



UNIVERSIDADE FEDERAL DA BAHIA
ESCOLA POLITÉCNICA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA ELÉTRICA

RONNEY DAS MERCES CERQUEIRA

**ALGORITMO BASEADO EM MACHINE LEARNING
PARA PROJETO DE DISPOSITIVOS DE
INTERFERÊNCIA MULTIMODO**

SALVADOR - BA
Junho de 2024

RONEY DAS MERCES CERQUEIRA

**ALGORITMO BASEADO EM MACHINE LEARNING
PARA PROJETO DE DISPOSITIVOS DE
INTERFERÊNCIA MULTIMODO**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, PPGEE da Universidade Federal da Bahia, como parte dos requisitos necessários para obtenção do grau de Mestre em Engenharia Elétrica.

Orientadores:

Prof. Dr. Vitaly Félix Rodríguez Esquerre

Prof. Dr. Anderson Dourado Sisnando

SALVADOR - BA

Junho de 2024

das Mercês Cerqueira, Roney

Algoritmo Baseado em Machine Learning para Projeto De Dispositivos De Interferência Multimodo / Roney das Mercês Cerqueira. Salvador, 2024.

88 f.

Orientador: Vitaly Félix Rodriguez Esquerre.

Coorientador: Anderson Dourado Sisnando.

Dissertação (Mestrado - Mestrado em Engenharia Elétrica) -- Universidade Federal da Bahia, UFBA, 2024.

1. Interferência Multimodo. 2. Machine Learning. 3. Design Inverso. I. Rodriguez Esquerre, Vitaly Félix. II. Dourado Sisnando, Anderson. III. Título.



UNIVERSIDADE FEDERAL DA BAHIA
ESCOLA POLITÉCNICA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**ALGORITMO BASEADO EM MACHINE LEARNING
PARA PROJETO DE DISPOSITIVOS DE
INTERFERÊNCIA MULTIMODO**

DISSERTAÇÃO DE MESTRADO

Autor: Roney das Mercês Cerqueira

Orientador: Vitaly Félix Rodriguez Esquerre

Dissertação de Mestrado aprovada em 12 de junho de 2024 pela banca examinadora composta pelos seguintes membros:

Prof. Dr. Vitaly Félix Rodriguez Esquerre

Prof. Dr. Anderson Dourado Sisnando

Prof. Dr. Karcus Day Rosário Assis

Prof. Dr. Cosme Eustáquio Rubio Mercedes

E disse Deus: Haja luz... e houve luz

AGRADECIMENTOS

Primeiramente agradeço a Deus por ter me dado forças, ânimo e capacidade para concluir este trabalho de dissertação. A minha esposa Tatiane Mercês dos Santos Cerqueira, que sempre esteve ao meu lado com palavras de incentivo, carinho e apoio incondicional, sendo uma pedra fundamental nessa caminhada: Muito obrigado, a sua presença muda tudo em minha vida!!! Aos Meus pais, Sra. Nilzete das Mercês Cerqueira e Sr. Roberval Silva Cerqueira (*in memoriam*) que sempre me mostraram que a jornada dos estudos engrandece e transforma o homem. Estendo meus agradecimentos aos meus amigos e irmãos de trabalho da CETO do 16º BPM, que sempre estiveram dispostos a ajudar em quaisquer circunstâncias, em especial ao turno DELTA e aos meus amigos de parceria acadêmica. Estender também meus agradecimentos aos professores que foram peças fundamentais para essa jornada, os quais contribuíram substancialmente pra trilhar este novo caminho, especialmente aos professores orientadores Dr. Anderson Dourado Sisnando e Vitaly Félix Rodríguez Esquerre: Muito obrigado professores pela paciência, por todo o conhecimento transmitido e por confiar a mim esse trabalho tão significativo pra minha vida!!!! Hoje eu venço mais uma batalha, com honra e orgulho de ter ao meu lado pessoas tão especiais, por isso merecem minha eterna gratidão.

Simplesmente Obrigado.

RESUMO

O aprimoramento dos Circuitos Fotônicos Integrados tornando-os mais flexíveis, reconfiguráveis e compactos tem sido via de regra para os sistemas de telecomunicações. Diante deste cenário, os olhares tem se voltado para os dispositivos de Interferência Multimodo devido as vantagens apresentadas em suas propriedades e funcionalidades. Neste projeto de dissertação, propõe utilizar técnicas de *Machine Learning* para design inverso de dispositivos de Interferência Multimodo, como divisor de potência na faixa de comprimento de ondas (λ) de 1,25 ~ 1,7 μ m integrando as bandas O, E, S, C, L e U. Para tal, além de realizar um levantamento bibliográfico, foi necessário simular dispositivos de interferência Multimodo em software específico e desenvolver um algoritmo utilizando Redes Neurais Artificiais. A arquitetura da rede neural foi configurada com os seguintes parâmetros de entrada: comprimento de ondas (λ), índice de refração do núcleo (n_1), índice de refração do substrato (n_2), largura do dispositivo (W_{MMI}) e eficiência de transmissão (%) e como parâmetro de saída teve como alvo as coordenadas do eixo- x_1 , eixo- x_2 referente a posição das portas de saída (1 x M) do dispositivo e a coordenadas do eixo-y referente ao comprimento (L_{MMI}) de corte do dispositivo para maior potência acoplada das portas de saída. Como resultados, a Rede Neural Artificial desenvolvida apresentou o Erro Quadrático Médio de validação cruzada equivalente a $6,39410 \times 10^{-5}$, regressão linear de 0,99997 e tempo de processamento computacional de 7,46 segundos, capaz de fornecer dados para design dos dispositivos de Interferência Multimodal compacto com dimensões a partir $2,00 \times 5,32\mu$ m e perdas que variam entre 0,32 a 0,47dB para o dispositivo mais eficiente.

Palavras-Chaves: Interferência Multimodo; Machine Learning; Design Inverso

ABSTRACT

The improvement of Photonic Integrated Circuits to make them more flexible, reconfigurable and compact has been the norm for telecommunications systems. Against this backdrop, attention has turned to Multimode Interference devices due to the advantages presented by their properties and functionalities. This dissertation project proposes using Machine Learning techniques for the inverse design of Multimode Interference devices, as a power divider in the wavelength range of $1.25 \sim 1.7\mu\text{m}$, integrating the O, E, S, C, L and U bands. To do this, in addition to carrying out a literature survey, it was necessary to simulate Multimode interference devices in specific software and develop an algorithm using Artificial Neural Networks. The neural network architecture was configured with the following input parameters: wavelength (λ), core refractive index (n_1), substrate refractive index (n_2), device width (W_{MMI}) and transmission efficiency (%) and as output parameters it targeted the coordinates of the x-axis1, x-axis2 referring to the position of the output ports ($1 \times M$) of the device and the coordinates of the y-axis referring to the cut-off length (L_{MMI}) of the arrangement for the highest coupled power of the output ports. As a result, the Artificial Neural Network developed presented a cross-validation Mean Square Error equivalent to 6.39410×10^{-5} , a linear regression of 0.99997 and a computational processing time of 7.46 seconds, capable of providing data for the design of compact Multimodal Interference devices with dimensions from $2.00 \times 5.32\mu\text{m}$ and losses ranging from 0.32 to 0.47dB for the most efficient device.

Key Words: Multimode Interference; Machine Learning; Inverse Design

LISTAS DE FIGURAS

Figura 1.1 – Linha do tempo de publicações de projeto de dispositivos fotônicos utilizando Redes Neurais	5
Figura 1.2 – Resultados das buscas nas bases de dados	7
Figura 1.3 – Diagrama de seleção de artigos.....	8
Figura 1.4 – Análise dos Artigos Selecionados.....	10
Figura 2.1– Espectro eletromagnético	14
Figura 2.2 – Guia de Onda 2D.....	15
Figura 3.1a – Esquema 3D guia de ondas MMI 1×2	20
Figura 3.1b – Esquema 2D guia de ondas MMI 1×2	21
Figura 3.2 – Representação do perfil do índice de refração e vista superior do guia de ondas multimodo	22
Figura 3.3 – Perfis de campo lateral com amplitude normalizada	22
Figura 4.1 – Subáreas da Inteligência Artificial	28
Figura 4.2 – Classificação do <i>Machine Learning</i>	30
Figura 4.3 – Representação gráfica de Regressão e Classificação	32
Figura 4.4 – Neurônio biológico	33
Figura 4.5 – Neurônio artificial	34
Figura 4.6 – Rede Neural Artificial <i>Perceptron</i> Multicamadas	37
Figura 6.1 – Campo formado pela <i>Self-image</i> das simulações com dados analíticos de dispositivos MMI (1×2) para o comprimento de ondas de $1,55\mu\text{m}$. a) $2,00 \times 5,33\mu\text{m}$; b) $3,00 \times 11,35\mu\text{m}$; c) $4,00 \times 19,60\mu\text{m}$; d) $5,00 \times 30,03\mu\text{m}$	43
Figura 6.2 – Campo formado pela <i>Self-image</i> das simulações com dados analíticos de dispositivos MMI (1×2) para o comprimento de ondas de $1,55\mu\text{m}$, com a proporção das dimensões conservadas. a) $2,00 \times 5,33\mu\text{m}$; b) $3,00 \times 11,35\mu\text{m}$; c) $4,00 \times 19,60\mu\text{m}$; d) $5,00 \times 30,03\mu\text{m}$	44
Figura 6.3 – Eficiência do dispositivo MMI (1×2) com LMMI de $2,20\mu\text{m}$ simulado analiticamente, para a faixa de comprimento de ondas de $1,25 \sim 1,7\mu\text{m}$ integrando as bandas O, E, S, C, L e U	45

Figura 6.4 – Arquitetura da Rede Neural Artificial	46
Figura 6.5 – Regressão Linear da Melhor Rede Neural Artificial (RNA 27- <i>Tan-Tan-Log</i>) para treinamento, teste, validação e geral com Regressão $\cong 1$, Y=Saída e T=Target	50
Figura 6.6 – Performance da Melhor Rede Neural Artificial (RNA 27- <i>Tan-Tan-Tan</i>) com melhor performance de validação em 114 épocas e Erro Médio Quadrático de $2,05336 \times 10^{-4}$	51
Figura 6.7 – Histograma de erro Melhor Rede Neural Artificial (RNA 01- <i>Tan-Tan-Tan</i>) com a maior quantidade de dados com a diferença entre a predição e o alvo próximo de $2,052 \times 10^{-3}$	51
Figura 6.8 – Campo formado pela <i>Self-image</i> das simulações com dados obtidos através da RNA 01- <i>Tan-Tan-tan</i> de dispositivos MMI (1 × 2) para o comprimento de ondas de 1,55 μm , com a proporção das dimensões conservadas. a) 2,00×5,3276 μm ; b) 3,00×11,3518 μm ; c) 4,00×19,5989 μm ; d) 5,00×29,9786 μm (Parâmetros da tabela 6.7).....	53
Figura 6.9 – Campo formado pela <i>Self-image</i> das simulações com dados obtidos através da RNA 01- <i>Tan-Tan-tan</i> de dispositivos MMI (1 × 2) para o comprimento de ondas de 1,55 μm . a) 2,00×5,3276 μm ; b) 3,00×11,3518 μm ; c) 4,00×19,5989 μm ; d) 5,00×29,9786 μm (Parâmetros da tabela 6.7)	54
Figura 6.10 – Coordenadas dos eixo-x e eixo-y dos dados analíticos vs. valores obtidos pelas RNA 01- <i>Tan-Tan-Tan</i>	55
Figura 6.11 – Comparação da Eficiência entre dispositivos MMI (1 × 2) com LMMI de 2,20 μm simulado analiticamente e predito pela Rede Neural Artificial de melhor desempenho, para a faixa de comprimento de ondas de 1,25 ~ 1,7 μm integrando as bandas O, E, S, C, L e U.....	56
Figura 6.12 – Comparação da Eficiência entre dispositivos MMI (1 × 2) com LMMI de 2,20 μm simulado analiticamente e predito pela Rede Neural Artificial de melhor desempenho, para a faixa de comprimento de ondas de 1,25 ~ 1,7 μm integrando as bandas O, E, S, C, L e U.....	57
Figura 6.13 – Campo formado pela <i>Self-image</i> da simulação com dados obtidos através da RNA 01- <i>Tan-Tan-Tan</i> de dispositivos MMI (1 × 2) com 2,13×5,3276 μm , para o comprimento de ondas de 1,55 μm	57

LISTAS DE TABELAS

Tabela 1.1 – Resultado das buscas através dos termos para as bases de dados	7
Tabela 1.2 – Lista de periódicos, quantidade de artigos e fator de impacto.....	8
Tabela 1.3 – Análise dos Artigos Selecionados	9
Tabela 6.1 – Parâmetros de configurações globais para <i>RSoft-BeamPROP</i>	41
Tabela 6.2 – Parâmetros Ópticos e Geométricos de quatro dispositivos MMI's (1 × 2) simulados analiticamente	42
Tabela 6.3 – Arquitetura e Desempenho das Redes Neurais Artificiais baseadas na média do Erro Quadrático Médio da validação cruzada, regressão e tempo de processamento	47
Tabela 6.4 – Desempenho das 10 melhores Redes Neurais Artificiais utilizando como critério de avaliação o Erro Quadrático Médio da validação cruzada.....	48
Tabela 6.5 – Desempenho das Redes Neurais Artificiais utilizando como critério de avaliação o tempo de processamento computacional.....	49
Tabela 6.6 – Desempenho das 10 melhores Redes Neurais Artificiais utilizando como critério de avaliação a regressão linear	50
Tabela 6.7 – Parâmetros Ópticos e Geométricos de 4 dispositivos MMI's (1 × 2) obtidos através da Rede Neural Artificial de melhor desempenho com nomenclatura RNA 01-Tan-Tan-Tan vs. Dispositivos simulados analiticamente.....	52
Tabela 6.8 – Comparação dos resultados de trabalhos disponível na análise bibliográfica com os resultados apresentado neste trabalho de dissertação	58

LISTAS DE ABREVEATURAS E SIGLAS

AIP	American Institute of Physics
BPM	Método de Propagação de Feixe (BPM)
EQM	Erro Quadrático Médio
FiO LS	Frontiers in Optics – Laser Science
IA	Inteligência Artificial
IEEE	International Electrical and Electronics Engineering
LM	Levenberg-Marquardt
L _{MMI}	Comprimento do dispositivo de Interferência Multimodo
Logsig	Função Logística
MCP	Neurônio Artificial McCulloch & Pitts
ML	Machine Learning
MLP	Multilayer Perceptron
MMI	Interferência Multimodal
MZI	Mach-Zehnder
OSA	Optica publishing Group
PCI's	Photonic Integrated Circuits
Purelin	Função Linear
ReLU	Rectified Linear Unit
RNA	Redes Neurais Artificiais (RNA)
Si	Silício
SiO ₂	Sílica
SPIE	International Society for Optics and Photonics
Tansig	Função Tangente Hiperbólica
TE	Transversal Elétrica
TM	Transversal Magnética
WG	Waveguides
W _{MMI}	Largura do dispositivo de Interferência Multimodo

LISTAS DE SÍMBOLOS

A	Coeficiente da Amplitude
B	Densidade de Fluxo Magnético
c	Coeficiente do modo de Excitação do Campo
D	Densidade de Fluxo Elétrico
d	Distância
E	Campo Elétrico
H	Campo Magnético
J	Densidade de corrente
k	Vetor de onda
M	PORTA DE SAÍDA
m	Metros
N	PORTA DE ENTRADA
n	Índice d Refração do Meio
n_1	Índice de Refração do núcleo
n_2	Índice de Refração do revestimento
n_{clad}	Índice de Refração da casca
n_{eff}	Índice de Refração Efetivo
n_s	Índice de Refração do Substrato
r	Vetor de Posição
s	Segundos
v	Velocidade
W_e	Largura efetiva
W_M	largura geométrica
β_v	Constante de propagação
ε	Permissividade
ε_0	Permissividade Dielétrica no Espaço Livre
ε_r	Permissividade Dielétrica Relativa
λ_0	comprimento de onda no espaço livre
ρ	Densidade Volumétrica de Carga
Ψ	Perfil do campo
ψ	Distribuição do Campo Modal
μ	Permeabilidade
μ_0	Permeabilidade Magnética no Espaço Livre
Θ_i	Ângulo de Incidência
Θ_r	Ângulo de Reflexão

SUMÁRIO

AGRADECIMENTOS	VI
RESUMO.....	VII
ABSTRACT	VIII
LISTAS DE FIGURAS	IX
LISTAS DE TABELAS.....	XI
LISTAS DE ABREVEATURAS E SIGLAS.....	XII
LISTAS DE SÍMBOLOS	XIII
CAPITULO I	1
1. INTRODUÇÃO	1
1.1. Justificativa	2
1.2. Objetivos	3
1.2.1. <i>Objetivo Geral</i>	3
1.2.2. <i>Objetivos Específicos</i>	4
1.3. Análise Bibliográfica	4
1.3.1. <i>Método de pesquisa</i>	5
1.4. Organização da Dissertação	11
CAPITULO II.....	12
2. DISPOSITIVOS FOTÓNICOS	12
2.1. Breve histórico	12
2.2. Aplicações.....	14
2.3. Guia de Ondas	15
CAPITULO III	20
3. DISPOSITIVO DE INTERFERÊNCIA MULTIMODO (MMI).....	20
3.1. Constantes de propagação	22
CAPITULO IV.....	27
4. INTELIGENCIA ARTIFICIAL	27
4.1. Machine Learning	29
4.1.1. <i>Métodos de Machine Learning</i>	29
4.1.2. <i>Técnicas e Algoritmos de Machine Learning</i>	31
4.2. Redes Neurais Artificiais	33
4.2.1. <i>Treinamento Supervisionado</i>	35

4.2.2. <i>Backpropagation e Rede Feedforward</i>	36
4.2.3. <i>Multilayer Perceptron (MLP)</i>	36
4.2.4. <i>Levenberg-Marquardt</i>	38
CAPITULO V	39
5. METODOLOGIA	39
CAPITULO VI	41
6. RESULTADOS E DISCUSSÕES	41
CONCLUSÃO	59
ARTIGOS ACEITOS E PUBLICADOS	62
REFERÊNCIAS BIBLIOGRÁFICAS	63

CAPITULO I

1. INTRODUÇÃO

A crescente demanda das comunicações ópticas com maior taxa de bits para transmissão de grande volume de dados, tem sido um fator importante na corrida por dispositivos miniaturizados, o que suscitou o aprimoramento de Circuitos Fotônicos Integrados (*Photonic Integrated Circuits-PCI's*), assim tornando-os mais flexíveis e reconfiguráveis [1]. Novas estruturas, materiais, dispositivos e sistemas fotônicos têm impulsionado tecnologias transformadoras, incluindo comunicações ópticas, computação de alta velocidade, bem como processamento de informação quântica [2].

Para desenvolver pesquisas que visam projetar dispositivos fotônicos, é necessário atender duas abordagens principais. A primeira abordagem leva em consideração os métodos baseados na física com modelos analíticos simplificados. Embora essa abordagem ofereça bases importantes, não é simples projetar as estruturas certas para realizar as propriedades fotônicas desejadas, principalmente quando se trata de geometrias complexas. A segunda abordagem está fundamentada na modelagem eletromagnética baseada em métodos de simulação numérica com ou sem algoritmo de otimização através de simulações computacionais, pois oferece um maior grau de liberdade para ajustes geométricos com o objetivo de obter uma resposta óptica desejada [1]. Graças às técnicas avançadas de simulação numérica é possível projetar, fabricar e demonstrar micro e nanoestruturas com geometrias e arranjos sofisticados.

A capacidade de manipular a luz de maneira desejada a volumes menores que o seu comprimento de onda no espaço livre impulsionaram importantes pesquisas e aplicações práticas de dispositivos fotônicos em diversas áreas do conhecimento [3], em especial para os dispositivos de Interferência Multimodal (MMI) devido às suas funcionalidades, aplicações e excelentes propriedades e integração. A partir dos dispositivos de interferência multimodo é possível extrair várias funcionalidades, entre eles o *Mach-Zehnder modulators* [4], *Mach-Zehnder Switch* [5], *MMI (De)Multiplexing* [6], *MMI Couplers* [7], *MMI sensors* [8], *Power Combiners* [9], *Power Splitters* [10].

A construção de dispositivos fotônicos com estruturas complexas ainda é um processo difícil. E devido à essa complexidade estrutural, a Inteligência Artificial (IA), através das Redes Neurais Artificiais (RNA) servem como uma solução potencial para contornar o complicado

processo de design [11], além de oferecer uma nova e poderosa abordagem para projetos fotônicos [12], [13], [14] permitindo-nos prever diretamente as respostas ópticas ou realizar o projeto inverso com alta eficiência e precisão [15], [16] em comprimentos de ondas da banda convencional (c) de estruturas fotônicas [17].

A literatura revela que há uma busca incessante na tentativa de aumentar o desempenho e reduzir custos de fabricação de dispositivos compactos e com características complexas que envolvem mecanismos tecnológicos sofisticados, testes e impressão [18]. Uma solução que pode oferecer vantagens para reduzir custos de fabricação e teste sem abrir mão de mecanismos computacionais é a utilização de algoritmos de aprendizado através técnicas de *Machine Learning (ML)* [19], [20] para projetar dispositivos fotônicos baseado no designer inverso [21], [22], [23].

Neste trabalho além de discutir os desafios experimentais, as perspectivas atuais desta área de pesquisa interdisciplinar [24], [25] e seu potencial, em um novo paradigma de ciência e engenharia no qual a fotônica e a Inteligência Artificial estão interligadas, buscou-se realizar uma revisão na literatura e demonstrar arquiteturas de RNA's mais importantes, enfatizando suas habilidades para projetar dispositivos MMI capaz de mapear características geométricas e ópticas permitindo a previsão e análise direta.

1.1. Justificativa

As equações de *Maxwell* descrevem completamente os fenômenos eletromagnéticos bem como são as bases fundamentais para modelagem das equações de guiamento de sinais ópticos. No entanto, sua solução analítica torna-se complexa quando aplicadas em dispositivos com estruturas e geometrias complexas. Devido à complexidade dos conceitos que modelam os dispositivos fotônicos com essas características, este trabalho propõe desenvolver e disponibilizar modelos utilizando Redes Neurais Artificiais como técnica de *Machine Learning*, para obter uma solução aproximada e de simples manipulação, no intuito de auxiliar profissionais atuando na área de telecomunicações para que possam projetar dispositivos sem a necessidade de possuir conhecimentos da teoria eletromagnética, das técnicas de análise convencionais e dos métodos de aprendizado de máquinas.

Os dispositivos de Interferência Multimodo têm se tornado um elemento muito importante nos sistemas de comunicações ópticas. Desenvolver mecanismos que facilitem

projetar esses dispositivos que possuem características peculiares e fazê-lo do seu processo de projeção algo mais simples é um fator preponderante para torna-lo acessível para o maior público possível, além de contribuir para novas pesquisas e desenvolvimento desses dispositivos com geometrias diferentes das convencionais e materiais com outras propriedades ópticas.

As Redes Neurais Artificiais são um dos métodos mais utilizados para aprendizado de máquinas e constantemente tem sido apresentado novas técnicas que oferecem resultados cada vez melhor com redes mais eficientes. Neste trabalho, busca investigar o aprimoramento das técnicas utilizadas nas RNA's, testando novas configurações e novas funções para serem aplicadas e auxiliar nos projetos de dispositivos de Interferência Multimodo.

Este trabalho torna-se relevante no que diz respeito ao desenvolvimento de um algoritmo eficiente, capaz estabelecer parâmetros ideais suficiente para reduzir custos quando pretende-se fabricar dispositivos fotônicos. Uma vez que, a fabricação de um dispositivo físico com essas características e dimensões nanométricas possui um custo muito elevado, descartando qualquer possibilidade de erro no designer do projeto.

Outro fator relevante para o desenvolvimento deste trabalho, é a possibilidade de utilizar o algoritmo na experimentação de novos materiais ópticos com geometrias diferentes das convencionais, para projeção de dispositivo utilizando de baixo custo e com boa eficiência de transmissão.

O tema proposto é de interesse atual e os resultados obtidos são passíveis de publicação. Além disso, este projeto será o ponto de partida para outros nesta mesma linha de pesquisa.

1.2. Objetivos

1.2.1. Objetivo Geral

O objetivo deste trabalho é realizar uma análise utilizando Redes Neurais Artificiais através de técnicas de *Machine Learning* para projetar dispositivos de Interferência Multimodo compacto, com boa tolerância de fabricação e baixas perdas ópticas, atuando como divisor de potência (*Power Splitters*) para operar nas faixas de comprimentos de ondas das bandas O, E, S, C, L e U.

1.2.2. *Objetivos Específicos*

- Realizar uma análise bibliográfica para identificar as diversas tipologias de dispositivos de Interferência Multimodo em base de dados específicos de publicações relacionados a fotônica e óptica integrada;
- Pesquisar e relacionar as técnicas de *Machine Learning* utilizadas em projetos de dispositivo de Interferência Multimodo;
- Realizar simulações de dispositivo MMI compactos com geometrias diferentes;
- Catalogar dados ópticos e geométricos gerados a partir das simulações para alimentar as Redes Neurais Artificiais;
- Desenvolver um algoritmo capaz de prever as coordenadas para o eixo- x das portas de saídas M ($1 \times M$) de maior potência e as coordenadas para eixo- y responsável pelo limite de corte para acoplamento das portas de saída para um dispositivo MMI compacto;
- Realizar simulações de dispositivos MMI compactos com parâmetros obtidos do algoritmo de aprendizado e analisar suas características de propagação;
- Elaborar conclusões que validem os resultados obtidos desta pesquisa, bem como sua contribuição na utilização de *Machine Learning* para design de dispositivos MMI e aplicação na área de óptica integrada

1.3. **Análise Bibliográfica**

As comunicações ópticas sofreram uma evolução significativa, revolucionando a forma de como dados são recebidos e transmitidos através de dispositivos mais sofisticados, demonstrando um grande impacto nas diversas indústrias do setor. Com avanços e pesquisas contínuas, esse campo se destaca com grande potencial para desenvolvimento futuro. No últimos anos, a área de comunicações ópticas valeu-se de novas ferramentas e métodos para fabricar dispositivos fotônicos, com tecnologia inovadora fundamentada na técnica de algoritmos computacionais inteligentes [18], [26].

A evolução da utilização de redes neurais para designer de dispositivos fotônicos data de 1990 com modelagem de dispositivos eletromagnéticos de micro-ondas, pouco tempo depois, redes profundas totalmente conectadas com pelo menos duas camadas foram usadas para modelar elementos de circuito de micro-ondas mais complexos. A partir de 2010 pesquisadores começaram a trabalhar com fotônica de Silício e fibras ópticas para modelagem de rede neural

de sistemas de ondas guiadas, período em que a *Deep Learning* tem crescimento significativo nesse campo. E finalmente, a partir de 2017 surgem as primeiras pesquisas utilizando esse método para modelar sistemas nanofotônicos em espaço livre. Esse processo histórico pode ser observado na figura 1.1 [27].

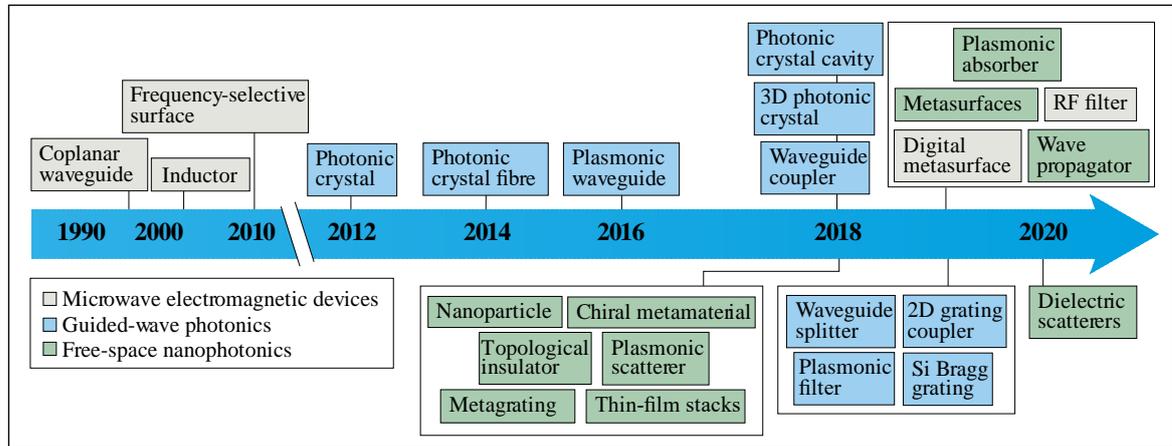


Figura 1.1 – Linha do tempo de publicações de projeto de dispositivos fotônicos utilizando Redes Neurais [27]

A Inteligência Artificial é um importante campo adotado por este setor para design de dispositivos fotônicos, dentre eles destaca-se a aplicação de técnicas de *Machine Learning*. O objetivo deste capítulo é apresentar uma revisão sistemática da aplicação de *ML* para design de dispositivos de interferência multimodo, e para isso foi investigado como a literatura tem lidado com relevantes desafios da aplicação da inteligência computacional nesta área de conhecimento.

Na literatura foi identificado outros trabalhos que realizaram revisões sobre a aplicação de *ML* em fotônica, relatando a evolução e designer de dispositivos desta natureza, bem como concepções e progressos recentes [27], [28]. No entanto, a presente revisão sistemática inova principalmente nos seguintes aspectos: metodologia de pesquisa observando os estudos mais citados e mais recentes de periódicos da área de comunicações ópticas; análise de dados das publicações; e análise de ferramentas computacionais para simulação e desenvolvimento de algoritmos de aprendizagem.

1.3.1. Método de pesquisa

O objetivo desta revisão foi analisar estudos e pesquisas com a temática de *Machine Learning* em design de dispositivos de interferência multimodo. Para tal, foi estipulado algumas perguntas para análise dos artigos:

1. Qual a principal área abordada do artigo?
2. Quais as principais aplicações do dispositivo projetado?
3. Quais as técnicas de *Machine Learning* foram aplicadas?
4. Quais softwares são utilizados para desenvolvimento do algoritmo de aprendizagem?
5. Quais softwares são utilizados para simulação dos dispositivos MMI?
6. Quais materiais são utilizados para fabricação dos dispositivos MMI?
7. Quais os principais resultados dos artigos?

Devido ao número restrito de artigos publicados com temática específica de acordo com tema tratado nesta dissertação, houve a necessidade de adotar mais de uma base de dados ampliando assim o número de periódicos. Para tal optou-se por consultar as bases de dados da *Optica Publish Group*, *Google Scholar* e *Capes Periódicos*.

Para realizar as buscas nas referidas bases de dados, antes de mais nada é preciso definir os termos (palavras chaves) que serão utilizados. A opção dos termos escolhidos para este procedimento foi determinado a partir de pesquisas preliminares nas bases de dados, onde pode constatar resultados com maior número de artigos mais adequado para o objetivo desta etapa. Dessa forma os termos utilizados foram:

- “*Interference Multimode*” AND “*Machine Learning*”;
- “*Interference Multimode*” AND “*Neural Networks*”;
- “*Power Splitter*” AND “*Machine Learning*”;
- “*Power Splitter*” AND “*Neural Networks*”;

A busca em língua inglesa possibilitou identificar um maior número de artigos e mais adequados para análise literária. A tabela 1.1 mostra os respectivos resultados considerando os termos de busca (palavras-chaves) utilizados em cada base de dados, onde “*r*” indica a quantidade de resultados por busca, “*s*” a quantidade de artigos selecionando e “*ar*” a quantidade de artigos repetidos.

Tabela 1.1 – Resultado das buscas através dos termos para as bases de dados

Termos de buscas	Optica Publish Group			Google Scholar			Capes Periódicos			Total		
	r	s	ar	r	s	ar	r	s	ar	r	s	ar
“Interference Multimode” AND “Machine Learning”	12	1	0	30	0	0	19	0	0	61	1	0
“Interference Multimode” AND “Neural Networks”	14	1	1	30	1	0	55	2	4	99	4	5
“Power Splitter” AND “Neural Networks”	32	7	1	30	3	7	51	4	8	113	14	16
“Power Splitter” AND “Machine Learning”	17	4	6	30	4	2	28	3	5	75	11	13
Total	75	11	8	120	7	9	153	9	17	348	27	34

Fonte: Próprio autor

A figura 1.2 sumariza a tabela 1.1, e como podemos observar, os trabalhos selecionados totalizaram 27 artigos dentre os 348 apresentados nos resultados das buscas. A base de dados da *Optica Publish Group* forneceu a maior quantidade de artigos, por se tratar de uma plataforma específica na área de conhecimento que integra este projeto.

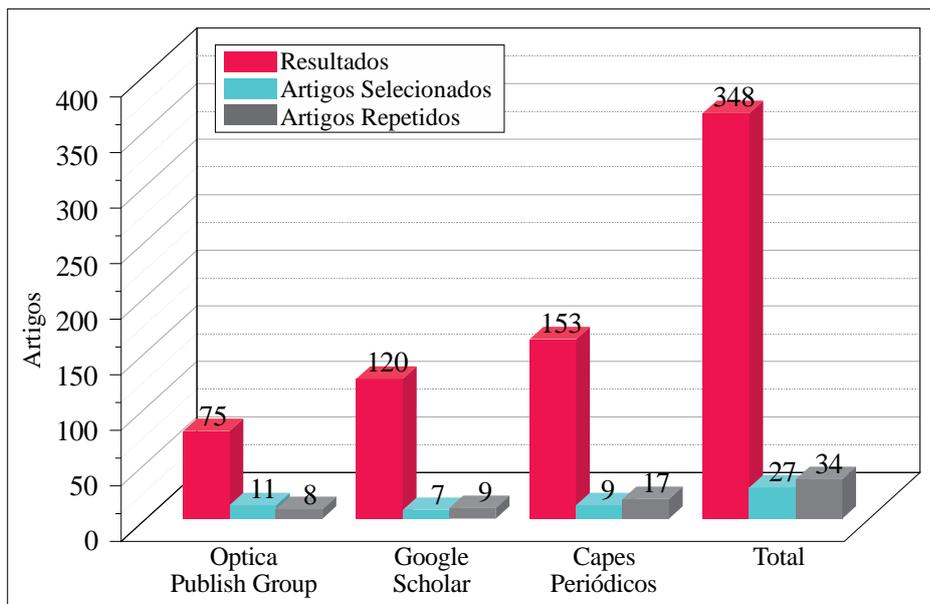


Figura 1.2 – Resultados das buscas nas bases de dados

A partir do estabelecimento dos termos de busca, foi definida a ordem das bases de dados e para iniciar as pesquisas optou-se escolher por afinidade ao tema. A primeira base de dados escolhida foi *Optica Publish Group* por conglomerar um maior número de periódicos em pesquisas relacionados a óptica. A ordem da segunda e terceira base de dados foram escolhidas de forma aleatória com o intuito de incluir outros periódicos com pesquisas na mesma área de conhecimento, assim ordenados como sendo *Google Scholar* e *Capes Periódicos* respectivamente. Nesta etapa, para os artigos encontrados nas buscas subsequentes idênticos

aos selecionados na busca anterior, foram considerados repetidos. Os periódicos encontrados e selecionados com as devidas publicações estão relacionados na tabela 1.2 bem como o fator de impacto.

Tabela 1.2 – Lista de periódicos, quantidade de artigos e fator de impacto

n	Periódicos	Fator de impacto	Qtd. de Artigo
1	<i>IEEE - Photonic journal</i> [29], [30]	2,4	2
2	<i>IEEE - Photonic e Technology</i> [31]	2,6	1
3	<i>IEEE - Journal of Selected Topics in Quantum Electronics</i> [32]	4,9	1
4	<i>IEEE - Journal of Lightwave Technology</i> [33]	4,7	1
5	<i>IEEE - Access</i> [34]	3,9	1
6	<i>FRANCIS & TAYLOR - Journal of Modern Optics</i> [35]	1,3	1
7	<i>OSA - Journal of the Optical Society of America B</i> [36]	1,9	1
8	<i>OSA - Optics Continuum</i> [37], [38]	1,6	2
9	<i>OSA - Optics Research</i> [39]	7,6	1
10	<i>OSA - Optics Xpress</i> [40], [41], [22], [42], [43], [44]	3,6	6
11	<i>OSA - Optics Letters</i> [45]	3,6	1
12	<i>SPRINGER - Optical and Quantum Electronics</i> [46], [47], [48]	3,0	3
13	<i>SPIE - Optical Engineering</i> [49]	1,3	1
14	<i>ELSEVIER - Optics & Laser Technology</i> [50]	3,8	1
15	<i>ELSEVIER - Optics Communications</i> [51], [52], [53]	2,4	3
16	<i>NATURE - Scientific Reports</i> [54]	4,6	1
Total			27

Fonte: Próprio autor

Figura 1.3 apresenta um diagrama sobre o procedimento para seleção dos artigos, composto pela fase de pesquisa, descartes e revisão. Ao final das buscas pode-se constatar que 348 artigos foram encontrados, sendo que 321 foram descartados sumariamente por não contemplar o objeto de estudo. Dessa forma, 27 publicações foram consideradas nesta revisão sistemática contemplando 16 periódicos no espaço temporal de 2006 a 2024.

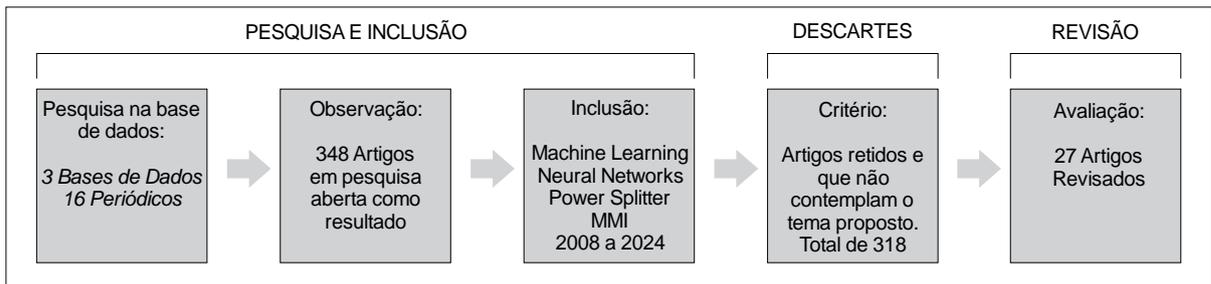


Figura 1.3 – Diagrama de seleção de artigos [55]

O processo de análise dos artigos teve por objetivo identificar quais mecanismos de inteligência computacional estão sendo empregados atualmente no desenvolvimento de dispositivos de Interferência Multimodo, bem como identificar os principais software de simulação mais utilizados, as aplicações dos respectivos dispositivos MMI, seu desempenho e

mecanismo de fabricação. A tabela 1.3 além de constar a análise descrita, apresenta os artigos de acordo com a quantidade de citações e ano de publicação, sendo que o artigo mais citado nos últimos 5 anos foi publicado em 2019 com 240 citações com referência [54] seguido do artigo com referência [33] publicado em 2021 com 79 citações.

Tabela 1.3 – Análise dos Artigos selecionados

Nº	[Referência]	Ano	Citação	Machine Learning	Simulador	Eficiência	Aplicação
1	[50]	2024	4	RNE	<i>Ansys Lumerical</i>	$3,6 \times 10^{-4}$	Beam Splitter
2	[47]	2024	0	RNC	<i>Ansys Lumerical</i>	0,45 dB	Power Splitter
3	[46]	2024	1	-	<i>COMSOL</i>	-	Sensor
4	[41]	2023	2	RNP	-	0,36dB	Multiplexer
5	[40]	2023	0	DBS	-	0,9dB	Power Splitter
6	[37]	2022	1	RNA	<i>Ansys Lumerical</i>	0,5dB	Power Splitter
7	[34]	2022	3	RNA	<i>Ansys Lumerical</i>	90%	Power Splitter
8	[39]	2021	42	RNP	-	0,36%	Power Splitter
9	[38]	2021	8	b-ARLA	<i>Ansys Lumerical</i>	0,92db	Beam Splitter
10	[35]	2021	0	-	-	0,45dB	Polarizador
11	[33]	2021	79	RNP	<i>Ansys Lumerical</i>	0,65dB	Power Splitter
12	[29]	2021	8	-	<i>Ansys Lumerical</i>	1dB	Multiplexer
13	[22]	2021	12	RNP	<i>Ansys Lumerical</i>	85%	Power Splitter
14	[49]	2020	5	-	<i>Ansys Lumerical</i>	1dB	Power Splitter
15	[31]	2020	45	-	-	1,3dB	Power Splitter
16	[30]	2020	1	-	<i>Ansys Lumerical</i>	0,967	Beam Splitter
17	[54]	2019	240	RNP	<i>Ansys Lumerical</i>	90%	Power Splitter
18	[36]	2019	19	RNA	<i>Ansys Lumerical</i>	87%	Power Splitter
19	[52]	2018	9	-	-	-	Coupler
20	[44]	2018	87	-	<i>Ansys Lumerical</i>	1,5dB	Power Splitter
21	[45]	2017	129	-	-	80%	Power Splitter
22	[43]	2013	456	-	<i>Ansys Lumerical</i>	0,46dB	Power Splitter
23	[42]	2013	305	-	<i>Ansys Lumerical</i>	0,6dB	Power Splitter
24	[32]	2011	131	-	<i>FIMProp</i>	1,5dB	Beam Splitter
25	[53]	2008	14	-	-	-	Switch
26	[51]	2008	18	-	-	-	Switch
27	[48]	2006	15	-	-	-	Switch

Fonte: Próprio autor

A figura 1.4 sintetiza através de gráfico os dados de análise da tabela 1.3 apresentado como estão distribuídos os artigos de acordo com o ano de publicação, aplicação, simulador e técnica de Machine Learning empregada. Nesta análise pode observar que 70% dos artigos coletados a aplicação destinada para os dispositivos de Interferência Multimodo é para divisores de potência (*Power Splitters e Beam Splitters*). Pode-se notar também que a maior quantidade publicados com o tema proposta data de 2021 e que as Redes Neurais é o método de inteligência computacional mais utilizado para realizar aprendizado para design inverso

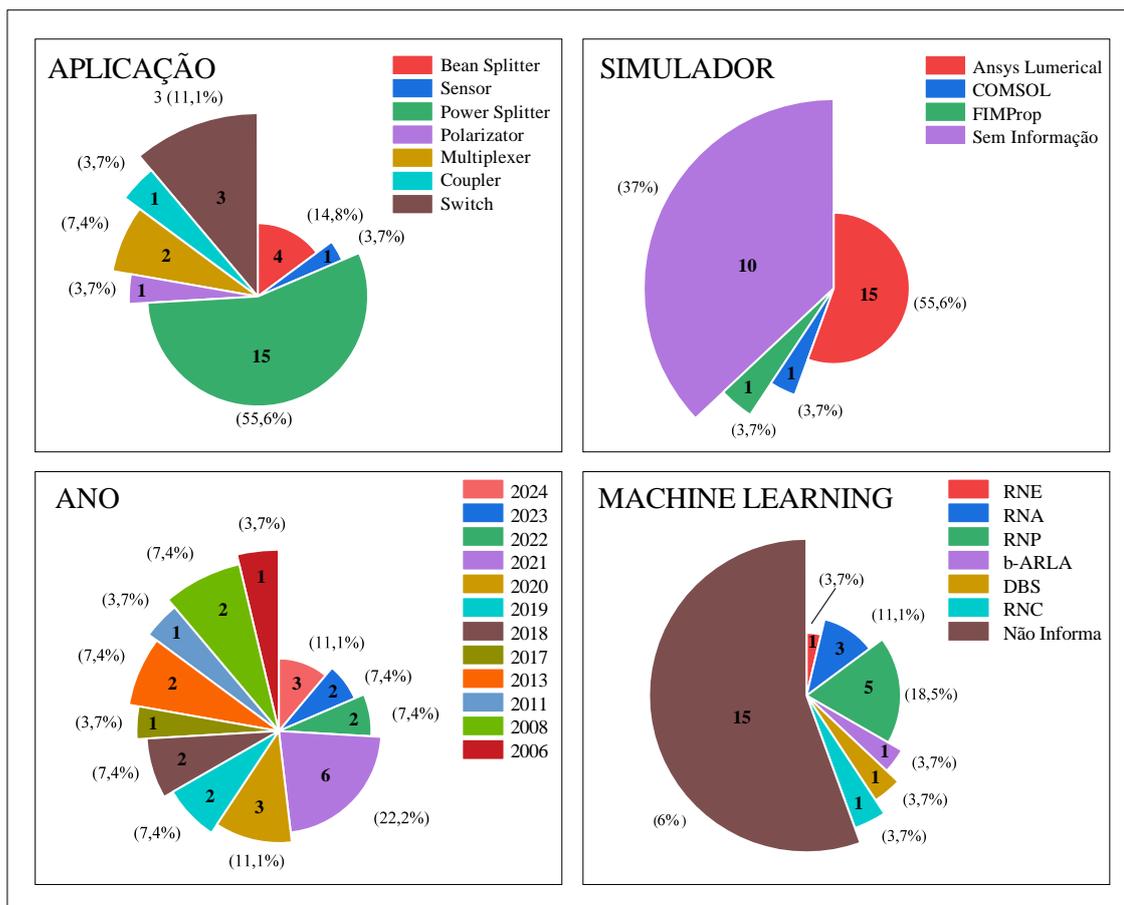


Figura 1.4 – Análise dos Artigos selecionados

Este levantamento bibliográfico demonstra as diversas aplicações de dispositivo de Interferência Multimodal e a importância de suas publicações no campo da fotônica. Foi possível observar que as técnicas de *Machine Learning* utilizadas para projetar esses dispositivos tem sido aperfeiçoada ao longo do tempo com técnicas mais robustas para aprendizado. Muito desse aperfeiçoamento deu-se à evolução dos *softwares* que apresentam excelentes performance para desenvolvimento de algoritmos inteligentes.

O uso da inteligência computacional associado aos *softwares* de simulações possibilitou realizar novas pesquisas introduzindo materiais com propriedades ópticas diferentes das convencionais. O Silício (Si) e a Sílica (SiO₂) se consolidaram como materiais que possuem as propriedades ópticas com melhores respostas para desempenho de dispositivo MMI e por oferecer boa tolerância de fabricação. Porém, foram identificadas publicações recentes que demonstram a eficiência de novos materiais com propriedades ópticas semelhantes e que oferecem uma resposta compatível ao Si e SiO₂, são os casos dos materiais poliméricos e do Nitreto de Silício (Si₃N₄).

1.4. Organização da Dissertação

Este trabalho está organizado em 7 capítulos. No primeiro capítulo contém a introdução, onde é transmitido o que foi estudado e feito no trabalho de investigação, bem como a razão da proposta deste trabalho, além de uma análise bibliográfica. Em seguida, nos capítulos dois, três e quatro, são apresentadas as fundamentações teóricas que embasam esta dissertação. No quinto capítulo é desenvolvida a metodologia adotada assim como a sua justificação. No sexto capítulo, contém a apresentação de resultados e discussão do que foi desenvolvido neste trabalho de dissertação. E por fim, no sétimo capítulo, conclui-se o projeto respondendo às questões de investigação e objetivos delineados, e apresentando as contribuições, limitações e investigação futura.

CAPITULO II

2. DISPOSITIVOS FOTÓNICOS

2.1. Breve histórico

No século XVII, o mais importante defensor de uma teoria das partículas da luz foi Sir Isaac Newton. Em seu tratado intitulado de *Opticks: or, A treatise of the reflections, refractions, inflexions and colours of light* [56], Newton considerava claramente os raios de luz como fluxos de partículas muito pequenas emitidas por uma fonte de luz e viajando em linha reta.

Já Christian Huygens, contemporâneo de Isaac Newton, defendeu uma visão em seu trabalho *Treatise on Light* [57] que considerava a luz como uma onda, espalhando-se a partir de uma fonte de luz em todas as direções e propagando-se através de um meio elástico onipresente chamado éter.

Em 1821, Augustin Fresnel publicou em *Annales de Chimie et de Physique* [58] os resultados de seus experimentos e análises, que estabeleciam que a luz fosse uma onda transversal, pois a dupla refração no cristal de calcita poderia ser entendida como um fenômeno que envolve luz polarizada, o que deu surgimento às equações de Fresnel, que fornece a amplitude da luz refletida e transmitida em uma interface plana que separa dois meios ópticos.

Entre os anos de 1861 e 1862, com trabalhos e publicações no campo da eletricidade e do magnetismo intitulados *A treatise on electricity and magnetism* [59], *A dynamical theory of the electromagnetic field* [60], e no artigo intitulado *On Physical Lines of Force* [61], James Clerk Maxwell sintetizou princípios conhecidos em um conjunto contendo quatro equações, conhecidas como Equações de Maxwell. Essas equações produziram uma previsão para a velocidade de uma onda eletromagnética, sendo a velocidade medida da luz, sugerindo seu caráter eletromagnético, assim, a luz passou a ser vista como uma região do espectro eletromagnético de radiação. Conceitualmente, as Equações de Maxwell são um grupo de equações diferenciais parciais que compõem a base do eletromagnetismo. O desenvolvimento das equações de Maxwell e o entendimento do eletromagnetismo contribuíram significativamente para toda uma revolução tecnológica iniciada no final do século XIX e continuada durante as décadas seguintes [62].

No início da década de 1960, a fotônica surge como resultado dos primeiros semicondutores emissores de luz, e então é definida como a ciência que envolve a geração de um fóton (luz), sua detecção, como manipulação via transmissão, emissão, processamento de sinal, modulação, comutação, amplificação e sensoriamento. Esse campo da ciência envolve o uso adequado da luz como uma das ferramentas mais importante para humanidade. O termo “fotônica” foi derivado da raiz da palavra “fóton”, que conota a menor entidade de luz análoga a um elétron na eletricidade, e foi usado pela primeira vez para designar um campo de pesquisa responsável pela utilização da luz para realizar tarefas convencionalmente relacionadas à esfera tradicional da eletrônica, como telecomunicações, processamento de informações e assim por diante. Assim como a eletrônica revolucionou o século XX, a fotônica tem revolucionado o século XXI [63], [64].

A fotônica é composta de muitas tecnologias diversas, incluindo fibras ópticas [65], lasers [66], detectors [67], eletrônica quântica [68] e materiais [18], [69]. Ela também se relaciona com a óptica quântica [70], eletro-óptica [71] e optoeletrônica [72]. Uma das maiores contribuições da fotônica para o mundo moderno é sua relação com a ciência da informação quântica ou computação quântica nos casos em que emprega métodos fotônicos [64].

As equações de Maxwell que governam a eletricidade e o magnetismo identifica a luz como uma onda eletromagnética com uma frequência específica para cada faixa espectral definida pela frequência da fonte da onda. Todas as ondas eletromagnéticas viajam no espaço livre com a mesma velocidade que é definida através das Equações de Maxwell. A distribuição de energia entre as várias ondas constituintes é chamada de espectro da radiação e implica uma dependência do comprimento de onda. Várias regiões do espectro eletromagnético são identificadas por nomes específicos, como: ondas de rádio, micro-ondas, raio x, raios gama, radiação ultravioleta e infravermelho, isso devido a diferenças na forma como são produzidas ou detectadas. A maioria das descrições comuns das várias faixas de frequência são fornecidas na Figura 2.1, na qual o espectro eletromagnético é exibido em termos de frequência e comprimento de onda [62].

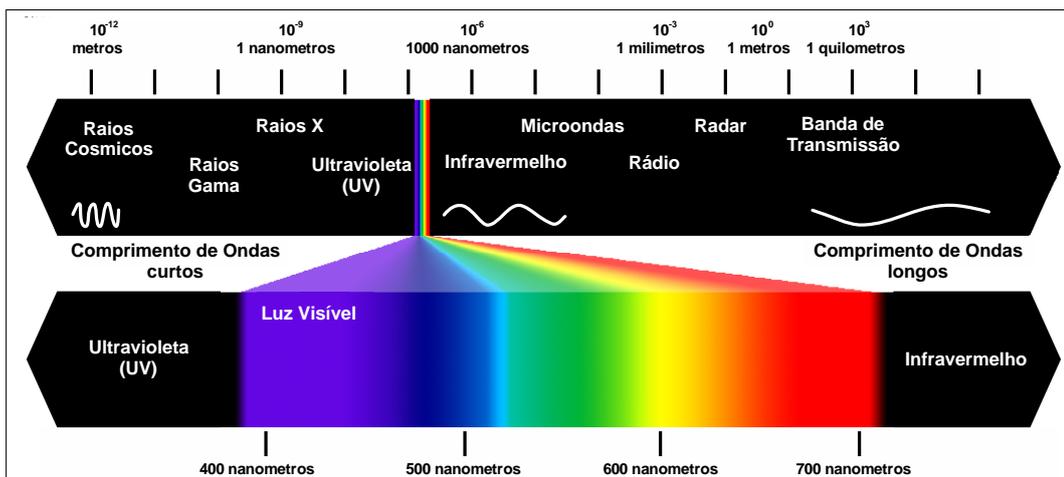


Figura 2.1 – Espectro eletromagnético [62]

2.2. Aplicações

A fotônica tem utilização em quase todos os aspectos e setores da humanidade, desde a vida cotidiana até a ciência altamente inovadora. Por exemplo, processamento de dados [73], telecomunicações, metrologia [74], iluminação, espectroscopia [75], holografia [76], medicina [77], armas bélicas [78], artes [79], agricultura [80], eficiência energética [81], processamento de material a laser [66], robótica [82] e biofotônica [83].

Economicamente, usos significativos de dispositivos fotônicos incluem telecomunicações de fibra óptica [65], armazenamento de dados ópticos, monitores, bombeamento óptico de lasers de alta potência e impressão a laser. As Aplicações na perspectiva da fotônica são praticamente ilimitadas e incluem diagnósticos médicos [84], síntese orgânica, informação, comunicação, bem como energia de fusão [64].

Pode-se verificar que as aplicações alcançadas e potencialmente viáveis, estão se tornando um grande estímulo para o interesse e desenvolvimento da fotônica. A presença cotidiana da fotônica está na internet, que não poderia funcionar sem comunicações de fibra óptica. A câmera, o smartphone e a crescente eficiência e versatilidade das tecnologias de iluminação dependem dos recentes avanços em fotônica. Muitos aspectos do diagnóstico médico, microscopia e terapêutica médica dependem da fotônica. A absorção óptica é o indicador das mudanças climáticas. Os processos de monitoramento que medem as mudanças em nossa atmosfera, bem como as soluções de longo prazo destinadas a conduzir nossas fontes de energia diretamente da luz, todos precisam de uma compreensão da fotônica e sua exploração [85].

2.3. Guia de Ondas

Guias de onda ópticos ou *waveguides* (WG) são estruturas dielétricas físicas capazes de transportar energia em comprimento de ondas do espectro eletromagnético [86]. São dispositivos que possibilitam a transmissão do sinal óptico confinado dentro do núcleo devido ao fenômeno da reflexão interna [87].

Fibras ópticas e guias de ondas ópticas são caracterizadas por uma região central, denominada núcleo, onde a luz fica confinada, composta por um índice de refração (n_1) que é superior ao da região circundante de índice normalmente uniforme (n_2), esta segunda região é chamada de revestimento [88].

A condição que estabelece o guiamento relacionando ao fenômeno de reflexão, está fundamentado na *Lei de Snell* apresentada na equação 2.1, onde θ_i e θ_r são os ângulos de incidência e refração de um raio que cruza a interface entre dois meios com índices de refração n_1 e n_2 [85].

$$n_1 \sin \theta_i = n_2 \sin \theta_r \quad (2.1)$$

Como pode ser observado na figura 2.2, mostra a geometria que pode ser considerada como um guia de ondas. A região central ou núcleo (*core*) é cercada por um revestimento chamado de *cladding*. Obviamente, o núcleo deve ter um índice de refração mais alto do que o revestimento para prender a luz dentro da estrutura após a reflexão interna total.

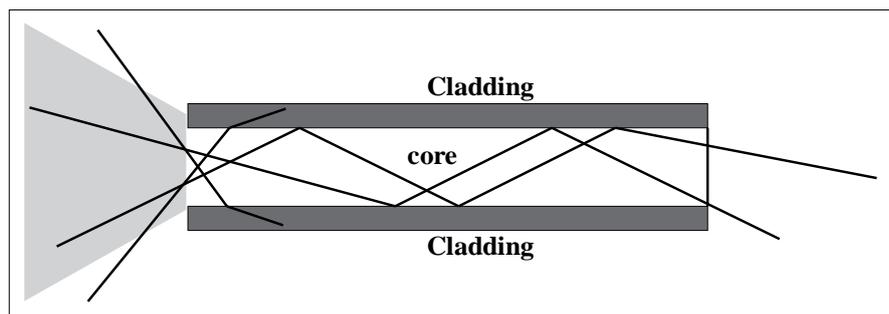


Figura 2.2 – Guia de Onda 2D [89]

A análise dos guias de ondas é feita utilizando as Equações de Maxwell escritas no sistema de coordenadas adequando a sua forma geométrica da seção transversal e impondo-lhe as condições de contorno. As Equações de Maxwell determinam os modos de propagação, respeitando as condições de contorno inerente ao guia. Os modos de propagação possuem um dado perfil de campo elétrico e de campo magnético e obedecem a certas condições de corte.

Para analisar a propagação de ondas em um meio isotrópico, é fundamental tomá-las como base [89].

$$\nabla \times E(r,t) = \frac{\partial}{\partial t} B(r,t) \quad (2.2)$$

$$\nabla \times H(r,t) = \frac{\partial}{\partial t} D(r,t) + J(r,t) \quad (2.3)$$

$$\nabla \cdot D(r,t) = \rho(r,t) \quad (2.4)$$

$$\nabla \cdot B(r,t) = 0. \quad (2.5)$$

Nas equações relacionadas acima, as grandezas de campo E e H representam, respectivamente, os campos elétrico e magnético, e D e B as densidades de fluxo elétricos e magnéticos. J e ρ representam densidade de corrente e densidade volumétrica de carga respectivamente. Além disso, o vetor de posição r define uma localização particular no espaço (x, y, z) na qual o campo está sendo medido.

A forma dada das Equações de Maxwell é bastante complicada por causa da inclusão da dependência do tempo e do espaço, porém é possível simplificar apenas trabalhando como campo harmônico no domínio do tempo, onde:

$$A(r,t) = A(r) \cos[\omega t + \phi(r)]. \quad (2.6)$$

Aplicando a função de Euler que relaciona a exponencial com funções seno e cosseno, a equação 2.6 pode ser reescrita da seguinte maneira:

$$A(r,t) = \text{Re} \left[A(r) e^{j\omega t} \right]. \quad (2.7)$$

Reescrevendo as equações de Maxwell na forma fasorial harmônica, fazendo as devidas substituições e manipulações, temos:

$$\nabla \times E(r) = j\omega B(r) \quad (2.8)$$

$$\nabla \times H(r) = j\omega D(r) + J(r) \quad (2.9)$$

$$\nabla \cdot D(r) = \rho(r) \quad (2.10)$$

$$\nabla \cdot B(r) = 0. \quad (2.11)$$

Através das relações constitutivas, representadas nas equações (2.12) e (2.13), pode-se obter as equações de ondas, pois elas eliminam $B(r)$ e $D(r)$. Essas relações caracterizam um determinado material em um nível macroscópico e são descritas em termos de duas grandezas escalares, a permeabilidade μ e a permissividade ε .

$$B = \mu H \quad (2.12)$$

$$D = \varepsilon E, \quad (2.13)$$

no espaço livre, a permeabilidade $\mu = \mu_0 = 4\pi \cdot 10^{-7} \text{ (H/m)}$ e a permissividade $\varepsilon = \varepsilon_0 = 8,85 \cdot 10^{-12} \text{ (F/m)}$. Fazendo as devidas substituições, é possível obter as equações de ondas para as duas primeiras equações de Maxwell na forma fasorial.

$$\nabla \times E(r) = j\omega\mu H(r) \quad (2.14)$$

$$\nabla \times H(r) = j\omega \varepsilon E(r). \quad (2.15)$$

Para estabelecer um significado físico para a equação é necessário definir o vetor de onda que é dado por k , onde:

$$k = \hat{x}k_x + \hat{y}k_y + \hat{z}k_z \quad (2.16)$$

$$|k| = \omega\sqrt{\mu\varepsilon}. \quad (2.17)$$

A partir destas equações é possível estabelecer as condições de guiamento para o modo TE de um campo elétrico polarizado na direção y , considerando um guia de ondas simétrico definida por:

$$E_y(x, z) = \begin{cases} A_1 e^{-\alpha_x x} & x > d/2 \\ A_2 \begin{bmatrix} \cos(k_{2x} x) \\ \text{sen}(k_{2x} x) \end{bmatrix} & |x| \leq d/2 \\ \pm A_3 e^{-\alpha_x x} & x < -d/2 \end{cases} \cdot e^{-jk_z z} \quad (2.18)$$

onde as opções superior e inferior para os campos correspondem a soluções pares e ímpares, respectivamente. As constantes d é a largura do núcleo do guia na direção do eixo x e k_{2x} e α_x são obtidas a partir das relações de dispersão no núcleo e regiões vizinhas e são dadas por:

$$k_{2x} = \sqrt{\omega^2 \mu \varepsilon - k_z^2} \quad (2.20)$$

$$\alpha_x = \sqrt{k_z^2 - \omega^2 \mu \epsilon_0}, \quad (2.21)$$

onde $\mu = \mu_0$ é a permeabilidade no espaço livre, A_1 e A_2 são os coeficientes de amplitude, k é o vetor de onda, ϵ é a permissividade e k_z é a constante de propagação.

A solução da equação que representa uma onda que se propaga com uma velocidade v (velocidade de fase) dada por [63]:

$$v = \frac{1}{\sqrt{\mu \epsilon}}. \quad (2.22)$$

A solução completa das equações vetoriais representa uma onda eletromagnética, onde cada uma das componentes cartesianas dos campos elétrico e magnético se propaga na forma de ondas de igual velocidade v . Para propagação no espaço livre, e usando os valores para ϵ_0 e μ_0 obtemos:

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \approx 3 \times 10^8 \text{ m/s}, \quad (2.23)$$

onde c corresponde à velocidade da luz no espaço livre medida experimentalmente. Vale a pena notar que aqui a velocidade da luz foi obtida apenas usando valores de constantes elétricas e magnéticas.

Normalmente é conveniente expressar a velocidade de propagação das ondas eletromagnéticas em um meio v em função da velocidade da luz no espaço livre c , através da relação:

$$v \equiv \frac{c}{n}, \quad (2.24)$$

onde n representa o índice de refração do meio dielétrico. O índice de refração está relacionado com a constante óptica do meio material e a permissividade dielétrica e a permeabilidade magnética do espaço livre por, dado desta forma:

$$n = \sqrt{\frac{\epsilon \mu}{\epsilon_0 \mu_0}}. \quad (2.25)$$

O índice de refração de um meio é o parâmetro mais importante na tecnologia fotônica integrada. Na maioria dos materiais e em particular em meios dielétricos, a permeabilidade magnética é muito próxima à do espaço livre: $\mu \approx \mu_0$. Com esta aproximação, o índice de refração pode ser simplificado para obter:

$$n \approx \sqrt{\frac{\epsilon}{\epsilon_0}} = \sqrt{\epsilon_r}, \quad (2.26)$$

onde a permissividade dielétrica relativa de magnitude ϵ_r , também frequentemente chamada de constante dielétrica, definida como a relação entre a permissividade dielétrica do meio material e a do espaço livre [63].

Neste capítulo, foi apresentado a fundamentação teórica contendo um breve histórico acerca de dispositivos fotônicos e equações importantes que sustentam a teoria eletromagnética necessária para embasar os tópicos discutidos ao longo desta dissertação. O principal objetivo foi estabelecer uma compreensão teórica robusta relacionada ao fenômeno do guiamento de ondas utilizadas para investigar e analisar o desenvolvimento deste projeto.

CAPITULO III

3. DISPOSITIVO DE INTERFERÊNCIA MULTIMODO (MMI)

O dispositivo de interferência multimodo é um guia de onda retangular, com uma estrutura em microescala onde um sinal óptico é confinado e guiado, de modo que a potência óptica seja dividida ou combinada de maneira previsível [90].

A arquitetura de um dispositivo MMI, como podem ser vistos nas figuras 3.1a e 3.1b, é composta por um conjunto de guias de ondas contendo portas de entradas e saídas em condições monomodo com largura $2a$, as quais são responsáveis pela entrada e saída do sinal de luz e que definem o formato $N \times M$, onde N e M são o número de portas de entradas e saídas respectivamente. As portas são acopladas na estrutura principal do dispositivo MMI, que por sua vez tem um formato retangular representados por W_{MMI} , L_{MMI} e $2d$, projetado para suportar vários modos, onde acontece o processo de interferometria [2]. Os índices de refração do núcleo n_1 e do revestimento n_2 são iguais para os guias de ondas monomodo e multimodo.

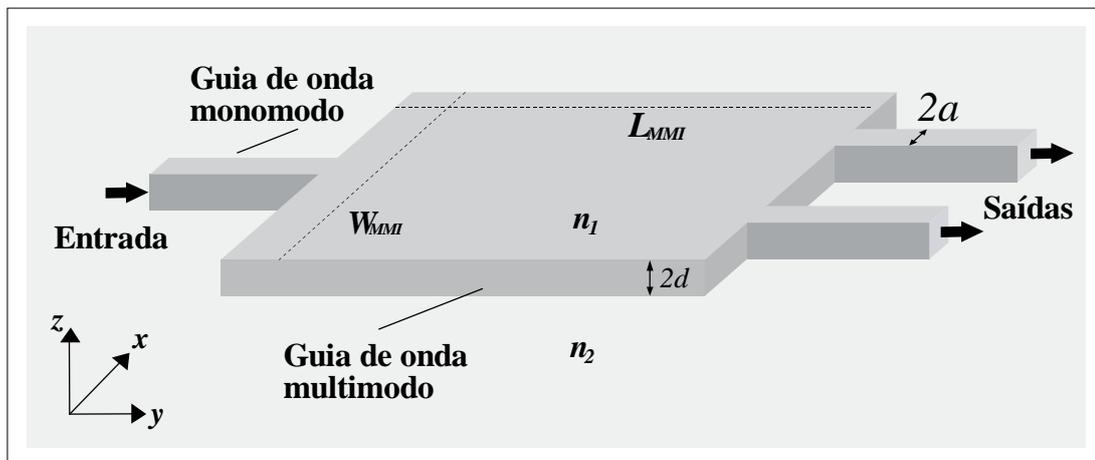


Figura 3.1a – Esquema 3D guia de ondas MMI 1×2

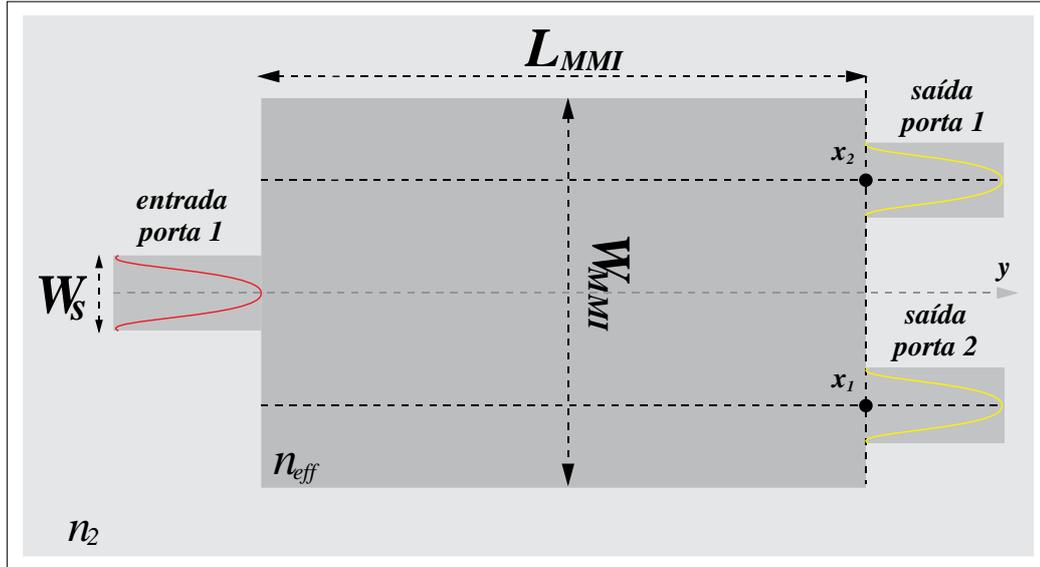


Figura 3.1b – Esquema 2D guia de ondas MMI 1×2

A estrutura do guia de ondas pode ser reduzida a um problema bidimensional como mostra a figura 1b, utilizando o método de índice efetivo, onde o n_1 no guia de onda 3d se torna n_{eff} no guia de onda 2d [91].

A operação de qualquer dispositivo MMI óptico é baseada no princípio de *self-image* de objetos periódicos, este princípio é descrito como uma propriedade de guias de onda multimodo pelo qual um perfil de campo de entrada é reproduzido em imagens únicas ou múltiplas em intervalos periódicos ao longo da direção de propagação da guia de ondas [2].

A análise de propagação através do fenômeno de auto imageamento dos modos de propagação fornece bases para modelagem numérica e informações sobre o processo de interferência multimodal.

O componente mais importante para obter os efeitos de MMI e da *self-image* é a partir de um guia de ondas de Interferência Multimodo projetado para suportar um grande número de modos. Quando um campo é lançado na região de interferometria, usando um guia de onda de monomodo como entrada, excitamos todos os modos suportados por este dispositivo. À medida que os modos se propagam na direção de guiamento, a interferência entre eles dá origem à formação das *selfies-images* do campo de entrada ao longo do guia de ondas. Como a formação das *selfies-images* ocorre em intervalos periódicos, o comprimento (L_{MMI}) do dispositivo MMI pode ser calculado de modo que pode obter as *selfies-images* no limite de corte e acoplar toda luz na saída do dispositivo onde a *selfies-images* são formadas.

3.1. Constantes de propagação

O guia de ondas multimodo é composto de índice efetivo da região de guiamento n_{eff} , de largura geométrica W_M e um revestimento com índice de refração efetivo n_{clad} conforme apresentado na figura 3.2.

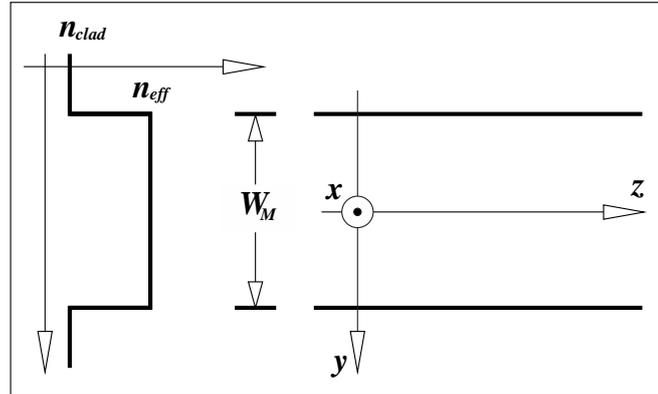


Figura 3.2 – Representação do perfil do índice de refração e vista superior do guia de ondas multimodo [1]

O guia de ondas suporta m modos laterais com números de modos $v = 0, 1, 2, \dots, (m - 1)$ em um comprimento de onda no espaço livre (λ_0) como pode ser visto na figura 3.3.

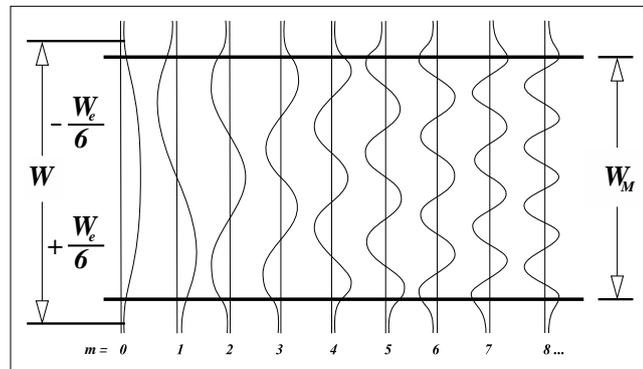


Figura 3.3 – Perfis de campo lateral com amplitude normalizada [2]

A constante de propagação β_v e o número de onda lateral k_{yv} estão relacionados ao índice da região de guiamento pela equação de dispersão:

$$k_{yv}^2 + \beta_v^2 = k_0^2 n_{eff}^2 \quad \text{onde } k_0 = \frac{2\pi}{\lambda_0} \quad \text{e } k_{yv} = \frac{(v+1)\pi}{W_{ev}}. \quad (3.1)$$

Assim, temos que a constante de propagação β_v pode ser expressa por,

$$\beta_v = \sqrt{k_0^2 n_{eff}^2 - k_y^2}, \quad (3.2)$$

dividindo os dois lados da equação (3.2) por $\sqrt{k_0^2 n_{eff}^2}$, temos que:

$$\beta_v = k_0 n_{eff} \sqrt{1 - \frac{k_y^2}{k_0^2 n_{eff}^2}} \equiv \beta_v = k_0 n_{eff} \sqrt{1 - \left(\frac{k_y}{k_0 n_{eff}}\right)^2}. \quad (3.3)$$

Usando a expansão de Taylor onde $(1 - x)^n \cong (1 - nx)$ para um x muito pequeno, considerando que $k_0 \gg k_y$ e fazendo $x = \left(\frac{k_y}{k_0 n_{eff}}\right)^2$ a constante de propagação β_v pode ser deduzida como,

$$\beta_v = k_0 n_{eff} \sqrt{(1 - x)^2}, \text{ logo } \beta_v = k_0 n_{eff} \sqrt{(1 - 2x)}, \quad (3.4)$$

expandindo $\sqrt{(1 - 2x)}$, temos os termos da expansão sendo:

$$1 - \frac{x^2}{2} - \dots, \quad (3.5)$$

substituindo a equação (3.5) expandida em (3.4) e retornando os valores de x , temos:

$$\beta_v = k_0 n_{eff} \left(1 - \frac{k_y^2}{2k_0^2 n_{eff}^2}\right) \equiv \beta_v = k_0 n_{eff} - \frac{k_y^2}{2k_0 n_{eff}}. \quad (3.6)$$

Substituindo os valores de k_0 e k_y da equação (3.1) em (3.6) definimos a constante de propagação β_v sendo:

$$\beta_v = k_0 n_{eff} - \frac{(v+1)^2 \pi \lambda}{4 n_{eff} W_{ev}^2}, \quad (3.7)$$

onde a largura efetiva W_{ev} leva em consideração a profundidade de penetração lateral dependente da polarização de cada modo do campo associado às mudanças de Goos-Hähnchen na interface do núcleo/revestimento. Em geral, as larguras efetivas W_{ev} podem ser aproximadas pela largura efetiva W_{e0} correspondente ao modo fundamental. A largura efetiva pode ser aproximada por,

$$W_{ev} \approx W_{e0} = W_M + \left[\frac{\lambda_0}{\pi} \right] \left[\frac{n_{clad}}{n_{eff}} \right]^{2\alpha} \frac{1}{\sqrt{n_{eff}^2 - n_{clad}^2}} \quad \alpha = \begin{cases} TM = 1 \\ TE = 0 \end{cases} \quad (3.8)$$

A diferença entre as constantes de propagação do modo fundamental (β_0) e o próximo modo (β_v) é dado por:

$$\beta_0 - \beta_v = \frac{3\pi\lambda}{4n_{eff}W_{ev}^2}. \quad (3.9)$$

O *beat length* na direção de propagação é dado por definição como:

$$L_\pi = \frac{\pi}{\beta_0 - \beta_v}, \quad (3.10)$$

assim, substituído a equação (3.9) em (3.10), L_π pode ser expressa por:

$$L_\pi = \frac{4n_{eff}W_{ev}^2}{3\lambda}. \quad (3.11)$$

Dessa forma podemos representar $\beta_0 - \beta_v$ como sendo:

$$\beta_0 - \beta_v = \frac{v(v+2)\pi\lambda_0}{4n_{eff}W_{ev}^2}, \quad (3.12)$$

no entanto, substituindo a equação (3.11) em (3.12), podemos reescrever $\beta_0 - \beta_v$ por:

$$\beta_0 - \beta_1 = \frac{\pi v(v+2)}{3L_\pi}. \quad (3.13)$$

constituindo o fator de fase para a formação das *selfies-images*.

A distribuição do campo de entrada, em $z = 0$, quando todos os modos são considerados e que o perfil do campo $\Psi(y,0)$ está contido totalmente dentro da largura efetiva W_{ev} é dado por

$$\Psi_v(y, 0) = \sum_{v=0}^{m-1} c_v \psi_v(y). \quad (3.14)$$

Onde c_v é o coeficiente do modo de excitação do campo e $\psi_v(y)$ é a distribuição do campo modal de todos os modos. Fazendo $z=L_{MMI}$, o perfil do campo será uma superposição das distribuições de campo de todos os modos

$$\Psi_v(y, z) = \sum_{v=0}^{m-1} c_v \psi_v(y) e^{j(\beta_0 - \beta v)z}. \quad (3.15)$$

Assim temos

$$\Psi_v(y, L_{MMI}) = \sum_{v=0}^{m-1} c_v \psi_v(y) e^{j\left(\frac{\pi v(v+2)}{3L\pi} L_{MMI}\right)}. \quad (3.16)$$

A forma de $\Psi(y, L_{MMI})$, e conseqüentemente os tipos de imagens formadas, será determinada pela excitação modal c_m , e as propriedades do fator de fase do modo dada por $e^{j\left(\frac{\pi v(v+2)}{3L\pi} L_{MMI}\right)}$. Assim, o campo $\Psi(y, L_{MMI})$ será uma *self-image* do campo de entrada $\Psi(y, 0)$.

Dos fenômenos de interferência existente em dispositivos MMI para análise baseada na posição lateral, onde o guia de onda de entrada é posicionado, este trabalho utiliza o fenômeno de interferência simétrica restrita, que ocorre quando a luz é acoplada exatamente no centro do eixo- x do dispositivo. Neste caso, apenas os modos simétricos $v = 0, 2, 4, \dots$ são excitados e a *self-image* é encontrada uma distância de $3L\pi/4$, que é quatro vezes menor que a interferência geral. N imagem dobradas é encontrada em um comprimento de $3L\pi/4N$. Levando em consideração a natureza periódica do efeito de autoimagem ao longo do dispositivo MMI, a distância para as autoimagens individuais é dando por:

$$L = p \left(\frac{3L\pi}{4} \right) \quad \text{para} \quad p = 0, 1, 2, 3 \dots \quad (3.17)$$

Neste trabalho foram considerados como materiais para composição do dispositivo a Sílica (SiO_2) e o Silício (Si) para o comprimento de onda $\lambda=1,55\mu\text{m}$. Os índices de refração para esta condição de guiamento foram definidos através da equação de *Sellmeier* apresentada nas equações (3.18) e (3.19), onde $n_{\text{Silício}}$ e $n_{\text{Sílica}}$ refere-se aos índices de refração do Silício e da Sílica respectivamente [92], [93].

$$n_{\text{Silício}} = \sqrt{\frac{0,6961663\lambda^2}{\lambda^2 - 0,0684043^2} + \frac{0,4079426\lambda^2}{\lambda^2 - 0,1162414^2} + \frac{0,8974794\lambda^2}{\lambda^2 - 9,896161^2} + 1} \quad (3.18)$$

$$n_{\text{Síllica}} = \sqrt{\frac{10,6684293\lambda^2}{\lambda^2 - 0,301516485^2} + \frac{0,0030434748\lambda^2}{\lambda^2 - 1,13475115^2} + \frac{1,54133408\lambda^2}{\lambda^2 - 1104^2} + 1} \quad (3.19)$$

Devido a condição de não-linearidade existente entre os parâmetros de entrada para obter uma resposta na saída do dispositivo, faz-se necessário utilizar mecanismos computacionais capazes de fornecer uma resposta rápida e eficiente, assim as Redes Neurais Artificiais (RNA) se revelam como uma proposta adequada para tal solução usando técnicas de *Machine Learning*.

Neste capítulo, foi apresentado a fundamentação teórica contendo elementos importantes sobre dispositivos de Interferência Multimodo e as equações de ondas que o modelam, bem como suas propriedades ópticas, necessária para embasar os tópicos discutidos nos próximos capítulos desta dissertação. O objetivo foi estabelecer uma compreensão teórica e prática relacionada ao funcionamento de um dispositivo MMI e as condições de guiamento, objeto principal deste trabalho.

CAPITULO IV

4. INTELIGENCIA ARTIFICIAL

Graças ao crescimento exponencial do poder computacional a Inteligência Artificial apresentou desenvolvimento significativo e notável nos últimos anos, se tornando uma ferramenta quase onipresente na vida cotidiana [94].

Não é fácil definir o conceito de IA com precisão, no entanto existem quatro linhas de pensamentos que juntos contemplam uma definição bastante abrangente, conceituando a Inteligência Artificial como: sistemas que pensam como seres humanos, aonde máquinas possuem mentes no sentido total e literal, na visão de Haugeland [95]; sistemas que atuam como seres humanos, que realizam funções executadas por pessoas no entendimento de Kurzweil [96]; sistemas que pensam racionalmente, possuidora de faculdades mentais pelo uso de modelo computacionais como afirmam Charniak e McDermott [97]; e sistemas que atuam racionalmente, compreendido por Poole como agentes inteligentes [98]. A IA, portanto pode ser concebida como uma simulação da capacidade de pensamento abstrato, criativo e dedutivo, e particularmente da capacidade de aprender, usando a lógica computacional.

A Inteligência Artificial é intitulada como uma das mais novas áreas de conhecimento na atualidade no campo das engenharias e da tecnologia da informação. O desenvolvimento desta área de estudo começou logo após a Segunda Guerra Mundial em 1956, com o artigo intitulado *Computing Machinery and Intelligence* do matemático inglês Alan Turing [99]. Posteriormente, a máquina foi utilizada para implementação no aprendizado e para realização de tarefas simples com jogar xadrez, demonstrações de teoremas matemáticos, direção de um carro e diagnóstico de doenças [100].

Atualmente, devido a abrangência interdisciplinar que esse área de conhecimento proporciona, outras áreas como, a psicologia, a biologia, a lógica matemática, a linguística, a engenharia, a filosofia, arte, música, agricultura, política, economia, esportes, entre outras tem usufruído dos benefícios desse recurso tecnológico, fazendo da Inteligência Artificial foco de muitos estudos e pesquisas [101].

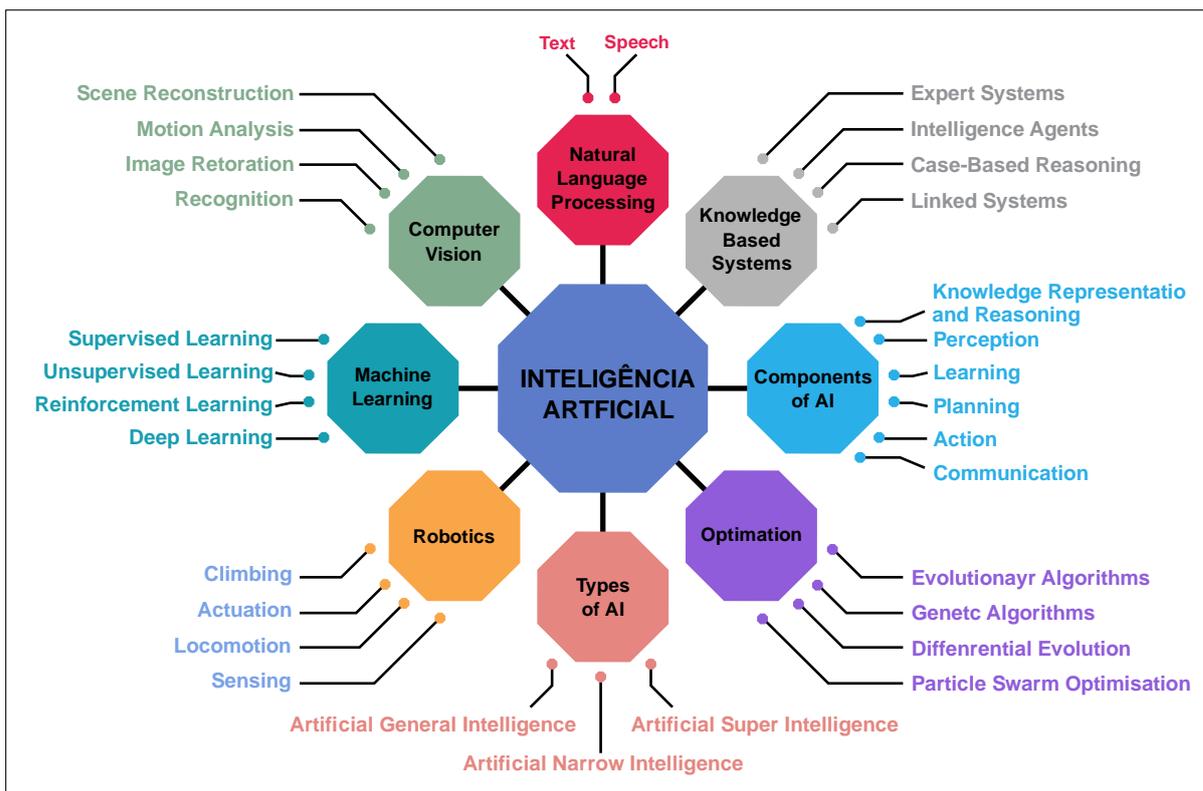


Figura 4.1 – Subáreas da Inteligência Artificial [102]

As principais subáreas da IA e suas ramificações são mais conhecidas por suas nomenclaturas em língua inglesa, que são: aprendizado de máquina (*Machine Learning*), sistemas baseados em conhecimento (*knowledge-based systems*), visão computacional (*computer vision*), robótica (*robotics*), processamento de linguagem natural (*natural language processing*), planejamento e programação automatizados (*automated planning and scheduling*) e otimização (*optimization*), como pode ser visto na figura 4.1.

A inteligência artificial tem visto uma evolução acelerada nos últimos anos com o surgimento do *Machine Learning*. Esta ferramenta tecnológica tornou-se cada vez mais popular devido à sua capacidade de analisar grandes conjuntos de dados, fazer previsões e fornecer uma nova compreensão que antes eram impossíveis de obter. À medida que a quantidade de dados gerados e o poder computacional aumentam, o potencial do *Machine Learning* tem se tornando cada vez mais aparente e necessário [103].

4.1. Machine Learning

A *Machine Learning* ou aprendizado de máquinas, tem concedido grandes possibilidades aos seres humanos nas diversas áreas do conhecimento, se revelando uma ferramenta poderosa capaz de executar tarefas de forma automatizada, tornando a vida do homem mais confortável devido sua condição contínua de aperfeiçoamento.

Conceitualmente, *Machine Learning* é um ramo de Inteligência Artificial cujo objetivo é o desenvolvimento de técnicas computacionais sobre o aprendizado bem como a construção de sistemas capazes de adquirir conhecimento automaticamente, que utiliza técnicas estatísticas para desenvolver tarefas através de algoritmos. São esses algoritmos que ensinam o computador a encontrar padrões de comportamento olhando através de milhares de exemplos, então usa essa experiência adquirida para resolver problemas em novas situações. [104].

A *ML* permite que um modelo computacional composto por várias camadas de unidades de processamento aprenda vários níveis de abstração a partir de determinados dados. As vantagens exclusivas dessa técnica estão em sua metodologia orientada a dados, que permite que o modelo descubra informações úteis automaticamente a partir de uma grande quantidade de dados [105].

O sistema de aprendizado é realizado computacionalmente, baseado na tomada de decisões e de experiências adquiridas a partir de soluções bem sucedidas de problemas. Apensar da *ML* ser uma ferramenta poderosa para realizar aprendizado automático, não existe um único algoritmo que apresente o melhor desempenho para todos os problemas. No entanto, é de fundamental importância compreender a limitação dos inúmeros algoritmos de aprendizado de máquinas para cada problema que serão aplicados [106].

4.1.1. Métodos de Machine Learning

O *Machine Learning* é classificado de acordo como um algoritmo aprende a fazer previsões em níveis de acurácia, pois a escolha do método depende do tipo de dados que se desejam prever. A literatura divide esse processo de aprendizagem em quatro metodologias fundamentais: Aprendizagem Supervisionada (*Supervised Learning*), Aprendizagem Não-Supervisionada (*Unsupervised Learning*), Aprendizagem Semi-supervisionada (*Semi-*

supervised Learning) e Aprendizagem Por Reforço (*Reinforcement Learning*) [107], [108], como ilustrado na figura 4.2.

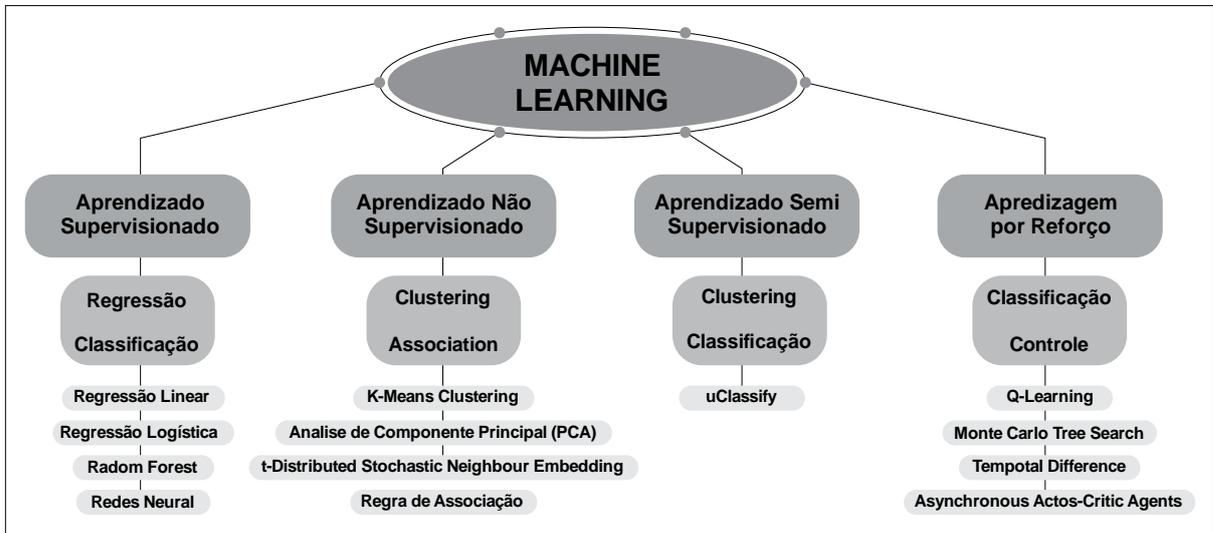


Figura 4.2 – Classificação do *Machine Learning* [108]

No aprendizado supervisionado é fornecido ao algoritmo, um conjunto de exemplos de treinamento para os quais o rótulo da classe associada é conhecido. Para esse algoritmo, o objetivo é construir um classificador capaz de identificar a classe de cada novo exemplo não rotulados. As tarefas supervisionadas mais comuns são identificados como classificação que separa os dados e regressão que ajusta os dados [109]. Este método é caracterizado pelo conceito da supervisão, cuja principal tarefa é fornecer uma medida precisa do seu erro. Um conjunto de treinamento composto por pares de entrada-saída é fornecido ao algoritmo e a partir desta informação, o algoritmo pode corrigir seus parâmetros através de iterações até que a diferença entre os valores previstos e esperados torna-se próxima de zero [110].

O aprendizado não supervisionado analisa conjuntos de dados não rotulados sem a necessidade de interferência humana, ou seja, é um processo baseado em dados. Este modelo é amplamente utilizado para extrair características generativas, identificar tendências e estruturas significativas, agrupamentos em resultados e fins exploratórios. As tarefas de aprendizagem não supervisionadas mais comuns são agrupamento (*Clustering*), estimativa de densidade, aprendizagem de recursos, redução de dimensionalidade, localização de regras de associação, detecção de anomalias, entre outros [106], [111].

O método de aprendizagem por reforço visa utilizar observações colhidas na interação com o ambiente para realizar ações que maximizem a recompensa ou minimizem o risco

[112]. Mesmo que não existam supervisores reais, a aprendizagem por reforço se baseia no *feedback*, geralmente chamado de recompensa, fornecido pelo ambiente. Porém, neste caso, a informação é mais qualitativa e não auxilia o agente na determinação de uma medida precisa do seu erro. Este método, é particularmente eficiente quando o ambiente não é completamente determinístico [110]. É uma ferramenta poderosa para treinar modelos de IA que pode ajudar a aumentar a automação ou otimizar a eficiência operacional de sistemas sofisticados, como robótica, tarefas de direção autônoma, logística de fabricação e cadeia de suprimentos, no entanto, não é preferível usá-la para resolver problemas básicos ou diretos [111].

Existem muitos problemas em que a quantidade de amostras rotuladas é muito pequena em comparação com o número potencial de elementos. Uma abordagem supervisionada não é suficiente porque os dados usados para treinar o modelo não podem ser representativos de toda a distribuição, portanto é necessário encontrar um *trade-off* entre uma estratégia supervisionada e não supervisionada. O Aprendizado Semi Supervisionado consiste em solucionar esses problemas [110]. Neste modelo, são fornecidos dados classificados e não classificados, ou seja, rotulados e não rotulados, com o intuito de gerar um modelo apropriado para a classificação dos dados. Na maioria dos casos, os dados rotulados são escassos e os dados não rotulados são abundantes. O objetivo da classificação Semi Supervisionada é aprender um modelo que irá prever classes de dados de teste futuros melhor do que o modelo gerado usando apenas os dados rotulados, sendo utilizada comumente em tradução automática, detecção de fraudes, rotulagem de dados e classificação de texto [112].

4.1.2. Técnicas e Algoritmos de Machine Learning

No *Machine Learning* os algoritmos mais utilizados para o método de aprendizado supervisionado incluem técnicas para análise de classificação e regressão. A Figura 4.3 mostra como uma representação gráfica da classificação é diferente do modelo de regressão.

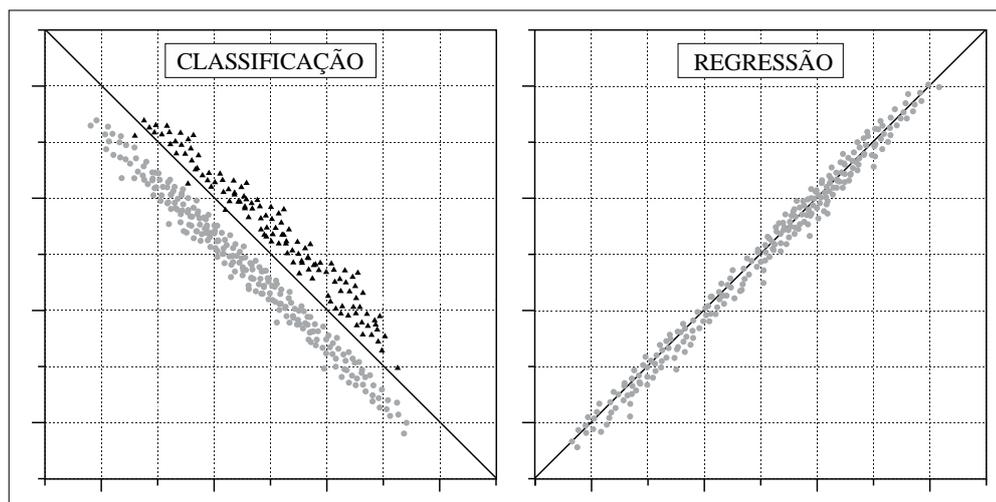


Figura 4.3 – Representação gráfica de Regressão e Classificação

A classificação é considerada um método de aprendizagem supervisionada em *Machine Learning*, referindo-se a um problema de modelagem preditiva, onde um rótulo de classe é previsto para um determinado exemplo. Na análise de classificação mapeia uma função f de variáveis de entrada x para variáveis de saída y como alvo através de um algoritmo capaz separar com precisão os dados de teste em categorias específicas. Ele reconhece entidades específicas dentro do conjunto de dados e tenta desenhar algumas conclusões sobre como essas entidades devem ser rotuladas ou definidas. Algoritmos de classificação mais comuns são: *Naive Bayes-NB* [113], *k-nearest neighbours-kNN* (k-vizinhos mais próximos) [114], *Support Vector Machine - SVM* (Máquina de vetores de suporte) [115], *Random Forest - RF* (Floresta aleatória) [116], *Linear Discriminant Analysis - LDA* (Análise Discriminante Linear) [117], *Decision Tree - DT* (Árvore de decisão) [118], [119], *Adaptive Boosting - AdaBoost* [120], [121].

A regressão inclui vários métodos de aprendizado de máquina que permitem prever uma variável de resultado contínua y com base no valor de uma ou mais variáveis preditoras x . A distinção mais significativa entre classificação e regressão é que a classificação prevê rótulos de classes distintas, enquanto a regressão facilita a projeção, modelagem, previsão, predição, etc. Algoritmos de classificação mais comuns são: *Linear Regression* [122], *Logistic Regression - LR* (Regressão Logística)[123] e *Polynomial Regression - PR* (regressão polinomial) [122].

4.2. Redes Neurais Artificiais

A ideia de construir uma máquina ou mecanismo autônomo, que seja dotado de inteligência, se constitui um sonho antigo dos pesquisadores das áreas de ciências e engenharias. Embora o primeiro trabalho em Redes Neurais Artificiais tenha sido publicado há mais de 50 anos, esta temática começou a ser fortemente pesquisada a partir dos anos 90 com grande potencial. As aplicações que envolvem a utilização de sistemas considerados inteligentes são das mais variadas possíveis, por exemplo: Avaliação de imagens captadas por satélites, classificação de padrões de fala e escrita, reconhecimento de faces, previsões de ações no mercado financeiro, identificações de perfis de doenças, controle de aparelhos eletrônicos, dentre outros [124].

As RNA's estão cada vez mais presente na vida do homem. As redes mais comuns que são acessadas diariamente são redes inteligentes utilizadas nos sistemas de operação de celulares e smartphones, entre elas as mais conhecidas são a assistente da *Google* presente no sistema *Android* e a assistente *Siri* da *IOS* presente nos dispositivos da *Apple*. Também são encontrados assistentes virtuais fundamentados em Redes Neurais nos *Call Centers* e sistemas bancários [19].

As redes neurais biológicas, formam-se a partir de um conjunto de neurônios, que possuem como função principal coletar, processar e dispersar os impulsos elétricos. O processamento das informações do cérebro surge principalmente das redes e conexões entre os neurônios [100]. A origem das Redes Neurais Artificiais está fundamentada na estrutura de Neurônios Biológicos, representado na figura 4.4, os quais serviram como espelho para a criação das RNA's.

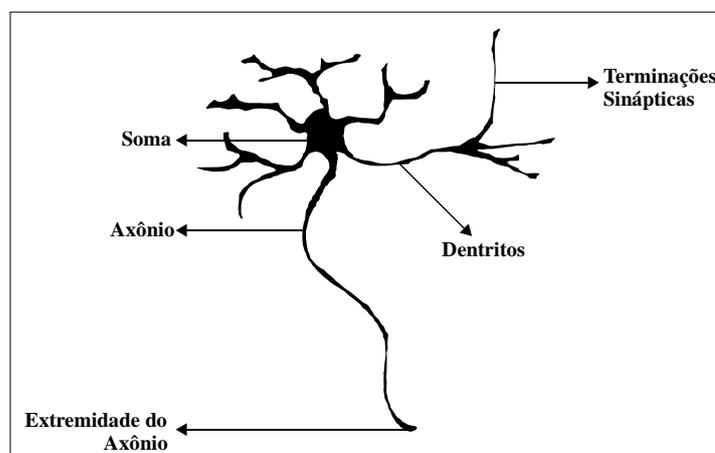


Figura 4.4 – neurônio biológico [124]

Redes Neurais Artificiais são sistemas computacionais adaptativos baseados em modelos matemáticos, as quais possuem sistemas paralelos distribuídos formados por unidades de processamento simples (nós) que calculam determinadas funções matemáticas, normalmente não-lineares [27]. Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões, geralmente unidirecionais, buscando simular o funcionamento das redes neurais biológicas, inspirado em uma estrutura física concebida por natureza: o cérebro humano, devido suas características de processamento de informação e nas particularidades das interconexões [125], [126].

O processamento da informação em RNA é feito nos neurônios artificiais, conhecidos como neurônio McCulloch e Pitts (ou modelo MCP) [127]. Ela é unidade fundamental de uma rede neural artificial, que recebe valores de entrada x e saída y , como representado na figura 4.5. O neurônio realiza duas operações matemáticas importantes, um somatório ponderado e um mapeamento não-linear.

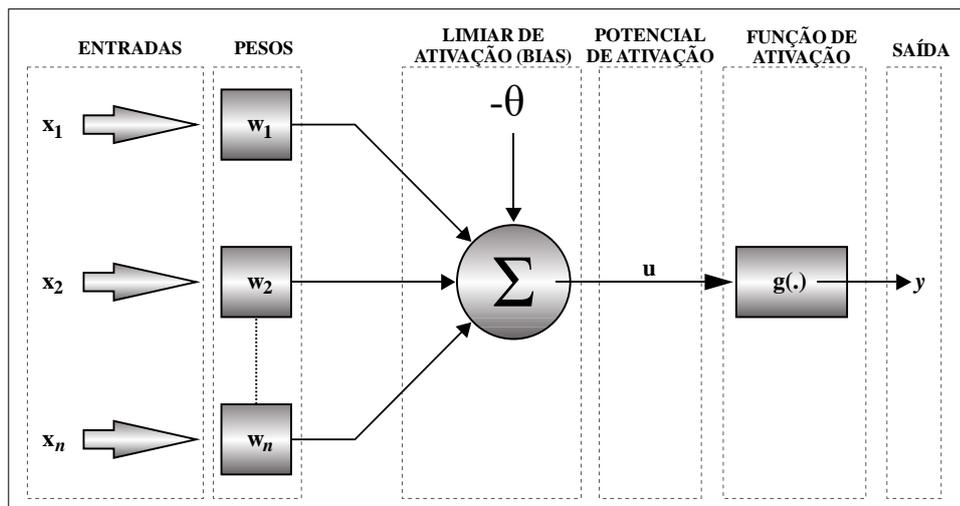


Figura 4.5 – Neurônio artificial [124]

As Redes Neurais são modelos computacionais inspirados no sistema nervoso de seres vivos. Possuem capacidade de aquisição e manutenção do conhecimento, baseado em informações, e podem ser definidas como um conjunto de unidades de processamento, caracterizadas por neurônios artificiais, que são interligados por um grande número de interconexões (sinapses artificiais), sendo representadas por vetores/matrizes de pesos sinápticos [124].

As equações (4.1) e (4.2) descrevem os resultados produzidos por uma Rede Neural Artificial, proposto por McCulloch e Pitts [127]:

$$u = \sum_{i=1}^n w_i x p - \theta \quad (4.1)$$

$$y = g(u) \quad (4.2)$$

As equações propostas estabelecem o modelo de funcionamento de um neurônio artificial. De maneira ordenada elas podem ser melhor compreendidas da seguinte maneira.

- i. Apresentação de um conjunto de valores que representam as variáveis de entrada;
- ii. Multiplicação de cada entrada de saída pelo seu respectivo peso;
- iii. Obtenção do potencial de ativação produzido pela soma ponderada dos sinais de entrada, subtraindo-se pelo limiar de ativação ou bias;
- iv. Aplicação de uma função apropriada, tendo-se como objetivo limitar a saída do neurônio;
- v. Compilação da saída a partir da aplicação da função de ativação neural em relação ao seu potencial de ativação.

As RNA's se dividem em duas categorias: as supervisionadas e as não-supervisionadas. Nos modelos supervisionados, dados históricos são utilizados para treinamento. Nos modelos não-supervisionados, nenhum treinamento ocorre, no entanto, o software em execução pesquisa e identifica padrões existentes em um determinado aglomerado de dados podendo lhe retornar resultado aleatórios [125].

4.2.1. Treinamento Supervisionado

O treinamento é a base de funcionamento de uma RNA, no qual é inserido na rede uma série de padrões e as suas respectivas respostas. Desta forma, a rede possui a capacidade de aprender a partir de exemplos e fazer interpolações e extrapolações do que aprenderam.

O Treinamento supervisionado de uma RNA consiste basicamente em apresentar pares entrada-saída à rede, calcular as saídas em função das entradas, calcular o erro entre a saída desejada e a saída calculada e alterar os valores sinápticos por algum tipo de algoritmo.

À medida que os pares são inseridas e os pesos atualizados, o erro médio quadrático (EQM) da saída da rede tende a diminuir. Dentre os algoritmos utilizados para executar esse processo, o mais utilizado e eficiente é o *Backpropagation*.

O processo de treinamento da RNA por este algoritmo calcula a retropropagação de erros das saídas em direção às entradas, através de um mecanismo de correção de erros e ajustes de pesos (*wieghts*) das camadas intermediárias da rede por meio de treinamento. O erro calculado na saída de um neurônio é multiplicado pela derivada da função de ativação daquele neurônio e propagado para a sua entrada. Este valor é enviado para todos os neurônios da camada anterior ponderado pelo peso da camada. O treinamento pode ser por época, onde todo conjunto de treinamento utilizado é apresentado à rede e só após isso, as sinapses são atualizadas[128].

4.2.2. *Backpropagation e Rede Feedforward*

O algoritmo *backpropagation* foi complementado em 1986, por Rumelhart, Hinton e Willian e consiste em um procedimento de aprendizagem de redes *Feedforward*. Esse procedimento faz uso de vetores capazes de mapear um conjunto de entradas para um conjunto específico de saídas. *Backpropagation* é o algoritmo utilizado no aprendizado supervisionado, através de um mecanismo de correção de erros e ajustes de pesos (*wieghts*) das camadas intermediárias na rede por meio de treinamento. O treinamento acontece em duas fases, *forward* e *backward*. A fase *forward* é utilizada para definir a saída da rede para um determinado padrão de entrada. A fase *backward* aproveita a saída desejada e a saída fornecida pela rede para atualizar os pesos de suas conexões [129].

4.2.3. *Multilayer Perceptron (MLP)*

A rede *Multilayer Perceptron (MLP)*, que também pode ser chamada de *Perceptron* Multicamadas, é considerada uma alternativa atraente para uma fórmula empírica, pois imita a relação não-linear entre as variáveis de entrada e saída de uma forma mais simplificada. O modelo visa obter os pesos otimizados da rede usando um algoritmo de treinamento projetado para minimizar o erro entre as variáveis de saída e de destino, modificando os pesos mutuamente conectados [128].

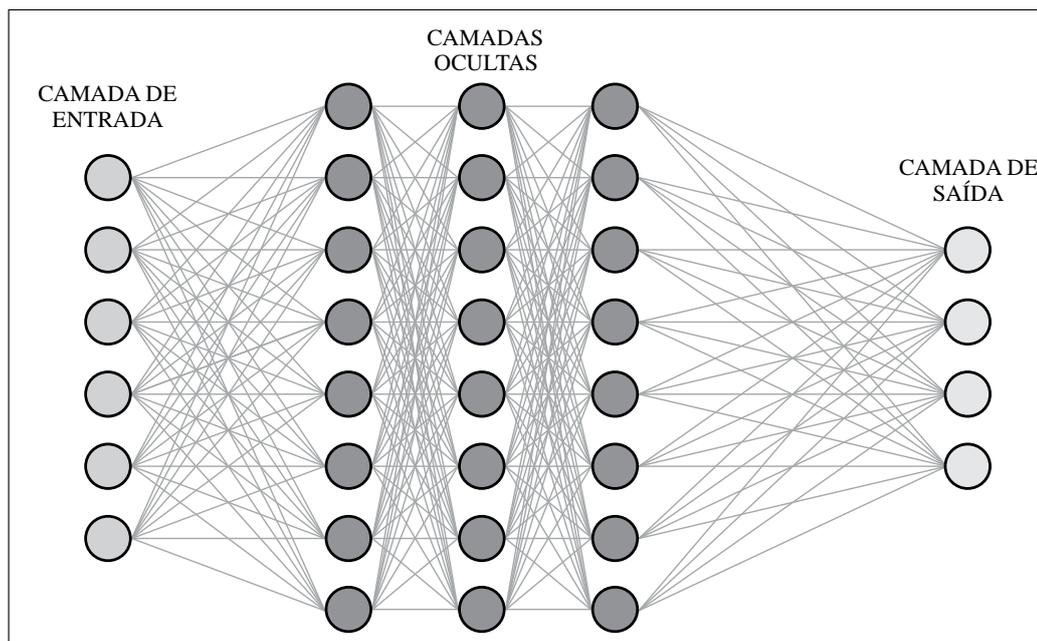


Figura 4.6 – Rede Neural Artificial *Perceptron* Multicamadas [130]

Uma rede neural aprende por meio de ajustes repetidos desses pesos. O traço característico dessas redes é que elas consideram o conhecimento acumulado adquirido durante o treinamento e respondem a novos eventos da maneira mais adequada, proporcionando a experiência adquirida como solução durante o processo de treinamento. A figura 4.6 representa um modelo típico de RNA *Perceptron* Multicamadas [124].

A *MLP* é uma rede neural em avanço com uma camada de entrada de neurônios fonte, pelo menos uma camada intermediária ou oculta de neurônios e uma camada de saída. A camada de entrada recebe sinais de entrada e redistribui esses sinais para todos os neurônios na camada oculta. Os pesos dos neurônios nas camadas ocultas representam os recursos nos padrões de entrada. A camada de saída estabelece o padrão de saída de toda a rede.

Para treinar a rede, um número adequado de exemplos representativos do fenômeno relevante deve ser selecionado de modo que a rede possa aprender as características fundamentais do problema. Mais de uma centena de diferentes algoritmos de aprendizado estão disponíveis, porém o mais popular é a propagação reversa. Em uma rede neural de propagação reversa, o algoritmo de aprendizado possui duas fases. Primeiro, um conjunto de dados de entrada de treinamento é apresentado à camada de entrada da rede. Em seguida, a rede propaga o conjunto de dados da camada de entrada para camada de saída até que o conjunto de dados seja gerado pela camada de saída. Se esse conjunto de dados for diferente da saída desejada,

um erro será calculado e, em seguida, propagado de volta pela rede a partir da camada de saída até a camada de entrada. Os pesos são modificados conforme o erro é propagado [124].

4.2.4. *Levenberg-Marquardt*

O algoritmo de aprendizado *Levenberg-Marquardt (LM)* foi desenvolvido para resolver iterativamente problemas de minimização de funções não-lineares pelo método de mínimos quadrados [131]. É o algoritmo de otimização mais utilizado hoje em dia, superando outros algoritmos para solução de uma grande variedade de problemas. De maneira geral, este algoritmo propõe uma solução o qual é utilizado para melhorar a performance de treinamento do *MLP*. O treinamento da rede pelo algoritmo *LM* ocorre por iterações, as quais consistem em aplicar todos os dados de entrada de treinamento da rede, checar o erro entre a saída obtida e a saída desejada para cada conjunto e ajustar os pesos das conexões pra reduzir o erro médio [128].

Neste capítulo, foram apresentados as bases teóricas sobre inteligência computacional com um breve histórico de cada passo importante para essa ciência, fundamentadas em referencias pioneiras com aplicações atuais, suficiente para embasar o desenvolvimento desta dissertação de mestrado. O principal objetivo foi estabelecer uma compreensão teórica consistente, relacionada a *Machine Learning* como subárea da IA, seus principais algoritmos e técnicas de aprendizados que serão utilizados no desenvolvimento desta pesquisa para o cumprimento estabelecido nos objetivos deste trabalho.

CAPITULO V

5. METODOLOGIA

A primeira etapa desta dissertação consiste em realizar uma revisão bibliográfica especializada nas bases de dados mais importantes para o campo da fotônica. Esta etapa de análise bibliográfica é sustentação para estado da arte e serve como base norteadora para o desenvolvimento desta pesquisa. O objetivo desta revisão inicial é identificar, destacar e catalogar os principais artigos cujo a temática abarca a utilização de técnicas de *Machine Learning* para designer de dispositivos de Interferência Multimodo.

Na segunda etapa de desenvolvimento deste trabalho, a metodologia empregada propõe realizar simulações de dispositivos MMI com geometrias compactas com o objetivo de gerar dados de treinamentos para alimentar a Rede Neural Artificial. Para tal, é necessário definir parâmetros geométricos e ópticos, tais como: intervalo da largura do dispositivo (W_{MMI}) padronizado com eixo- x , quantidade de portas de entradas e saídas ($N \times M$), faixa de comprimento de ondas (λ) e potência óptica nas portas de saídas M ($1 \times M$) do dispositivo. O intervalo de comprimento do dispositivo (L_{MMI}), padronizado com eixo- y , é obtido pelo fenômeno de interferometria em função dos parâmetros anteriores onde é obtido o limite de corte.

Para geração dos dados de treinamento utilizou o software *BeamPROP* desenvolvido pela *RSoft Photonics CAD Suit, Version 5.1.0.2, copyright© 1993-2003*. *BeamPROP™* é a ferramenta de projeto para área de fotônica, baseada no Método de Propagação de Feixe (BPM) para projeto e simulação de dispositivos e circuitos de guia de onda integrados e de fibra óptica [132].

A terceira etapa desta pesquisa tem como proposta metodológica desenvolver, treinar, testar e validar diversas arquiteturas de Redes Neurais Artificiais. Para este processo optou-se por utilizar o *MATLAB Mathworks R2016a, copyright© 1994-2024 The MathWorks, Inc*. De acordo com o desenvolvedor, a *MATLAB (MATrix LABoratory)* é o ambiente de computação produtivo, interativo e de alta performance para engenheiros e cientistas. Em sua plataforma é possível contar com ferramentas de aplicação para Redes Neurais Artificiais com linguagem de programação específica. O objetivo deste processo é o desenvolvimento de um algoritmo capaz de prever as coordenadas para o eixo- x das portas de saídas M ($1 \times M$) de maior potência e as

coordenadas para eixo-y responsável pelo limite de corte para acoplamento das portas de saída para um dispositivo MMI compacto.

Na quarta etapa, a metodologia empregada consiste em identificar, classificar por ordem de desempenho e selecionar as melhores RNA's, desenvolvidas na etapa anterior. O próximo passo é utilizar os dados das coordenadas (x, y) preditos pelas melhores Redes Neurais Artificiais para realizar novas simulações no software utilizado na segunda etapa. Nesta etapa propõe-se fazer uma comparação entre dados geométrico e ópticos gerados da simulação na segunda etapa e os dados gerados a partir das simulações preditas pelas RNA's desta fase, dos seguintes parâmetros: coordenadas para o eixo- x das portas de saídas M ($1 \times M$), coordenadas para eixo- y e potência óptica nas portas de saídas M ($1 \times M$) do dispositivo.

A metodologia utilizada para quinta e última etapa propõe analisar os resultados obtidos, fundamentada tanto nos resultados práticos realizado neste trabalho quanto nos resultados extraídos da literatura na primeira etapa. A análise empregada visa atender os seguintes aspectos avaliativos: performance das RNA's (EQM e custo computacional), performance do dispositivo projetado (desempenho óptico) e comparação dos resultados apresentados pelos dispositivos frutos desta pesquisa com os dispositivos catalogados na revisão bibliográfica.

O computador utilizado neste trabalho de dissertação dispõe de um Processador *12th Gen Intel® Core™ i7-1255U, 1,70 GHz e RAM 16,0 Gb*.

CAPITULO VI

6. RESULTADOS E DISCUSSÕES

Os dispositivos MMI a ser projetado nesta etapa da dissertação compreenderá o fenômeno de *self-image* para o mecanismo de Interferência Simétrica. As características analíticas de acordo com equações do capítulo 3 podem ser utilizadas para modelar dispositivos MMI com $(1 \times M)$ portas de saída. Porém, neste trabalho utilizaremos o simulador desenvolvido pela Synopsys® através do *software RSoft-BeamPROP 2D*, que utiliza o Método de Propagação de Feixe (*Beam Propagation Method - BPM*) para simular e projetar dispositivos ópticos.

O dispositivo é composto por um núcleo de Silício (Si), um substrato de Dióxido de Silício (SiO₂) com uma espessura de 0,220µm. Concebemos um divisor de potência (1×2) baseado em MMI com índices de refração do núcleo (3,476) e do revestimento (1,444) no comprimento de onda de 1,550µm, apresentando um índice de refração efetivo (n_{eff}) de 2,848.

As configurações dos parâmetros globais para modelar um guia de ondas no *software RSoft-BeamPROP* podem ser conferidos na tabela 6.1

Tabela 6.1 – Parâmetros de configurações globais para *RSoft-BeamPROP*

Parâmetros	Valor
Índice de Refração do Núcleo	3,476
Índice de Refração do Revestimento	1,444
Diferença do índice	2,032
Índice de Refração Efetivo	2,848
Comprimento de ondas no espaço livre (µm)	1,550
Espessura do Guia de ondas (µm)	0,220
Polarização	TE

Fonte: Próprio autor

Os parâmetros de configuração globais apresentados na tabela 6.1 foram utilizados para simular os dispositivos para geração de dados de treinamento. Nesta fase, foram simulados 301 dispositivo MMI com características de (1×2) portas de entrada e saídas. A largura W_{MMI} do dispositivo está compreendidos entre 2,0 a 5,0µm sendo simulados a cada intervalo de 0,01µm. Após cada simulação foi possível catalogar parâmetros ópticos e geométricos como: potência óptica na saída do guia para acoplamento das portas, expressa em porcentagem (%), coordenada do eixo-y para limite de corte da *self-image* com maior intensidade expresso em micrometros

(μm) e coordenadas no eixo- x onde as portas de saídas serão colocadas para acoplamento do sinal.

As simulações foram responsáveis por gerar 2.408 dados independentes distribuídos entre: comprimento de onda (μm), largura W_{MMI} (μm), comprimento L_{MMI} (μm), índice de refração do núcleo ($n_{\text{Silício}}$), índice de refração do substrato (n_{Silica}), coordenada para o eixo- x_1 (μm), coordenada para o eixo- x_2 (μm) e potência na saída do dispositivo (%).

Tabela 6.2 – Parâmetros Ópticos e Geométricos de quatro dispositivos MMI's (1×2) simulados analiticamente

PARÂMETROS ÓPTICOS E GEOMÉTRICOS	Dispositivo 6.1a	Dispositivo 6.1b	Dispositivo 6.1c	Dispositivo 6.1d
Comprimento de Onda (λ)	1,550 μm	1,550 μm	1,550 μm	1,550 μm
Índice do Substrato (n_{Silica})	1,4440	1,4440	1,4440	1,4440
Índice do Núcleo ($n_{\text{Silício}}$)	3,4777	3,4777	3,4777	3,4777
Largura (W_{MMI})	2,00 μm	3,00 μm	4,00 μm	5,00 μm
Coordenada eixo- y (L_{MMI})	5,33 μm	11,35 μm	19,60 μm	30,03 μm
Coordenada eixo- x_1	-0,5420 μm	-0,8000 μm	-1,0526 μm	-1,3020 μm
Coordenada eixo- x_2	0,5420 μm	0,8000 μm	1,0526 μm	1,3020 μm
Eficiência	91,46%	88,90%	86,95%	85,52%

Fonte: Próprio autor

Dispositivos MMI's (1×2) foram simulados com campos formado pela *self-image*, como mostra a figura 6.1, utilizando as configurações da tabela 6.2. A figura 6.1a representa o dispositivo mais compacto dentre os simulados com Largura (W_{MMI}) de 2,00 μm e comprimento (L_{MMI}) de 5,33 μm , enquanto a figura 6.1d representa o dispositivo com maior dimensão medindo 5,00 μm de Largura (W_{MMI}) e com comprimento (L_{MMI}) de 30,03 μm . A potência do sinal de entrada foi dividida e acoplada em suas portas de saídas (x_1 e x_2) e totalizaram a transferência da potência equivalente a 91,46 e 85,52%, contabilizando uma perda menor que 0,38 e 0,67dB respectivamente.

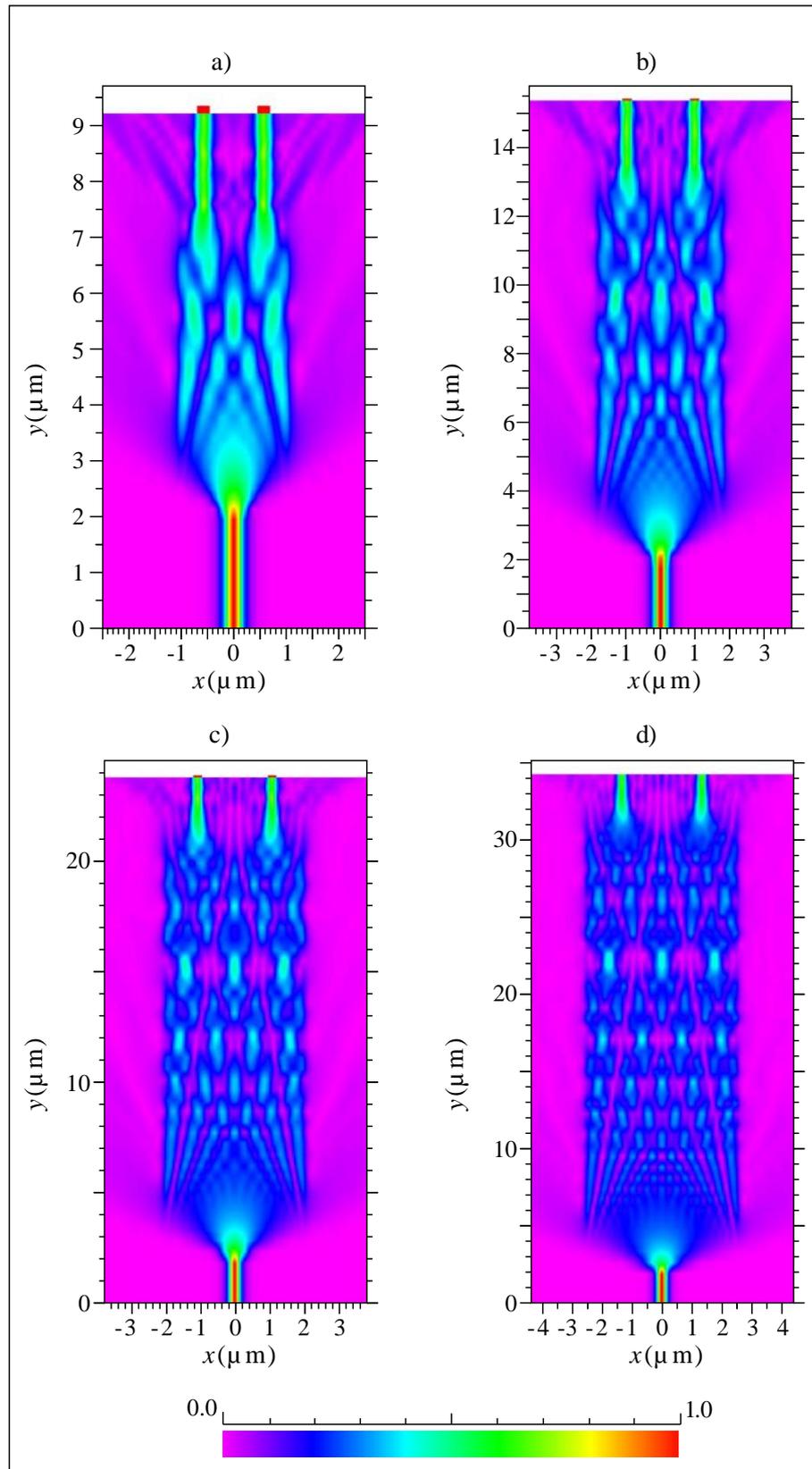


Figura 6.1 – Campo formado pela *Self-image* das simulações com dados analíticos de dispositivos MMI (1×2) para o comprimento de ondas de $1,55 \mu\text{m}$. a) $2,00 \times 5,33 \mu\text{m}$; b) $3,00 \times 11,35 \mu\text{m}$; c) $4,00 \times 19,60 \mu\text{m}$; d) $5,00 \times 30,03 \mu\text{m}$

A figura 6.2 apresenta os dispositivos simulados com as mesmas configurações da tabela 6.2, porém com a proporção das dimensões conservadas para fins de visualização.

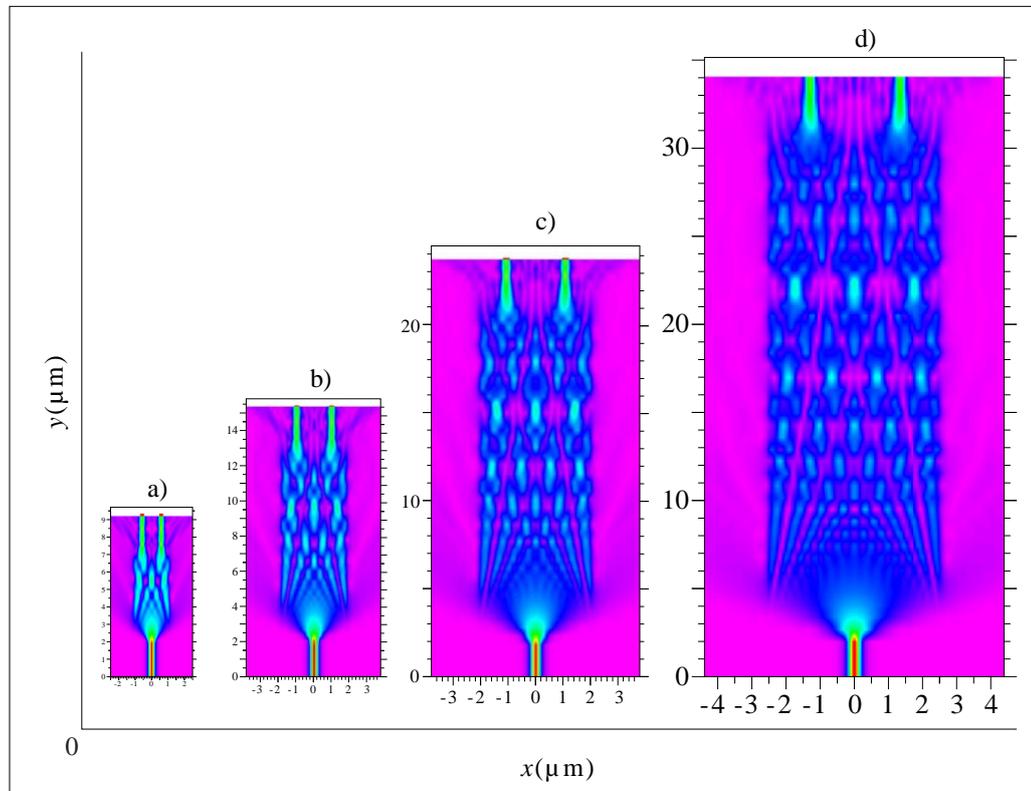


Figura 6.2 – Campo formado pela *Self-image* das simulações com dados analíticos de dispositivos MMI (1×2) para o comprimento de ondas de $1,55\mu\text{m}$, com a proporção das dimensões conservadas. a) $2,00 \times 5,33\mu\text{m}$; b) $3,00 \times 11,35\mu\text{m}$; c) $4,00 \times 19,60\mu\text{m}$; d) $5,00 \times 30,03\mu\text{m}$

A figura 6.3 mostra a eficiência da transferência de potência óptica de um dispositivo dentre os simulados, com largura (W_{MMI}) de $2,20\mu\text{m}$, sendo que tanto os índices de refração do núcleo e do substrato quanto o comprimento (L_{MMI}) dependem do comprimento de ondas para qual serão projetados. A eficiência foi calculada para a faixa de comprimento compreendida entre $1,25$ e $1,7\mu\text{m}$, abrangendo as bandas de comunicação O, E, S, C, L e U, com perdas que variam entre $0,31$ e $0,41\text{dB}$. A partir da leitura do gráfico, pode-se concluir as duas bandas de comunicação que oferecem melhores respostas aos dispositivos MMI são as bandas O e C.

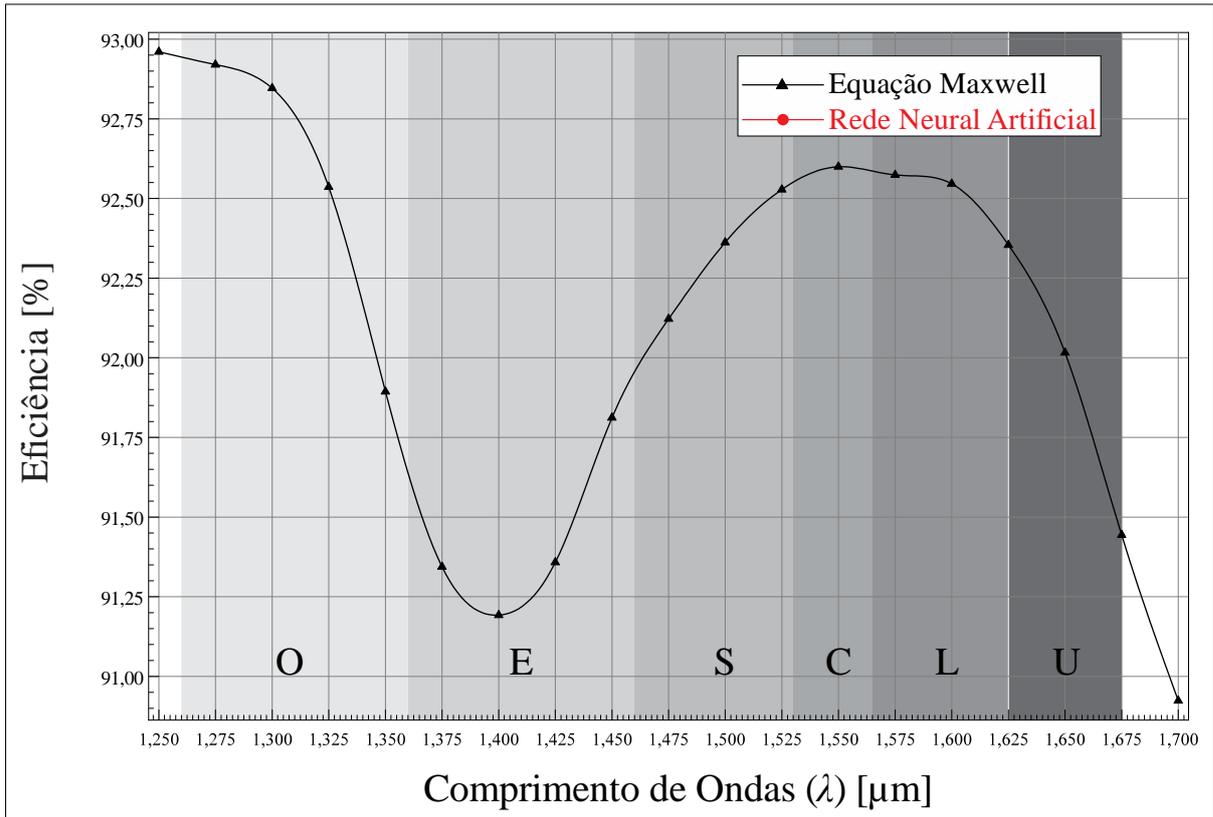


Figura 6.3 – Eficiência do dispositivo MMI (1×2) com L_{MMI} de $2,20\mu\text{m}$ simulado analiticamente, para a faixa de comprimento de ondas de $1,25 \sim 1,7\mu\text{m}$ integrando as bandas O, E, S, C, L e U

A construção de dispositivos fotônicos com estruturas complexas ainda é um processo difícil. O *Machine Learning* serve como uma solução eficiente para o complicado processo de design, pois através das Redes Neurais Artificiais, além de oferecer uma nova e poderosa abordagem para projetos fotônicos, permite realizar o projeto inverso com alta eficiência e precisão.

Nesta etapa foi desenvolvido um algoritmo baseado em técnicas de *Machine Learning* capaz de prever os parâmetros geométricos de dispositivos de Interferência Multimodo compacto. Para o processo de implementação das RNA's optou-se por utilizar o *MATLAB Mathworks R2016a*, copyright© 1994-2024 The MathWorks, Inc.

O objetivo desta fase foi o desenvolvimento de um algoritmo capaz de prever as coordenadas para o eixo- x_m das portas de saídas M ($1 \times M$) de maior potência e a coordenada para eixo-y responsável pelo limite de corte para acoplamento das respectivas portas.

As arquiteturas da RNA's foram configuradas com 3 camadas ocultas, alternadas entre si com as funções de ativação *Tansig* (tangente hiperbólica), *Logsig* (logística) e a *Poslin* (*ReLU - Rectified Linear Unit*), e na camada de saída a função *Purelin* (linear), como pode ser observado

na figura 6.4. Cada RNA foi treinada com [10, 8 e 5] neurônios nas respectivas camadas e na segunda etapa de treinamento inverteu-se a ordem de quantidade de neurônios da primeira e da última camada [5, 8 e 10], totalizando 52 redes neurais.

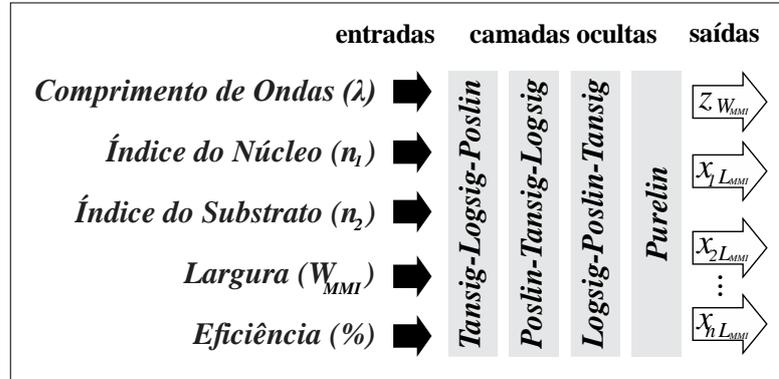


Figura 6.4 – Arquitetura da Rede Neural Artificial

A RNA foi configurada com 5 parâmetros de entradas sendo eles: comprimento de ondas (λ), índice de refração do núcleo (n_1), índice de refração do substrato (n_2), largura (W_{MMI}) e eficiência (%). Como o objetivo é desenvolver um algoritmo que estabeleça uma relação entre os parâmetros de entrada com as características geométricas e ópticas, os alvos da rede estabelecidos foram: a posição dos guias de ondas nas portas de saídas referente ao eixo-x compreendido no intervalo da largura (W_{MMI}) e o comprimento de corte no eixo-y (L_{MMI}), assim fornecendo as coordenadas (x_m, y) suficiente para projetar um dispositivo compacto que apresente a maior eficiência do sinal nas portas de saídas.

As 52 RNA's treinadas pelo algoritmo apresentaram excelente desempenho. Nesta fase foram oferecidos 301 pares de dados (entradas-saídas), treinadas pelo método da validação cruzada através da técnica *k-fold* para $k=10$. Os dados de entrada foram divididos em 10 blocos de dados k , aonde cada etapa de treinamento, o algoritmo usa todos os blocos de dados para treinamentos, menos em um ($k-1$), em seguida, avalia o modelo com o bloco de dados que não foi usado para treinamento.

Tabela 6.3 –Arquitetura e Desempenho das Redes Neurais Artificiais baseadas na média do Erro Quadrático Médio da validação cruzada, regressão e tempo de processamento

Arquitetura das camadas ocultas (Função de Ativação)	Quantidade de Neurônios por camada oculta	EQM Validação Cruzada	EAM Validação Cruzada	EAPM Erro Absoluto Percentual Médio	Regressão	Tempo (s)
RNA 01- Tan-Tan-Tan	[10 8 5]	$6,39410 \times 10^{-5}$	$4,89013 \times 10^{-3}$	0,49%	0,999997	7,46
RNA 02- Tan-Tan-Tan	[5 8 10]	$7,02583 \times 10^{-5}$	$4,92453 \times 10^{-3}$	0,49%	0,999997	14,40
RNA 03- Log-Log-Log	[10 8 5]	$1,08222 \times 10^{-4}$	$6,59319 \times 10^{-3}$	0,66%	0,999996	6,29
RNA 04- Log-Log-Log	[5 8 10]	$1,41045 \times 10^{-4}$	$7,79456 \times 10^{-3}$	0,78%	0,999994	9,33
RNA 05- Pos-Pos-Pos	[10 8 5]	$2,64873 \times 10^{-4}$	$1,01749 \times 10^{-2}$	1,02%	0,999994	4,10
RNA 06- Pos-Pos-Pos	[5 8 10]	$4,47628 \times 10^{-4}$	$1,09581 \times 10^{-2}$	1,10%	0,999991	3,87
RNA 07- Tan-Log-Pos	[10 8 5]	$1,63198 \times 10^{-4}$	$8,31821 \times 10^{-3}$	0,83%	0,999993	4,65
RNA 08- Tan-Log-Pos	[5 8 10]	$1,63198 \times 10^{-4}$	$8,31821 \times 10^{-3}$	0,83%	0,999990	4,38
RNA 09- Pos-Log-Tan	[10 8 5]	$1,48882 \times 10^{-4}$	$8,38973 \times 10^{-3}$	0,84%	0,999994	5,79
RNA 10- Pos-Log-Tan	[5 8 10]	$1,44911 \times 10^{-4}$	$7,73398 \times 10^{-3}$	0,77%	0,999994	7,45
RNA 11- Tan-Pos-Tan	[10 8 5]	$3,36524 \times 10^{-1}$	$1,21872 \times 10^{-1}$	12,19%	0,996455	5,71
RNA 12- Tan-Pos-Tan	[5 8 10]	$1,31423 \times 10^{-4}$	$7,34432 \times 10^{-3}$	0,73%	0,999995	4,83
RNA 13- Log-Pos-Log	[10 8 5]	$6,46915 \times 10^{-4}$	$1,43221 \times 10^{-2}$	1,43%	0,999974	5,46
RNA 14- Log-Pos-Log	[5 8 10]	$1,41158 \times 10^{-4}$	$7,76258 \times 10^{-3}$	0,78%	0,999995	5,45
RNA 15- Pos-Log-Pos	[10 8 5]	$2,76814 \times 10^{-4}$	$1,05615 \times 10^{-2}$	1,06%	0,999989	4,38
RNA 16- Pos-Log-Pos	[5 8 10]	$1,45731 \times 10^{-4}$	$7,91920 \times 10^{-3}$	0,79%	0,999994	4,88
RNA 17- Pos-Tan-Pos	[10 8 5]	$1,35359 \times 10^{-4}$	$7,86647 \times 10^{-3}$	0,79%	0,999995	4,20
RNA 18- Pos-Tan-Pos	[5 8 10]	$1,42448 \times 10^{-3}$	$1,42286 \times 10^{-2}$	1,42%	0,999973	4,05
RNA 19- Pos-Pos-Tan	[10 8 5]	$2,60067 \times 10^1$	$3,31291 \times 10^{-1}$	33,13%	0,982490	3,94
RNA 20- Pos-Pos-Tan	[5 8 10]	$2,13920 \times 10^{-4}$	$8,58644 \times 10^{-3}$	0,86%	0,999994	3,68
RNA 21- Log-Log-Pos	[10 8 5]	$2,59967 \times 10^1$	$3,29816 \times 10^{-1}$	32,98%	0,987833	4,48
RNA 22- Log-Log-Pos	[5 8 10]	$2,33113 \times 10^{-4}$	$9,84271 \times 10^{-3}$	0,98%	0,999991	4,59
RNA 23- Log-Log-Tan	[10 8 5]	$1,11553 \times 10^{-4}$	$6,55894 \times 10^{-3}$	0,66%	0,999995	7,55
RNA 24- Log-Log-Tan	[5 8 10]	$1,20493 \times 10^{-4}$	$7,04970 \times 10^{-3}$	0,70%	0,999995	12,03
RNA 25- Pos-Tan-Tan	[10 8 5]	$1,64211 \times 10^{-4}$	$8,40628 \times 10^{-3}$	0,84%	0,999985	4,80
RNA 26- Pos-Tan-Tan	[5 8 10]	$1,13645 \times 10^{-4}$	$6,72375 \times 10^{-3}$	0,67%	0,999996	6,24
RNA 27- Tan-Tan-Log	[10 8 5]	$7,57674 \times 10^{-5}$	$5,42460 \times 10^{-3}$	0,54%	0,999998	7,17
RNA 28- Tan-Tan-Log	[5 8 10]	$9,57148 \times 10^{-3}$	$1,88255 \times 10^{-2}$	1,88%	0,999239	9,08
RNA 29- Tan-Pos-Pos	[10 8 5]	$2,74206 \times 10^1$	$3,42363 \times 10^{-1}$	34,24%	0,987073	3,94
RNA 30- Tan-Pos-Pos	[5 8 10]	$2,72044 \times 10^{-4}$	$1,05949 \times 10^{-2}$	1,06%	0,999990	4,13
RNA 31- Log-Pos-Pos	[10 8 5]	$3,20043 \times 10^{-4}$	$1,14715 \times 10^{-2}$	1,15%	0,999990	5,24
RNA 32- Log-Pos-Pos	[5 8 10]	$1,80111 \times 10^{-4}$	$9,25905 \times 10^{-3}$	0,93%	0,999991	5,36
RNA 33- Pos-Log-Log	[10 8 5]	$1,46056 \times 10^{-4}$	$8,38945 \times 10^{-3}$	0,84%	0,999993	6,10
RNA 34- Pos-Log-Log	[5 8 10]	$2,06648 \times 10^{-4}$	$8,91202 \times 10^{-3}$	0,89%	0,999994	5,10
RNA 35- Tan-Pog-Log	[10 8 5]	$7,58158 \times 10^{-5}$	$5,43222 \times 10^{-3}$	0,54%	0,999997	7,48
RNA 36- Tan-Log-Log	[5 8 10]	$1,47491 \times 10^{-4}$	$7,94200 \times 10^{-3}$	0,79%	0,999995	10,72
RNA 37- Tan-Tan-Pos	[10 8 5]	$2,43626 \times 10^{-4}$	$9,89357 \times 10^{-3}$	0,99%	0,999992	5,43
RNA 38- Tan-Tan-Pos	[5 8 10]	$1,35164 \times 10^{-4}$	$7,63262 \times 10^{-3}$	0,76%	0,999994	4,98
RNA 39- Log-Tan-Tan	[10 8 5]	$1,33181 \times 10^{-4}$	$7,02014 \times 10^{-3}$	0,70%	0,999996	9,70
RNA 40- Log-Tan-Tan	[5 8 10]	$8,87151 \times 10^{-5}$	$5,85212 \times 10^{-3}$	0,59%	0,999996	13,34
RNA 41- Pos-Log-Tan	[10 8 5]	$1,31649 \times 10^{-4}$	$7,68753 \times 10^{-3}$	0,77%	0,999994	6,59
RNA 42- Pos-Log-Tan	[5 8 10]	$1,44911 \times 10^{-4}$	$7,73398 \times 10^{-3}$	0,77%	0,999994	6,77
RNA 43- Tan-Log-Pos	[10 8 5]	$1,48882 \times 10^{-4}$	$8,38973 \times 10^{-3}$	0,84%	0,999993	4,45
RNA 44- Tan-Log-Pos	[5 8 10]	$1,63198 \times 10^{-4}$	$8,31821 \times 10^{-3}$	0,83%	0,999990	4,26
RNA 45- Pos-Tan-Log	[10 8 5]	$2,69579 \times 10^{-4}$	$9,43302 \times 10^{-3}$	0,94%	0,999985	5,10
RNA 46- Pos-Tan-Log	[5 8 10]	$1,39467 \times 10^{-4}$	$7,63722 \times 10^{-3}$	0,76%	0,999995	4,93
RNA 47- Tan-Pos-Log	[10 8 5]	$2,45374 \times 10^{-4}$	$9,77989 \times 10^{-3}$	0,98%	0,999995	5,32
RNA 48- Tan-Pos-Log	[5 8 10]	$1,49527 \times 10^{-4}$	$8,13746 \times 10^{-3}$	0,81%	0,999994	5,02
RNA 49- Log-Pos-Tan	[10 8 5]	$6,42951 \times 10^{-4}$	$1,40564 \times 10^{-2}$	1,41%	0,999973	9,29
RNA 50- Log-Pos-Tan	[5 8 10]	$1,71987 \times 10^{-4}$	$8,28721 \times 10^{-3}$	0,83%	0,999989	7,12
RNA 51- Log-Tan-Pos	[10 8 5]	$2,88645 \times 10^{-4}$	$1,00876 \times 10^{-2}$	1,01%	0,999976	4,92
RNA 52- Log-Tan-Pos	[5 8 10]	$1,67953 \times 10^{-4}$	$7,95916 \times 10^{-3}$	0,80%	0,999994	4,78

Fonte: Próprio autor (Pos = Poslin; Tan = Tansig; Log=Logsig)

Como apresentado na tabela 6.3, as arquiteturas foram configuradas com 3 camadas ocultas e uma camada de saída. Para as camadas ocultas foram alternadas três funções de ativação, sendo elas a *Tansig* (tangente hiperbólica), a *Logsig* (Função Logística) e a *Poslin* (*ReLU - Rectified Linear Unit*). Assim, foram arquitetadas 26 redes com funções de ativação

alternadas entre si, sendo que cada uma foi testada com [10, 8 e 5] e [5, 8 e 10] neurônios nas respectivas camadas, totalizando 52 redes neurais.

Para avaliar as redes neurais treinadas adotou-se prioritariamente o critério baseado no Erro Quadrático Médio, seguido da regressão linear e tempo com processamento computacional. A regressão representa quão próximo os valores da predição obtidos da RNA estão dos valores reais, sendo que o ideal é o valor da regressão estar mais próxima de 1. Já O EQM calcula a média dos quadrados das diferenças entre os valores previstos pela RNA e os valores reais enquanto o EAM calcula a média das diferenças entre os valores absolutos previsto e valores simulados, quais podem ser representados pelas equações (6.1) e (6.2) respectivamente.

$$EQM = \frac{1}{n} \left(\sum_{i=1}^n Y_i - \hat{Y}_i \right)^2 \quad (6.1)$$

$$EAM = \frac{1}{n} \left(\sum_{i=1}^n |Y_i - \hat{Y}_i| \right) , \quad (6.2)$$

onde n é o número total de observações, Y_i são os valores reais e \hat{Y}_i são os valores previstos pelo modelo.

Tabela 6.4 – Desempenho das 10 melhores Redes Neurais Artificiais utilizando como critério de avaliação o Erro Quadrático Médio da validação cruzada.

Arquitetura das camadas ocultas (Função de Ativação)	Quantidade de Neurônios por camada oculta	EQM Validação Cruzada	EAM Validação Cruzada	Regressão	Tempo (s)
RNA 01- Tan-Tan-Tan	[10 8 5]	$6,39410 \times 10^{-5}$	$4,89013 \times 10^{-3}$	0,999997	7,46
RNA 02- Tan-Tan-Tan	[5 8 10]	$7,02583 \times 10^{-5}$	$4,92453 \times 10^{-3}$	0,999997	14,40
RNA 27- Tan-Tan-Log	[10 8 5]	$7,57674 \times 10^{-5}$	$5,42460 \times 10^{-3}$	0,999998	7,17
RNA 35- Tan-Log-Log	[10 8 5]	$7,58158 \times 10^{-5}$	$5,43222 \times 10^{-3}$	0,999997	7,48
RNA 40- Log-Tan-Tan	[5 8 10]	$8,87151 \times 10^{-5}$	$5,85212 \times 10^{-3}$	0,999996	13,34
RNA 03- Log-Log-Log	[10 8 5]	$1,08222 \times 10^{-4}$	$6,59319 \times 10^{-3}$	0,999996	6,29
RNA 23- Log-Log-Tan	[10 8 5]	$1,11553 \times 10^{-4}$	$6,55894 \times 10^{-3}$	0,999995	7,55
RNA 26- Pos-Tan-Tan	[5 8 10]	$1,13645 \times 10^{-4}$	$6,72375 \times 10^{-3}$	0,999996	6,24
RNA 24- Log-Log-Tan	[5 8 10]	$1,20493 \times 10^{-4}$	$7,04970 \times 10^{-3}$	0,999995	12,03
RNA 12- Tan-Pos-Tan	[5 8 10]	$1,31423 \times 10^{-4}$	$7,34432 \times 10^{-3}$	0,999995	4,83

Fonte: Próprio autor (Pos = Poslin; Tan = Tansig; Log=Logsig)

A tabela 6.4 apresenta 10 RNA's com os melhores rendimentos baseado no Erro Quadrático Médio como critério de avaliação. Pode-se constatar que a função de ativação *Tansig* está presente em 9 das 10 melhores redes, contribuindo significativamente para o bom desempenho do algoritmo em função da natureza do projeto. Foi possível verificar também que 5 entre as 10 melhores redes tem a função *Tansig* na primeira camada escondida e 7 redes tem

a mesma função na terceira camada. A melhor rede com nomenclatura RNA 01- *Tan-Tan-Tan* possui [10 8 5] neurônios nas três camadas ocultas respectivamente, apresentando um EQM de $6,39410 \times 10^{-5}$, EAM de $4,89013 \times 10^{-5}$, regressão linear de 0,99997 e tempo de processamento computacional de 7,46 segundos. Para este critério de avaliação, 50% das redes tiveram 10 e 5 neurônios na primeira e última camada oculta respectivamente.

Tabela 6.5 – Desempenho das Redes Neurais Artificiais utilizando como critério de avaliação o tempo de processamento computacional

Arquitetura das camadas ocultas (Função de Ativação)	Quantidade de Neurônios por camada oculta	EQM Validação Cruzada	EAM Validação Cruzada	Regressão	Tempo (s)
RNA 20- Pos-Pos-Tan	[5 8 10]	$2,13920 \times 10^{-4}$	$8,58644 \times 10^{-3}$	0,999994	3,68
RNA 06- Pos-Pos-Pos	[5 8 10]	$4,47628 \times 10^{-4}$	$1,09581 \times 10^{-2}$	0,999991	3,87
RNA 19- Pos-Pos-Tan	[10 8 5]	$2,60067 \times 10^1$	$3,31291 \times 10^{-1}$	0,982490	3,94
RNA 29- Tan-Pos-Pos	[10 8 5]	$2,74206 \times 10^1$	$3,42363 \times 10^{-1}$	0,987073	3,94
RNA 18- Pos-Tan-Pos	[5 8 10]	$1,42448 \times 10^{-3}$	$1,42286 \times 10^{-2}$	0,999973	4,05
RNA 05- Pos-Pos-Pos	[10 8 5]	$2,64873 \times 10^{-4}$	$1,01749 \times 10^{-2}$	0,999994	4,10
RNA 30- Tan-Pos-Pos	[5 8 10]	$2,72044 \times 10^{-4}$	$1,05949 \times 10^{-2}$	0,999990	4,13
RNA 17- Pos-Tan-Pos	[10 8 5]	$1,35359 \times 10^{-4}$	$7,86647 \times 10^{-3}$	0,999995	4,20
RNA 44- Tan-Log-Pos	[5 8 10]	$1,63198 \times 10^{-4}$	$8,31821 \times 10^{-3}$	0,999990	4,26
RNA 08- Tan-Log-Pos	[5 8 10]	$1,63198 \times 10^{-4}$	$8,31821 \times 10^{-3}$	0,999990	4,38

Fonte: Próprio autor (Pos = Poslin; Tan = Tansig; Log=Logsig)

A tabela 6.5 apresenta 10 RNA's que apresentaram melhores rendimentos baseado no tempo de processamento computacional. Pode-se constatar que a função de ativação *ReLU* está presente em todas as 10 redes que tiveram o menor tempo de processamento. Foi possível notar também que 6 redes tem a função *ReLU* na primeira camada escondida e 8 redes com a mesma função na terceira camada. A melhor rede para esse critério de avaliação foi identificada como RNA 20- *Pos-Pos-Tan* que tem [5 8 10] neurônios nas camadas ocultas, apresentou um EQM de $2,139920 \times 10^{-4}$, EAM de $8,58644 \times 10^{-5}$, regressão linear de 0,99994 e tempo de processamento computacional de 3,68 segundos. Devido à natureza da função de ativação *ReLU*, nesta etapa de avaliação, o tempo de processamento computacional das 10 melhores redes nessa etapa foi um pouco menor que as demais redes, sendo que a diferença entre as redes com maior e menor tempo é de 10,72 segundos, onde a RNA com maior tempo de processamento foi identificada como RNA 02- *Tan-Tan-Tan* considerada a segunda melhor rede pelo critério de EQM, com 14,40 segundos.

A tabela 6.6 consta 10 RNA's que apresentaram melhores desempenhos utilizando como critério de avaliação a regressão linear. Pode-se verificar mais uma vez que a função de ativação *Tansig* é responsável pelo bom desempenho da rede neural nesta fase, constando sua presença em 9 das 10 redes. Foi observado que 8 redes com melhores índices de regressão linear coincidiram entre as 10 melhores no critério de avaliação Erro Quadrático Médio, isto porque a regressão linear tem relação direta com a performance do EQM. A melhor rede para esse

critério de avaliação foi identificada como RNA 27- *Tan-Tan-Log* com [10 8 5] neurônios nas respectivas camadas ocultas, apresentando um EQM de $7,57674 \times 10^{-5}$, EAM de $5,42460 \times 10^{-3}$, regressão linear de 0,99998 e tempo de processamento computacional de 7,17 segundos e tem seu gráfico de regressão linear apresentado na figura 6.5.

Tabela 6.6 – Desempenho das 10 melhores Redes Neurais Artificiais utilizando como critério de avaliação a regressão linear.

Arquitetura das camadas ocultas (Função de Ativação)	Quantidade de Neurônios por camada oculta	EQM Validação Cruzada	EAM Validação Cruzada	Regressão	Tempo (s)
RNA 27- <i>Tan-Tan-Log</i>	[10 8 5]	$7,57674 \times 10^{-5}$	$5,42460 \times 10^{-3}$	0,999998	7,17
RNA 35- <i>Tan-Log-Log</i>	[10 8 5]	$7,58158 \times 10^{-5}$	$5,43222 \times 10^{-3}$	0,999997	7,48
RNA 01- <i>Tan-Tan-Tan</i>	[10 8 5]	$6,39410 \times 10^{-5}$	$4,89013 \times 10^{-3}$	0,999997	7,46
RNA 02- <i>Tan-Tan-Tan</i>	[5 8 10]	$7,02583 \times 10^{-5}$	$4,92453 \times 10^{-3}$	0,999997	14,40
RNA 40- <i>Log-Tan-Tan</i>	[5 8 10]	$8,87151 \times 10^{-5}$	$5,85212 \times 10^{-3}$	0,999996	13,34
RNA 26- <i>Pos-Tan-Tan</i>	[5 8 10]	$1,13645 \times 10^{-4}$	$6,72375 \times 10^{-3}$	0,999996	6,24
RNA 03- <i>Log-Log-Log</i>	[10 8 5]	$1,08222 \times 10^{-4}$	$6,59319 \times 10^{-3}$	0,999996	6,29
RNA 39- <i>Log-Tan-Tan</i>	[10 8 5]	$1,33181 \times 10^{-4}$	$7,02014 \times 10^{-3}$	0,999996	9,70
RNA 46- <i>Pos-Tan-Log</i>	[5 8 10]	$1,39467 \times 10^{-4}$	$7,63722 \times 10^{-3}$	0,999995	4,93
RNA 23- <i>Log-Log-Tan</i>	[10 8 5]	$1,11553 \times 10^{-4}$	$6,55894 \times 10^{-3}$	0,999995	7,55

Fonte: Próprio autor (*Pos* = Poslin; *Tan* = Tansig; *Log*=Logsig)

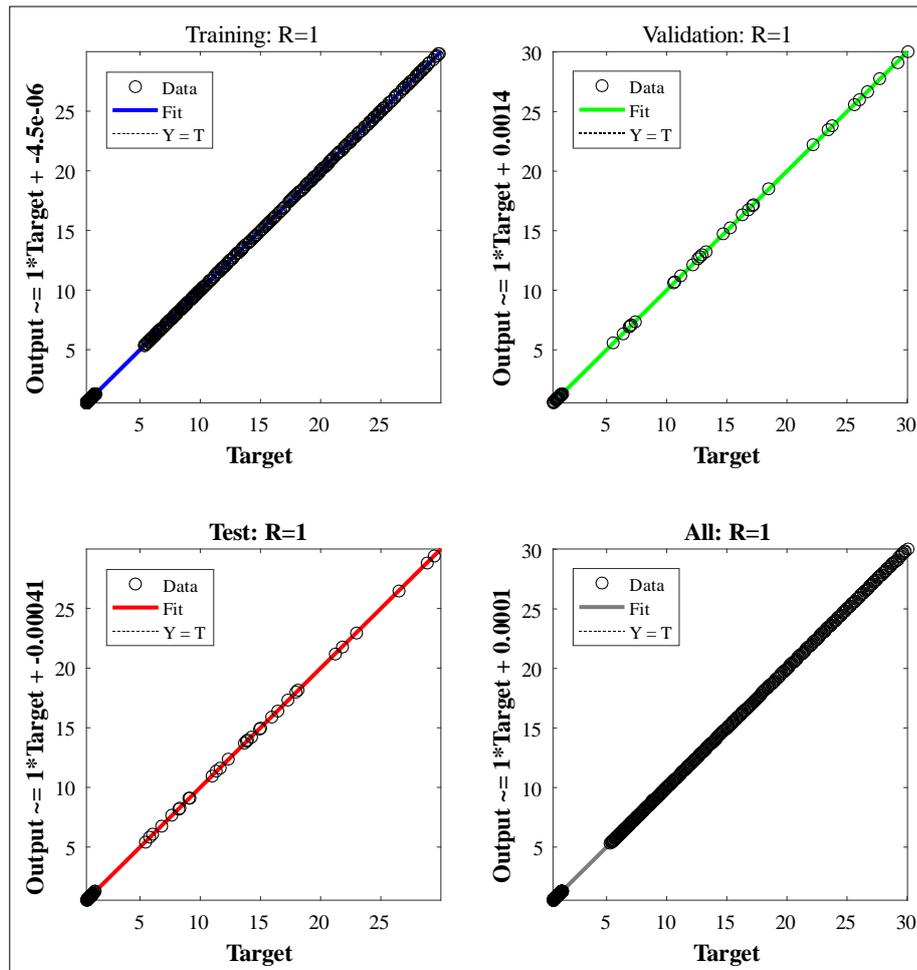


Figura 6.5 – Regressão Linear da Melhor Rede Neural Artificial (RNA 27- *Tan-Tan-Log*) para treinamento, teste, validação e geral com Regressão $\cong 1$, Y=Saída e T=Target

A figura 6.6 mostra o desempenho da *RNA 01- Tan-Tan-Tan*, com a melhor performance de validação alcançada em 114 épocas com EQM de $2,05336 \times 10^{-4}$.

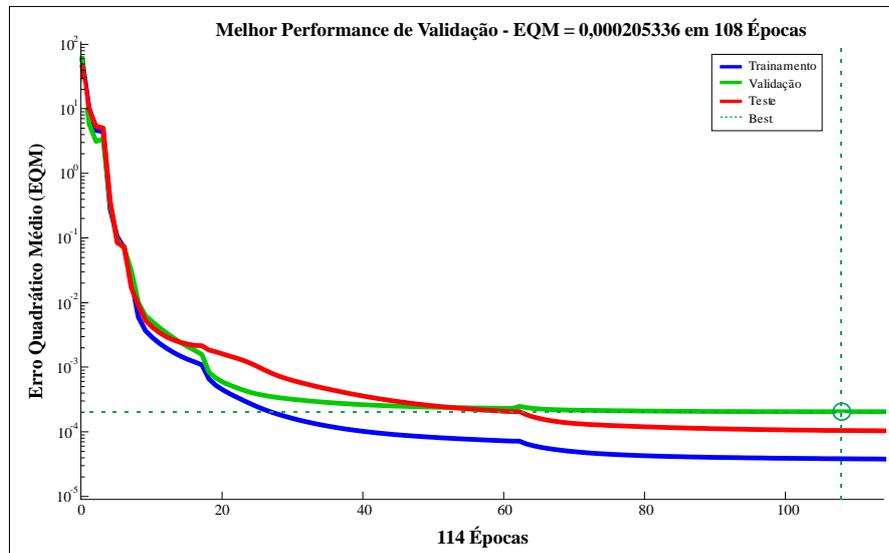


Figura 6.6 – Performance da Melhor Rede Neural Artificial (*RNA 01- Tan-Tan-Tan*) com melhor performance de validação em 114 épocas e Erro Médio Quadrático de $2,05336 \times 10^{-4}$

O histograma de erro referente a *RNA 01- Tan-Tan-Tan* de melhor desempenho entre as 52 RNA's treinadas revela um bom desempenho a partir desse critério de avaliação. Como pode ser observado, todas as barras tem um erro muito próximo de zero e a barra mais próxima de zero está no intervalo de $(-0,005155; 0,009259)$. Ou seja, a maior parte de dados (instâncias) preditos estão neste intervalo próximo de zero, enquanto a quantidade de dados para as demais barras é significativamente menor, como poder ser visto na figura 6.7.

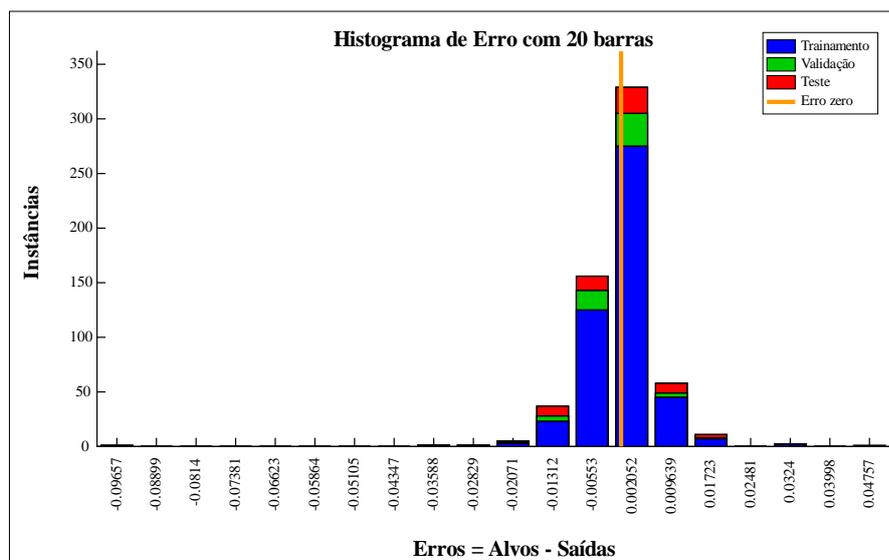


Figura 6.7 – Histograma de erro Melhor Rede Neural Artificial (*RNA 01- Tan-Tan-Tan*) com a maior quantidade de dados com a diferença entre a predição e o alvo próximo de $2,052 \times 10^{-3}$

Após constatar o bom desempenho do algoritmo baseado em *Machine Learning* foi possível obter as coordenadas para o eixo-x das portas de saídas M (1 × M) de maior potência e as coordenadas para eixo-y responsável pelo limite de corte para acoplamento das portas de saída para um dispositivo MMI compacto. Para tal, foram simulados dispositivos com dados preditos pela RNA, analisados, comparados com os dispositivos simulados e confrontados com os desempenhos dos dispositivos catalogados na revisão bibliográfica.

Tabela 6.7 – Parâmetros Ópticos e Geométricos de 4 dispositivos MMI's (1 × 2) obtidos através da Rede Neural Artificial de melhor desempenho com nomenclatura *RNA 01-Tan-Tan-Tan* vs. Dispositivos simulados analiticamente

PARÂMETROS ÓPTICOS E GEOMÉTRICOS	DISPOSITIVOS							
	6.1a (Maxwell)	6.9a (RNA)	6.1b (Maxwell)	6.9b (RNA)	6.1c (Maxwell)	6.9c (RNA)	6.1d (Maxwell)	6.9d (RNA)
Comprimento de Onda (λ)	1,550 μ m							
Índice do Substrato (n_{Silica})	1,4440	1,4440	1,4440	1,4440	1,4440	1,4440	1,4440	1,4440
Índice do Núcleo (n_{Silicio})	3,4777	3,4777	3,4777	3,4777	3,4777	3,4777	3,4777	3,4777
Largura (W_{MMI})	2,00 μ m	2,00 μ m	3,00 μ m	3,00 μ m	4,00 μ m	4,00 μ m	5,00 μ m	5,00 μ m
Coordenada eixo-y (L_{MMI})	5,330 μ m	5,327 μ m	11,350 μ m	11,351 μ m	19,600 μ m	19,598 μ m	30,030 μ m	29,978 μ m
Coordenada eixo- x_1	-0,542 μ m	-0,544 μ m	-0,800 μ m	-0,796 μ m	-1,052 μ m	-1,049 μ m	-1,302 μ m	-1,292 μ m
Coordenada eixo- x_2	0,542 μ m	0,5444 μ m	0,800 μ m	0,796 μ m	1,052 μ m	1,049 μ m	1,302 μ m	1,292 μ m
Eficiência	91,46%	90,59%	89,95%	89,81%	86,95%	86,32%	85,52%	85,47%

Fonte: Próprio autor

A tabela 6.7 apresentam informações dos dispositivos simulados com dados analíticos e os dispositivos obtidos através da melhor rede neural de nome *RNA 01-Tan-Tan-Tan*. Através os dados da RNA, foi possível simular 4 dispositivos MMI's (1 × 2) com campos formado pela *self-image*, como mostra a figura 6.8, utilizando as configurações da tabela 6.7.

A figura 6.9a representa o dispositivo mais compacto com Largura (W_{MMI}) de 2,00 μ m comprimento (L_{MMI}) de corte medindo 5,3276 μ m, computado uma diferença de $2,4 \times 10^{-3}$ μ m em relação ao dispositivo simulado com os dados analíticos. As coordenadas do eixo- x_1 e eixo- x_2 foram -0,5443 μ m e 0,5443 μ m respectivamente, contabilizando uma diferença de $2,3 \times 10^{-3}$ μ m quando comparado com o dispositivo simulado com os dados analíticos. No que diz respeito a transferência de potência na saída do guia de ondas, foi calculado com uma eficiência de 90,59%, constatado uma diferença de 0,87%, ou seja, equivalente a perda menor que 0,42dB.

A figura 6.9d representa o dispositivo com maior dimensão dentre os simulados com os dados da RNA, com Largura (W_{MMI}) de 5,00 μ m e comprimento (L_{MMI}) de corte medindo 29,9786 μ m, computando uma diferença de $5,1 \times 10^{-2}$ μ m em relação ao dispositivo simulado

com os dados analíticos. Para este dispositivo, as coordenadas do eixo- x_1 e eixo- x_2 foram $-1,2929\mu\text{m}$ e $1,2929\mu\text{m}$ respectivamente, contabilizando uma diferença de $9,1\times 10^{-3}\mu\text{m}$ quando comparado com o dispositivo simulado com os dados analíticos. No que diz respeito a transferência de potência na saída do guia de ondas, foi calculado com uma eficiência de 85,47%, constatado uma diferença de 0,05%, configurando uma perda menor que 0,68dB.

A figura 6.8 apresenta os dispositivos simulados com as mesmas configurações da tabela 6.7, porém com a proporção das dimensões conservadas para fins de visualização.

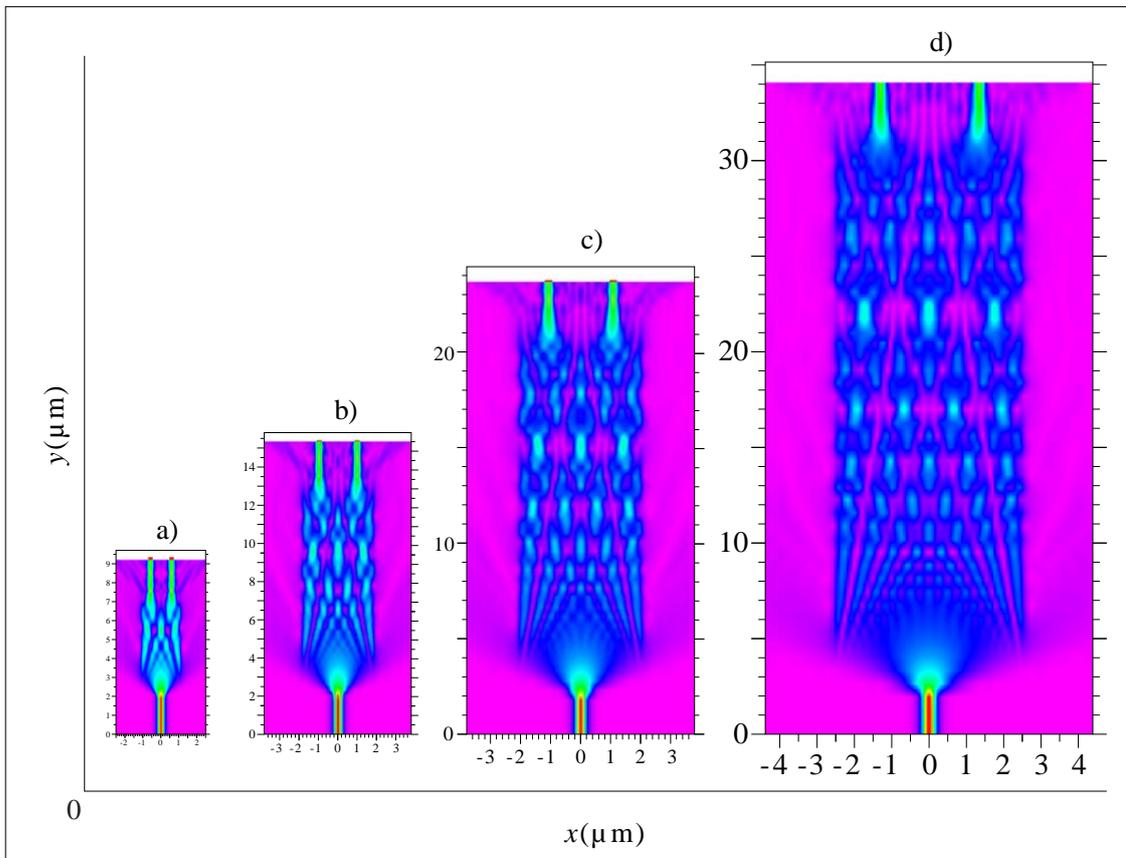


Figura 6.8 – Campo formado pela *Self-image* das simulações com dados obtidos através da *RNA 01-Tan-Tan-tan* de dispositivos MMI (1×2) para o comprimento de ondas de $1,55\mu\text{m}$, com a proporção das dimensões conservadas. a) $2,00 \times 5,3276\mu\text{m}$; b) $3,00 \times 11,3518\mu\text{m}$; c) $4,00 \times 19,5989\mu\text{m}$; d) $5,00 \times 29,9786\mu\text{m}$ (Parâmetros da tabela 6.7)

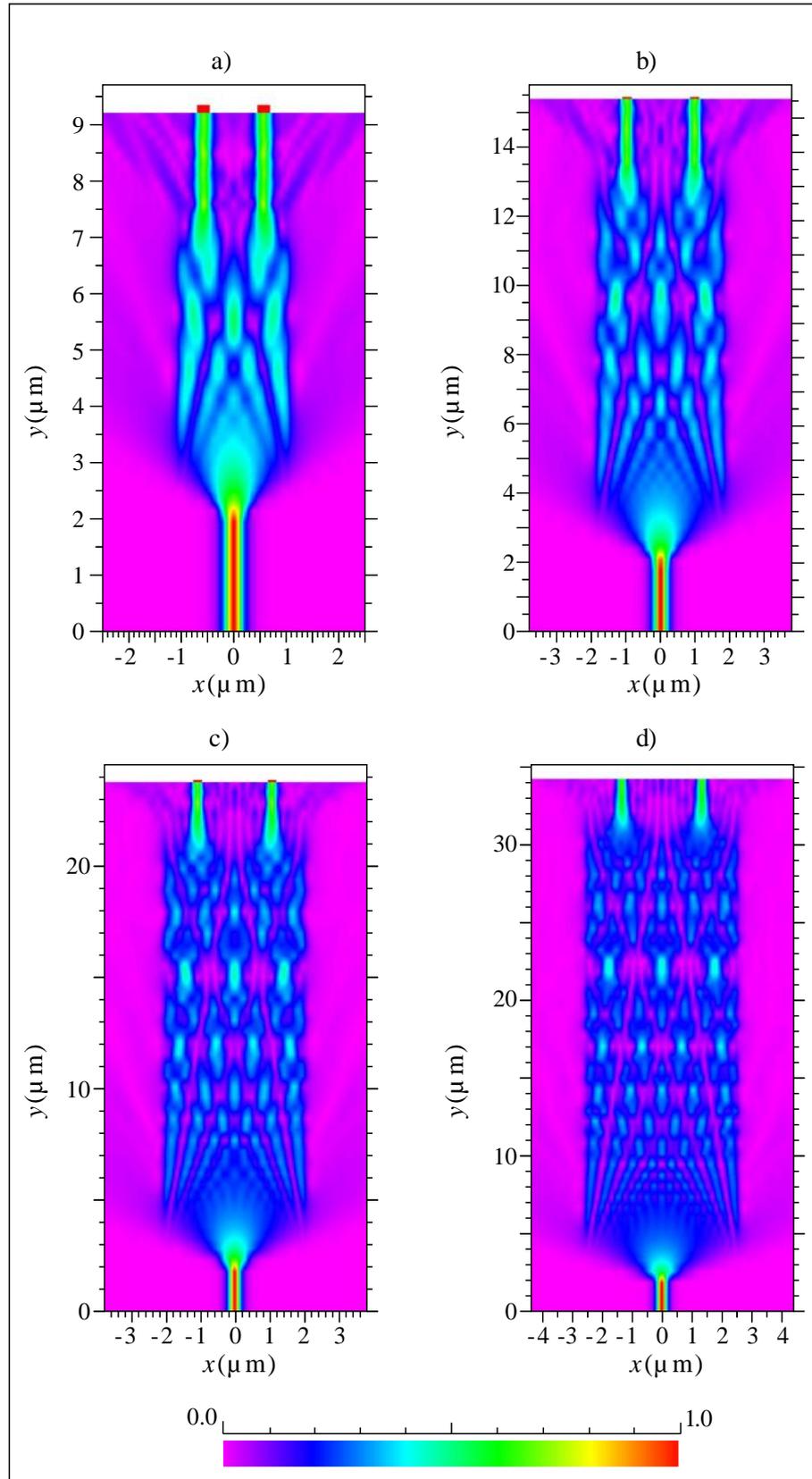


Figura 6.9 – Campo formado pela *Self-image* das simulações com dados obtidos através da *RNA 01-Tan-Tan-tan* de dispositivos MMI (1×2) para o comprimento de ondas de $1,55\mu\text{m}$. a) $2,00 \times 5,3276\mu\text{m}$; b) $3,00 \times 11,3518\mu\text{m}$; c) $4,00 \times 19,5989\mu\text{m}$; d) $5,00 \times 29,9786\mu\text{m}$ (Parâmetros da tabela 6.7)

A figura 6.10 apresenta uma comparação entre os dados analíticos e os dados obtidos através da Rede Neural Artificial de melhor desempenho dentre as 52 redes desenvolvidas. O gráfico mostra os pontos do triângulo preto (coordena x e y referente aos dados analíticos) sobre o triângulo branco (coordena x e y obtidas pela *RNA 01-Tan-Tan-tan*), indicando uma diferença muito pequena, contabilizando como a menor e maior diferença entre os eixos- x de $2,46 \times 10^{-5} \mu\text{m}$ e $1,17 \times 10^{-2} \mu\text{m}$ respectivamente, e para os eixos- y com de $8,96 \times 10^{-5} \mu\text{m}$ e $1,00 \times 10^{-2} \mu\text{m}$ respectivamente.

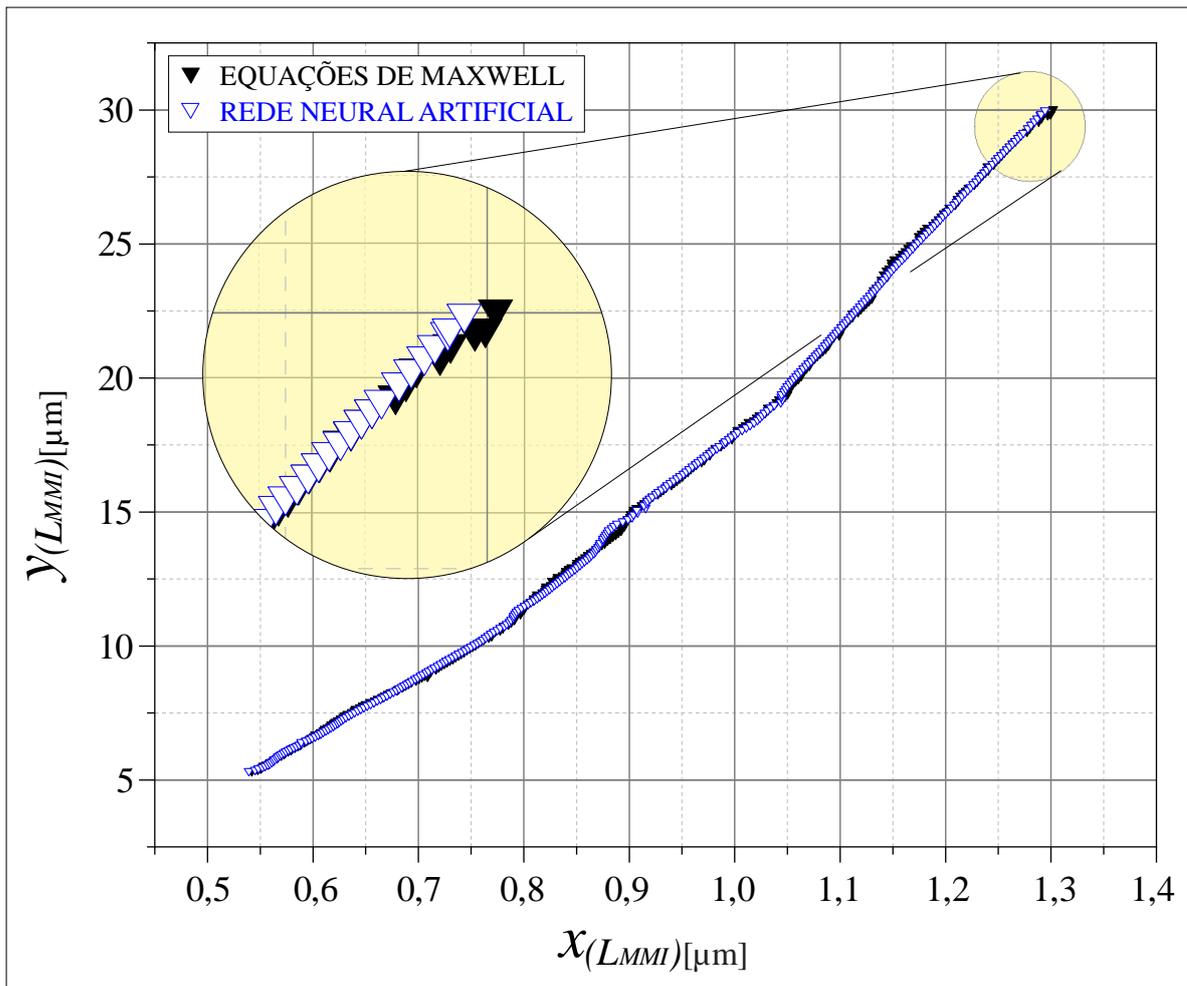


Figura 6.10 – Coordenadas dos eixo- x e eixo- y dos dados analíticos vs. valores obtidos pelas *RNA 01-Tan-Tan-Tan*

A figura 6.11 compara a eficiência da transferência de potência óptica entre um dispositivo com largura (W_{MMI}) de $2,20 \mu\text{m}$ simulado analiticamente e um dispositivo com a mesma largura simulado com dados da *RNA 01-Tan-Tan-Tan*. A eficiência foi calculada para a faixa de comprimento compreendida entre $1,25$ e $1,7 \mu\text{m}$, abrangendo as bandas de comunicação O, E, S, C, L e U, mostraram perdas que variam entre $0,31$ e $0,41 \text{ dB}$ para o

dispositivo simulado analiticamente representado no gráfico com a linha preta e perdas que variam entre 0,32 a 0,45dB para o dispositivo simulado pelo algoritmo desenvolvido utilizado a RNA, que pode ser identificado no gráfico com a linha vermelha. A partir da leitura do gráfico, pode-se concluir as duas bandas de comunicação que oferecem melhores respostas aos dispositivos MMI com as características apresentadas são as bandas O e C.

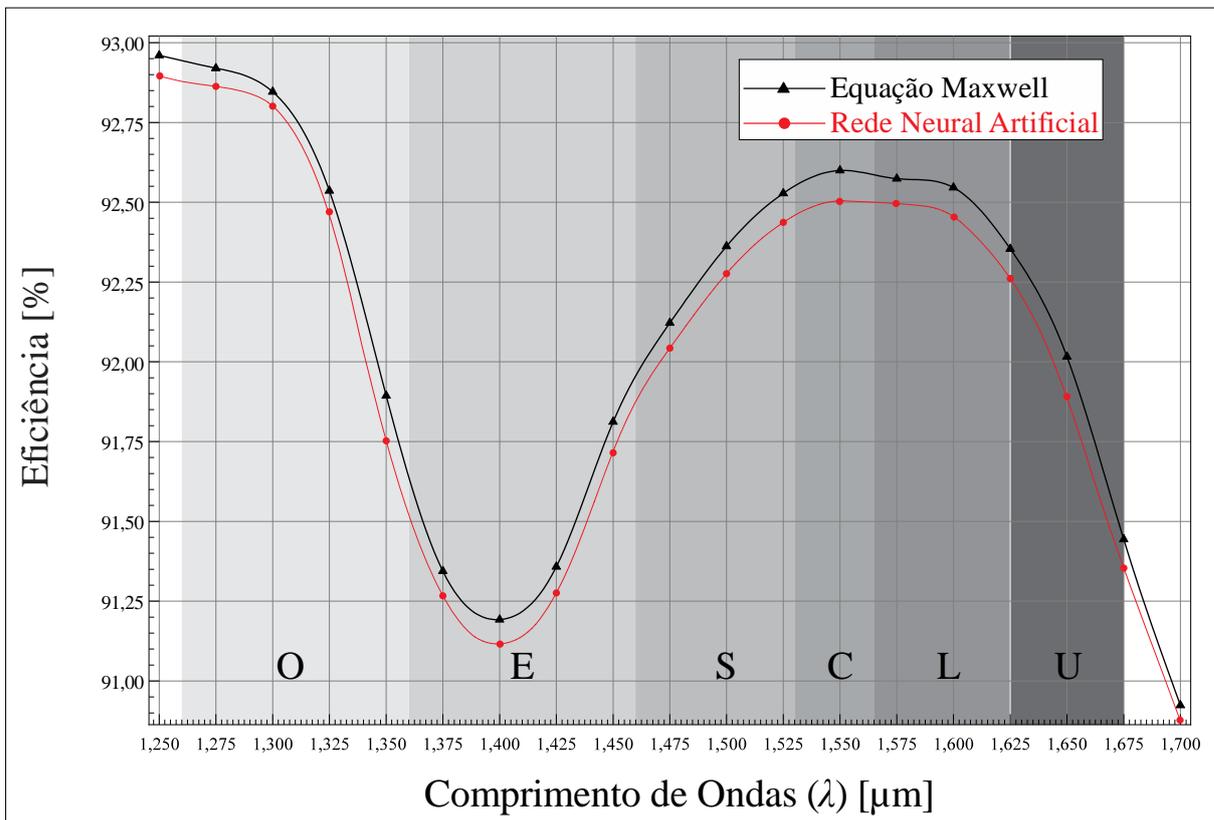


Figura 6.11 – Comparação da Eficiência entre dispositivos MMI (1×2) com W_{MMI} de $2,20\mu\text{m}$ simulado analiticamente e predito pela Rede Neural Artificial de melhor desempenho, para a faixa de comprimento de ondas de $1,25 \sim 1,7\mu\text{m}$ integrando as bandas O, E, S, C, L e U

A figura 6.12 compara a eficiência da transferência de potência óptica entre um dispositivo que apresentou a melhor transferência de potência com largura (W_{MMI}) de $2,13\mu\text{m}$ simulado analiticamente e um dispositivo com as mesmas características, porém simulado com dados da RNA *01-Tan-Tan-Tan*. A eficiência foi calculada para a faixa de comprimento compreendida entre $1,25$ e $1,7\mu\text{m}$, abrangendo as bandas de comunicação O, E, S, C, L e U, com perdas que variam entre 0,30 e 0,45dB para o dispositivo simulado analiticamente representado no gráfico com a linha preta e perdas que variam entre 0,32 a 0,47dB para o dispositivo simulado pelo algoritmo desenvolvido utilizado a RNA, que pode ser identificado no gráfico com a linha vermelha. A partir da leitura do gráfico, pode-se concluir as duas bandas

de comunicação que oferecem melhores respostas aos dispositivos MMI com as características apresentadas são as bandas O e C.

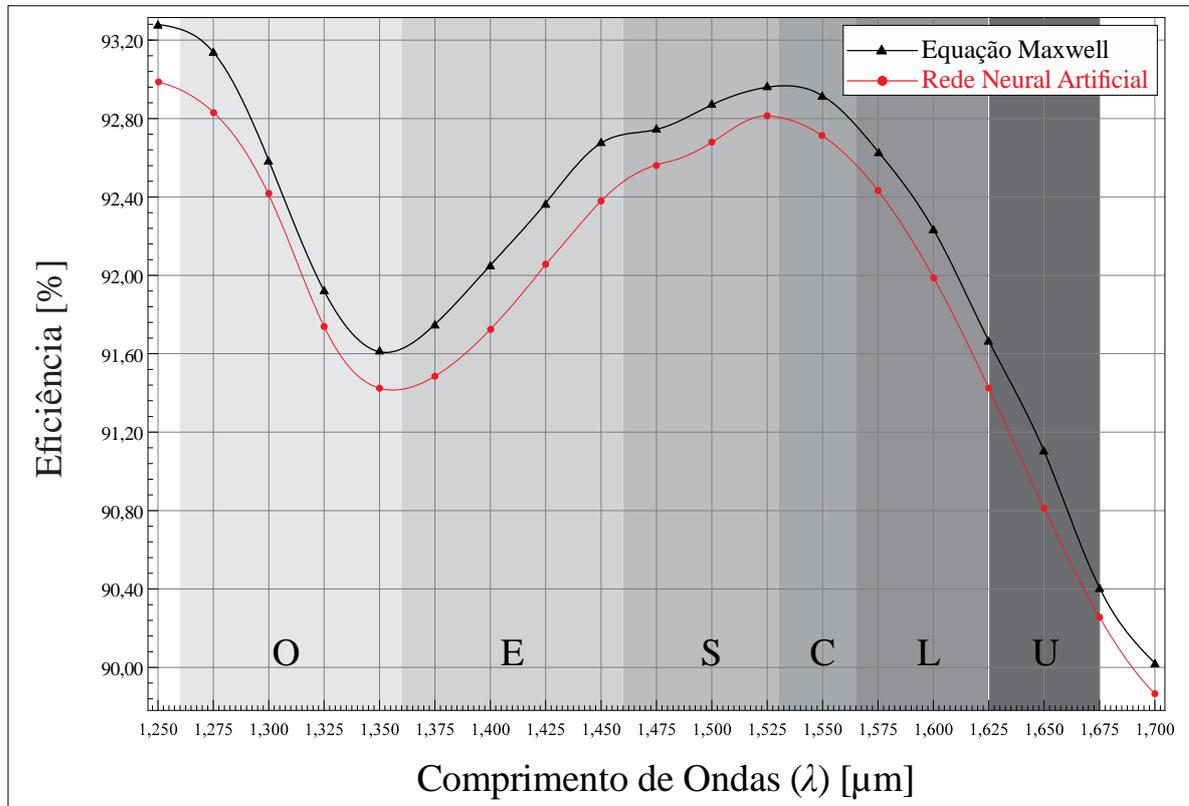


Figura 6.12 – Comparação da Eficiência entre dispositivos MMI (1×2) com L_{MMI} de $2,13\mu\text{m}$ simulado analiticamente e previsto pela Rede Neural Artificial de melhor desempenho, para a faixa de comprimento de ondas de $1,25 \sim 1,7\mu\text{m}$ integrando as bandas O, E, S, C, L e U

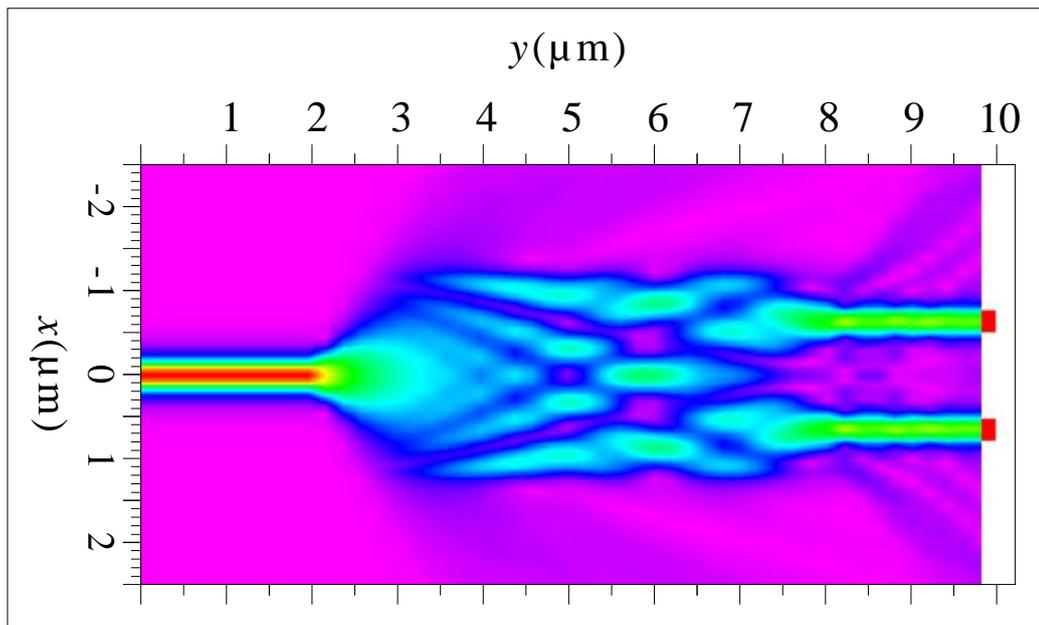


Figura 6.13 – Campo formado pela *Self-image* da simulação com dados obtidos através da RNA *01-Tan-Tan-Tan* de dispositivos MMI (1×2) com $2,13 \times 5,3276\mu\text{m}$, para o comprimento de ondas de $1,55\mu\text{m}$.

A tabela 6.8 compara as técnicas, recursos computacionais, técnicas de *Machine Learning*, aplicação e eficiência ou perdas dos dispositivos simulados como resultado do algoritmo. Baseado no levantamento bibliográfico, podemos afirmar que as Redes Neurais tem sido utilizada como técnica para aprendizado de máquinas em pesquisas com aplicações semelhante a este trabalho, a partir de 2019 até o ano corrente. Assim, legitimado as RNA's como uma ferramenta eficiente e poderosa quando se pretende realizar o design inverso de dispositivos de Interferência Multimodo atuando como divisores de potência.

Tabela 6.8 – Comparação dos resultados de trabalhos disponível na análise bibliográfica com os resultados apresentado neste trabalho de dissertação

[Referência]	Ano	Machine Learning	Simulador	Aplicação	Eficiência (%) ou Perda (dB)
[47]	2024	RNC	<i>Ansys Lumerical</i>	Power Splitter	0,45 dB
[40]	2023	DBS	-	Power Splitter	0,9dB
[37]	2022	RNA	<i>Ansys Lumerical</i>	Power Splitter	0,5dB
[34]	2022	RNA	<i>Ansys Lumerical</i>	Power Splitter	90%
[38]	2021	b-ARLA	<i>Ansys Lumerical</i>	Beam Splitter	0,92db
[33]	2021	RNP	<i>Ansys Lumerical</i>	Power Splitter	0,65dB
[54]	2019	RNP	<i>Ansys Lumerical</i>	Power Splitter	90%
[36]	2019	RNA	<i>Ansys Lumerical</i>	Power Splitter	87%
[44]	2018	-	<i>Ansys Lumerical</i>	Power Splitter	1,5dB
[45]	2017	-	-	Power Splitter	80%
[43]	2013	-	<i>Ansys Lumerical</i>	Power Splitter	0,46dB
[42]	2013	-	<i>Ansys Lumerical</i>	Power Splitter	0,6dB
[32]	2011	-	<i>FIMProp</i>	Beam Splitter	1,5dB
Este trabalho	2024	RNA	<i>BeamProp</i>	Beam Splitter	92,98 a 89,66% ou 0,32 a 0,47dB

Fonte: Próprio autor

Comparando à eficiência de transmissão ou perdas entre os resultados obtidos neste trabalho de dissertação e os trabalhos relacionados na análise bibliográfica, pode-se concluir que os resultados do referido projeto estão dentro do escopo exigidos para projetos de dispositivos fotônico desta natureza com bom desempenho e boa tolerância de fabricação, fornecendo valores de eficiência ou perdas de 92,98 a 89,66% ou 0,32 a 0,47dB respectivamente, enquanto os valores disponíveis na literatura apresentam eficiência de transmissão que variam de 80 a 90% e perdas entre 0,45 e 1,5dB.

CONCLUSÃO

Utilizar recurso de modelagem para design de dispositivo MMI através de técnicas de *Machine Learning* mostraram a capacidade dos algoritmos de IA de aprender as relações complexas entre estruturas compactas e suas respostas ópticas associadas, assim assumindo uma solução capaz de resolver problemas complexos de design inverso. Várias melhorias importantes permitiriam o uso generalizado de métodos em aplicações práticas. O primeiro e mais importante é melhorar a robustez dos métodos de projeto para lidar com variações de processo na fotolitografia, o que permitiria uma fabricação de alto rendimento.

A primeira fase deste projeto de dissertação foi realizada um levantamento bibliográfico em bases de dados importantes com o intuito de investigar o desenvolvimento de pesquisas na área de dispositivos fotônicos, mais especificamente, utilização de técnicas de *Machine Learning* para design de dispositivos de Interferência Multimodo. A partir do levantamento literário pode concluir que os dispositivos desenvolvidos utilizando as técnicas propostas nesse trabalho estão compatíveis com as técnicas empregadas atualmente. Os resultados alcançados no desempenho e nos critérios de avaliação da Rede Neural Artificial também se mostraram eficientes e aceitáveis quando comparados com aqueles que foram encontrados na literatura. Pode concluir também nessa etapa baseado na análise bibliográfica que as características de comportamentos dos dispositivos projetados pelo algoritmo satisfazem as condições necessárias tanto em desempenho óptico quanto para processo de fabricação, por apresentarem baixas perdas ópticas e boa tolerância de fabricação.

O referencial teórico e revisão bibliográfica possibilitaram analisar, interpretar e comparar as respostas obtida pelo algoritmo com maior propriedade, suficiente para fazer inferências e estabelecer uma conclusão sólida do estudo realizado. Assim, pode-se concluir que, a abordagem apresentada demonstra alta precisão para a previsão de projeto de guias de ondas utilizando as Rede Neurais Artificiais e que abordagem proposta pode ser adaptada e aplicada para a análise e projeto de uma infinidade de dispositivos fotônicos integrados complexos.

A segunda etapa deste trabalho consistiu nas simulações utilizando software comercial *BeamPROP* responsável por gerar dados de treinamento. Para esse estágio, conclui-se que o processo de simulação foi bastante exaustivo, que além de trazer experiência e conhecimento

na manipulação deste software, pode-se gerar dados de qualidade destinados a alimentação da Rede Neural Artificial. De acordo com a literatura, a qualidade dos dados de treinamentos implica no bom desenvolvimento de algoritmo de aprendizado. Nesta fase, para os dispositivos MMI com configuração (1×2) das portas, foram realizadas 301 simulações, totalizando 2408 dados independentes e 301 pares entrada/saída de dados para as RNA. Já para os dispositivos MMI com configuração (1×3) das portas, foram realizados também 301 simulações, totalizando 2709 dados independentes, ou seja, 301 pares entrada/saída de dados para alimentar a Rede Neural Artificial.

A etapa seguinte deste trabalho de dissertação teve como objetivo desenvolver um algoritmo de aprendizado com técnicas de Machine Learning capaz de prever as coordenadas para o eixo- x das portas de saídas M $(1 \times M)$ de maior potência e as coordenadas para eixo- y responsável pelo limite de corte para acoplamento das portas de saída para um dispositivo MMI compacto. Para este processo optou-se por utilizar o *MATLAB Mathworks R2016a*, copyright© 1994-2024 The MathWorks, Inc. Nessa etapa pode-se concluir que, baseado na análise bibliográfica, o referido *software* é uma ferramenta muito eficiente utilizada para desenvolver algoritmos quando se pretende trabalhar com Redes Neurais Artificiais. Apesar de sua interface de aplicação para desenvolvimento das RNA's oferecer um acompanhamento visual, o que facilita a sua compreensão, seu script é facilmente compilado para outras linguagens de programação.

Nesta fase de criação do algoritmo, pode-se concluir também que todas as 52 RNA's desenvolvidas apresentaram Erro Quadrático Médio e avaliação de regressão compatível com o que a literatura descrever como aceitável, com regressão próximo de 1 e EQM com erros menores que 10^{-5} . Esse desempenho deve-se às escolhas e testes exaustivos das funções de ativação e através de inúmeros testes e ajustes baseados no empirismo. Dessa forma chegou-se à definição da arquitetura das redes neurais. Esse processo faz jus o que afirma Chollet, que não há metodologia específica para definição da arquitetura de uma RNA na literatura. Tal etapa é dita "mais arte do que ciência" [133]. A variedade de configurações testadas até alcançar resultados coerentes, levando em consideração a natureza do problema, trouxe experiências significativas à medida que eram feitas as alterações, correções e implementações. Pois, cada erro encontrado e solucionado servira como desafios que aprimoraram o conhecimento prático e teórico.

A etapa seguinte deste trabalho consistiu na aplicação do algoritmo de aprendizado para extrair dados preditos dos parâmetros geométricos e ópticos de um dispositivo de Interferência Multimodal. Para esta fase do trabalho, pode concluir que as simulações realizadas utilizando dados do algoritmo trouxe resultados surpreendentes quando comparadas com os dispositivos simulados com os dados analíticos. Como pode ser verificado nos resultados deste trabalho, os dispositivos preditos apresentaram respostas ópticas aceitáveis, com perdas que variam entre 0,32 a 0,47dB. Considerando que existem pesquisas recentes que relatam perdas ópticas que variam entre 0,45 e 1,5dB, utilizando o mesmo método de aprendizado de máquina proposto nesta dissertação.

Este trabalho torna-se relevante no que diz respeito ao desenvolvimento de um algoritmo eficiente, capaz de estabelecer parâmetros ideais, suficiente para reduzir custos quando pretende-se fabricar dispositivos fotônicos. Uma vez que, a construção de um dispositivo físico com essas características e dimensões manométricas possui um valor muito elevado de fabricação, descartando qualquer possibilidade de erro no designer do projeto. Dessa forma, esse algoritmo contribui para que designers de dispositivos de Interferência Multimodal seja projetado sem erro nas dimensões, respeitando as tolerâncias de fabricação, conseqüentemente, reduzindo custos.

De modo geral podemos concluir que ao utilizar recurso de modelagem para design de dispositivo MMI através de técnicas de *Machine Learning* foi possível mostrar a capacidade dos algoritmos de inteligência computacional de aprender as relações complexas existentes entre estruturas nanofotônicas e suas respostas ópticas associadas, assim, assumindo uma solução capaz de resolver problemas complexos de design inverso. Com este trabalho, pode-se desenvolver modelos de dispositivos MMI eficientes, que apresentaram resultados compatíveis com os dispositivos comercialmente projetados.

ARTIGOS ACEITOS E PUBLICADOS

Artigo aceito em 20 de maio de 2024

Título: Algorithm Based in Machine Learning for Design of Multimode Interference Devices

Evento: International Conference on Optical MEMS and Nanophotonics - 2024

Local: Denostia-San Sebastian, Spain

Data: July 28th – August 1st

Título: Metamaterial waveguide modeling by a genetic algorithm-based artificial neural network [134]

Event: SPIE OPTO - 2022,

Local: San Francisco, California, United States

Data: 5 March 2022

DOI: <https://doi.org/10.1117/12.2612161>

Referência: R. Das Mercês Cerqueira, A. Dourado Sisnando, e V. F. Rodriguez-Esquerre, “Metamaterial waveguide modeling by a genetic algorithm-based artificial neural network”, em *Integrated Optics: Devices, Materials, and Technologies XXVI*, S. M. García-Blanco e P. Cheben, Orgs., San Francisco, United States: SPIE, mar. 2022, p. 71. doi: 10.1117/12.2612161.

Título: Metamaterial Waveguide Modelling by an Artificial Neural Network with Genetic Algorithm [19]

Evento: FiO LS – Frontiers in Optics + Laser Science – 2021

Local: Washington, DC United States

Data: 1 a 4 November 2021

DOI: <https://doi.org/10.1364/FIO.2021.JTh5A.141>

Referência: R. M. Cerqueira, A. D. Sisnando, e V. F. R. Esquerre, “Metamaterial Waveguide Modelling by an Artificial Neural Network with Genetic Algorithm”, em *Frontiers in Optics + Laser Science 2021*, Washington, DC: Optica Publishing Group, 2021, p. JTh5A.141. doi: 10.1364/FIO.2021.JTh5A.141.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, e Y. Liu, “Deep learning for the design of photonic structures”, *Nat. Photonics*, vol. 15, nº 2, p. 77–90, fev. 2021, doi: 10.1038/s41566-020-0685-y.
- [2] L. B. Soldano e E. C. M. Pennings, “Optical multi-mode interference devices based on self-imaging: principles and applications”, *J. Light. Technol.*, vol. 13, nº 4, p. 615–627, abr. 1995, doi: 10.1109/50.372474.
- [3] J. D. Caldwell *et al.*, “Sub-diffractive volume-confined polaritons in the natural hyperbolic material hexagonal boron nitride”, *Nat. Commun.*, vol. 5, nº 1, p. 5221, out. 2014, doi: 10.1038/ncomms6221.
- [4] F. Valdez, V. Mere, X. Wang, e S. Mookherjea, “Integrated O- and C-band silicon-lithium niobate Mach-Zehnder modulators with 100 GHz bandwidth, low voltage, and low loss”, *Opt. Express*, vol. 31, nº 4, p. 5273, fev. 2023, doi: 10.1364/OE.480519.
- [5] S. Sun *et al.*, “1 × 2 mode-independent polymeric thermo-optic switch based on a Mach–Zehnder interferometer with a multimode interferometer”, *Opt. Express*, vol. 31, nº 8, p. 12049, abr. 2023, doi: 10.1364/OE.483578.
- [6] Z. Zhang *et al.*, “Vertical Fibre Interfacing Interleaved Angled MMI for Thermal-Tuning-Free Wavelength Division (de)Multiplexing and Low-Cost Fibre Packaging”, *J. Light. Technol.*, vol. 39, nº 19, p. 6260–6268, out. 2021, doi: 10.1109/JLT.2021.3098020.
- [7] S. Sun *et al.*, “Mode-insensitive and mode-selective optical switch based on asymmetric Y-junctions and MMI couplers”, *Photonics Res.*, vol. 12, nº 3, p. 423, mar. 2024, doi: 10.1364/PRJ.509773.
- [8] H. Fukano e T. Mukai, “High-sensitivity multipoint refractive index measurement system composed of multimode interference sensors and arrayed waveguide wavelength filters”, *Jpn. J. Appl. Phys.*, vol. 61, nº SK, p. SK1004, ago. 2022, doi: 10.35848/1347-4065/ac5d87.
- [9] E. Samoi, Y. Benezra, e D. Malka, “An ultracompact 3×1 MMI power-combiner based on Si slot-waveguide structures”, *Photonics Nanostructures - Fundam. Appl.*, vol. 39, p. 100780, maio 2020, doi: 10.1016/j.photonics.2020.100780.
- [10] M. K. Alam, N. Afsary, Md. S. Sikder, Md. S. Parvez, P. Roy, e M. O. Faruk Rasel, “Near-infrared Y-branch polymer splitters realized with compact MMI structures for efficient power splitting”, *Opt. Contin.*, vol. 3, nº 3, p. 413, mar. 2024, doi: 10.1364/OPTCON.506248.
- [11] B. Liang *et al.*, “Physics-Guided Neural-Network-Based Inverse Design of a Photonic – Plasmonic Nanodevice for Superfocusing”, *ACS Appl. Mater. Interfaces*, vol. 14, nº 23, p. 27397–27404, jun. 2022, doi: 10.1021/acsami.2c05083.

- [12] K. Yao, R. Unni, e Y. Zheng, “Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale”, *Nanophotonics*, vol. 8, n° 3, p. 339–366, mar. 2019, doi: 10.1515/nanoph-2018-0183.
- [13] Q. Zhang, H. Yu, M. Barbiero, B. Wang, e M. Gu, “Artificial neural networks enabled by nanophotonics”, *Light Sci. Appl.*, vol. 8, n° 1, p. 42, maio 2019, doi: 10.1038/s41377-019-0151-0.
- [14] R. S. Hegde, “Deep learning: a new tool for photonic nanostructure design”, *Nanoscale Adv.*, vol. 2, n° 3, p. 1007–1023, 2020, doi: 10.1039/C9NA00656G.
- [15] Y. Xu, X. Zhang, Y. Fu, e Y. Liu, “Interfacing photonics with artificial intelligence: an innovative design strategy for photonic structures and devices based on artificial neural networks”, *Photonics Res.*, vol. 9, n° 4, p. B135, abr. 2021, doi: 10.1364/PRJ.417693.
- [16] J. Peurifoy *et al.*, “Nanophotonic particle simulation and inverse design using artificial neural networks”, *Sci. Adv.*, 2018.
- [17] X. He *et al.*, “Total net-rate of 27.88 Tb/s full C-band transmission over 4,550 km using 150 km span length and high-gain EDFA amplification”, *Opt. Express*, vol. 32, n° 8, p. 13500, abr. 2024, doi: 10.1364/OE.517730.
- [18] A. M. Al-Amri, “Recent Progress in Printed Photonic Devices: A Brief Review of Materials, Devices, and Applications”, *Polymers*, vol. 15, n° 15, p. 3234, jul. 2023, doi: 10.3390/polym15153234.
- [19] R. M. Cerqueira, A. D. Sisnando, e V. F. R. Esquerre, “Metamaterial Waveguide Modelling by an Artificial Neural Network with Genetic Algorithm”, em *Frontiers in Optics + Laser Science 2021*, Washington, DC: Optica Publishing Group, 2021, p. JTh5A.141. doi: 10.1364/FIO.2021.JTh5A.141.
- [20] C. Yeung *et al.*, “Elucidating the Behavior of Nanophotonic Structures through Explainable Machine Learning Algorithms”, *ACS Photonics*, vol. 7, n° 8, p. 2309–2318, ago. 2020, doi: 10.1021/acsp Photonics.0c01067.
- [21] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, e A. W. Rodriguez, “Inverse design in nanophotonics”, *Nat. Photonics*, vol. 12, n° 11, p. 659–670, nov. 2018, doi: 10.1038/s41566-018-0246-9.
- [22] X. Xu, Y. Li, e W. Huang, “Inverse design of the MMI power splitter by asynchronous double deep Q-learning”, *Opt. Express*, vol. 29, n° 22, p. 35951, out. 2021, doi: 10.1364/OE.440782.
- [23] M. Butz, A. S. Abazi, R. Ross, B. Risse, e C. Schuck, “Inverse design of nanophotonic devices using dynamic binarization”, *Opt. Express*, vol. 31, n° 10, p. 15747, maio 2023, doi: 10.1364/OE.484484.

- [24] Y. Liu, T. Hao, W. Li, J. Capmany, N. Zhu, e M. Li, “Observation of parity-time symmetry in microwave photonics”, *Light Sci. Appl.*, vol. 7, n° 1, p. 38, jul. 2018, doi: 10.1038/s41377-018-0035-8.
- [25] S. Iezekiel, “Microwave Photonics – an Introductory Overview”, em *Microwave Photonics*, 1° ed, S. Iezekiel, Org., Wiley, 2009, p. 1–37. doi: 10.1002/9780470744857.ch1.
- [26] T. Kim, “Analysis of Optical Communications, Fiber Optics, Sensors and Laser Applications”, *J. Mach. Comput.*, p. 115–125, abr. 2023, doi: 10.53759/7669/jmc202303012.
- [27] J. Jiang, M. Chen, e J. A. Fan, “Deep neural networks for the evaluation and design of photonic devices”, *Nat. Rev. Mater.*, vol. 6, n° 8, p. 679–700, dez. 2020, doi: 10.1038/s41578-020-00260-1.
- [28] K. Hamamoto e H. Jiang, “Active MMI devices: concept, proof, and recent progress”, *J. Phys. Appl. Phys.*, vol. 48, n° 38, p. 383001, set. 2015, doi: 10.1088/0022-3727/48/38/383001.
- [29] A. Taha *et al.*, “Compact MMI-Based AWGs in a Scalable Monolithic Silicon Photonics Platform”, *IEEE Photonics J.*, vol. 13, n° 4, p. 1–6, ago. 2021, doi: 10.1109/JPHOT.2021.3099436.
- [30] N. P. Dipta, R. W. Purnamaningsih, M. R. Gumelar, e S. E. Rijal, “Design of 2 x 2 Multimode Interference Beam Splitter Based on GaN/Si Structure”, em *2020 Asia Conference on Computers and Communications (ACCC)*, Singapore, Singapore: IEEE, set. 2020, p. 29–33. doi: 10.1109/ACCC51160.2020.9347900.
- [31] H. Xie *et al.*, “An Ultra-Compact 3-dB Power Splitter for Three Modes Based on Pixelated Meta-Structure”, *IEEE Photonics Technol. Lett.*, vol. 32, n° 6, p. 341–344, mar. 2020, doi: 10.1109/LPT.2020.2975128.
- [32] A. Hosseini, D. N. Kwong, Y. Zhang, H. Subbaraman, X. Xu, e R. T. Chen, “1× N Multimode Interference Beam Splitter Design Techniques for On-Chip Optical Interconnections”, *IEEE J. Sel. Top. Quantum Electron.*, vol. 17, n° 3, p. 510–515, maio 2011, doi: 10.1109/JSTQE.2010.2099210.
- [33] K. Kojima *et al.*, “Deep Neural Networks for Inverse Design of Nanophotonic Devices”, *J. Light. Technol.*, vol. 39, n° 4, p. 1010–1019, fev. 2021, doi: 10.1109/JLT.2021.3050083.
- [34] L. T. Rahman *et al.*, “A 1.55 μm Wideband 1 × 2 Photonic Power Splitter With Arbitrary Ratio: Characterization and Forward Modeling”, *IEEE Access*, vol. 10, p. 20149–20158, 2022, doi: 10.1109/ACCESS.2022.3151722.
- [35] J. Wang, L. Huangfu, e H. Chen, “Design of compact polarization-insensitive multimode interference triplexer”, *J. Mod. Opt.*, vol. 68, n° 9, p. 496–506, maio 2021, doi: 10.1080/09500340.2021.1924887.

- [36] A. M. Gabr, C. Featherston, C. Zhang, C. Bonfil, Q.-J. Zhang, e T. J. Smy, “Design and optimization of optical passive elements using artificial neural networks”, *J. Opt. Soc. Am. B*, vol. 36, n° 4, p. 999, abr. 2019, doi: 10.1364/JOSAB.36.000999.
- [37] S. Kim, N. Kim, I. Park, e H. Han, “Ultra-compact terahertz 50:50 power splitter designed by a perceptron-based algorithm”, *Opt. Contin.*, vol. 1, n° 7, p. 1565, jul. 2022, doi: 10.1364/OPTCON.464459.
- [38] S. Banerji, A. Majumder, A. Hamrick, R. Menon, e B. Sensale-Rodriguez, “Ultra-compact integrated photonic devices enabled by machine learning and digital metamaterials”, *OSA Contin.*, vol. 4, n° 2, p. 602, fev. 2021, doi: 10.1364/OSAC.417729.
- [39] Y. Ren *et al.*, “Genetic-algorithm-based deep neural networks for highly efficient photonic device design”, *Photonics Res.*, vol. 9, n° 6, p. B247, jun. 2021, doi: 10.1364/PRJ.416294.
- [40] H. Ma *et al.*, “Different-mode power splitters based on a multi-dimension direct-binary-search algorithm”, *Opt. Express*, vol. 31, n° 17, p. 27393, ago. 2023, doi: 10.1364/OE.495918.
- [41] A. Sun *et al.*, “Inverse design of an ultra-compact dual-band wavelength demultiplexing power splitter with detailed analysis of hyperparameters”, *Opt. Express*, vol. 31, n° 16, p. 25415, jul. 2023, doi: 10.1364/OE.493866.
- [42] Y. Zhang *et al.*, “A compact and low loss Y-junction for submicron silicon waveguide”, *Opt. Express*, vol. 21, n° 1, p. 1310, jan. 2013, doi: 10.1364/OE.21.001310.
- [43] C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, e E. Yablonovitch, “Adjoint shape optimization applied to electromagnetic design”, *Opt. Express*, vol. 21, n° 18, p. 21693, set. 2013, doi: 10.1364/OE.21.021693.
- [44] W. Chang *et al.*, “Inverse design and demonstration of an ultracompact broadband dual-mode 3 dB power splitter”, *Opt. Express*, vol. 26, n° 18, p. 24135, set. 2018, doi: 10.1364/OE.26.024135.
- [45] K. Xu *et al.*, “Integrated photonic power divider with arbitrary power ratios”, *Opt. Lett.*, vol. 42, n° 4, p. 855, fev. 2017, doi: 10.1364/OL.42.000855.
- [46] Y. H. Isayama, “Design and numerical demonstration of a multimode interference sensor using engineered slot-waveguides”, *Opt. Quantum Electron.*, vol. 56, n° 3, p. 304, mar. 2024, doi: 10.1007/s11082-023-05963-7.
- [47] J. Wen *et al.*, “Inverse design of high efficiency and large bandwidth power splitter for arbitrary power ratio based on deep residual network”, *Opt. Quantum Electron.*, vol. 56, n° 4, p. 512, jan. 2024, doi: 10.1007/s11082-023-06165-x.
- [48] D. A. May-Arrijoja, N. Bickel, e P. Likamwa, “Robust 2×2 Multimode Interference Optical Switch”, *Opt. Quantum Electron.*, vol. 38, n° 7, p. 557–566, maio 2006, doi: 10.1007/s11082-005-4699-y.

- [49] D. Chack e S. Hassan, “Design and experimental analysis of multimode interference-based optical splitter for on-chip optical interconnects”, *Opt. Eng.*, vol. 59, n° 10, out. 2020, doi: 10.1117/1.OE.59.10.105102.
- [50] Z. Deng, Z. Dang, T. Chen, Z. Ding, e Z. Zhang, “3-Port beam splitter of arbitrary power ratio enabled by deep learning on a multimode waveguide”, *Opt. Laser Technol.*, vol. 169, p. 109950, fev. 2024, doi: 10.1016/j.optlastec.2023.109950.
- [51] R. Yin, J. Teng, e S. Chua, “A 1×2 optical switch using one multimode interference region”, *Opt. Commun.*, vol. 281, n° 18, p. 4616–4618, set. 2008, doi: 10.1016/j.optcom.2008.05.042.
- [52] H. Yang, P. Zheng, P. Liu, G. Hu, B. Yun, e Y. Cui, “Design of polarization-insensitive 2 × 2 multimode interference coupler based on double strip silicon nitride waveguides”, *Opt. Commun.*, vol. 410, p. 559–564, mar. 2018, doi: 10.1016/j.optcom.2017.11.003.
- [53] X. Jia, S. Luo, e X. Cheng, “Design and optimization of novel ultra-compact SOI multimode interference optical switch”, *Opt. Commun.*, vol. 281, n° 5, p. 1003–1007, mar. 2008, doi: 10.1016/j.optcom.2007.10.074.
- [54] M. H. Tahersima *et al.*, “Deep Neural Network Inverse Design of Integrated Photonic Power Splitters”, *Sci. Rep.*, vol. 9, n° 1, p. 1368, fev. 2019, doi: 10.1038/s41598-018-37952-2.
- [55] A. L. C. Ottoni, “Métodos para Recomendação de Hiperparâmetros de Aprendizado de Máquina na Classificação de Imagens da Construção Civil”, Doctor’s Thesis, Universidade Federal da Bahia, Departamento de Engenharia Elétrica, Salvador, Bahia, 2022.
- [56] I. Newton, G. W. Hemming, e Burndy Library, *Opticks: or, A treatise of the reflections, refractions, inflexions and colours of light*. London: Printed for Sam. Smith, and Benj. Walford, 1704. doi: 10.5479/sil.302475.39088000644674.
- [57] C. Huygens, *Treatise on Light*. Leyden: Books on Demand, 1690.
- [58] A. J. Fresnel e J. L. Gay_Lussac, *Annales de Chimie et de Physique*. Paris: Chez Crochard, 1821. Acesso em: 12 de abril de 2024. [Online]. Disponível em: <https://archive.org/details/s3id13207870/page/90/mode/2up>
- [59] J. C. Maxwell, *A treatise on electricity & [and] magnetism*, 3rd ed. New York: Dover, 1873.
- [60] J. C. Maxwell, “VIII. A dynamical theory of the electromagnetic field”, *Philos. Trans. R. Soc. Lond.*, vol. 155, p. 459–512, dez. 1865, doi: 10.1098/rstl.1865.0008.
- [61] J. C. Maxwell, Org., “On Physical Lines of Force”, em *The Scientific Papers of James Clerk Maxwell*, 1° ed, Cambridge University Press, 1961, p. 451–513. doi: 10.1017/CBO9780511698095.026.
- [62] F. L. Pedrotti, L. M. Pedrotti, e L. S. Pedrotti, *Introduction to Optics*., 3° ed. Cambridge University Press, 2017. doi: 10.1017/9781108552493.

- [63] G. Lifante, *Integrated photonics: fundamentals*. Hoboken, NJ: J. Wiley, 2003.
- [64] I. Amiri *et al.*, “Introduction to Photonics: Principles and the Most Recent Applications of Microstructures”, *Micromachines*, vol. 9, n° 9, p. 452, set. 2018, doi: 10.3390/mi9090452.
- [65] C. Yang *et al.*, “Portable optical fiber biosensors integrated with smartphone: technologies, applications, and challenges [Invited]”, *Biomed. Opt. Express*, vol. 15, n° 3, p. 1630, mar. 2024, doi: 10.1364/BOE.517534.
- [66] I. S. Panyaev, P. A. Itrin, D. A. Korobko, e A. A. Fotiadi, “Sub-100-Hz DFB Laser Injection-Locked to PM Fiber Ring Cavity”, *J. Light. Technol.*, vol. 42, n° 8, p. 2928–2937, abr. 2024, doi: 10.1109/JLT.2023.3348994.
- [67] M. Protte, T. Schapeler, J. Sperling, e T. J. Bartley, “Low-noise balanced homodyne detection with superconducting nanowire single-photon detectors”, *Opt. Quantum*, vol. 2, n° 1, p. 1, fev. 2024, doi: 10.1364/OPTICAQ.502201.
- [68] H.-H. Lu, N. A. Peters, A. M. Weiner, e J. M. Lukens, “Characterization of Quantum Frequency Processors”, *IEEE J. Sel. Top. Quantum Electron.*, vol. 29, n° 6: Photonic Signal Processing, p. 1–12, nov. 2023, doi: 10.1109/JSTQE.2023.3266662.
- [69] M. Murray, I. Naydenova, e S. Martin, “Review of recent advances in photosensitive polymer materials and requirements for transmission diffractive optical elements for LED light sources”, *Opt. Mater. Express*, vol. 13, n° 12, p. 3481, dez. 2023, doi: 10.1364/OME.502234.
- [70] Z. Zhang *et al.*, “Entanglement-based quantum information technology: a tutorial”, *Adv. Opt. Photonics*, vol. 16, n° 1, p. 60, mar. 2024, doi: 10.1364/AOP.497143.
- [71] J. Peltier *et al.*, “High-speed silicon photonic electro-optic Kerr modulation”, *Photonics Res.*, vol. 12, n° 1, p. 51, jan. 2024, doi: 10.1364/PRJ.488867.
- [72] G. Zhang *et al.*, “Improving optoelectronic properties of InP/InAs nanowire p-i-n devices with telecom-band electroluminescence”, *Opt. Contin.*, vol. 3, n° 2, p. 176, fev. 2024, doi: 10.1364/OPTCON.511645.
- [73] K. Zhang, Y. He, K. Yu, e Y. Liu, “Fast data processing method for multispectral radiation thermometry based on Euclidean distance optimization”, *Opt. Express*, vol. 32, n° 2, p. 1342, jan. 2024, doi: 10.1364/OE.510084.
- [74] J. Wang, J. Huang, Q. Liu, W. Du, F. Zhang, e T. Zhu, “Linear group delay spectral interferometry for full-range precision absolute length metrology”, *Photonics Res.*, vol. 12, n° 2, p. 313, fev. 2024, doi: 10.1364/PRJ.506474.
- [75] H. Lou *et al.*, “High-SNR mid-infrared dual-comb spectroscopy using active phase control cooperating with CWs-dependent phase correction”, *Opt. Express*, vol. 32, n° 4, p. 5826, fev. 2024, doi: 10.1364/OE.514809.

- [76] L.-W. Chen, S.-Y. Lu, F.-C. Hsu, C.-Y. Lin, A.-S. Chiang, e S.-J. Chen, “Deep-computer-generated holography with temporal-focusing and a digital propagation matrix for rapid 3D multiphoton stimulation”, *Opt. Express*, vol. 32, n° 2, p. 2321, jan. 2024, doi: 10.1364/OE.505956.
- [77] C.-L. Lai *et al.*, “Lung cancer cells detection by a photoelectrochemical MoS₂ biosensing chip”, *Biomed. Opt. Express*, vol. 15, n° 2, p. 753, fev. 2024, doi: 10.1364/BOE.511900.
- [78] L. Kulhánek, “Laser Weapons as a New Entity to Small Firearms”, *Vojen. Rozhl.*, vol. 32, n° 3, p. 048–062, set. 2023, doi: 10.3849/2336-2995.32.2023.03.048-062.
- [79] V. P. Veiko *et al.*, “Laser paintbrush as a tool for modern art”, *Optica*, vol. 8, n° 5, p. 577, maio 2021, doi: 10.1364/OPTICA.420074.
- [80] F. Chen *et al.*, “Pulsed-lighting LED luminaire for agriculture with a geometrical optical solution”, *Opt. Express*, vol. 31, n° 4, p. 5609, fev. 2023, doi: 10.1364/OE.483071.
- [81] Md. E. Karim e A. S. M. Mohsin, “Metasurface absorber based single junction thin film solar cell exceeding 30% efficiency”, *Opt. Express*, vol. 32, n° 5, p. 8214, fev. 2024, doi: 10.1364/OE.510421.
- [82] H. Zhang *et al.*, “Omnidirectional optical communication system designed for underwater swarm robotics”, *Opt. Express*, vol. 31, n° 11, p. 18630, maio 2023, doi: 10.1364/OE.490076.
- [83] I. Meglinski, A. Dunn, T. Durduran, D. Postnov, e D. Zhu, “Dynamic Light Scattering in Biomedical Applications: feature issue introduction”, *Biomed. Opt. Express*, vol. 15, n° 5, p. 2890, maio 2024, doi: 10.1364/BOE.525699.
- [84] K. Liu e J. Zhang, “Glaucoma detection model by exploiting multi-region and multi-scan-pattern OCT images with dynamical region score”, *Biomed. Opt. Express*, vol. 15, n° 3, p. 1370, mar. 2024, doi: 10.1364/BOE.512138.
- [85] B. Culshaw, *Introducing photonics*. Cambridge; New York, NY: Cambridge University Press, 2020.
- [86] A. W. Snyder e J. D. Love, *Optical Waveguide Theory*. Boston, MA: Springer US, 1984. doi: 10.1007/978-1-4613-2813-1.
- [87] P. Lima, “Análise e Projeto de Acopladores de Guia de Onda Plasmônico”, Master’s Thesis, Universidade Federal da Bahia, Departamento de Engenharia Elétrica, Salvador, Bahia, 2018.
- [88] K. Okamoto, *Fundamentals of Optical Waveguides*, Second Edition. Okamoto Laboratory Ltd Ibaraki, Japan: Academic Press, 2006.
- [89] D. L. Lee, *Electromagnetic principles of integrated optics*. New York Chichester Brisbane [etc.]: J. Wiley & sons, 1986.
- [90] K. Cooney e F. H. Peters, “Analysis of multimode interferometers”, *Opt. Express*, vol. 24, n° 20, p. 22481, out. 2016, doi: 10.1364/OE.24.022481.

- [91] K. Okamoto, “Beam Propagation Method”, em *Fundamentals of Optical Waveguides*, Elsevier, 2022, p. 339–407. doi: 10.1016/B978-0-12-815601-8.50007-0.
- [92] I. H. Malitson, “Interspecimen Comparison of the Refractive Index of Fused Silica*,†”, *J. Opt. Soc. Am.*, vol. 55, n° 10, p. 1205, out. 1965, doi: 10.1364/JOSA.55.001205.
- [93] C. D. Salzberg e J. J. Villa, “Infrared Refractive Indexes of Silicon Germanium and Modified Selenium Glass*”, *J. Opt. Soc. Am.*, vol. 47, n° 3, p. 244, mar. 1957, doi: 10.1364/JOSA.47.000244.
- [94] S. N. Ajani, P. Khobragade, M. Dhone, D. B. Ganguly, N. Shelke, e N. Parati, “Advancements in Computing: Emerging Trends in Computational Science with Next-Generation Computing”, *Int. J. Intell. Syst. Appl. Eng.*, dez. 2023.
- [95] J. Haugeland, *Artificial Intelligence: The Very Idea*, Cambridge, Massachusetts, p. 86–87, 1985.
- [96] R. Kurzweil, *The age of spiritual machines: when computers exceed human intelligence*. New York: Viking, 1990.
- [97] E. Charniak e R. P. Goldman, “A Bayesian model of plan recognition”, *Artif. Intell.*, vol. 64, n° 1, p. 53–79, nov. 1993, doi: 10.1016/0004-3702(93)90060-O.
- [98] D. L. Poole, A. K. Mackworth, e R. Goebel, *Computational intelligence: a logical approach*. New York: Oxford University Press, 1998.
- [99] A. M. Turing, “Computing Machinery and Intelligence”, *Mind New Ser.*, vol. 59, n° 236, p. 433–460, 1950.
- [100] P. Norvig, *Inteligência Artificial*, 4° ed. Rio de Janeiro, RJ: Grupo Gen, 2022.
- [101] J. Bajorath, “Specific contributions of artificial intelligence to interdisciplinary life science research – exploring and communicating new opportunities”, *Artif. Intell. Life Sci.*, vol. 3, p. 100052, dez. 2023, doi: 10.1016/j.aillsci.2022.100052.
- [102] M. Regona, T. Yigitcanlar, B. Xia, e R. Y. M. Li, “Opportunities and Adoption Challenges of AI in the Construction Industry: A PRISMA Review”, *J. Open Innov. Technol. Mark. Complex.*, vol. 8, n° 1, p. 45, mar. 2022, doi: 10.3390/joitmc8010045.
- [103] K. Sharifani e M. Amini, “Machine Learning and Deep Learning: A Review of Methods and Applications”, *World Information Technology and Engineering Journal*, vol. 10, n° 07, p. 3897–3904, 2023.
- [104] S. M. Weiss e C. A. Kulikowski, *Computer systems that learn: classification and prediction methods from statistics, neural nets, machine learning, and expert systems*. San Mateo, Calif: M. Kaufmann Publishers, 1991.
- [105] Y. LeCun, Y. Bengio, e G. Hinton, “Deep learning”, *Nature*, vol. 521, n° 7553, p. 436–444, maio 2015, doi: 10.1038/nature14539.

- [106] M. C. Monard e J. A. Baranauskas, *Sistemas Inteligentes: Fundamentos e Aplicações*. Barueri, São Paulo: Editora Manole, 2005.
- [107] L. Deng, “A tutorial survey of architectures, algorithms, and applications for deep learning”, *APSIPA Trans. Signal Inf. Process.*, vol. 3, n° 1, 2014, doi: 10.1017/atsip.2013.9.
- [108] M. M. Taye, “Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions”, *Computers*, vol. 12, n° 5, p. 91, abr. 2023, doi: 10.3390/computers12050091.
- [109] P. Cheeseman e J. Stutz, *Bayesian classification (AutoClass): theory and results*, 5. print., vol. 17. em *Advances in knowledge discovery and data mining*, no. 3, vol. 17. American Association for Artificial Intelligence, California, United States: AAAI Press [u.a.], 1996.
- [110] G. Bonaccorso, *Machine Learning Algorithms - Second Edition*, 2nd edition. Packt Publishing, 2018. [Online]. Disponível em: https://books.google.com.br/books?hl=pt-BR&lr=&id=_EZsDwAAQBAJ&oi=fnd&pg=PP1&dq=machine+learning+algorithms+article&ots=cJP15Xe87G&sig=9-xd_fxiiTEWOJeoUQjgvx0Sazg#v=onepage&q=machine%20learning%20algorithms%20article&f=false
- [111] I. H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions”, *SN Comput. Sci.*, vol. 2, n° 3, p. 160, maio 2021, doi: 10.1007/s42979-021-00592-x.
- [112] M. Mohammed, M. B. Khan, E. B. M. Bashier, e M. B. Khan, *Machine learning: algorithms and applications*. Boca Raton London New York: CRC Press, Taylor & Francis Group, 2017.
- [113] A. McCallum e K. Nigam, “A Comparison of Event Models for Naive Bayes Text Classification”, em *Proceedings in Workshop on Learning for Text Categorization*, 1998, p. 41–48.
- [114] E. Fix e J. Hodges, *Discriminatory analysis: nonparametric discrimination, consistency properties*. em USAF school of Aviation Medicine. University of California, Berkeley, 1951.
- [115] V. N. Vapnik, “The Support Vector method”, *Artif. Neural Netw. — ICANN97*, p. 261–271, 1997.
- [116] Tin Kam Ho, “Random decision forests”, em *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Montreal, Que., Canada: IEEE Comput. Soc. Press, 1995, p. 278–282. doi: 10.1109/ICDAR.1995.598994.
- [117] R. A. Fisher, “THE USE OF MULTIPLE MEASUREMENTS IN TAXONOMIC PROBLEMS”, *Ann. Eugen.*, vol. 7, n° 2, p. 179–188, set. 1936, doi: 10.1111/j.1469-1809.1936.tb02137.x.
- [118] J. R. Quinlan, “Induction of decision trees”, *Mach. Learn.*, vol. 1, n° 1, p. 81–106, mar. 1986, doi: 10.1007/BF00116251.

- [119] H. Yi, A. K. Goyal, e Y. Massoud, “Performance optimization of Bloch surface wave based devices using a XGBoost machine learning model”, *Opt. Contin.*, vol. 3, n° 5, p. 693, maio 2024, doi: 10.1364/OPTCON.520638.
- [120] Y. Freund e R. E. Schapire, “A Short Introduction to Boosting”, *Journal of Japanese Society for Artificial Intelligence*, vol. 14, p. 771–780, set. 1999, doi: https://doi.org/10.11517/jjsai.14.5_771.
- [121] R. E. Schapire e Y. Singer, “Improved boosting algorithms using confidence-rated predictions”, em *Proceedings of the eleventh annual conference on Computational learning theory*, Madison Wisconsin USA: ACM, jul. 1998, p. 80–91. doi: 10.1145/279943.279960.
- [122] S. M. Stigler, *The history of statistics: the measurement of uncertainty before 1900*, 9. print. Cambridge, Mass.: Belknap Pr. of Harvard Univ. Pr, 2003. [Online]. Disponível em: <https://archive.org/details/historyofstatist00stig/page/n15/mode/2up>
- [123] P. F. Verhulst, *Recherches Mathématiques Sur La Loi d’Accroissement de La Population*. Belgium, 1844. [Online]. Disponível em: https://gdz.sub.uni-goettingen.de/download/pdf/PPN129323640_0018/PPN129323640_0018.pdf
- [124] I. N. da Silva, *Redes neurais artificiais para engenharia e ciências aplicadas*. Artliber, 2014.
- [125] A. D. P. Braga, *Redes Neurais Artificiais: Teoria e Aplicações*. Ltc, 1999.
- [126] S. S. Haykin, *Neural networks: a comprehensive foundation*, 2nd ed. Upper Saddle River, N.J: Prentice Hall, 1999. Acesso em: 29 de abril de 2024. [Online]. Disponível em: https://books.google.com.br/books?id=bhMwDwAAQBAJ&printsec=frontcover&hl=pt-BR&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false
- [127] W. S. McCulloch e W. Pitts, “A logical calculus of the ideas immanent in nervous activity”, *Bull. Math. Biophys.*, vol. 5, n° 4, p. 115–133, dez. 1943, doi: 10.1007/BF02478259.
- [128] A. D. Sisnando, “Análise e síntese de dispositivos fotônicos utilizando redes neurais artificiais”, Master’s Thesis, Universidade Federal da Bahia, Departamento de Engenharia Elétrica, Salvador, Bahia, 2011.
- [129] D. E. Rumelhart, G. E. Hinton, e R. J. Williams, “Learning representations by back-propagating errors”, *Nature*, vol. 323, n° 6088, p. 533–536, out. 1986, doi: 10.1038/323533a0.
- [130] E. Ghafari, M. Bandarabadi, H. Costa, e E. Júlio, “Design of Ultra-high performance concrete using artificial neural networks”, em *Brittle Matrix Composites 10*, Elsevier, 2012, p. 61–69. doi: 10.1533/9780857099891.61.
- [131] D. W. Marquardt, “An Algorithm for Least-Squares Estimation of Nonlinear Parameters”, *J. Soc. Ind. Appl. Math.*, vol. 11, n° 2, p. 431–441, jun. 1963, doi: 10.1137/0111030.

- [132] *BeamPROP Product Overview*. em The Source for Photonic and Network Design Software. RSoft Design Group, Inc. [Online]. Disponível em: <https://www.synopsys.com/photonic-solutions/rsoft-photonic-device-tools/passive-device-beamprop.html>
- [133] F. Chollet, *Deep learning with Python*. Shelter Island, New York: Manning Publications Co, 2018.
- [134] R. Das Mercês Cerqueira, A. Dourado Sisnando, e V. F. Rodriguez-Esquerre, “Metamaterial waveguide modeling by a genetic algorithm-based artificial neural network”, em *Integrated Optics: Devices, Materials, and Technologies XXVI*, S. M. García-Blanco e P. Cheben, Orgs., San Francisco, United States: SPIE, mar. 2022, p. 71. doi: 10.1117/12.2612161.