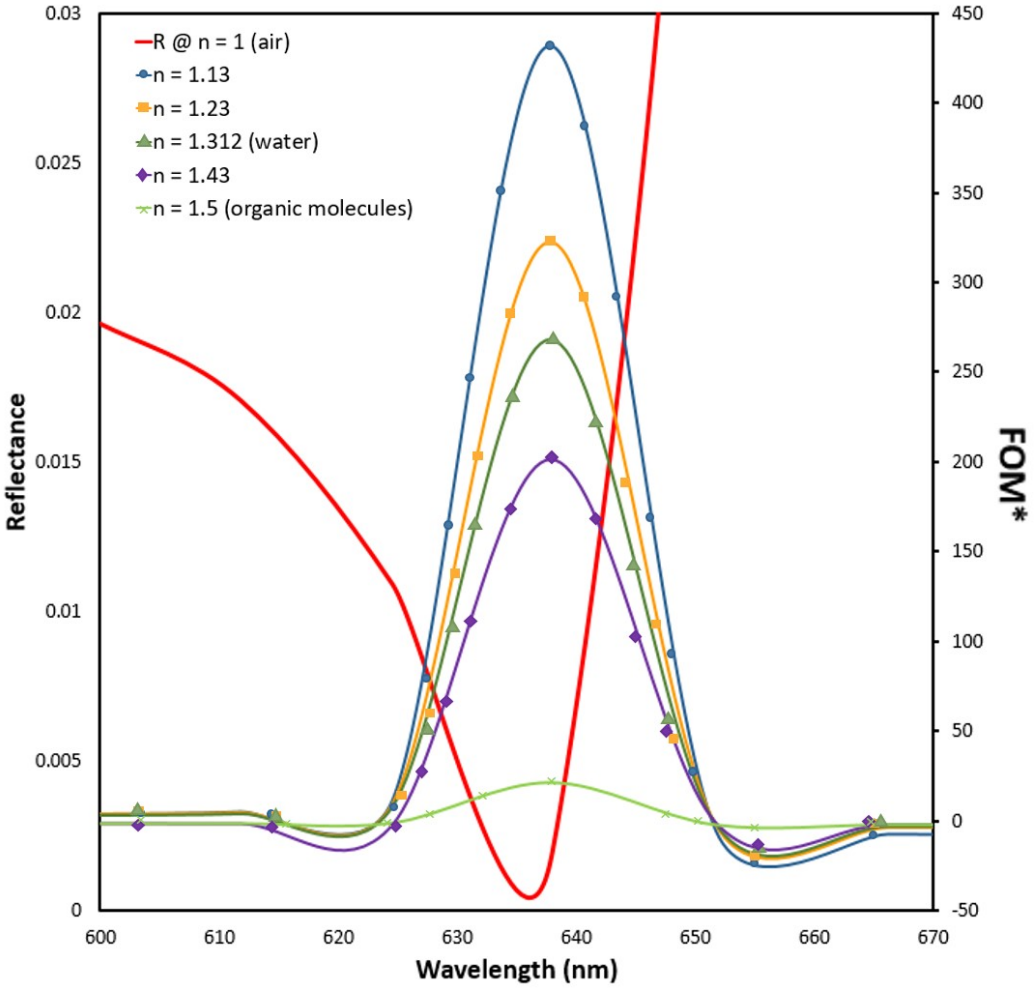


Figure 4.14: Calculated FOM* as a function of wavelength in various media at ε=120%.

As previously mentioned, LSP resonance strongly depends on the geometrical parameters, shape, and surrounding dielectric environment of nanostructures [30]. So, we have simultaneously adjusted both periodicity (by altering the strain level) and the surrounding medium to open a route toward RI sensing enabling the detection of small concentrations of target molecules. Furthermore, this sensor could be a practical solution for crucial challenges of the classical sensing methods like low resolution, limited dynamic range, and low efficiency of the reconfiguration process [31, 32].

Unlike common LSPR sensors which detect the spectral resonance shift upon

a RI change of various surrounding medium, the proposed one measures a relative intensity change at a fixed wavelength induced by a RI change. A figure of merit (FOM*), which is introduced by J. Becker et al. [33], can be calculated according to Equation 4.4.1, where dI(λ)/I(λ) is the relative intensity change at a fixed wavelength induced by a RI change dn. λo is chosen where FOM* has a maximum value.

$$FOM^* = max \left| \frac{dI(\lambda)/d_n(\lambda)}{I(\lambda)} \right| \tag{4.4.1}$$

The variation of the RI of the surrounding medium in the extreme mode gives rise to nonzero reflectance and consequently provides the extremely sensitive detection of the intensity variation in reflectance at a specific frequency. Figure 4.14 presents the calculated FOM* of the sensor in various media with different refractive indexes ranging from n=1 (air) to n=1.5 (common organic molecules). The reflectance reaches a minimum of 0.001 at 637 nm in air. The maximum achieved FOM* is around 431 by measuring the variation of intensity with different dielectric materials (n=1 (air) and n=1.13). Moreover, the FOM* for water (n=1.312) is almost four times larger than the previous sensor, where a passive plasmonic sensor was designed to operate in the near-infrared regime. This design was developed by placing gold disks with a diameter of 352 nm and a thickness of 20 nm on a multi-layered structure consisting of MgF_2 spacer, gold, and glass [34]. In the calculation of the FOM*, the reflectance difference between various mediums (I_{medium} (λ)) and air (I_{air} (λ)) was taken for the derivative in Equation 4.4.1.

Switchable Reflector/Absorber

Last but not least, the potential of developed ultra-stretchable metasurfaces in the design of a switch from a high reflector to a high absorber in the visible regions was theoretically demonstrated. The proposed metasurface operates in a wide range of strains, in which the exciting ED and MD are coupled, and form hybridised modes. The high absorption and high reflection arise from destructive multipolar interferences in the opposite directions, respectively through the generalised Kerker effect [35]. The energy level of the hybridised modes can be

dynamically controlled by the strain level, allowing these metasurfaces to have anisotropic optical responses.

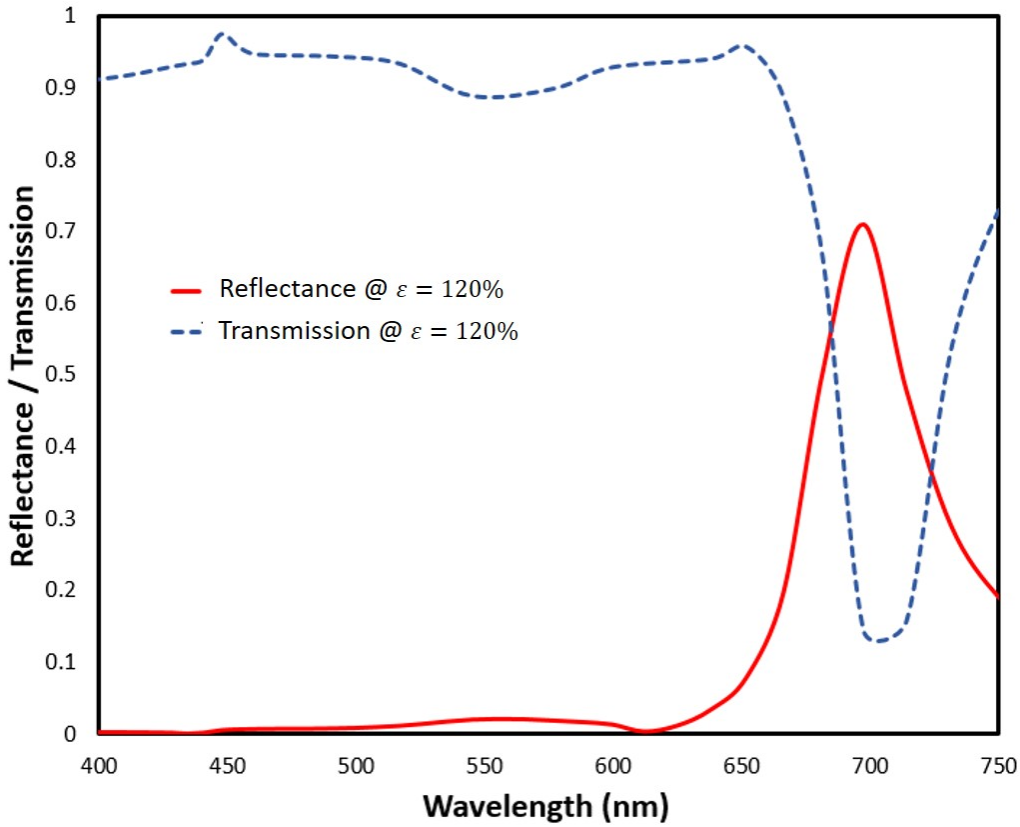


Figure 4.15: a) Reflection and Transmission spectra of metasurface at the different strain levels. The unit cell is comprised of Au nanodisk with a height of 25 and a diameter of 200 nm.

Figure 4.15 shows the reflectance and transmission of a new stretchable metasurface based on a square array of Au nanodisks at extreme mode (ε=120%). Here, the unit cell is comprised of an Au nanodisk with a height of 25 and a diameter of 200 nm. As shown in Figure 4.15, the proposed design can support reflection/absorbtion occurring at different wavelengths through multipolar interference. More interestingly, the operational wavelength can be altered by tuning the strain level with no change in the shape and dimension of resonators.

In the extreme mode (cf. Figure 4.15), the maximum reflectance occurs in λ = 695nm where the transmission is almost 20%. However, when the strain level hits the maximum (ε=120%), the transmission reaches 96%. This huge difference can be used to design an absorption switch. In addition, they can be fine-tuned

to the point where the wavelengths of reflection and absorption nearly coincide. So, the metasurface switches from a highly reflective state to a highly transparent one. For instance, a low reflection (lower than10%) was observed in a wide wavelength range of 400–651 nm with a minimum of 0.05 at 443 nm. However, the transmission value changes in the range of 90% to 100% in the same range with a maximum of 97.6% at 447 nm. As a result, strong transparency is supported in this wavelength range.

4.5 Conclusion

In this result chapter, the focus shifted towards overcoming the limitations of conventional mechanically tunable metasurfaces by introducing two main geometries: pillar-based and kirigami-inspired patterns. These designs aim to achieve ultra-stretchable metasurfaces, pushing the maximum stretchability from the typical 10-35% range to an impressive 120%. The key strategy involves relocating stress concentration from the interface of rigid-soft materials to safer parts of the structure. Mechanical simulations were conducted to explore alternative arrangements of resonators, facilitating a combination of in-plane and out-of-plane deformation, leading to 2D-to-3D shape transformations.

In the pillar-based structures, it was found that the height of pillars significantly influences stress redistribution. Increasing pillar height from 0 to 150 nm can reduce the stress ratio by almost 80%. Conversely, in kirigami-inspired geometries, a minimum width ratio (w/d $\geq$ 0.05) is essential to induce out-of-plane deformation in the interconnections, otherwise, the ribbons deform within the plane (2D). The optimised structure with maximum stretchability underwent optical simulation, revealing a peak shift of 143 nm (from 508 nm to 651 nm), nearly twice that of the normal geometry. This extensive tunability enhances the potential of these geometries in various applications discussed, such as rewritable metadevices, plasmonic sensors with high accuracy, and Switchable reflector/absorber systems.

Furthermore, optical simulations were utilised for generating datasets to train ma-

chine learning algorithms discussed in Chapter 5. Parameter optimisation resulted in a minimal dataset for training the geometry optimisation network, which was then applied for structural colour design and optimisation.

Bibliography

[1] A. Romeo, Q. Liu, Z. Suo, and S. P. Lacour, "Elastomeric substrates with embedded stiff platforms for stretchable electronics," *Applied Physics Letters*, vol. 102, no. 13, 2013.

[2] T. Badloe, J. Lee, J. Seong, and J. Rho, "Tunable metasurfaces: the path to fully active nanophotonics," *Advanced Photonics Research*, vol. 2, no. 9, p. 2000205, 2021.

[3] H.-S. Ee and R. Agarwal, "Tunable metasurface and flat optical zoom lens on a stretchable substrate," *Nano letters*, vol. 16, no. 4, pp. 2818–2823, 2016.

[4] O. A. Abdelraouf, Z. Wang, H. Liu, Z. Dong, Q. Wang, M. Ye, X. R. Wang, Q. J. Wang, and H. Liu, "Recent advances in tunable metasurfaces: materials, design, and applications," *ACS nano*, vol. 16, no. 9, pp. 13 339–13 369, 2022.

[5] T. Lu, C. Wang, C. Hsiao, and P. Lee, "Tunable nanoblock lasers and stretching sensors," *Nanoscale*, vol. 8, no. 37, pp. 16 769–16 775, 2016.

[6] K. Bashandeh, J. Lee, Q. Wu, Y. Li, X. Wang, Y. Shi, X. Guo, Y. Huang, J. A. Rogers, and A. A. Polycarpou, "Mechanics and deformation of shape memory polymer kirigami microstructures," *Extreme Mechanics Letters*, vol. 39, p. 100831, 2020.

[7] B. Kim, S. B. Lee, J. Lee, S. Cho, H. Park, S. Yeom, and S. H. Park, "A comparison among neo-hookean model, mooney-rivlin model, and ogden model for chloroprene rubber," *International Journal of Precision Engineering and Manufacturing*, vol. 13, pp. 759–764, 2012.

[8] Y. Zheng, K. Chen, W. Yang, L. Wu, K. Qu, J. Zhao, T. Jiang, and Y. Feng, "Kirigami reconfigurable gradient metasurface," *Advanced Functional Materials*, vol. 32, no. 5, p. 2107699, 2022.

[9] S. Chen, J. Chen, X. Zhang, Z.-Y. Li, and J. Li, "Kirigami/origami: unfolding the new regime of advanced 3d microfabrication/nanofabrication with "folding"," *Light: Science & Applications*, vol. 9, no. 1, p. 75, 2020.

[10] A. Ghasemi, R. Fang, D. A. Zeze, and M. K. Hedayati, "Ultra-stretchable active metasurfaces for high-performance structural color," *AIP Advances*, vol. 13, no. 8, 2023.

[11] Y. Zhang, Z. Yan, K. Nan, D. Xiao, Y. Liu, H. Luan, H. Fu, X. Wang, Q. Yang, J. Wang *et al.*, "A mechanically driven form of kirigami as a route to 3d mesostructures in micro/nanomembranes," *Proceedings of the National Academy of Sciences*, vol. 112, no. 38, pp. 11 757–11 764, 2015.

[12] J. Chen, Z. Li, X. Zhang, J. Xiao, and Q. Gong, "Submicron bidirectional all-optical plasmonic switches," *Scientific reports*, vol. 3, no. 1, p. 1451, 2013.

[13] E. D. Palik, *Handbook of optical constants of solids*. Academic press, 1998, vol. 3.

[14] E. Prodan, C. Radloff, N. Halas, and P. Nordlander, "A hybridization model for the plasmon response of complex nanostructures," in *APS March Meeting Abstracts*, vol. 2004, 2004, pp. P34–005.

[15] S. Chen, W. Wei, Z. Liu, X. Liu, S. Feng, H. Guo, and J. Li, "Reconfigurable nano-kirigami metasurfaces by pneumatic pressure," *Photonics Research*, vol. 8, no. 7, pp. 1177–1182, 2020.

[16] W. Li, K. Ren, and J. Zhou, "Aluminum-based localized surface plasmon resonance for biosensing," *TrAC Trends in Analytical Chemistry*, vol. 80, pp. 486–494, 2016.

[17] M. L. Tseng, J. Yang, M. Semmlinger, C. Zhang, P. Nordlander, and N. J. Halas, "Two-dimensional active tuning of an aluminum plasmonic array for full-spectrum response," *Nano letters*, vol. 17, no. 10, pp. 6034–6039, 2017.

[18] Y. Lee, J. Kwon, J. Lim, W. Shin, S. Park, E. Hwang, J. Shin, H. Cho, J. Jung, H.-J. Kim *et al.*, "Digital laser micropainting for reprogrammable optoelectronic applications," *Advanced Functional Materials*, vol. 31, no. 1, p. 2006854, 2021.

[19] N. I. Zheludev, "Obtaining optical properties on demand," *Science*, vol. 348, no. 6238, pp. 973–974, 2015.

[20] N. I. Zheludev and E. Plum, "Reconfigurable nanomechanical photonic metamaterials," *Nature nanotechnology*, vol. 11, no. 1, pp. 16–22, 2016.

[21] J.-Y. Ou, E. Plum, J. Zhang, and N. I. Zheludev, "Giant nonlinearity of an optically reconfigurable plasmonic metamaterial," *Advanced Materials*, vol. 28, no. 4, pp. 729–733, 2016.

[22] Q. Wang, E. T. Rogers, B. Gholipour, C.-M. Wang, G. Yuan, J. Teng, and N. I. Zheludev, "Optically reconfigurable metasurfaces and photonic devices based on phase change materials," *Nature photonics*, vol. 10, no. 1, pp. 60–65, 2016.

[23] M. Z. Alam, S. A. Schulz, J. Upham, I. De Leon, and R. W. Boyd, "Large optical nonlinearity of nanoantennas coupled to an epsilon-near-zero material," *Nature Photonics*, vol. 12, no. 2, pp. 79–83, 2018.

[24] G. Qu, W. Yang, Q. Song, Y. Liu, C.-W. Qiu, J. Han, D.-P. Tsai, and S. Xiao, "Reprogrammable meta-hologram for optical encryption," *Nature communications*, vol. 11, no. 1, pp. 1–5, 2020.

[25] Z. Sun, J. Zhou, and R. Ahuja, "Unique melting behavior in phase-change materials for rewritable data storage," *Physical review letters*, vol. 98, no. 5, p. 055505, 2007.

[26] K. Dong, S. Hong, Y. Deng, H. Ma, J. Li, X. Wang, J. Yeo, L. Wang, S. Lou, K. B. Tom *et al.*, "A lithography-free and field-programmable photonic metacanvas," *Advanced Materials*, vol. 30, no. 5, p. 1703878, 2018.

[27] S. Abdollahramezani, O. Hemmatyar, M. Taghinejad, H. Taghinejad, Y. Kiarashinejad, M. Zandehshahvar, T. Fan, S. Deshmukh, A. A. Eftekhar, W. Cai *et al.*, "Dynamic hybrid metasurfaces," *Nano letters*, vol. 21, no. 3, pp. 1238–1245, 2021.

[28] Y. Zhang, J. B. Chou, J. Li, H. Li, Q. Du, A. Yadav, S. Zhou, M. Y. Shalaginov, Z. Fang, H. Zhong *et al.*, "Broadband transparent optical phase change materials for high-performance nonvolatile photonics," *Nature communications*, vol. 10, no. 1, pp. 1–9, 2019.

[29] M. Wuttig and N. Yamada, "Phase-change materials for rewriteable data storage," *Nature materials*, vol. 6, no. 11, pp. 824–832, 2007.

[30] M. E. Stewart, C. R. Anderton, L. B. Thompson, J. Maria, S. K. Gray, J. A. Rogers, and R. G. Nuzzo, "Nanostructured plasmonic sensors," *Chemical reviews*, vol. 108, no. 2, pp. 494–521, 2008.

[31] J. B. Lassiter, J. Aizpurua, L. I. Hernandez, D. W. Brandl, I. Romero, S. Lal, J. H. Hafner, P. Nordlander, and N. J. Halas, "Close encounters between two nanoshells," *Nano letters*, vol. 8, no. 4, pp. 1212–1218, 2008.

[32] T.-H. Park, N. Mirin, J. B. Lassiter, C. L. Nehl, N. J. Halas, and P. Nordlander, "Optical properties of a nanosized hole in a thin metallic film," *Acs Nano*, vol. 2, no. 1, pp. 25–32, 2008.

[33] J. Becker, A. Trügler, A. Jakab, U. Hohenester, and C. Sönnichsen, "The optimal aspect ratio of gold nanorods for plasmonic bio-sensing," *Plasmonics*, vol. 5, no. 2, pp. 161–167, 2010.

[34] N. Liu, M. Mesch, T. Weiss, M. Hentschel, and H. Giessen, "Infrared perfect absorber and its application as plasmonic sensor," *Nano letters*, vol. 10, no. 7, pp. 2342–2348, 2010.

[35] T. Shibanuma, S. A. Maier, and P. Albella, "Polarization control of high transmission/reflection switching by all-dielectric metasurfaces," *Applied Physics Letters*, vol. 112, no. 6, p. 063103, 2018.

Chapter 5

Result Chapter: Machine Learning

5.1 Chapter Overview

This chapter presents the results of machine learning (ML) algorithms. The initial section focuses on the optimised neural network (NN) employed for the design and optimisation of kirigami-inspired ultra-stretchable metasurfaces as introduced in Chapter 4. Detailed discussions include the complexity of data generation, the training process, and the model's performance. The accuracy of the NN is checked for the design and optimisation of a structural colour as a practical application of ultra-stretchable metasurfaces where one single geometry can provide a broad range of colours in various levels of strain.

Unlike the first network that fixed the properties of the material by considering a thin layer of one pure material for the optimisation of a structural colour system, the subsequent part of this chapter focuses on the optimisation of materials in a

more complex structure. Here, a six-layer multilayer structure is considered to develop a universal algorithm for the inverse design (ID) and material discovery for a custom amorphous metasurface. Detailed information on data-generating techniques and an analysis of the accuracy of these networks are discussed. The ID network was trained to predict refractive indices of materials as the most significant material properties, enabling the design of multilayer metamaterials with customised optical properties with a 10^4 order faster than the simulation technique. The material discovery network simultaneously explores an infinite material space to propose new compositions for creating amorphous metamaterials with unique properties. In addition, a comparison study is done by comparing the model created in this thesis with other networks described in existing literature.
The last section of this chapter reports the details of the NN algorithm that was developed to optimise the structure in the third domain which focuses on free-form geometry. This network was trained with a small dataset generated by simulation on a common geometry with a random free-form shape which is coated with two random materials. Two forward and inverse design networks were trained in this part. The former was used for data augmentation, and the latter was tasked to search an infinite design space to propose the geometry to achieve a custom optical property. This comparative analysis highlights the significant potential and effectiveness of the developed networks, establishing them as valuable additions to the field of metamaterials and metasurfaces.

5.2 Optimisation of Geometric Design Parameters

The optimisation of the kirigami-based geometry (introduced in Chapter 4) can unlock a new generation of active structural colour. Traditional design optimisation methods such as numerical simulations typically demand either laborious data collection or intricate reconstruction algorithms for data processing. This challenge makes them inadequate for complex calculations where adjustment of both mechanical and optical variables is necessary.
The research methodology involved the development of a fully connected deep

neural network (DNN) designed to predict and optimise structural colour parameters. This encompassed tuning key geometrical parameters, including the diameter of the Al disk (D), the thickness of the resonator (H), and the mechanically induced deformation (ε), as outlined in Figure 5.1. The DNN was initially trained using a limited dataset generated through a validated finite element simulation model. This training process aimed to teach the network the complex nonlinear relationship between displacement induced by strain releases in a kirigami-inspired geometry and the corresponding design parameters, serving as a proof of concept. Subsequently, the trained network was tasked with generating a more extensive dataset to reduce parameter sweeping steps to a reasonable time. As reported in Table 5.1, the input values for this network are H, D, and ε, and the output values are two structural colour parameters (X, Y). The augmented dataset was then used to identify the optimised **"single geometry"** capable of providing the widest colour gamut and an efficient structural colour system.

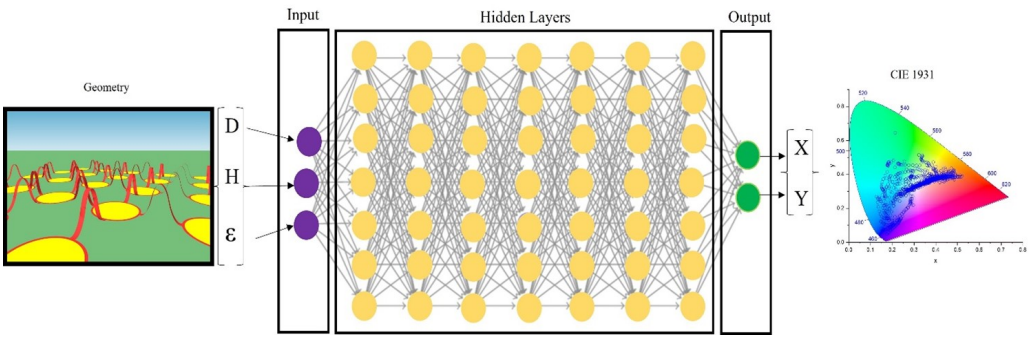


Figure 5.1: Schematic of a fully connected DNN which extracts the complex relationship between mechanical and geometrical parameters (D: diameter, H: height, and ε: strain level) and structural colour parameters (X, Y) plotted on International Commission on Illumination (CIE) 1931 gamut. Reprinted by permission from AIP Publishing [1], Copyright (2023).

Table 5.1 presents the range of variables considered for both the training and optimisation phases in this study. The small dataset comprising 2100 data points was extracted through a combination of optical and mechanical simulations, enabling the generation of a diverse array of colours for training the network. This small-scale data was carefully generated to guarantee sufficient variation for network training, while simultaneously encapsulating enough comparable data to en-

able complete input-to-output mapping. As discussed in Chapter 4, the proposed kirigami-inspired pattern operated within a maximum strain level of 120%, establishing this as the upper limit for the strain with a step range of 20%. Initial findings showed the significance of the resonator's thickness over the diameters, prompting the choice of steps at 10 and 20nm, respectively.

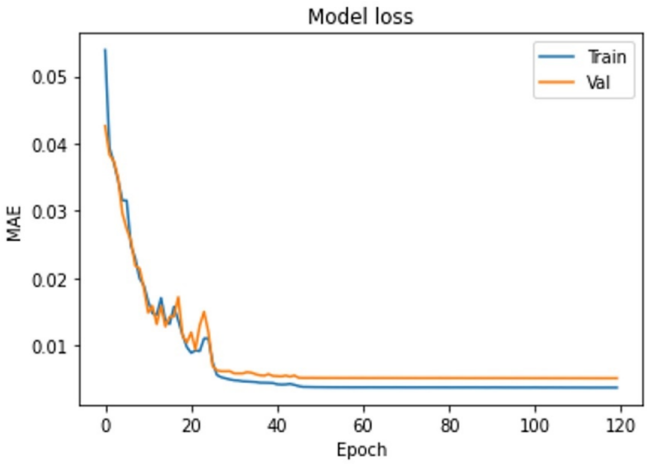
Table 5.1: Range of variables for training DNN, and optimisation

Variables	Training			Optimisation		
	Range	Step	No. of samples	Range	Step	No. of samples
Strain (ε) %	0-120	20	7	0-120	10	13
Height of Al (H) (nm)	5-145	10	15	5-145	5	29
Diameter of Disk (D) (nm)	10-295	20	20	10-295	5	58
Total	2100			21866		

For the training of the network, the dataset was divided into three distinct subsets: 80% for training, 10% for validation, and 10% for testing purposes. The hyperparameters governing the network architecture, including the number of layers, the number of neurons in each layer, and the choice of activation functions, were chosen through a process of trial and error. The optimised network architecture is composed of four hidden layers, with each layer using Leakyrectified linear unit (ReLU) activation functions to connect batch-normalised input values to the desired output. The final layer incorporates a sigmoid activation function to constrain prediction values within the range of 0 and 1. The mean absolute error (MAE) was selected as the loss function, and the Adam optimiser was employed for training. Notably, the initial learning rate was set at 0.01 and logarithmically reduced whenever the validation loss plateaued for more than 10 epochs, with a minimum threshold of 1×10^{-6}. This configuration ensures a robust and efficient training process, allowing the network to adapt its parameters effectively during optimisation.

In Figure 5.2 (a), the convergence of both training and validation curves is evident at a significantly higher level. This convergence results in a MAE of 0.0051 for the test dataset and 0.0038 for the training dataset. This accuracy is accompanied by a remarkable order of magnitude decrease in computational time, specifically by a factor of 10^4, when compared to finite element method (FEM) simulations traditionally employed in conventional design workflows.

a



b

No.	Input values			Target Color	Predicted Color	Absolute Error Value (1931 CIE)	
	D (nm)	H (nm)	Strain (%)			Δ CIE X	Δ CIE Y
1	145	145	120			0.0245	0.0189
2	25	160	80			0.0043	0.0039
3	45	205	20			0.0094	0.0214
4	75	205	60			0.0156	0.0103
5	75	235	40			0.0063	0.0089
6	15	190	80			0.1102	0.0687
7	85	220	80			0.0016	0.0039
8	65	220	60			0.0012	0.0021
9	105	205	80			0.0109	0.0084
10	115	295	0			0.0103	0.0114
11	145	160	0			0.0041	0.0035
12	145	70	0			0.0016	0.0048
13	105	265	80			0.0042	0.0063
14	145	280	60			0.121	0.143

Figure 5.2: a) The learning curve of the DNN model for training and validation cost as a function of epochs, b) Comparison of the prediction accuracy of 14 random target structural colours. Reprinted by permission from AIP Publishing [1], Copyright (2023).

A qualitative analysis is presented in Figure 5.2(b) which illustrates the accuracy of the model in predicting 14 random modes. The model accurately predicts a diverse range of colours. The outcomes of the qualitative and quantitative analysis demonstrate the model's potential in accurately interpolating the relationship between the geometrical parameters of the optomechanical metamaterials and the resulting colour across various levels of strain. We also considered more 50 random samples to check the accuracy of the network by comparing it with the data calculated by the simulation. In all samples, the colours were predicted accurately. The trained model was then used to expand the dataset by predicting the structural

colour parameters (X, Y) for 21866 samples as presented in Table 5.1.

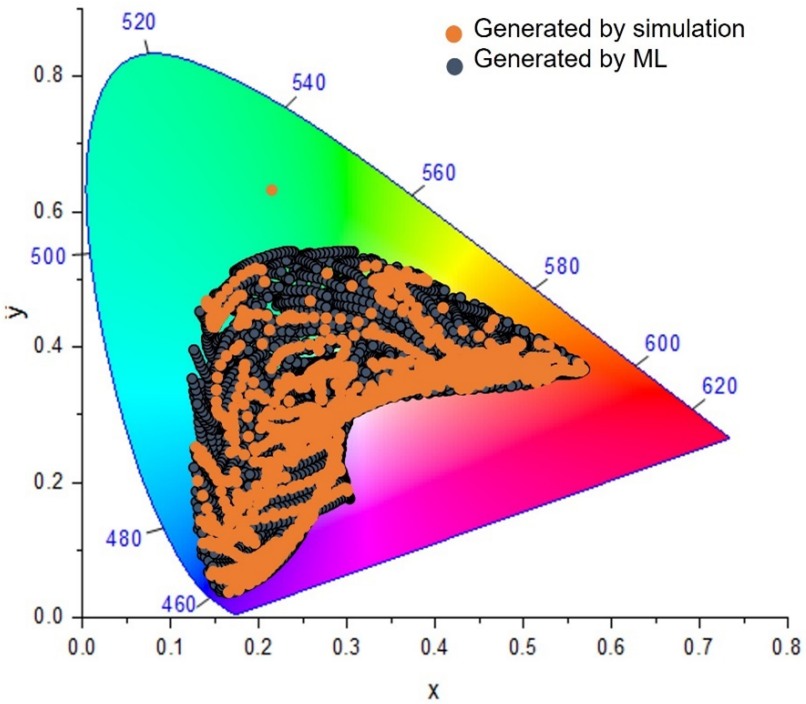


Figure 5.3: CIE chromaticity diagrams of the samples derived from FEM (orange) and ML (grey).Reprinted by permission from AIP Publishing [1], Copyright (2023).

Figure 5.3 presents the scattering patterns of samples obtained through FEM and ML. The orange dots on the plot represent the 2100 samples employed for training the model, while the grey dots correspond to the 21866 geometries generated by the model. As demonstrated, the sweeping of key parameters with smaller steps enables the identification of previously missing geometries, leading to a wider colour gamut. This effect is particularly noticeable in the green and blue regions, highlighting the efficacy of the ML approach in discovering novel and diverse colour possibilities in the structural design space.

The optimised single geometry is characterised by an Al nanodisk with a thickness of 130 nm and a diameter of 200 nm. This configuration provides the broadest areal colour space coverage, encompassing 27.65% of the standard RGB (sRGB) space (cf. Figure 5.4(a)). The scattering pattern of structural colour parameters (X, Y) for all examined samples is available in video format within the supplementary document of the published work [1].

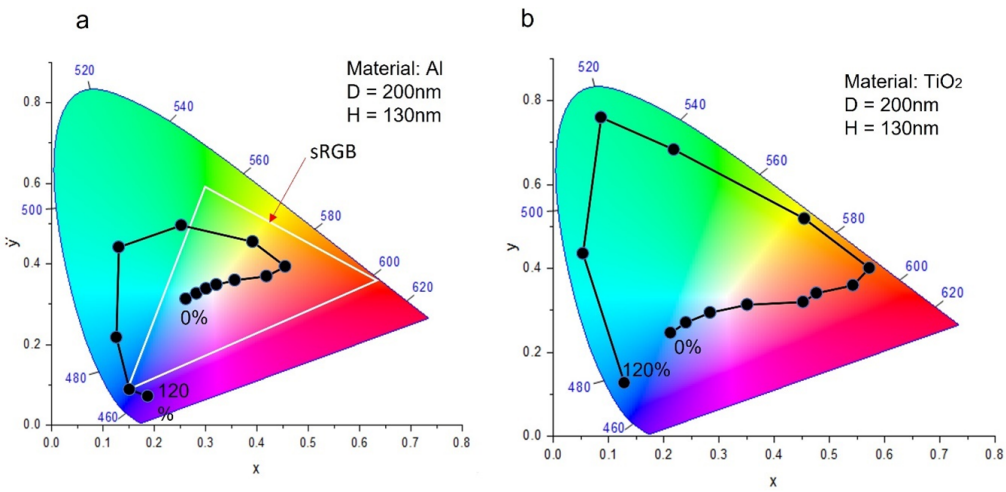


Figure 5.4: a) The optimised geometry, featuring an Al nanodisk with a thickness (H) of 130 nm and a diameter (D) of 200 nm, showcases a notable 27.65% coverage of the sRGB colour space. b) CIE 1931 chromaticity diagrams of the TiO_2 nanodisks with optimised geometry and varied strain (H=130 nm, D= 200 nm). Reprinted by permission from AIP Publishing [1], Copyright (2023).

The universal single-geometry demonstrates significant potential for expanding the colour coverage within dielectric-based structural colour systems. By substituting the plasmonic component with a high-dielectric material, such as TiO_2, a diverse range of colours is achievable in the CIE 1931 chromaticity diagram. Figure 5.4(b) visually represents the versatility of the proposed kirigami-inspired design, showcasing a comprehensive colour palette spanning from white to red, green, and blue. This observation substantiates the universal characteristics of the design, indicating its applicability to both plasmonic and Mie resonators.

5.3 AI-assisted Materials Development

A comprehensive design of an active metasurface needs consideration of both geometric parameters and structural material properties. In the previous phase, the geometric parameters of the proposed mechanically tunable metasurface were optimised for a structural colour as a proof of concept. However, during this process, the material properties remained fixed, relying on conventional materials typically used in metasurface development. This section shifts focus to optimising material properties, specifically the refractive index (RI) values, within a more complex

multilayer thin film structure. By integrating these optimised material properties with previously optimised geometric parameters, the potential of the proposed active metasurface can be significantly expanded.

The general structure is a six-layer metamaterial, where layers are systematically composed of a pair of randomly selected materials. The materials are chosen from a material library comprising 33 distinct materials categorised into four main groups, as outlined in Table 5.2. Although the fabrication possibility of various combinations of all main classes of materials was experimentally reported [2], the possibility of fabricating random samples was not considered in dataset generation. This structured configuration is situated on an infinite glass substrate and exposed to normally incident light waves. The thickness of each layer varies in the range of 10 to 100 nm. The resulting reflectance spectrum is calculated within the visible light wavelength range, as schematically illustrated in Figure 5.5 (a, b). This comprehensive approach allows for the exploration of a diverse array of material combinations and layer thicknesses to understand the optical characteristics of the multilayer metamaterial under consideration.

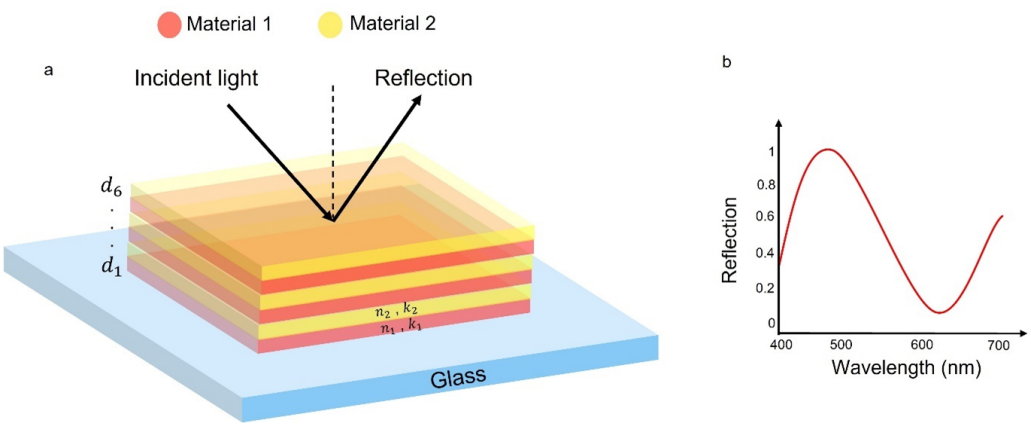


Figure 5.5: a) Schematic of the randomly designed thin-film structures consisting of two different materials, which are arranged periodically with each other. d1 to d6 corresponds to the thickness of the first and last layer respectively. b) Reflectance spectra calculated within the visible light wavelength range.

5.3.1 Data Generation

The selection of materials for fabricating multilayer thin films is diverse and depends on the specific application requirements. Materials are systematically clas-

sified into four main groups, each designated by a two-digit binary code (cf. Table 5.2):

1. Metals (00): Metals are commonly used in multilayer thin metamaterials due to their high reflectivity and electrical conductivity. Silver (Ag), aluminium (Al), and gold (Au) are common choices for applications like mirrors, optical filters, and electrodes [3].

2. Dielectrics and Oxides (01): Dielectric materials find extensive use in multilayer thin films owing to their high refractive indices, low absorption coefficients, and minimal scattering losses. Silicon dioxide (SiO_2), titanium dioxide (TiO_2), and tantalum pentoxide (Ta_2O_5) are examples employed in mirrors, filters, and antireflection coatings [4].

3. Fluoride-Based Materials (10): This group offers unique properties, including a high refractive index, low dispersion, and excellent transparency in the UV and IR regions. For instance, CaF_2-based multilayer thin films find applications in antireflection coatings, beam splitters, optical filters, and mirrors [5, 6].

4. Polymers (11): The combination of flexibility, cost-effectiveness, ease of processing, and the ability to tailor properties makes polymers a versatile and widely used choice in the fabrication of multilayer thin film structures across various industries. Polyimide, polyethylene terephthalate (PET), and polycarbonate are examples used to create coatings, barriers, and adhesives [7].

Table 5.2: Classified materials library with a binary coding system

Metals	**Oxides**	**Fluorides**	**Polymers**
00	**01**	**10**	**11**
Ag	$KTaO_3$	CaF_2	PC
Au	MoO_3	BaF_2	PMMA
Bi	SiO_2	MgF_2	PET
Co	Ta_2O_5		PVC
Fe	$Al2O_3$		PS
Mg	GeO_2		PVA
Mn	MgO		
Mo	$Nb2O_5$		
Ni	$Sc2O_3$		
Ta	TiO_2		
Zn	ZnO		
Pt	ZrO_2		

As shown in Figure 5.6, the methodology has two interconnected neural networks: the "inverse design" and "materials discovery" networks. These networks were trained using distinct datasets to capture geometrical and material parameters. The initial network, the inverse design, was trained with a dataset generated by the transfer matrix method (TMM). Initially, 70,000 random six-layer metamaterial structures were created, with layer thicknesses constrained to multiples of 10 nm. A 36-point discretisation of the reflectance spectrum is represented by

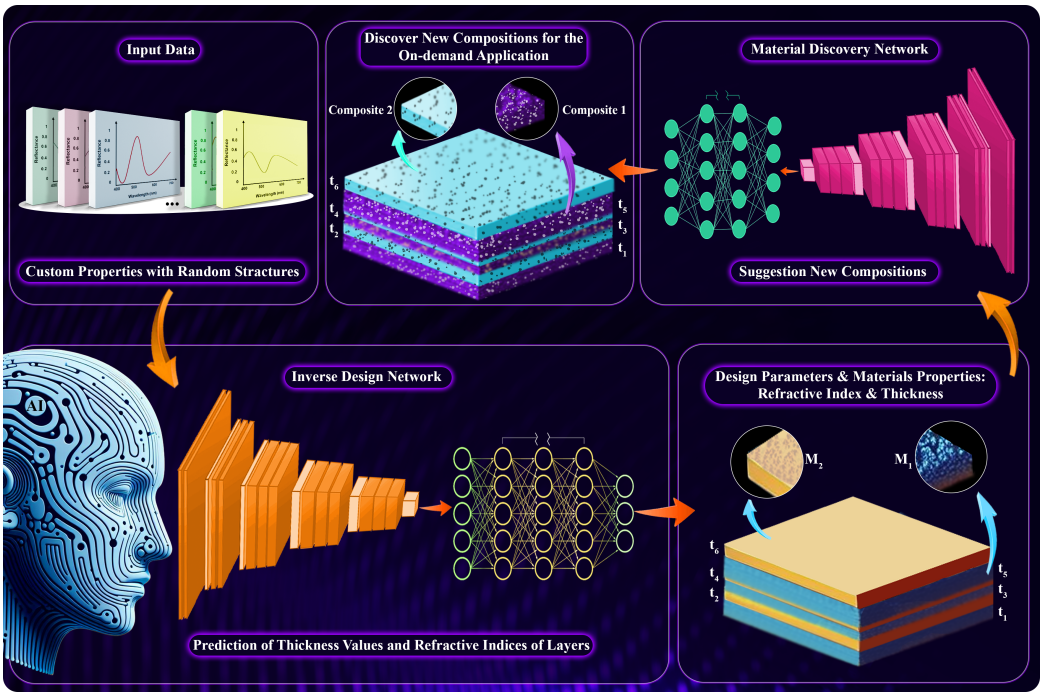


Figure 5.6: Schematic illustration of the process flow.

the input array $R^{36} = \{R_{400}, R_{410}, \ldots, R_{750}\}$. However, the output array has 150 elements that represent the materials and geometry of a layered construction. It provides details on each layer's thickness and reflectance at various wavelengths:

- The first six elements of the array correspond to the thickness of each layer
- The following elements, after thicknesses in the output array, represent the refractive indices of each layer for wavelengths ranging from 400 to 750 nm with a step size of 10 nm. Extinction coefficients represented by K_i, which specifies the order of the corresponding material, and RIs, represented by N_i, were extracted from an online RI database for wavelengths ranging from

400 to 750 nm. The refractive indices at each wavelength are specifically listed in the following order: N_1, K_1, N_2, K_2, N_1 and K_1. The RI and extinction coefficient of the first layer are represented by N_1 and K_1, respectively, and the RI and extinction coefficient of the second layer are represented by N_2 and K_2, which are chosen from every 36 points in the input.

However, the second network—materials discovery—was trained with a dataset produced by effective medium approximation (EMA). Here, the considered composites were simplified to consist of just two materials to streamline the problem. This method is a theoretical model in optics that is used to describe the behaviour of light in materials that are composed of a mixture of two or more different types of particles or structures. EMA assumes that the material can be approximated as a homogeneous medium with effective optical properties that depend on the properties of the individual components and their spatial arrangement. There are several approximations, but two of the most common ones are the Maxwell–Garnett (MG) approximation and the Bruggeman approximation.

- **Maxwell-Garnett Approximation:**

  The MG approximation is a widely used effective medium approximation in optics and material science. It is used to calculate the effective optical properties of composite materials consisting of a host medium and small, homogeneously distributed inclusions with a different dielectric constant. The MG approximation assumes that the inclusions are small compared to the wavelength of light and that their spatial arrangement is random [8].

  The MG approximation provides an explicit formula for the effective dielectric constant of the composite material, which can be related to the refractive index of the material through the relation $n_{eff}^2 = \varepsilon_{eff}$. Therefore, the MG approximation can be used to calculate the effective RI of a composite material. The effective dielectric constant in the MG approximation is given by:

$$\varepsilon_{eff} = \varepsilon_m[1 + (\varepsilon_i - \varepsilon_m)\varphi/(\varepsilon_i + 2\varepsilon_m)] \tag{5.3.1}$$

  where ε_m and ε_i are the dielectric constants of the matrix and the inclusion,

respectively, and φ is the volume fraction of the inclusions.

While the MG approximation is a powerful tool for calculating the effective optical properties of composite materials, it has some limitations. For example, it assumes that the inclusions are small compared to the wavelength of light and that they are randomly distributed in the host medium. In addition, the MG approximation is only valid for low-volume fractions of inclusions, typically less than 10% [9]. For higher volume fractions, the MG approximation can overestimate the effective dielectric constant, and more complex models, such as the Bruggeman approximation, may be needed.

- **Bruggeman Approximation:**

  Bruggeman approximation is another commonly used theory for estimating the effective RI of a composite material, which is a material made up of two or more different components with different refractive indices. The effective RI is the one that a composite material would have if it were a homogeneous material with the same optical properties [10].

  Bruggeman approximation assumes that the composite material is randomly distributed and that the size of the different materials in the composite is much smaller than the wavelength of the light being used. The effective dielectric constant of the material can be calculated using the Bruggeman approximation as follows:

$$(\varepsilon_eff - \varepsilon_i)/(\varepsilon_eff + 2\varepsilon_i) = \varphi * (\varepsilon_i - \varepsilon_m)/(\varepsilon_i + 2\varepsilon_m) \tag{5.3.2}$$

  where ε_{eff} is the effective dielectric constant of the material, ε_i and ε_m are the dielectric constants of the inclusion and matrix respectively, and φ is the volume fraction of the inclusion. The effective refractive index, (n_{eff}), can then be calculated from the effective dielectric constant using the relation $n_{eff}^2 = \varepsilon_{eff}$.

The advantages of the Bruggeman approximation are that it is relatively simple to use and does not require detailed information about the microstructure of

the composite material. It is also a useful tool for predicting the optical properties of composite materials, which can be important in a variety of applications.

However, the Bruggeman approximation has some limitations. For example, it assumes that the composite material is randomly distributed and that the size of the different materials in the composite is much smaller than the wavelength of the light being used, which may not be true for all composite materials. Additionally, the approximation may not be accurate for materials with large volume fractions of one of the materials or materials with complex microstructures.

Using two materials from a material library, we created 70,000 random datasets for this investigation. We then used these two models to determine the effective RI for various volume fractions. Since the MG model implies that the composite is composed of scattered spherical inclusions in a homogenous medium, it was applied to volume fractions less than 0.5%. We employed the Bruggeman model, which accounts for the inclusions' random distribution and their interfacial interactions with the surrounding medium, for volume fractions ranging from 0.5% to 30%. Although this model can explain a greater range of volume fractions than the MG model, it is not without limits, especially in cases when the inclusions are very irregular or exhibit notable differences in size and shape. The second network then used the complex RI of each layer that the ID network had predicted as its input. The result was two materials (matrix and inclusion) together with their corresponding volume fractions (F_1 and F_2). The data imbalance was checked by examining the distribution of class labels within the dataset. This involved analysing the frequency of each class to determine if there were disproportionately more instances of certain classes compared to others. Additionally, common techniques such as visualisations, class distribution plots, and calculating class frequencies were considered to gain insights into the imbalance. These assessments helped us understand if any class was underrepresented or overrepresented, enabling us to address potential biases during model training and evaluation.

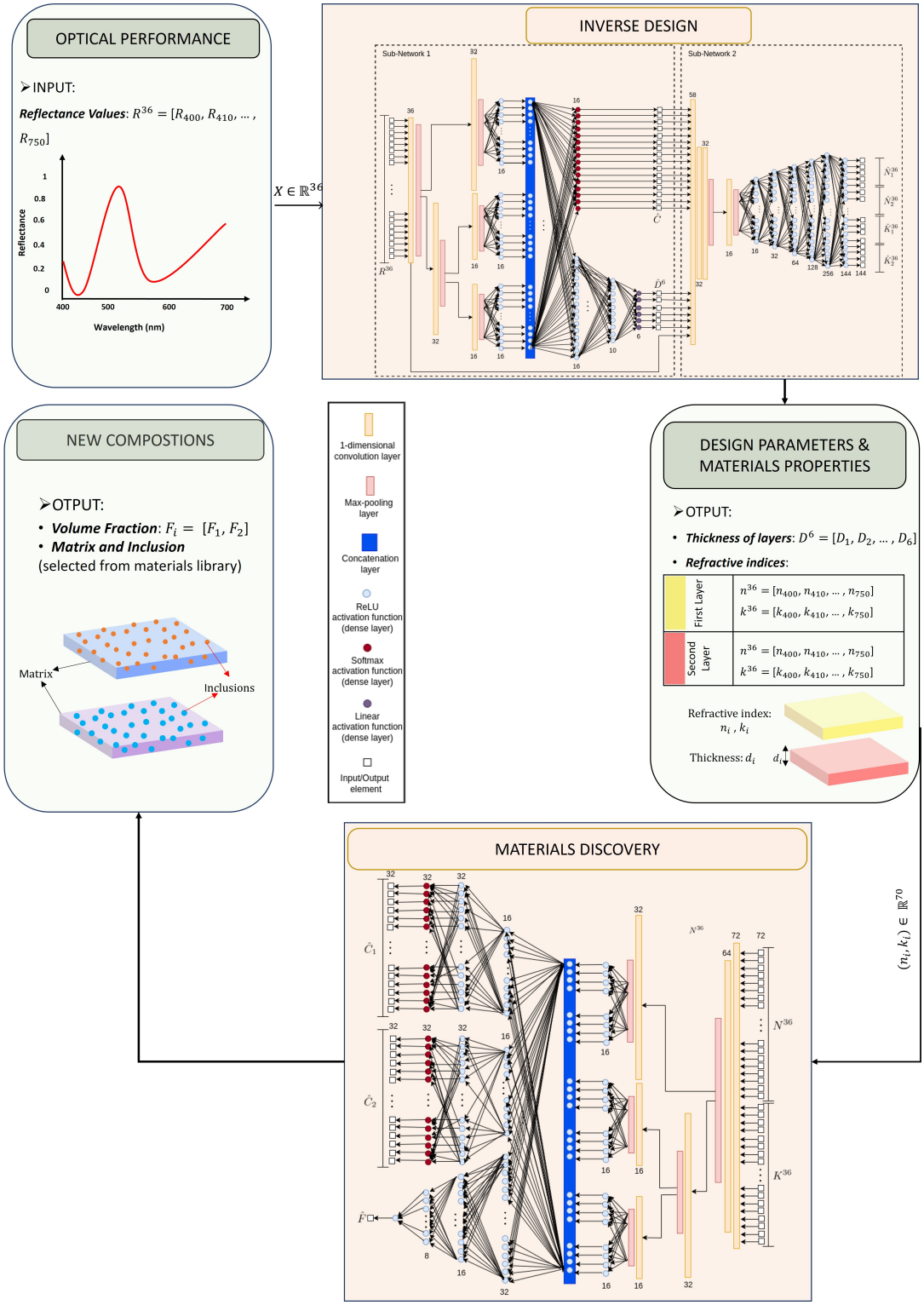


Figure 5.7: Detailed process flow for the inverse design of multilayer metamaterial. Inputting the desired reflectance spectrum leads to outputs comprising the refractive indices of each layer and the structural parameters of nanostructures. The material discovery network, employing the EMA, determines the compositions of layers in this schematic process.

5.3.2 Methodology

Figure 5.7 shows the proposed learning methodology consisting of two sequential networks: the inverse design network (comprising two sub-networks) and the materials discovery network. Similar to the previous network, in the training process of the following networks, the dataset was split into 80% for network training, 10% for validation, and 10% for testing.

Forward Model

A five-layered fully connected neural network (FCNN) was implemented, with a batch normalisation layer to enhance convergence. The input and output layers contained 10 and 71 neurons respectively, representing the thicknesses and refractive indices of the randomly selected materials, and the reflectance values. Like the previous study [11], the MAE loss function was selected to address the regression task. Moreover, the Adam optimiser was chosen for its expedited convergence rate. The learning rate was initially set to 0.01 and was decreased logarithmically whenever the validation loss remained stagnant for more than 10 epochs until it reached 1×10^{-6}. The LeakyReLU activation function is used for all layers, except the output one, which uses a sigmoid function to constrain the predictions between 0 and 1. The hyperparameters governing the network architecture, including the number of layers, the number of neurons in each layer, and the choice of activation functions, were chosen through a process of grid search.
The learning curves for both training and dataset are presented in Figure 5.8. The MAE for the training and the test dataset is 0.00071, and 0.00075 respectively.

Inverse Design Network

The objective of the ID network is to predict the suitable refractive indices, represented as N_1^{36}, and N_2^{36} for two consecutive layers, along with the corresponding extinction coefficients denoted as K_1^{36} and K_2^{36}. Additionally, the network aims to predict the thickness of their respective layers (denoted as d_1 to d_6). These predictions are made based on a given array of reflectance spectra (i.e.,

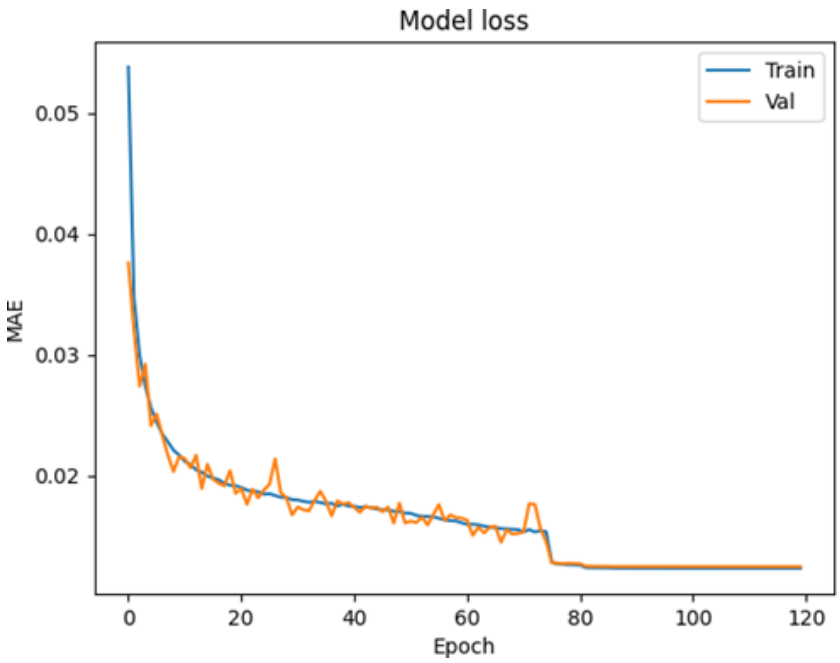


Figure 5.8: Learning curve of the forward model in prediction of reflectance values.

$R^{36} = \{R_{400}, R_{410}, \ldots, R_{750}\}$). As mentioned earlier, the dimensions of the input and output for this network are 36 and 150, respectively, as per our previous explanation. Our initial attempts to train a convolutional neural network (CNN) on the produced data using the Bayesian optimisation method for hyper-parameter tuning revealed a difficult learning task within a reasonable amount of time (24 hours) and a reasonable number of epochs (10,000) with a desired validation and test error. To address this challenge, we divided the prediction task into three stages, each of which was based on a combination of feedforward and CNN neural networks. The proposed ID network comprises two subsequent sub-networks. The first sub-network aims to predict the thicknesses of six layers (denoted by $D^6 = \{d_1, d_2, \ldots, d_6\}$) and classes of two corresponding materials, based on the given array of reflectance spectrum (i.e., R^{36}). All possible combinations of pairs of alternative layers using the binary codes of metals (00), oxides (01), fluorides(10), and polymers (11) are categorised into 16 groups which are listed in Table 5.3. Each pair will have a unique four-digit binary code based on the combination of the individual layer binary codes and its decimal number presents the group number (classes of materials).

Table 5.3: Coding system of all available classes

First layer + Second layer	Binary Code	Group Number
Metal + Metal	0000	**0**
Metal + Oxide	0001	**1**
Metal + Fluoride	0010	**2**
Metal + Polymer	0011	**3**
Oxide + Metal	0100	**4**
Oxide + Oxide	0101	**5**
Oxide + Fluoride	0110	**6**
Oxide + Polymer	0111	**7**
Fluoride + Metal	1000	**8**
Fluoride + Oxide	1001	**9**
Fluoride + Fluoride	1010	**10**
Fluoride + Polymer	1011	**11**
Polymer + Metal	1100	**12**
Polymer + Oxide	1101	**13**
Polymer + Fluoride	1110	**14**
Polymer + Polymer	1111	**15**

As a classification task, this subnetwork uses a distinct number, C, that corresponds to the 4-digit binary number, an integer ranging from 0 to 15, to make predictions about these two classes. This sub-network produces one-hot encoded output for the class number, which consists of 16 binary outputs, only one of which is 1 and the rest are zeros. As a result, this network completes a regression job and a classification task simultaneously. The objective of classification is to predict the classes of materials, while the task of regression is to predict the thicknesses. The mean squared error (MSE) in the regression task and the sparse categorical entropy in the classification task are added to form the training process's loss function. The loss functions of two different tasks operated separately without sharing the neutrons and weights. Then, the input and output layers of the first sub-network were connected to the input layer of the second sub-network. The second sub-network aims to predict the refractive indices and extinction coefficients at each wavelength for two materials.

Subsequently, the combined network was trained end-to-end, using the initial weights obtained from the two trained sub-networks. This training process was successful, leading to the convergence of the network within a reasonable time-

frame and a number of epochs, along with achieving the desired validation and test error. By dividing the network into two sub-networks, the task becomes more manageable: the first network predicts thickness values, while the second predicts refractive indices. Furthermore, training the sub-networks allowed for obtaining initial weights, which were then frozen for the subsequent end-to-end training process. The loss functions of two different tasks operated separately without sharing the neutrons and weights.

Materials Discovery Network

The material discovery network is designed to predict two materials (denoted by C_1 and C_2) and their respective fractions (where their sum equals one) within a composite material. This prediction is based on the provided refractive indices and extinction coefficients of a specific material at each wavelength. Similar to the ID network, this network combines both CNN and feedforward NN. It simultaneously performs three learning tasks: one regression and two classifications. The regression task aims to predict the fraction value (denoted by F) of one material, with the fraction of the second material in the composite being one minus this obtained value. The two classification tasks involve classifying the materials in the composite among 32 materials (CF. Table 5.2). The network employs the same loss functions as the inverse design network, using MSE for the regression task and sparse categorical entropy for the classification tasks.

5.3.3 Networks Performance

Predicting a multilayer metamaterial's refractive indices, thicknesses, and material classes involves both classification and regression, all of which are critical to the accurate design and best possible enhancement of its optical properties. With an F1-score of 0.391 for the classification part and an root mean square error (RMSE) of 0.135 for the regression section, the sub-networks created in this study show promising results. With a score of 0.904, the ROC-AUC-curve with one versus rest also shows excellent performance for the classification task. ROC-AUC measures the area under the receiver operating characteristic curve, which plots

the true positive rate (sensitivity) against the false positive rate (1 - specificity) at various threshold settings. It provides a comprehensive evaluation of the model's ability to discriminate between classes across all possible thresholds. ROC-AUC is particularly useful when the relative costs of false positives and false negatives are not known or when you want to assess the model's performance across different decision thresholds.

Spearman's correlation coefficient (SCC) was employed to measure the relationship between two variables (the target and predicted values) among all samples. This non-parametric technique has no assumptions about the data distribution. Instead, it reports how closely two variables' ranks are related to one another linearly [12]. To compute the SCC, each variable should be ranked separately, from smallest to largest. Then, the differences between the ranks for each observation on each variable are calculated. Finally, the correlation between the two sets of ranks is investigated. The SCC spans from -1 to 1. A value of -1 signifies a perfect negative correlation, meaning as one variable increases, the other decreases. A correlation-free state is indicated by a value of 0, while a value of 1 indicates a perfect positive correlation, signifying that as one variable increases, the other variable also increases. The closer the coefficient is to -1 or 1, the stronger the correlation between the two variables. The formula for the SCC, denoted as "ρ", is:

$$\rho = 1 - 6\sum d^2/n(n^2-1) \tag{5.3.3}$$

where $\sum d^2$ is the sum of the squared differences between the ranks of the two variables and n is the number of observations (or data points). Figure 5.9 compares the SCC for all available combinations.

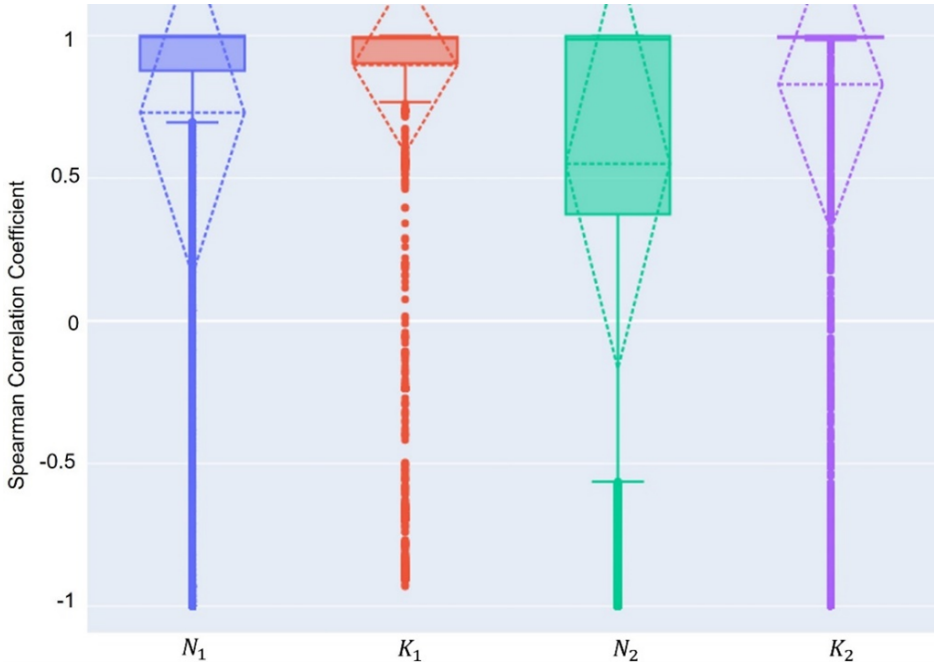


Figure 5.9: SCC graph to compare the refractive indices of target values and predicted ones among all samples.

As can be seen in Figure 5.9, the SCC values demonstrate a strong positive correlation between the target values of N_1, N_2, K_1, and K_2 and the predicted ones with the ID network. This indicates that the predicted values closely follow the trend of the target values, suggesting a high level of accuracy in the prediction model (i.e., the median in all box plots is higher than 0.995, in addition to the first quartile higher than 0.87 for N_1, K_1, and K_2). Additionally, the results of this measurement for all possible combinations of pairs of alternative layers, using the binary codes of metals (00), oxides (01), fluorides (10), and polymers (11), are presented in Figure 5.10. As shown, a strong positive correlation is evident among most combinations of layers. However, a noteworthy disparity in SCC values is observed in the vice-versa combination. This implies that the trained network tends to grasp the underlying pattern more effectively in one of these combinations. For instance, groups numbered 1 and 4 represent a combination of "metal and oxide" with different subsequent layers. Nevertheless, the accuracy of prediction in group 4 significantly outperforms its counterpart. This observation underscores the intricate nature of the ID in this particular problem, indicating the challenges posed by the complexity of the underlying patterns.

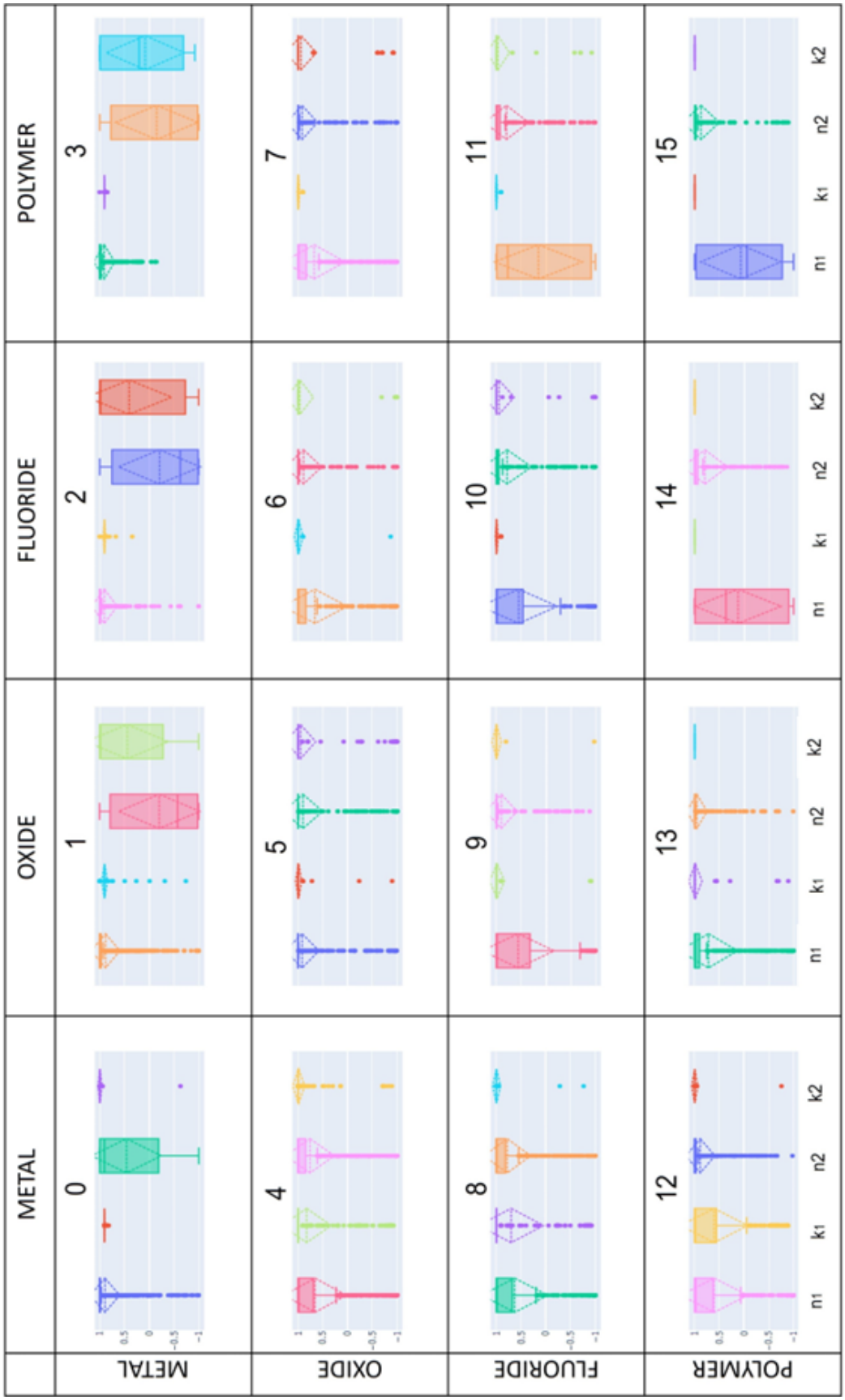


Figure 5.10: SCC values for each combination of materials.

5.4 Related Works

Optimising both structural and material parameters is a common challenge in combinatorial optimisation. Previous research, as outlined in Table 5.4, has mostly used ML algorithms to extract geometrical parameters and materials for multilayer metamaterials, resulting in diverse optical performance metrics. Although this is not a fair way to compare the performance of various models, Table 5.4 represents the input values, range of data, and complexity of the output. However, the model developed in this study differs from prior works in several key aspects. Firstly, while earlier research explored a limited number of materials, the materials library used in this project has expanded to 33 items, treating them as discrete variables, marking it as the most extensive reported in the literature. Furthermore, none of the prior studies attempted to predict the RI of layers as a continuous variable dependent on wavelength. Moreover, whereas earlier works relied on limited pre-existing material, my model explores practically infinite composition space to propose amorphous metamaterials.

Another notable difference is the optimisation approach. Most previous research employed global optimisation methods, such as genetic algorithm (GA) or particle swarm optimisation, to find the optimum designs. However, the method proposed in this thesis adopts a more local optimisation approach, employing a gradient-based technique to iteratively enhance designs. This facilitates a more focused exploration of the design space and may provide greater efficiency in discovering desired solutions. Additionally, while prior works concentrated on achieving specific optical performance metrics, such as reflectance or transmittance, the method proposed in this paper adopts a more comprehensive optimisation approach. Specifically, my model optimises the overall spectral response of the multilayer structure by considering both the optical extinction coefficient and refractive index. This approach surpasses what can be achieved by solely focusing on individual performance metrics.

Table 5.4: Comparison to state-of-the-art related works

Network	Applications	Materials	Input	Output	Number of data	Ref.
CNN	5-layer MMM	5 (Au, Ag, Al2O3, ITO, TiO2)	R/T spectra (450–950 nm)	Thickness, materials	200000	[13]
DNN (tandem model)	Three-layer Coreshell	7 (Ag, Au, Al, Cu, TiO2, SiO2, Si)	ED / MD (300 to 1000 nm)	shell thickness, materials	18000	[14]
DNN (tandem model)	6-layer MMM	2 (Si3N4, SiO2)	R spectra (380 - 780 nm)	Thickness	46656	[11]
DNN (tandem model)	20-layer MMM	2 (Si3N4, SiO2)	T spectra (450–750 nm)	Thickness	500,000	[15]
cINN	9-layer MMM	2 (Nb2O5, SiO2)	R spectra (400-700 nm)	Thickness	1000000	[16]
VAE-tSNE + GA	5-layer solar absorber	300 materials	A spectra (250-800 nm)	Thickness, materials	N/A	[17]
VAE	5-layer MMM	2 (glass , Silicon)	T spectra (1000–2000 nm)	Thickness	100000	[18]
Memetic Algorithm + DNN	2-layer CRC film	2 (Ag, SiO_x)	Structural colour (0.3–15 μm)	Dimension, materials	11000	[19]
CNN	6-layer MMM	32 Materials (supporting Inf)	R spectra (400-750nm)	RI, thickness, unknown compositions	70000	**Our work**

R: reflectance, T: transmittance, A: absorption, ED: electric dipole, MD: magnetic dipole, RI: refractive index, GA: genetic algorithm, CRC: Colored radiative cooling,

cINN: conditional invertible neural networks, VAE-tSNE: variational autoencoder stochastic neighbour embedding.

5.5 Optimisation of Free-form Shape Configuration

Up to now, two critical domains of design and optimisation have been completed using developed NN. However, another aspect of the proposed tunable metasurface is the shape of the pillars or kirigami-based patterns. While the previous networks focused on general geometries involving structures with non-continuous disk-shaped features, the design space for shape optimisation is virtually infinite. This section aims to benefit from the potential of artificial intelligence (AI) to inversely discover shapes without restrictions, allowing for the customisation of structures.

Here another ML-based algorithm was developed to discover the free-form structure. The concept of a global freeform design strategy has allowed the exploration of design candidates that were previously unattainable [20]. Freeform nanophotonic structures are essential components that possess a non-restricted geometric shape in real space, allowing for more diverse forms beyond traditional straight lines and polygons [21]. Freeform structures enable the exploration of limitless design possibilities that are difficult to attain with traditional rigid devices [21]. However, finding an optimised nanophotonic structure in the infinite design space that increases exponentially with the number of freeform geometries has been a long-standing challenge [22].

The studied geometry is presented in Figure 5.11 (a,b), depicting a common layering configuration consisting of a photoresist atop a bottom anti-reflective coating (BARC) layer, both positioned on a silicon wafer. Within this structure, a diverse nanohole array emerges in the photoresist layer, featuring circles, elliptical shapes, and exotic configurations. Subsequently, the entire structure undergoes random combinations of coating layers including no coating layer, a single layer, and double-layered metallic materials. The base of the various nanoholes is generated through a comprehensive approach using a two-dimensional (2D) origin circle as shown in Figure 5.11(c). A line passing through 12 strategically placed points, defined by coordinates along the x and y axes, shapes this circle. Introducing a crucial parameter denoted as "d", representing displacement, adds a dynamic element to the geometry. By considering the random value of "d", the positions

of these points undergo variation in the x-y plane, resulting in the formation of an array of infinite symmetrical shapes (cf. Figure 5.11 (c)). Subsequently, this newly generated shape serves as the basis for a three-dimensional (3D) electromagnetic simulation, where optical performance metrics such as reflectance. This innovative geometry generator facilitates a comprehensive exploration of diverse shapes with a high degree of freedom, leading to finding the optimised geometry. As shown in Figure 5.11 (d), the algorithm is developed to explore the infinite design space associated with a well-studied geometry comprising nanoholes of varied shapes. Its goal is to identify the structure capable of delivering customised optical performance. Initially, a compact dataset was carefully generated using 3D FEM to facilitate the swift training of a forward NN. This network serves as a tool for rapidly expanding the dataset, effectively augmenting its size. Subsequently, the new dataset is used to train the ID Network. Here, the network leverages the on-demand reflectance curve as its primary input to predict key design parameters such as shape, material composition, and layer thickness.

Table 5.5: Data training and data generation values

Variables	Training			Optimisation		
	Range	Step	No. of samples	Range	Step	No. of samples
Materials - First layer (M_1)	Al, Au, Pt, Ni, Ti	-	5	Al, Au, Pt, Ni, Ti	-	5
Materials - First layer (M_1)	Al, Au, Pt, Ni, Ti	-	5	Al, Au, Pt, Ni, Ti	-	5
Displacement (d) (nm)	0-102	34	4	0-320	5	65
Size of pitch (p)	400, 500, 600, 800	-	4	400-800	100	5
Thickness – Second layer (T_1)	0, 30, 50	-	3	0-60	10	7
Thickness – First layer (T_2)	0, 30, 50	-	3	0-60	10	7
Total	3,600			284,375		

The problem at hand involves the determination of a material specification capable of producing a predefined reflectance spectrum (R^{36}). This specification pertains to a two-layer structure and encompasses various parameters, namely: (i) material composition of the first layer, (ii) material composition of the second layer, (iii) pitch size, (iv) displacement value, (v) thickness of the first layer, and (vi) thickness of the second layer as shown in Table 5.5.

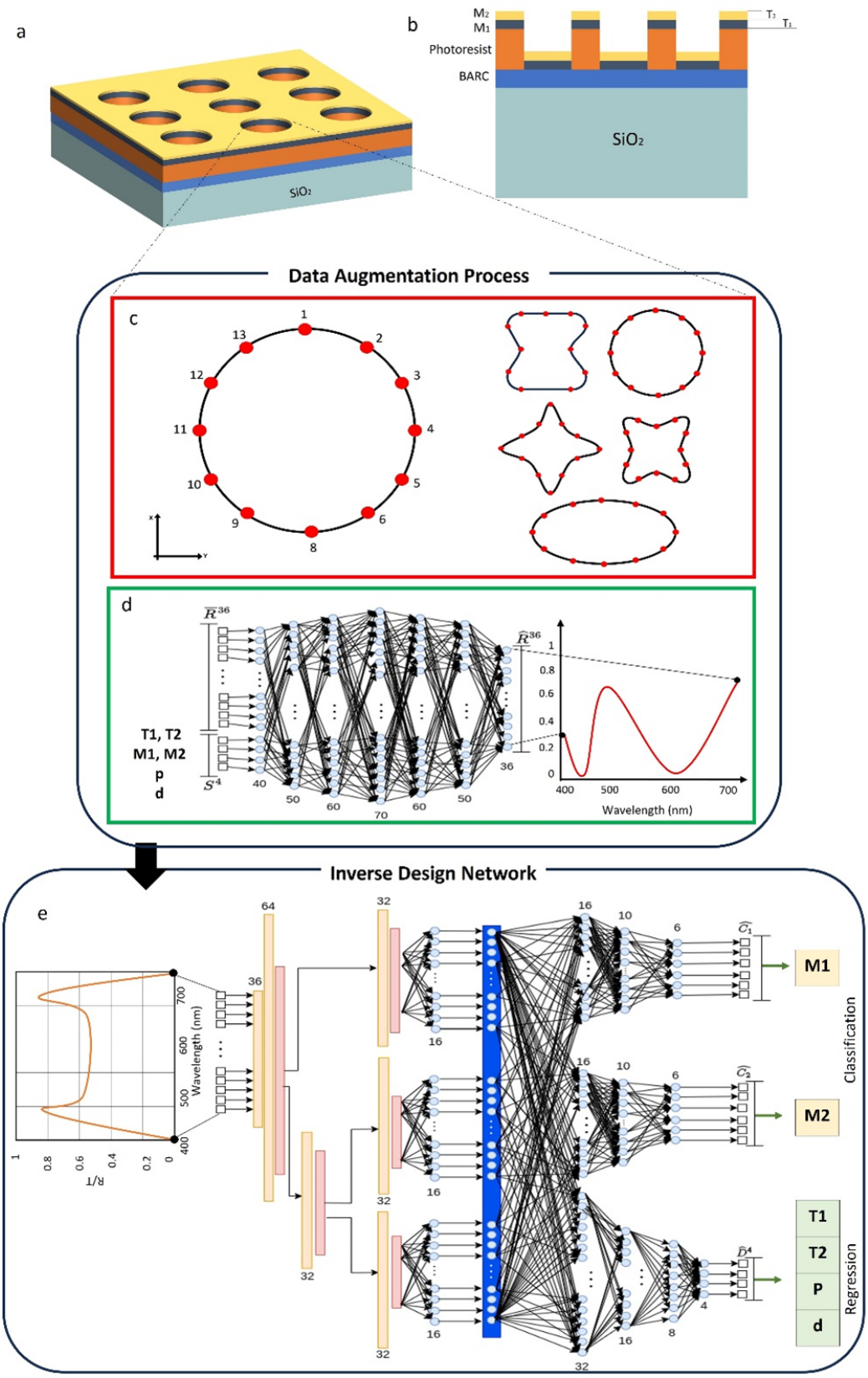


Figure 5.11: a) The 3D and b) side view of the general structure. c) multiple symmetric nanoholes' bases are created by altering the in-plane positions of the 12 connecting points, achieved by applying a random displacement parameter denoted as "d". d) The inputs represent the desired reflectance spectrum, and the output is the materials and thickness of each layer and the nanohole's shape, e) the schematic of the inverse design network.

To establish the relationship between the reflectance values and the aforementioned specifications, a supervised machine learning approach is employed. The data for this approach consists of supervised data, i.e., the desired reflectance spectrum corresponding to a specific specification. Given the impracticality of physically constructing all possible specifications, a series of simulations are conducted to gather reflectance spectrums for a subset of specifications.
The proposed methodology comprises two networks, namely the data augmentation network, and an ID network as shown in Figure 5.11 (d, e).

5.5.1 Data Augmentation

The simulation process for data collection is time-intensive and typically takes from 5 to 8 minutes per unit cell. The variables used for the creation of the training dataset are detailed in Table 5.5. These parameters include layer thickness (T_1 and T_2), coating materials (M_1 and M_1), size of the pitch (p), and displacement (d). The careful generation of this small-sized dataset aimed to provide a balance, ensuring diversity for robust network training while incorporating sufficient similar data to provide a comprehensive mapping of input-output relationships.
The forward neural network (FNN) is developed to generate additional data samples, as illustrated in Figure 5.11 (d). The input for this network comprises the values of specifications from the initial dataset. In preparation for the FNN, the following pre-processing steps are undertaken:

- Normalisation of Input Data: The values of pitch, displacement, and thickness of the first and second layers are (Min-Max) scaled and denoted as S^4. The significance of data normalisation is elaborated upon in [23].

- Encoding Categorical Material Information: Given that the materials of the layers are categorical, they are encoded using the leave-one-out method. The importance of encoding categorical features is discussed in [24]. The leave-one-out method in cross-validation is beneficial for assessing the performance and generalisation of a generative model. The process entails training the model on all data points except one, after which its performance

is assessed on the excluded point. For every data point in the dataset, this procedure is carried out again. The advantage lies in maximising the use of available data for training while still providing a comprehensive evaluation, which can lead to a more robust model [24]. In this research, materials categories are replaced with the average reflectance spectrum ($\bar{R}^{36}$) for instances that share the same material category for their layers. For instance, if the material categories for the first and second layers are 2 and 3, respectively, these categories are substituted in the dataset with the average reflectance spectrum of all other specifications with material categories 2 and 3 for their first and second layer, respectively.

After the preparation of input features in the initial dataset, the FNN is trained to predict the reflectance spectrum ($\hat{R}^{36}$). The MSE of the trained network on the validation data (10% of the training data) is 0.004.

5.5.2 Feature Importance

In the data augmentation model, it is essential to find the differential impact of various specifications on the predicted reflectance spectrum. A state-of-the-art metric for elucidating the contribution of each input feature to the predicted outcome of a neural network is the Shapley value, as outlined in [25]. Consequently, this measure can be effectively employed to ascertain the impact share of individual specifications. However, it is pertinent to note that our modelling involves multiple input features for each specification, exemplified by the presence of 36 input features serving as proxies for the used materials in the structure.
Given this complexity, adopting the Shapley value to explicate the data augmentation model based on specifications requires establishing a mapping from the impact share of input features to the higher abstraction level of each specification [26]. This mapping is accomplished through the following mechanisms:

- The impact share of the material of layers is explicated by the summation of Shapley values of features within the corresponding input vector, denoted as $\bar{R}^{36}$.

- The impact share of the thickness of layers is elucidated by the summation of Shapley values associated with thickness input features for each layer.

Through this mapping process, the influences of specifications on the reflectance spectrum can be understood in terms of material composition, thickness, displacement, and pitch size. Notably, given that displacement and pitch size serve as two input features of the neural network, no additional mapping is required for these parameters.

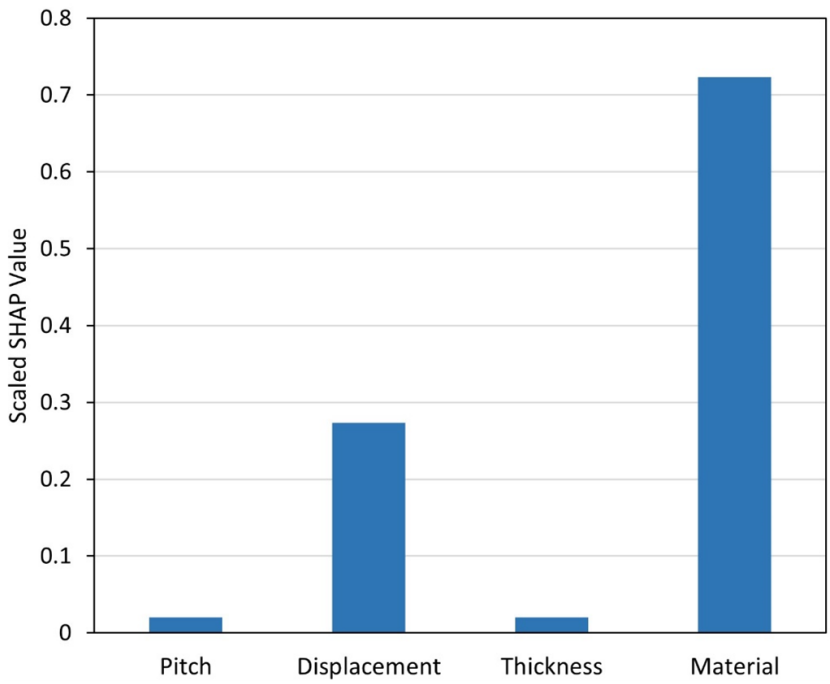


Figure 5.12: Feature importance in the studied dataset.

Figure 5.12 presents the scaled Shapley values for these specifications, ensuring that their cumulative sum equals 1. The outcomes of this analysis lead to the observation that material composition and displacement exert the most significant effects on the reflectance spectrum in our initial dataset (0.723 and 0.273, respectively). In contrast, the impact shares of pitch size and the thickness of layers are deemed negligible (0.002 for each of them). Nonetheless, the impact share of thickness and pitch size on the reflectance are quite similar.

5.5.3 Inverse Design

The objective in this phase is to devise a model for material structure specifications grounded in the reflectance spectrum, effectively employing an approach inverse to the earlier elucidated data augmentation methodology. To achieve this, we propose a NN architecture that integrates CNN and FCNN, as illustrated in Figure 5.11 (e). The tuning of hyper-parameters, such as the number of layers and neurons, was conducted through experimental procedures employing heuristics for selecting feasible configurations and Bayesian optimisation techniques.

This NN is designed to undertake three distinct learning tasks based on its input, which is the reflectance spectrum (R^{36}). Specifically, these tasks consist of one regression and two classifications. The regression task focuses on predicting the values of four specifications denoted as $\hat{D}^4$: (i) pitch size, (ii) displacement, (iii) thickness of the first layer, and (iv) thickness of the second layer. Simultaneously, the two classification tasks aim to predict the materials of the first and second layers denoted as $\hat{C}_1$ and $\hat{C}_2$, respectively.

In the training and testing of this learning model, the initial dataset derived from simulation with around 3600 data records, and the dataset generated by the data augmentation model, totalling approximately 284,375 data records. Subsequently, the new dataset was randomly split with ratios of 0.8 and 0.2 for the training/validation and testing datasets, respectively. Furthermore, the data designated for training/validation underwent additional random splits, with ratios of 0.8 and 0.2 for training and validation, respectively.

The model's performance on the test data is evaluated through the Mean Average Error (MAE) for the regression task, yielding a value of 0.21. Additionally, the F1-scores for the model's classification of the materials of the layers are 0.70 and 0.90 for $\hat{C}_1$ and $\hat{C}_2$, respectively. These metrics collectively provide insights into the effectiveness and accuracy of the proposed neural network model in predicting material structure specifications based on the reflectance spectrum.

5.6 Conclusion

This result chapter of the thesis focused on the development and implementation of universal ML algorithms for optimisation of the proposed tunable metasurface in three main domains:

Optimisation of Geometrical Design Parameters: This section introduces a FCNN trained to optimise the parameters of a kirigami-inspired metasurface for structural colour generation. The FCNN efficiently predicts the optimised structural parameters for a specific geometry, achieving high accuracy with a MAE of 0.0051 for the test dataset and 0.0038 for the training dataset. This algorithm significantly reduces computational time compared to traditional methods like FEM simulations by a factor of 10^4. The proposed optimised single geometry was characterised by an Al nanodisk with a thickness of 130 nm and a diameter of 200 nm. This configuration provides the broadest areal colour space coverage, encompassing 27.65% of the sRGB space.

Optimisation of Materials Properties: Another set of networks was trained using a dataset generated by TMM to optimise parameters and materials for custom structures. These subsequent networks, including ID and material discovery networks, predict thickness values and material's RI for a six-layer structure based on the desired reflectance spectrum. Predicting a multilayer metamaterial's refractive indices, thicknesses, and material classes involves both classification and regression. The results showed promising performance for the ID network including an F1-score of 0.391 for the classification part and an RMSE of 0.135 for the regression section. The materials discovery was combined with the EMA to predict the classes of two materials and their respective fractions within a composite material. This prediction was based on the provided refractive indices and extinction coefficients of a specific material at each wavelength.

Optimisation and Discovery of Free-from Geometry: The final network was tasked to discover optimised geometries meeting on-demand optical property requirements. This network combines data augmentation and ID networks to learn nonlinear relationships between structural parameters, geometrical features, and reflectance spectra. The augmentation network revealed that the material compo-

sition and shape of the unit cell have the greatest effects on reflectance spectra, with 72% and 27%, respectively. The ID model showed a MAE of 0.21 for the regression task, and the F1-scores of 0.90 for the classification task.

Bibliography

[1] A. Ghasemi, R. Fang, D. A. Zeze, and M. K. Hedayati, "Ultra-stretchable active metasurfaces for high-performance structural color," *AIP Advances*, vol. 13, no. 8, 2023.

[2] P. N. Hishimone, H. Nagai, and M. Sato, "Methods of fabricating thin films for energy materials and devices," in *Lithium-ion Batteries-Thin Film for Energy Materials and Devices*. IntechOpen, 2020.

[3] A. Banerjee, R. M. Heath, D. Morozov, D. Hemakumara, U. Nasti, I. Thayne, and R. H. Hadfield, "Optical properties of refractory metal based thin films," *Optical Materials Express*, vol. 8, no. 8, pp. 2072–2088, 2018.

[4] P. H. Lissberger, "Optical applications of dielectric thin films," *Reports on Progress in physics*, vol. 33, no. 1, p. 197, 1970.

[5] M. A. Zahid, S. Q. Hussain, Y. H. Cho, and J. Yi, "Optical properties of caf2 thin film deposited on borosilicate glass and its electrical performance in pv module applications," *Applied Sciences*, vol. 10, no. 16, p. 5647, 2020.

[6] T. Pilvi, K. Arstila, M. Leskelä, and M. Ritala, "Novel ald process for depositing caf2 thin films," *Chemistry of materials*, vol. 19, no. 14, pp. 3387–3392, 2007.

[7] A. Rasad, H. T. Yudistira, F. Qalbina, A. G. Saputro, and A. Faisal, "Multilayer flexible metamaterials based on circular shape with negative refractive index at microwave spectrum," *Sensors and actuators A: physical*, vol. 332, p. 113208, 2021.

[8] S. G. Moiseev, "Active maxwell–garnett composite with the unit refractive index," *Physica B: Condensed Matter*, vol. 405, no. 14, pp. 3042–3045, 2010.

[9] V. A. Markel, "Maxwell garnett approximation (advanced topics): tutorial," *JOSA A*, vol. 33, no. 11, pp. 2237–2255, 2016.

[10] M. Sela and C. Haspel, "Predicting the refractive index of amorphous materials using the bruggeman effective medium approximation," *Applied Optics*, vol. 59, no. 28, pp. 8822–8827, 2020.

[11] S. Head and M. Keshavarz Hedayati, "Inverse design of distributed bragg reflectors using deep learning," *Applied Sciences*, vol. 12, no. 10, p. 4877, 2022.

[12] P. Sedgwick, "Spearman's rank correlation coefficient," *Bmj*, vol. 349, 2014.

[13] A. Lininger, M. Hinczewski, and G. Strangi, "General inverse design of layered thin-film materials with convolutional neural networks," *ACS Photonics*, vol. 8, no. 12, pp. 3641–3650, 2021.

[14] S. So, J. Mun, and J. Rho, "Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core–shell nanoparticles," *ACS applied materials & interfaces*, vol. 11, no. 27, pp. 24 264–24 268, 2019.

[15] 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.

[16] A. Luce, A. Mahdavi, H. Wankerl, and F. Marquardt, "Investigation of inverse design of multilayer thin-films with conditional invertible neural networks," *Machine Learning: Science and Technology*, 2022.

[17] A. Jiang and O. Yoshie, "A reinforcement learning method for optical thin-film design," *IEICE Transactions on Electronics*, vol. 105, no. 2, pp. 95–101, 2022.

[18] J. Roberts and E. Wang, "Modeling and optimization of thin-film optical devices using a variational autoencoder," *University of Stanford*, 2018.

[19] Q. Guan, A. Raza, S. S. Mao, L. F. Vega, and T. Zhang, "Machine learning-enabled inverse design of radiative cooling film with on-demand transmissive color," *ACS Photonics*, 2023.

[20] C. Park, S. Kim, A. W. Jung, J. Park, D. Seo, Y. Kim, C. Park, C. Y. Park, and M. S. Jang, "Physics-informed reinforcement learning for sample-efficient optimization of freeform nanophotonic devices," *arXiv preprint arXiv:2306.04108*, 2023.

[21] J. Park, S. Kim, D. W. Nam, H. Chung, C. Y. Park, and M. S. Jang, "Free-form optimization of nanophotonic devices: from classical methods to deep learning," *Nanophotonics*, vol. 11, no. 9, pp. 1809–1845, 2022.

[22] J. Jiang, D. Sell, S. Hoyer, J. Hickey, J. Yang, and J. A. Fan, "Free-form diffractive metagrating design based on generative adversarial networks," *ACS nano*, vol. 13, no. 8, pp. 8872–8878, 2019.

[23] J. Sola and J. Sevilla, "Importance of input data normalization for the application of neural networks to complex industrial problems," *IEEE Transactions on nuclear science*, vol. 44, no. 3, pp. 1464–1468, 1997.

[24] J. T. Hancock and T. M. Khoshgoftaar, "Survey on categorical data for neural networks," *Journal of Big Data*, vol. 7, no. 1, pp. 1–41, 2020.

[25] B. Rozemberczki, L. Watson, P. Bayer, H.-T. Yang, O. Kiss, S. Nilsson, and R. Sarkar, "The shapley value in machine learning," *arXiv preprint arXiv:2202.05594*, 2022.

[26] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

Chapter 6

Result Chapter: Experimental Validation

6.1 Chapter Overview

The last result chapter outlines the experimental methodologies employed for the fabrication and characterisation of the structures offered by the inverse design (ID)neural network (NN) and materials discovery NN in Chapter 5. In the experimental phase, three distinct samples were fabricated to ascertain the accuracy of the networks. The first device was fabricated to verify the ID network, in which a custom optical filter was made according to the design parameters and pure materials suggested by the network. The material discovery network's predictions were then used to create an on-demand perfect reflector with an amorphous meta-

surface that operates in the visible spectrum. Finally, a free-form structure was fabricated to verify the performance of the machine learning (ML)-based algorithm developed for geometry optimisation. Every sample was characterised, and the mismatch between the properties of the fabricated prototypes and the expected ones is discussed.

6.2 Fabrication Process

6.2.1 Sample Preparation

The samples used to validate the ID and material discovery networks were grown on glassware (cover slides) that was purchased from VWR (Leicestershire, UK). Before any film deposition, the substrates were cleaned with the common wet processing procedure. It involved subjecting all glassware to ultrasonic treatment for 2 minutes in an acetone solution, followed by another 2 minutes in isopropanol to remove residual organics and contaminants. The glass substrates were dried using N_2 gas. The entire processing time, including film deposition and cleaning, was less than thirty minutes.

6.2.2 Magnetron Sputtering

Magnetron sputtering is a physical vapour deposition technique, in which the atoms released from the target (the material that is the source of the thin film to be deposited) travel through the vacuum chamber until they encounter a surface (substrate), where they condense. As time progresses, additional atoms accumulate on the surface, gradually forming the desired thin film [1]. Figure 6.1 illustrates a simplified overview of the sputtering process. The target material in a vacuum chamber is negatively charged, while the other parts stay grounded. When argon (Ar) gas is injected into the chamber, it undergoes ionisation. This means that some of the argon atoms in the gas become positively charged ions (Ar^+). The cathode, which is usually the desired target material to sputter, becomes exposed to these ionised argon ions. This energy transfer is forceful enough to dislodge or

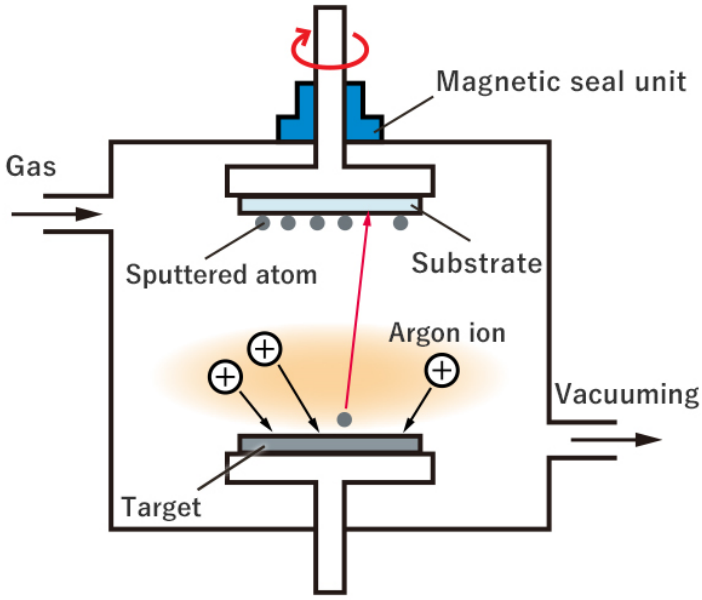


Figure 6.1: A schematic representation of the sputtering system. This chamber includes other equipment that is depicted (refer to Figure 6.2).

eject atoms from the surface of the target material. The ejected atoms become part of the sputtered material that travels through the vacuum chamber and eventually deposits onto a substrate, forming a thin film.

The sputter head's magnets, featuring alternating polarity, generate a potent magnetic field near the target surface. A plasma forms close to the target when the primary and secondary electrons rotate within the magnetic field. This concentrated increase in electron density speeds up the deposition rate by increasing the rate at which the argon gas surrounding the negatively charged cathode ionises. While a direct current (DC) bias effectively charges electrically conductive target materials, insulating materials rapidly accumulate a positive charge on their surface, rendering DC bias impractical. To address this, a radio frequenc (RF) source, typically operating at 13.56 MHz, is employed to facilitate the dissipation of the positive charge build-up.

A fully automated MiniLab 060 (Moorfield Nanotechnology Ltd.) sputtering system with up to three independent and symmetrical water-cooled magnetron sources for 3 inches of circular targets (Figure 6.2) was used in this research. The substrate holder is located in the centre of the 30° sputter cannons' field of view.

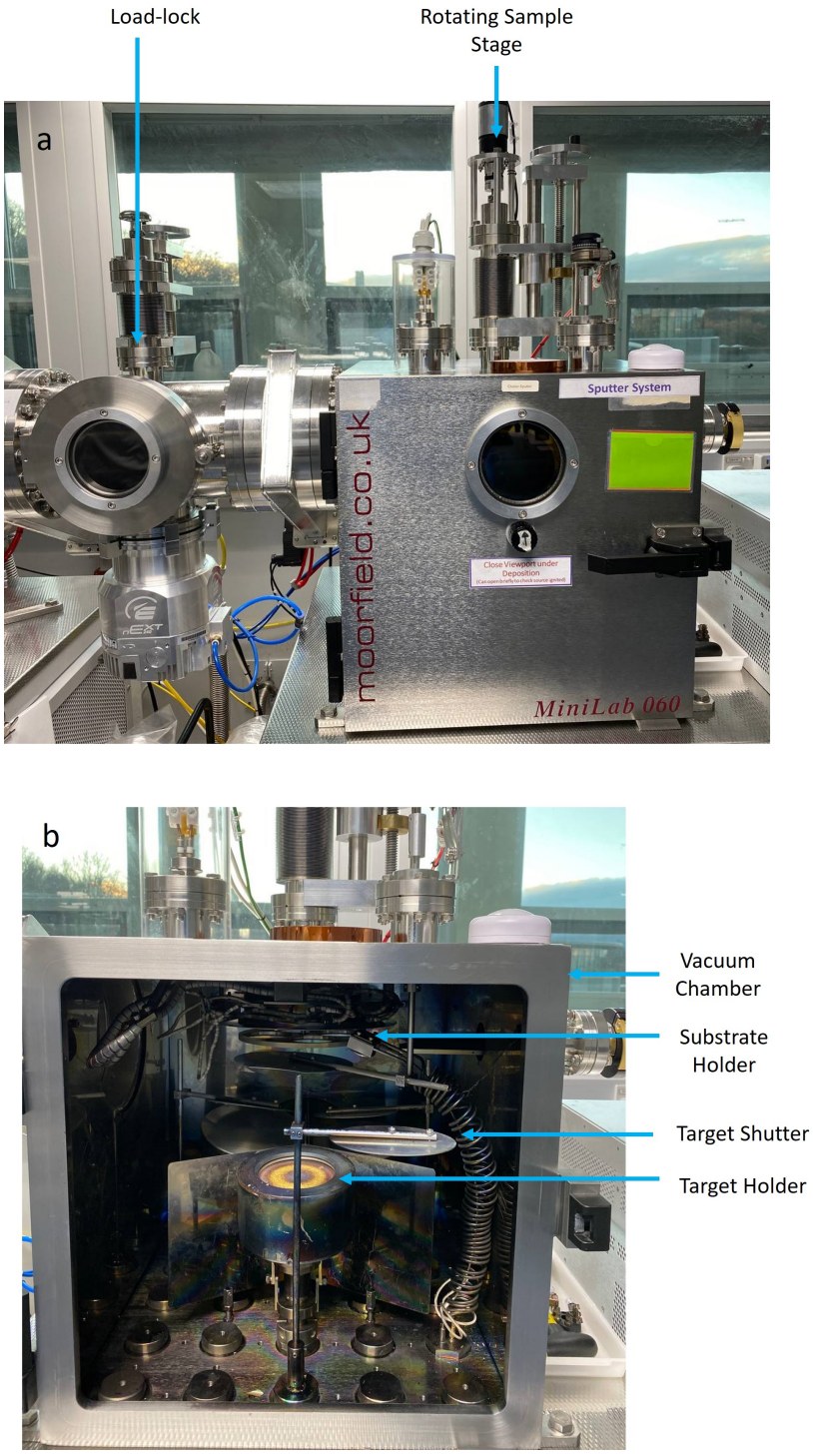


Figure 6.2: (a) MiniLab 060 sputtering system overview, (b) a close-up view of the magnetron sputtering chamber.

The machine is divided by a gate valve into two chambers: the main chamber on the right and the load-lock on the left. The load-lock achieves its base vacuum pressure in approximately three hours, a significant improvement compared

to the one day required for the main chamber, owing to its smaller volume. Consequently, the presence of the load-lock enhances the efficiency of sample fabrication.

As suggested by the material discovery network, TiO_2 (99.99%) and Au (99.95%) targets were inserted into the target sites to fabricate the composite layer. TiO_2 was sputtered using an RF magnetron system, while Au was sputtered using a DC magnetron sputter source. The sputtering rates of TiO_2 and Au for the proposed composition were 0.36 Å/s and 0.6 Å/s, respectively. The distance between the target and the substrate was kept at 90 mm. The standard base vacuum pressure is 5×10^{-6} Torr, with a deposition Ar gas pressure of 1×10^{-3} Torr. Each sputter head is equipped with a target shutter, which prevents the target from cross-contamination.

6.3 Characterisation Techniques

6.3.1 Reflectivity

X-ray reflectometry (XRR) is a technique used to study the structural characteristics of thin films and layered structures. It is a non-destructive method that provides information about the thickness, density, and roughness of thin films by analysing the reflection of X-rays incident on the sample. In XRR, a monochromatic X-ray beam is directed at the sample at a specific angle of incidence [2]. The X-rays penetrate the sample and undergo multiple reflections at the interfaces between different layers within the film. The resulting interference pattern is then recorded as a function of the angle of incidence. XRR is often employed alongside other microscopic characterisation and analytical methods to provide a comprehensive characterisation of thin film structures [3].

The electrons in a material's atomic shell oscillate at the identical frequency as the incident photons when the substance is exposed to X-rays. This oscillation serves as a source of radiation, producing an X-ray known as Thomson scattering that has the same energy as the input photon [4, 5]. XRR measurements analysed by considering reflection and transmission coefficients at the interfaces of materials

with different refractive index (RI). The Fresnel equations for transmission (t) and reflection (r) define these coefficients as follows:

$$r = \frac{n_1 \cos(\theta_i) - n_2 \cos(\theta_t)}{n_1 \cos(\theta_i) + n_2 \cos(\theta_t)} \tag{6.3.1}$$

$$t = \frac{2n_1 \cos(\theta_i)}{n_1 \cos(\theta_i) + n_2 \cos(\theta_t)} \tag{6.3.2}$$

where n_1 and n_2 are the refractive indices of the incident and transmitted media respectively, θ_i is the angle of incidence, and θ_t is the angle of transmission. In Equation 6.3.1, $n_1(1)$ denotes the RI of the incoming medium, and θ_i(t) represents the incident angle of the transmitted beam. Given the very small angle, the expression $\sin(\theta) = \theta$ is applicable. The transmitted and reflected beams' intensities are determined using $|r|^2$ and $|t|^2$, respectively. Although the proposed equations hold for the transmission and reflection at a single interface, real samples involving multiple interfaces require additional complexity to be taken into account. For example, a single-layer thin film provides air/film and film/substrate interfaces. In this structure, the initial reflection and further reflections of the initially transmitted wave make up the overall reflected intensity, as depicted in Figure 6.3.

Employing the approach proposed by Als-Nielsen [6], the overall reflectivity (r_{total}) can be measured as follows:

$$r_{\text{total}} = \left| \frac{r_{12} + r_{23} e^{i2\beta}}{1 + r_{12} r_{23} e^{i2\beta}} \right|^2 \tag{6.3.3}$$

where r_{12} and r_{23} are the reflection coefficients at the interfaces of the single layer, and β is the phase difference between consecutive interfaces. While Equation 6.3.3 holds for a single layer, its generalisation to encompass any two distinct layers is possible. Similarly, in a multi-layered thin film, the overall reflectivity can be calculated using Paratt's method [7], where the reflection at the lowest interface is calculated initially, progressively moving through each film interface:

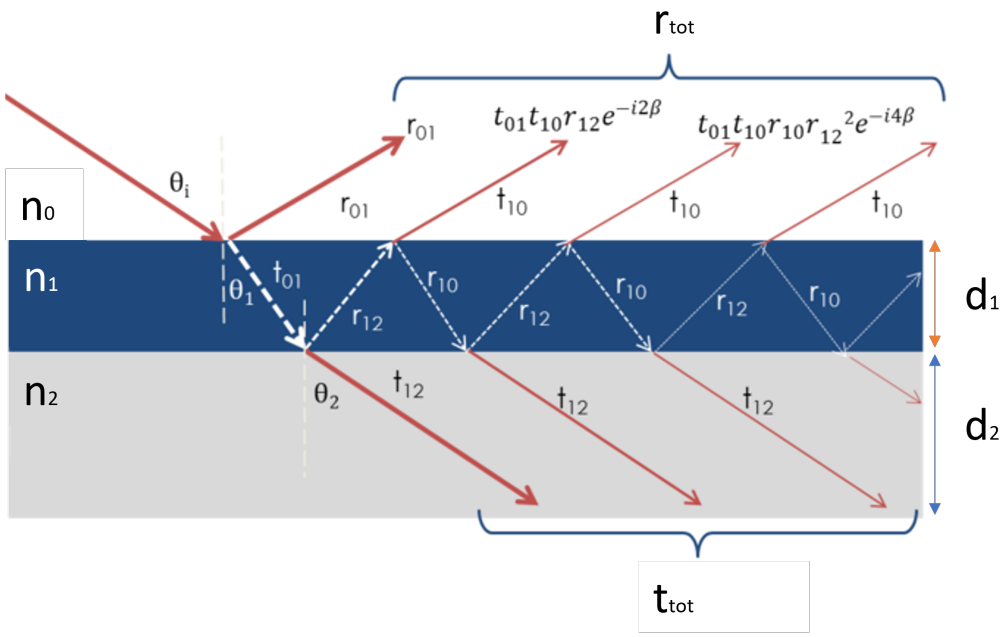


Figure 6.3: A schematic representation of the multiple paths of reflected and transmitted X-rays within a single-layer film.

$$r_{\text{total}} = \left| \frac{r_n + r_{n+1} e^{i2\beta_1} r_{n+2} e^{i2\beta_2}}{1 + r_n r_{n+1} e^{i2\beta_1} e^{i2\beta_2}} \right|^2 \tag{6.3.4}$$

where r_n, r_{n+1}, r_{n+2} are the reflection coefficients at the interfaces of the two layers, β_1, β_2 are the phase differences between consecutive interfaces within each layer.

Rough Interfaces

Equations 6.3.1-6.3.4 compute the reflection from perfect and sharp interfaces. However, in reality, all interfaces have imperfections. These imperfections can be characterised by an interface width, denoted as σ. This roughness significantly affects the optical performance, particularly the reflection properties of the device. XRR is capable of determining the interface width, however; it is unable to differentiate between intermixing and interface roughness as shown in Figure 6.4.

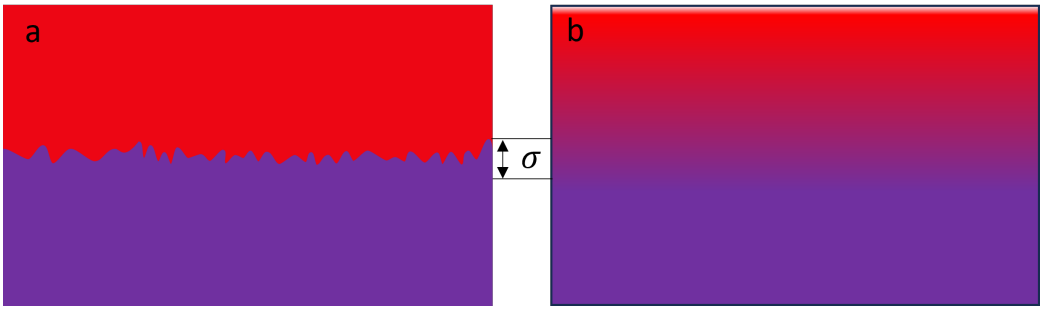


Figure 6.4: The schematic illustrates contributions to the overall interface width from a) interfacial roughness and b) intermixing.

A series of completely flat intermediate layers can be used to simulate both interfaces, with the chemical gradient following a Gaussian distribution with a width σ. The total interface reflectivity is calculated by integrating the reflectivity of each intermediate layer over the full interface width, provided that the thickness of these intermediate layers is much less than the overall interface width. This approach is extendable to cover the entire film, as illustrated in Figure 6.5 [8]. A Debye-Waller-like factor for the interface reflectivity is produced when a finite interface width is introduced. This factor is expressed as follows:

$$R_{\text{rough}} = R_{\text{smooth}} e^{(-\sigma^2 q_z^2)} \tag{6.3.5}$$

where σ is the root mean square (RMS) roughness of the interface and q_z is the component of the scattering vector perpendicular to the interface. Rocking curve measurement is a technique to distinguish between interface roughness and intermixing. This technique involves scanning the sample angle while keeping the angle between the X-ray source and detector fixed (i.e., $\theta_i + \theta_r$ = constant). This type of measurement is also referred to as off-specular. The distorted-wave Born approximation was initially developed by Sinha et al. [9] to model the resulting reflectivity profile. This model was then refined by Wormington [10].

All XRR measurements were carried out by a Bede D1 Diffractometer, as depicted in Figure 6.6. Using the Cu-Kα1 spectral line, the device produces X-rays with a wavelength of 1.54 Å. The generated ray is collimated two sequential slits. The sample holder has the capability to move along the three axes and rotate

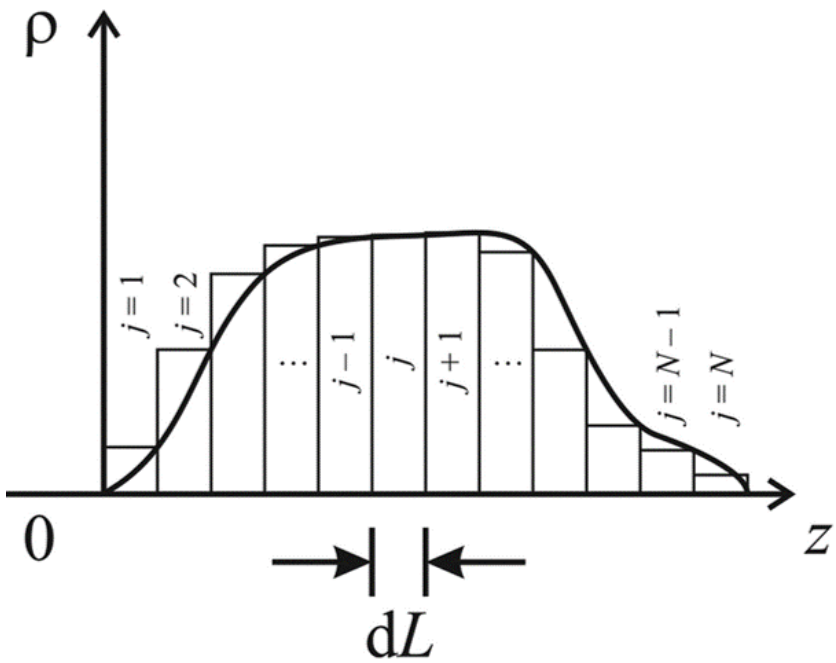


Figure 6.5: An example of discretising a sample depth (z) in a non-uniform film with density profile (ρ). Employing Paratt's recursive method, the reflectivity is initially computed at the lower j = 1 interface, followed by an iterative process through each subsequent j + 1 interface. Reprinted by permission from Royal Society of Chemistry [8], Copyright (2023).

within all three planes, providing a maximum of six degrees of freedom. A second set of collimating slits is located right before the detector as shown in Figure 6.6.

In the XRR measurements, the incident angle and reflected beam one should be equal. Firstly, the sample is positioned in line with the beam, guaranteeing that it is level in relation to the beam's shape and located precisely in the centre of the beam. The beam profile is generally rectangular due to the slit size of 0.5 mm $\times$ 15 mm. To confirm that the sample surface is level with the extended axis of the beam envelope, scans are carried out along each axis. During the process of aligning the sample, a half-cut is made. This means that the intensity detected at a detector angle of 0° is exactly half of what it would be if the sample was not present. This guarantees that the sample's surface is positioned precisely in the centre of the shorter axis of the envelope. The angle representing the sample is symbolised as θ, while the angle representing the detector is double the value of θ. XRR measurements typically entail scanning the 2θ angle within the range of 0° to 10°, using step increments of 0.01°.

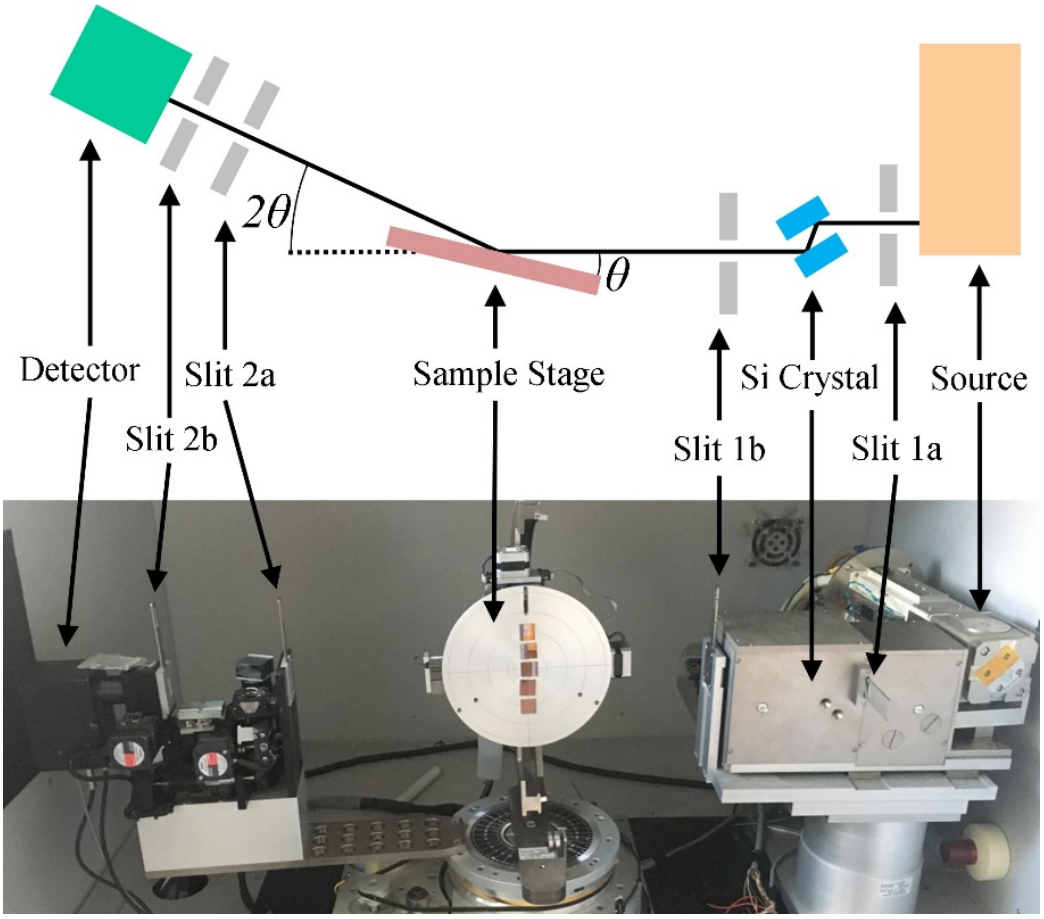


Figure 6.6: Different parts of the Bede D1 Diffractometer used in this project.

Data modelling

The model for each dataset was optimised using GenX, an open-source software. Nicholson's PhD thesis offers a detailed methodology for the fitting procedure [5], which mirrors the approach employed in this thesis and is outlined below.

The initial model was constructed using nominal structural details, where input parameters were categorised into three primary groups based on their initial specifications. Material parameters, such as density and scattering factors, were assigned values from existing literature. Instrument parameters, like the size of the sample and beam width, were obtained through direct measurement. Sample parameters, which encompassed factors like interface width and thickness, were initially assigned their nominal values. Typically, the category to which a parameter belonged determined the allowable range for its variation. In this thesis, measured values such as sample length and incident X-ray intensity were restricted to a ±

5% range to maintain stability.

Following the optimisation of the original model, the results were checked visually to ensure accuracy. Subsequent models were then hypothesised, which may include surface oxide layers and intermediate interface layers. Following additional optimisation, the nominal model and these models were contrasted. To discover the best answer, it may be necessary to include elements outside the nominal structure and allow for a wide range of parameters while optimising the model.

6.3.2 Microscopy Analysis

Atomic Force Microscopy

Throughout this thesis, the roughness of surfaces was assessed using an atomic force microscopy (AFM). In this technique, a sharp probe is meticulously scanned across a surface, with the probe height recorded to generate surface images. Typically, the probe tip is fixed to the cantilever's end, which undergoes deflection due to interatomic force interactions with the surface. Various forces, such as Van der Waals interactions and capillary effects, contribute to the tip-sample forces. Deflection sensors measure the cantilever's deflection by projecting a laser beam onto a position-sensitive detector and having it reflect off the cantilever as schematically shown in Figure 6.8 [11].

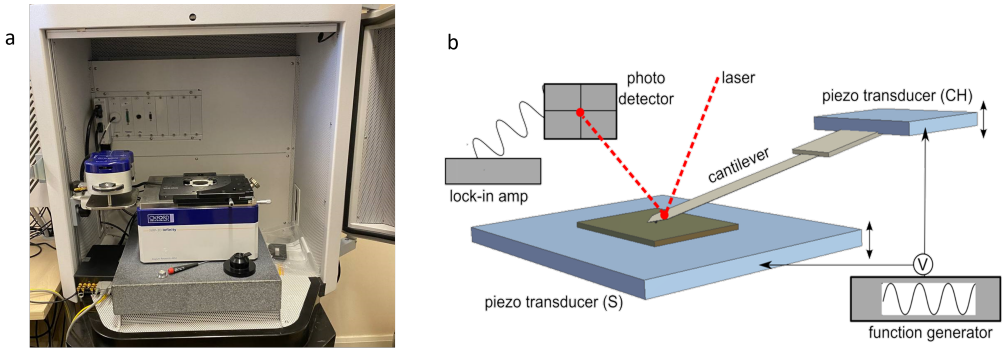


Figure 6.7: a) The AFM instrument used in this project, b) schematic diagram of a typical AFM set-up.

AFM operates in various modes including contact mode, non-contact mode, or tapping mode. In contact mode, the tip of the instrument makes gentle phys-

ical contact with the sample, encountering repulsive forces in the process. This operation mode is prone to surface damage, particularly for soft materials. In non-contact mode, the tip-sample separation is relatively large, and the probe experiences weak attractive forces. The cantilever oscillates at its resonant frequency, and changes in force alter the oscillating frequency, providing a feedback signal for surface imaging. Tapping mode (the mode that I used in this thesis) combines aspects of both contact and non-contact modes [12]. The AFM analysis was carried out using Oxford Instruments Asylum Research's MFP-3D Infinity AFM machine (cf. 6.7 (a,b)).

Scanning Electron Microscopy

Scanning electron microscopy (SEM) is an imaging method that provides detailed information about the morphology and composition of a sample at the micro/nanoscale. In this technique, a tungsten filament electron cannon is used to generate an electron beam, which is then accelerated to an energy of 0.2–40 keV. This electron beam is systematically scanned over the sample through the use of electromagnetic deflection coils, and the signal detected at each point is assigned as a pixel value to form the image. The use of electrons allows for smaller wavelengths compared to optical microscopes, enabling SEM to achieve highly impressive resolutions. The electron beam interactes with the surface and scatter from atoms inside an interaction volume that is shaped like a teardrop and spans several hundred nanometers. This scattering is then detected by the detector [13]. In this study, high-magnification SEM images of sample intersections were captured using a ZEISS Sigma 300 after meticulous preparation with a focused ion beam (FIB). These SEM images served to measure the thickness of developed multi-layer metamaterials.

Transmission Electron Microscopy

Transmission electron microscopy (TEM) is a powerful imaging technique that uses transmitted electrons to achieve high spatial resolution in the nanometer range. It enables the visualisation of fine details within a specimen, providing

insights into its structural and morphological features at a scale not easily attainable by other microscopy methods. The high-resolution capabilities of TEM allow researchers to discern the arrangement of atoms and study the crystallography of amorphous thin films. This detailed analysis contributes to a deeper understanding of the microstructure, defects, and composition of thin films, essential for optimising their properties in various applications [14].

In this project, TEM was specifically employed to study the dimensions and dispersion of particles within a matrix in the amorphous layer. TEM samples with a thickness of 10 nm were sputtered onto holey carbon grids. High-resolution images and elemental maps of the composite layer were captured using a JEOL 2100F field emission gun TEM operating at 200 kV. The elemental maps represent the spatial distribution of specific chemical elements within the nanocomposite layer. This map provides information about the elemental composition and its variation across different regions of the specimen at a high spatial resolution.

UV-VIS Spectroscopy

UV-VIS spectrophotometer is an analytical technique to measure the reflectance/transmittance of electromagnetic radiation through a material within the UV-Vis spectrum. The fundamental principle involves irradiating a material with a beam of light at a specific wavelength and intensity. The structure absorbs a certain amount of light through the excitation of electrons, while some light is reflected at the material's surfaces, and some is scattered, preventing it from reaching the detector. The remaining light passes through the material and reaches the detector on the opposite side. Reflectance/transmittance is then determined by the ratio of the light intensity reflecting (or reaching the detector) to the emitted light intensity [15].

To measure the reflectivity within the visible wavelength range, the Ocean Optics FLAME-UV-VIS spectrometer was used. The initial measurements involved recording the spectrum of a reference sample, which approximates a near-perfect mirror within the visible light spectrum. The reflectance of each sample was then determined as the ratio of its spectrum to the reference spectrum.

6.4 Validation of Inverse Design Network

To demonstrate the accuracy of the ID network as a practical solution for real-world applications, I tasked the optimised network with predicting the optical properties of a custom sample. This sample was an optical filter designed to selectively filter light with wavelengths below 550 nm, as illustrated in Figure 6.8 (a). The input values for the network are the reflectance spectrum of the target filter (depicted by the black dashed line). This input was used to find the refractive indices for two distinct layers along with their corresponding thicknesses. Subsequently, the predicted refractive indices were matched to available materials, resulting in the proposed structure characterised by alternating layers of SiO_2 and Au. The final configuration of the multi-layer metamaterial comprised six layers with thicknesses of 28, 24, 20, 34, 26, and 38 nm, respectively, as shown in Figure 6.8 (b).

The proposed structure was fabricated using the magnetron sputtering system. The sputtering rates for the Au and SiO_2 layers were 0.03 Å/s and 0.5 Å/s, respectively. The spectrometer was employed to measure the reflectance curve of the created sample, and it is depicted by the red line in Figure 6.8(a). Despite some mismatch between the predicted and actual results, there is an overall alignment with the anticipated trend, validating the effectiveness of the proposed inverse design methodology. The successful implementation of the custom-designed optical filter demonstrates the approach's ability to translate theoretical predictions into actual structures with the appropriate optical properties.

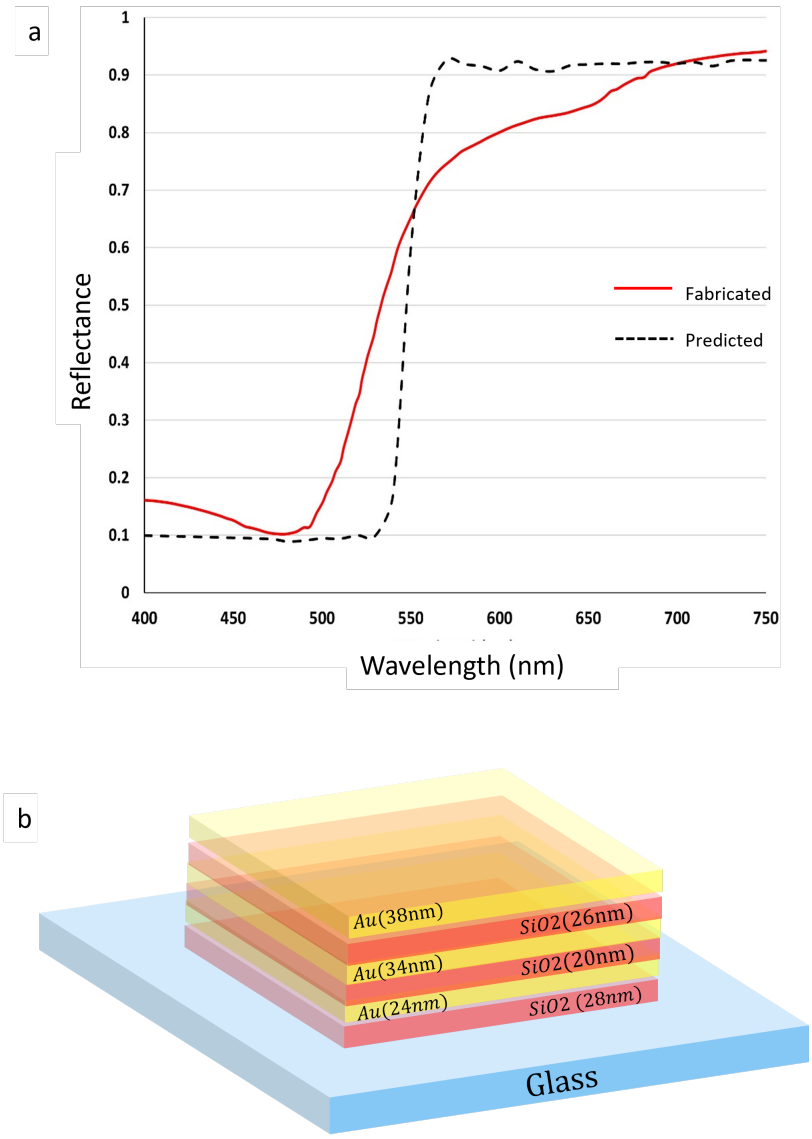


Figure 6.8: A custom optical filter consisting of six alternating layers. a) The reflectance spectrum of the fabricated filter (red line) and the predicted filter (black dashed line) by the inverse design network. b) predicted metamaterial consisting of SiO_2 and Au for odd and even layers with thicknesses of 28, 24, 20, 34, 26, and 38 nm.

The XRR measurements are presented in Table 6.1. The density of Au and SiO_2 were manually inserted in the model based on the literature [16, 17]. Figure 6.9 presents the best-fit XRR model to the developed sample.

Table 6.1: Results of the best-fit model with the XRR measurement

Layer	Density (g/cm^3)	Thickness (nm)	Interface width (nm)
Au	19.3	36	5.47
SiO_2	2.64	28	1.16
Au	19.3	31	4.63
SiO_2	2.64	22	2.66
Au	19.3	26	4.98
SiO_2	2.64	26	1.14
Substrate (glass)	-	-	-

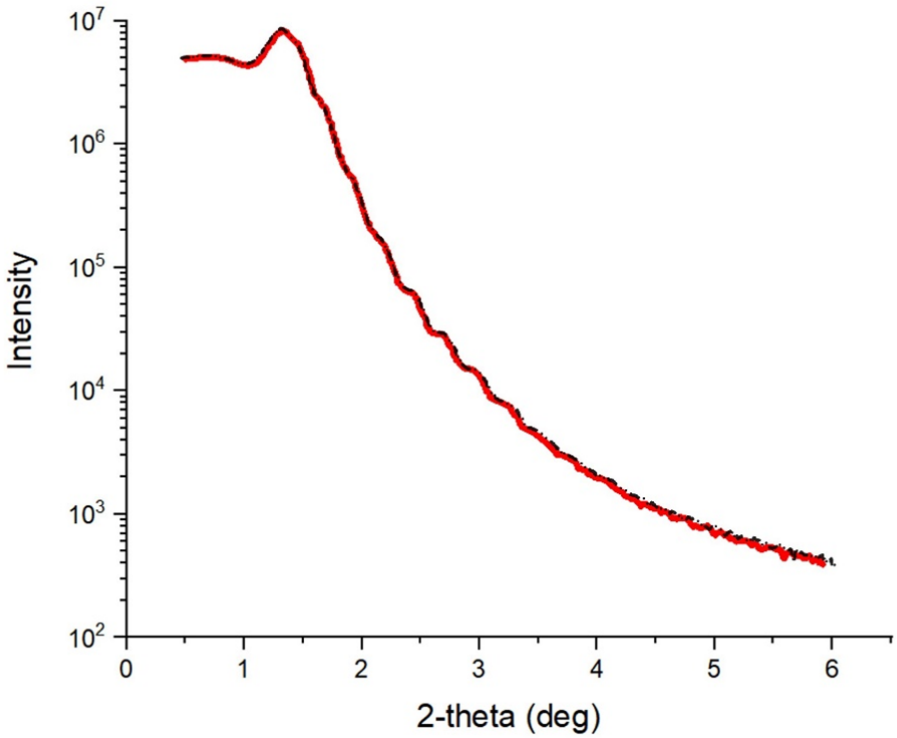


Figure 6.9: The best-fit models (dashed lines) with the XRR results.

The analysis of the best-fit model revealed that a notable interface width contributed to the deviation between the predicted curve and the actual fabricated one. This finding was further substantiated by microscopic analysis using SEM images and AFM. As illustrated in Figure 6.10 (a, b), the AFM results indicate that the roughness of the gold layers, brighter layers in the SEM images of the sample's cross-section, was approximately 5 nm, aligning with the results obtained from XRR measurements (cf. Table 6.1). This could prove that the observed mismatch in the predicted and fabricated curves is indeed influenced by the high interface

width.

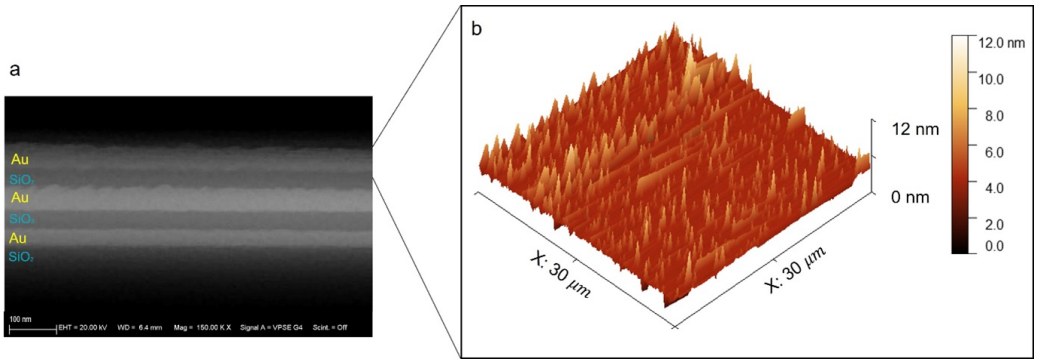


Figure 6.10: a) SEM image of the cross-section of the fabricated sample. The magnification of the image is 15000, and the scale bar indicating 100 nanometers is provided for reference., b) AFM image of the top layer (Au).

6.5 Validation of the Materials Discovery Network

A filter with a large and selective high reflectance zone extending from 510 nm to 656 nm in the visible spectrum was targeted as a proof of concept using the materials discovery network. As shown in Figure 6.11 (a), the custom device has two reflection dips due to absorption at approximately 439 nm and 717 nm.

By using the optimised material discovery network, the key design parameters were predicted. The model proposed layer thicknesses of 20, 26, 30, 42, 10, and 32 nm, as well as the composition of amorphous layers that form a pair of Cu and TiO_2-22% Au, with a high degree of accuracy. Figure 6.11 (a) compares the target reflectance curve (dashed line) with the fabricated one (red line), and Figure 6.11 (b) shows the schematic of the construction. The reflectance of the fabricated sample fairly matches the target curve, especially when it comes to duplicating the high reflectance zone and the minimum points. This shows the accuracy of the inverse design model.

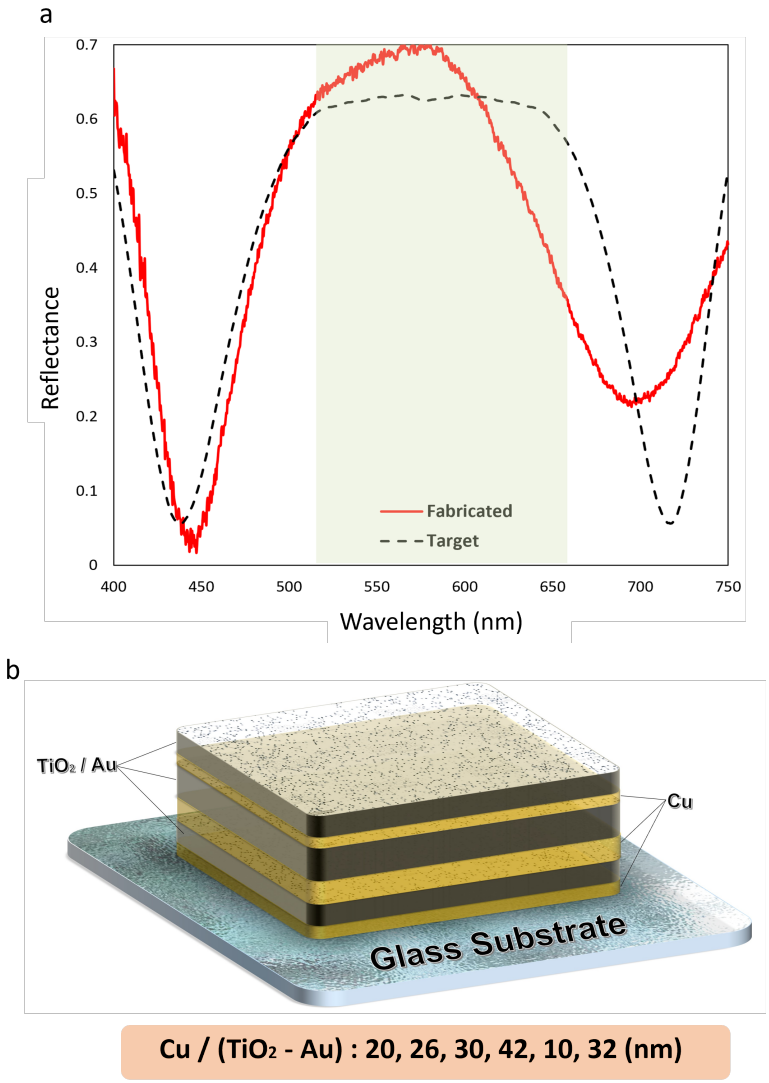


Figure 6.11: a) Comparison of the reflectance spectrum of the on-demand filter (black dashed line) and that of the fabricated (red line) provided by the material discovery network. The thicknesses of the layers are 20, 26, 30, 42, 10, and 32 nm respectively, b) Schematic of the optical filter manufactured by the co-sputtering system. The six-layer structure is made of alternative layers of Cu (green layers) and TiO_2-Au nanocomposite (yellow layers).

There are two primary reasons for the observed mismatch between the custom design and the fabricated sample. Firstly, a homogeneous dispersion of nanoparticles was assumed by the effective medium theory equations. However, some particle agglomerations and nonuniformity in the size of gold nanoparticles are seen in TEM images (cf. Figure 6.12 (a)). These deviations from the ideal dispersion could lead to variations in the material's optical properties [18, 19]. Furthermore, controlling the nanocomposite's surface roughness—as highlighted in the preced-

ing section—is seen as a further potential source of error in achieving the intended result. The surface roughness plays a critical role in influencing interactions between light and matter, thereby impacting the overall efficacy of the fabricated structure in comparison to the idealised design.

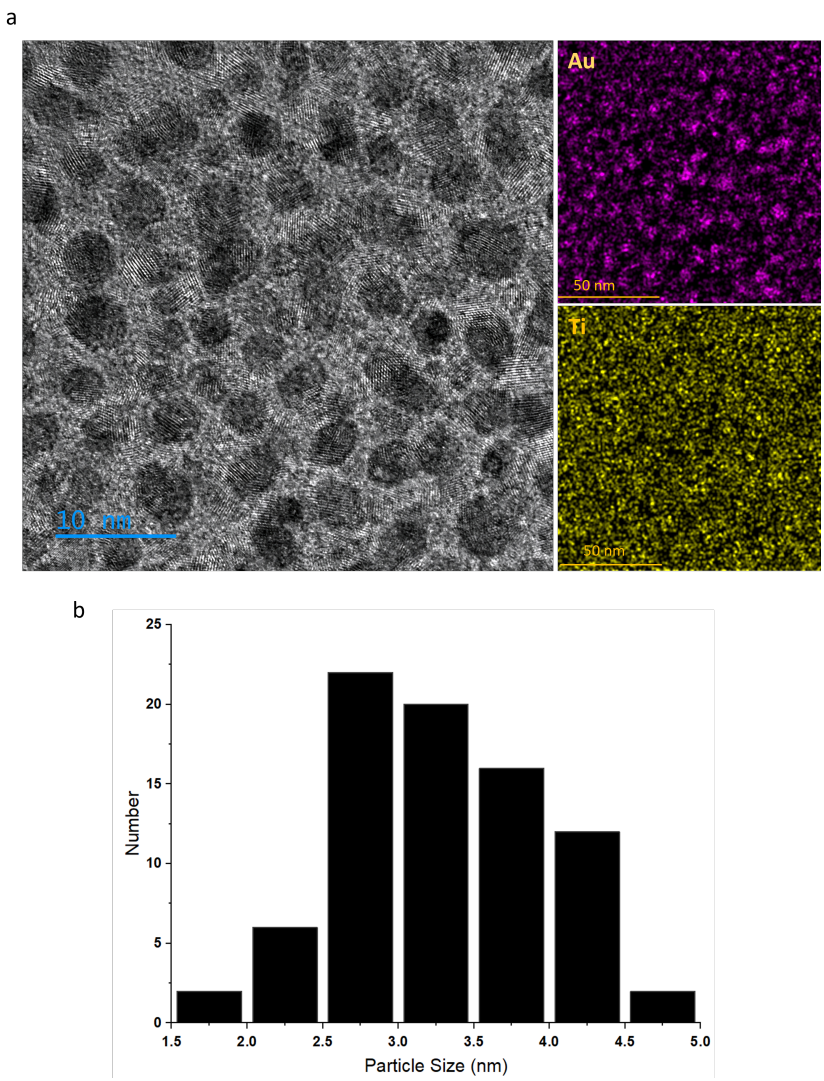


Figure 6.12: a) The top-view TEM image of the nanocomposite layer consisting of Au nanoparticles in TiO_2, and the elemental map of the nanocomposite layer, b) Size distribution of the nanoparticles in the nanocomposite layer.

6.6 Validation of the Free-form Discovery Network

To validate the performance of the inverse free-form geometry network, I used the old samples from previous projects provided by my supervisor. These samples

were used as a proof of concept due to our lack of access to the equipment for their fabrication. The samples had a diverse array of nonohole shapes, spanning from conventional circular shapes to more exotic and complex designs.

The accuracy of the network was evaluated using a random reflectance curve within the visible range, shown as the black dashed line in Figure 6.13 (a). This curve represents the desired performance target, aiming for high reflectivity beyond 550 nm wavelengths for a sensor application. The network's input comprised 36 points from the reflectance spectrum, enabling the extraction of design parameters and materials. The resulting suggested structure, depicted in Figure 6.13 (b), features free-form interconnected holes and is coated with a 20-nm single layer of aluminium. The reflectance curves of the fabricated sample and simulated one are illustrated by the solid red line with circle marker, and solid green line with square marker in Figure 6.13 (a), respectively. The top-view SEM image of the fabricated sample is displayed in Figure 6.13 (c). The residual photoresist portion exhibits dimensions, with a height of 322 nm and a width of 270 nm.

The fabricated sample remarkably matches the desired curve, showcasing the impressive capabilities of our inverse design network in the prediction of freeform structures for on-demand applications.

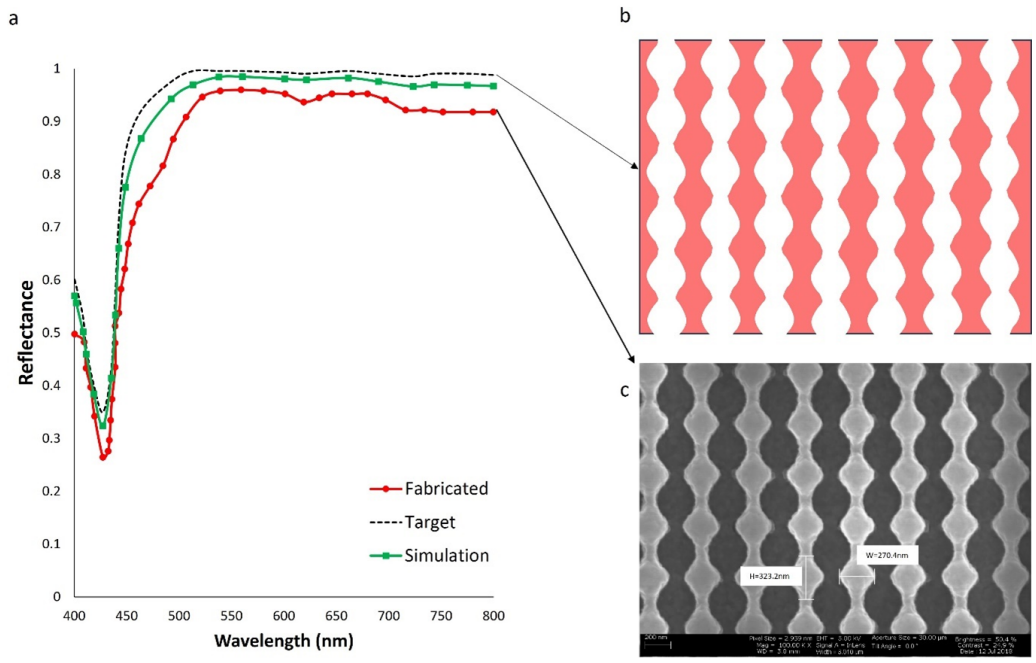


Figure 6.13: a) Comparison of the reflectance spectrum of the on-demand sample (black dashed line), simulated sample (solid green line with square), and that of the fabricated (solid red line with circle) provided by the inverse design network. c) the top view of the structure with exotic shapes of nanohole, which is coated by a 20nm Al, c) a top-view SEM image of the fabricated sample based on the inverse design network prediction.

6.7 Conclusion

The results chapter of this thesis encompasses the experimental validation of the accuracy of three distinct machine learning algorithms. Firstly, a network aimed at inversely designing a multilayer metamaterial for a specific optical filter demonstrated a close fit with predicted values, though some discrepancies arose due to layer roughness, particularly within metallic layers, which were not accounted for in the training dataset. Secondly, a material discovery network was evaluated using an on-demand perfect reflector, suggesting a 6-layer amorphous metamaterial with varied thicknesses. Although the fabricated sample exhibited similar performance, minor mismatches were attributed to layer roughness, dispersion of Au nanoparticles, and particle size uniformity. Lastly, the free-form geometry designer network was experimentally validated using a high-reflectivity sensor beyond 550 nm wavelengths. The network proposed unconventional shapes for the

unit cell, and the fabricated device closely matched the desired curve, showcasing promising results.

Bibliography

[1] P. J. Kelly and R. D. Arnell, "Magnetron sputtering: a review of recent developments and applications," *Vacuum*, vol. 56, no. 3, pp. 159–172, 2000.

[2] A. Gibaud, M. Chebil, and T. Beuvier, "X-ray reflectivity," *Surface Science Techniques*, pp. 191–216, 2013.

[3] J. Tiilikainen, J.-M. Tilli, V. Bosund, M. Mattila, T. Hakkarainen, J. Sormunen, and H. Lipsanen, "Accuracy in x-ray reflectivity analysis," *Journal of Physics D: Applied Physics*, vol. 40, no. 23, p. 7497, 2007.

[4] M. R. Sardela Jr, "X-ray diffraction and reflectivity," in *Practical materials characterization*. Springer, 2014, pp. 1–41.

[5] B. Nicholson, "Understanding interfacial and spin-orbit torque effects in thin film magnetic multilayers." Ph.D. dissertation, Durham University, 2022.

[6] J. Als-Nielsen, D. Jacquemain, K. Kjaer, F. Leveiller, M. Lahav, and L. Leiserowitz, "Principles and applications of grazing incidence x-ray and neutron scattering from ordered molecular monolayers at the air-water interface," *Physics Reports*, vol. 246, no. 5, pp. 251–313, 1994.

[7] L. G. Parratt, "Surface studies of solids by total reflection of x-rays," *Physical review*, vol. 95, no. 2, p. 359, 1954.

[8] C. Majkrzak, "The development of neutron reflectometry as a probe of the nanoscale structure of polymer thin film systems–founded on the pioneering work of professor thomas p. russell," *Nanoscale*, vol. 15, no. 10, pp. 4725–4737, 2023.

[9] S. Sinha, E. Sirota, Garoff, S, and H. Stanley, "X-ray and neutron scattering from rough surfaces," *Physical Review B*, vol. 38, no. 4, p. 2297, 1988.

[10] M. Wormington, "Evidence for grading at polished surfaces from grazing-incidence x-ray scattering," *Philosophical magazine letters*, vol. 74, no. 3, pp. 211–216, 1996.

[11] P. Eaton and P. West, *Atomic force microscopy*. Oxford university press, 2010.

[12] B. Voigtländer, *Atomic force microscopy*. Springer, 2019.

[13] A. Mohammed and A. Abdullah, "Scanning electron microscopy (sem): A review," in *Proceedings of the 2018 International Conference on Hydraulics and Pneumatics—HERVEX, Băile Govora, Romania*, vol. 2018, 2018, pp. 7–9.

[14] C. Tang and Z. Yang, "Transmission electron microscopy (tem)," in *Membrane characterization*. Elsevier, 2017, pp. 145–159.

[15] H.-H. Perkampus, *UV-VIS Spectroscopy and its Applications*. Springer Science & Business Media, 2013.

[16] Y. Mie, H. Okabe, C. Mikami, T. Motomura, and N. Matsuda, "Nanostructured gold thin film electrode derived from surfactant-free gold nanoparticles for enhanced electrocatalysis," *Electrochemistry Communications*, vol. 146, p. 107415, 2023.

[17] C. Liang, Y. Zhong, Q. Zhong, J. Li, W. Cao, X. Wang, S. Wang, X. Xu, J. Wang, and Y. Cao, "Low-temperature deposition of high-quality sio2 films with a sloped sidewall profile for vertical step coverage," *Coatings*, vol. 12, no. 10, p. 1411, 2022.

[18] S. Singh, "Impact of dispersion of nanoscale particles on the properties of nematic liquid crystals," *Crystals*, vol. 9, no. 9, p. 475, 2019.

[19] A. Giusto, R. Saija, M. A. Iatì, P. Denti, F. Borghese, and O. I. Sindoni, "Optical properties of high-density dispersions of particles: application to intralipid solutions," *Applied optics*, vol. 42, no. 21, pp. 4375–4380, 2003.

Chapter 7

Conclusions

7.1 Hypothesis & Chapter Overview

The significant Young's modulus mismatch between the soft and rigid components in a mechanically tunable metasurface leads to stress concentration at the interface upon mechanical loading, causing such issues as crack propagation, delamination, and plastic deformation at lower levels of strain (20-50%). To address this challenge, the hypothesis of this project was to explore geometries capable of redistributing stress concentration away from the interface and towards safer regions of the substrate. Proposed geometries, including kirigami-based and high-relief pillar-based designs, provide local large but low-energy deformation that can be engineered to relocate the stress accumulation. Additionally, machine learning (ML)-based method was used to optimise the structure across three main domains: geometric design parameters, material development, and free-form shape configuration of the unit cell.

This chapter summarises the most important results achieved in each result chapter and suggests some suggestions for further work.

7.2 Thesis Conclusion

We have proposed (Chapter 4) two universal geometries that allow the pairing of rigid and soft materials without limitations on stretchability and function. Mechanical and optical three-dimensional (3D) finite element method (FEM) simulations demonstrated the potential of pillar-based structures and kirigami-inspired patterns to minimise stress at the interface between flexible substrates and resonators, enabling stretchability over a wide range of strains, even exceeding 120%. The optimised structure with maximum stretchability provides a peak shift of 143 nm (from 508 nm to 651 nm), nearly twice that of the conventional geometries where the rigid components directly are located on the polymer. The potential of this wide range of tunability was shown in various applications like rewritable metadevices, plasmonic sensors with high accuracy, and switchable reflector/absorber systems.

The conventional process of discovery, design, and optimisation of the proposed structure is computationally expensive and time-intensive, relying on iterative trial-and-error approaches. In chapter 5 we have presented three deep learning (DL) algorithms as follow:

Design and optimisation network which determines the nonlinear relationship between input (geometrical parameters and strain levels) and output (reflectance) of the ultra-stretchable mechanically tunable metasurface. As a proof of concept, a structural colour system was designed using the kirigami-inspired pattern. The optimised single geometry was characterised by an Al nanodisk with a thickness of 130 nm and a diameter of 200 nm. This configuration provides the broadest areal colour space coverage, encompassing 27.65% of the standard RGB (sRGB) space. By substituting the plasmonic component with a high-dielectric material like TiO_2, a diverse range of colours was achieved in the Commission on Illumination (CIE) 1931 chromaticity diagram, providing a comprehensive colour palette spanning from white to red, green, and blue.

Inverse design and material discovery networks which proposed the composition and thickness of each layer in a 6-layer structure for on-demand optical performance. The optimised convolutional neural network (CNN) was trained with

a dataset generated by the transfer matrix method (TMM) to discover the real and imaginary parts of refractive index (RI) of each layer and the corresponding thickness. However, the material discovery was trained by a dataset generated by effective medium approximation (EMA) to navigate infinite composition space and discover RI values of compositions providing the desired reflectance. The network uses EMA to navigate composition space and discover optimal compositions.

Inverse design of a freeform nanophotonic structure was tasked to explore infinite geometry space to propose a free-from geometry which guarantees the design features of an on-demand reflectance. The mean absolute error (MAE) of the regression task was 0.21, and the F1-scores for the model's classification of the materials of the layers were 0.90.

In the final results section (chapter 6), we have provided the experimental validations of three DL algorithms developed in Chapter 5. Firstly, a network aimed at inversely designing a multilayer metamaterial for an optical filter demonstrated close agreement with predicted values, despite some discrepancies attributed to the layer roughness, particularly within metallic layers. Secondly, a material discovery network suggested a 6-layer amorphous metamaterial for a perfect reflector, with minor mismatches in performance linked to layer roughness, dispersion of Au nanoparticles, and particle size uniformity. Lastly, the free-form geometry discovery network proposed a random shape for a high-reflectivity sensor beyond 550 nm wavelengths, with the fabricated device closely matching the desired curve, indicating promising results.

To summarise, the thesis addressed the limited stretchability in mechanically tunable metasurfaces, where stress concentration at the interface causes plastic deformation and cracks. Through geometry engineering, stress has redistributed from the interface to the substrate, providing a remarkable stretchability exceeding 120%. This interdisciplinary research combined the design, optimisation, and processing of optical metamaterials, employing physics-based simulation, DL, and thin film fabrication techniques to discover new compositions and free-form geometries, along with their corresponding parameters, for custom mechanically

tunable metasurfaces.

7.3 Further Work and Recommendations

This thesis has advanced the AI-assisted design and optimisation of tunable metasurfaces, although further research is needed in this field. Here are some thoughts and suggestions for future study directions.

Consider Other Types of Stimulus

This research concentrated on mechanically tunable metasurfaces, which represent a well-studied and commonly explored example in the field as discussed in Chapter 4. However, it is important to consider certain limitations and complexity associated with mechanically tunable metasurfaces. A significant challenge is managing the interaction of various applied forces like bending, tension, and compression. Considering the combination of these forces requires a precise model to investigate the spatial position of the resonators in the out-of-plane deformation and calculate the optical performance accordingly. In addition, while the potential of the proposed geometries (pillar-based and kirigami-based structures) was discussed in detail, future research could focus on investigating the fatigue behaviour of the structures by subjecting them to cyclic forces [1]. This evaluates the influence of key parameters on structural integrity and performance over time. Moreover, exploring alternative forms of stimuli, such as electrical actuation, temperature responsiveness, and chemical reactivity, could broaden the scope of tunable metasurfaces. For instance, investigating the use of electroactive polymers [2] or phase change materials (PCMs) [3] in the proposed geometry may achieve higher tunability in metasurfaces. This multidisciplinary approach would facilitate a deeper understanding of the design principles and practical applications of tunable metasurfaces beyond mechanical mechanisms.

Transfer Learning Method

Although we showed that deep neural network (DNN)s offer advantages over expensive numerical simulations of Maxwell equations by significantly reducing overall design time, training DNNs requires large training datasets, which can be challenging to obtain. The effectiveness of knowledge transfer learning is worth trying to improve the performance of developed models in different optical scenarios [4]. Using a multi-scenario DNN model training approach with imbalanced datasets which involves training with small datasets (approximately 10^3 scale) for some optical scenarios and normal datasets (approximately 10^4 scale) can significantly reduce the data generation time [5]. By maintaining a suitable training dataset size, particularly for scenarios formed by combinations of materials selected from a material library, multi-scenario training can enhance model accuracy even when trained with small datasets.

Roughness Consideration in Data Generation

In future research, minimising the observed mismatch between reflectance values predicted by machine learning methods and those achieved by fabricated samples will be a priority. As highlighted in sections 6.4 and 6.5, surface roughness at layer interfaces emerged as a significant factor influencing this discrepancy. Surface roughness is known to significantly impact the external properties of thin films [6]. Therefore, future investigations could incorporate roughness as a parameter in the dataset generation process. A potential approach could involve using models such as the sophisticated roughness layer model proposed by Lehmann et al. [7] for multilayer structures, which uses empirical evaluation of atomic force microscopy (AFM) measurements and statistical considerations to propose a general model for various materials with high surface roughness. Integrating such models with EMA could enhance the accuracy of predicting roughness values at material interfaces, thereby refining the performance of DL algorithms in the inverse design (ID) and material discovery for the multilayer structures.

More Complex Multilayer Structures

A promising direction for further research in my PhD thesis would involve refining the developed algorithms discussed in Chapter 5 to accommodate more complex scenarios beyond the simplified assumptions of the initial model. While the algorithms exhibited promising results with general samples, there are some assumptions to simplify the training of the model. For instance, the basic studied structure in this thesis was a 6-layer distributed Bragg reflector under a normal incident light in the visible range wavelength for data generation via the TMM. Expanding the degree of freedom (DOF) in data generation for the networks could introduce complexities but also unlock new potential applications. By incorporating additional variables individually investigated in the previous research such as random incident angles [8], polarisation states [9], a random number of layers [10], and a broader range of wavelengths [11] the model could be used in a wider range of scenarios, thereby enhancing its versatility and effectiveness in optical design tasks. This advancement would enable the integration of the material discovery concept developed in this thesis with the optimisation of multiple variables, ultimately improving the performance and adaptability of the algorithms to real-world applications.

Bibliography

[1] I. Koch, G. Just, M. Brod, J. Chen, A. Doblies, A. Dean, M. Gude, R. Rolfes, C. Hopmann, and B. Fiedler, "Evaluation and modeling of the fatigue damage behavior of polymer composites at reversed cyclic loading," *Materials*, vol. 12, no. 11, p. 1727, 2019.

[2] S. Chen and M. P. Jonsson, "Dynamic conducting polymer plasmonics and metasurfaces," *ACS Photonics*, vol. 10, no. 3, pp. 571–581, 2023.

[3] D.-Q. Zhang, G.-M. Pan, Z.-W. Jin, F.-Z. Shu, X.-F. Jing, Z. Hong, and C.-Y. Shen, "Tunable dielectric metasurfaces by structuring the phase-change material," *Optics Express*, vol. 30, no. 3, pp. 4312–4326, 2022.

[4] P. J. Freire, D. Abode, J. E. Prilepsky, N. Costa, B. Spinnler, A. Napoli, and S. K. Turitsyn, "Transfer learning for neural networks-based equalizers in coherent optical systems," *Journal of Lightwave Technology*, vol. 39, no. 21, pp. 6733–6745, 2021.

[5] C. Qiu, X. Wu, Z. Luo, H. Yang, G. He, and B. Huang, "Nanophotonic inverse design with deep neural networks based on knowledge transfer using imbalanced datasets," *Optics Express*, vol. 29, no. 18, pp. 28 406–28 415, 2021.

[6] G. Song, Y. Wang, and D. Q. Tan, "A review of surface roughness impact on dielectric film properties," 2022.

[7] D. Lehmann, F. Seidel, and D. R. Zahn, "Thin films with high surface roughness: thickness and dielectric function analysis using spectroscopic ellipsometry," *SpringerPlus*, vol. 3, pp. 1–8, 2014.

[8] A. Lininger, M. Hinczewski, and G. Strangi, "General inverse design of layered thin-film materials with convolutional neural networks," *ACS Photonics*, vol. 8, no. 12, pp. 3641–3650, 2021.

[9] R. Unni, K. Yao, X. Han, M. Zhou, and Y. Zheng, "A mixture-density-based tandem optimization network for on-demand inverse design of thin-film high reflectors," *Nanophotonics*, vol. 10, no. 16, pp. 4057–4065, 2021.

[10] A. Luce, A. Mahdavi, H. Wankerl, and F. Marquardt, "Investigation of inverse design of multilayer thin-films with conditional invertible neural networks," *Machine Learning: Science and Technology*, 2022.

[11] Q. Pan, S. Zhou, S. Chen, C. Yu, Y. Guo, and Y. Shuai, "Deep learning-based inverse design optimization of efficient multilayer thermal emitters in the near-infrared broad spectrum," *Optics Express*, vol. 31, no. 15, pp. 23 944–23 951, 2023.