



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
UNIDADE ACADÊMICA DO CABO DE SANTO AGOSTINHO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA FÍSICA

ALLAN CHRISTOPHER RODRIGUES BARBOZA

Estudo e desenvolvimento de classificadores ópticos de imagens baseados no
algoritmo de Máquina de Aprendizado Extremo.

Cabo de Santo Agostinho

2024

ALLAN CHRISTOPHER RODRIGUES BARBOZA

Estudo e desenvolvimento de classificadores ópticos de imagens baseados no algoritmo de Máquina de Aprendizado Extremo.

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Física, Unidade Acadêmica do Cabo de Santo Agostinho, da Universidade Federal Rural de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Engenharia Física.

Área de Concentração: Optoeletrônica

Linha de pesquisa: Espectroscopia coerente e óptica não linear

Orientador(a): Prof^a. Dr^a. Martine Patricia Arlette Chevrollier

Coorientador(a): Prof^a. Dr^a. Amanda de Paula

Cabo de Santo Agostinho

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

- B239e Rodrigues Barboza, Allan Christopher
Estudo e desenvolvimento de classificadores ópticos de imagens baseados no algoritmo de Máquina de Aprendizado Extremo. / Allan Christopher Rodrigues Barboza. - 2023.
137 f. : il.
- Orientador: Dr^a. Martine Patricia Arlette Chevrollier.
Coorientador: Dr^a. Amanda de Paula.
Inclui referências e apêndice(s).
- Dissertação (Mestrado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Engenharia Física, Cabo de Santo Agostinho, 2024.
1. Classificação de imagens. 2. Sistemas ópticos. 3. Máquina de Aprendizado Extremo. I. Chevrollier, Dr^a. Martine Patricia Arlette, orient. II. Paula, Dr^a. Amanda de, coorient. III. TÃ-tulo

ALLAN CHRISTOPHER RODRIGUES BARBOZA

Estudo e desenvolvimento de classificadores ópticos de imagens baseados no algoritmo de Máquina de Aprendizado Extremo.

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Física da Unidade Acadêmica do Cabo de Santo Agostinho da Universidade Federal Rural de Pernambuco para obtenção do título de Mestre em Engenharia Física.

APROVADA EM: 30/11/2023

Banca Examinadora

Prof^ª. Dr^ª. Martine Patricia Arlette Chevrollier – UFRPE
Orientador

Prof. Dr. João Carlos de Aquino Carvalho- UFPE
Examinador externo

Prof. Dr. Marcos César Santos Oriá - UFRPE
Examinador interno

Dedico este trabalho às três mulheres incríveis da minha família que moldaram a minha vida e me ensinaram a importância da educação.

AGRADECIMENTOS

Gostaria de expressar minha sincera gratidão a todas as pessoas que, ao longo de minha jornada acadêmica, me guiaram e acreditaram em mim, mesmo quando eu próprio duvidava. Esta dissertação representa o resultado de meses de pesquisa, esforço e dedicação, e não teria sido possível sem o apoio de inúmeras pessoas, mas, acima de tudo, graças à minha fé e persistência.

Agradeço profundamente à minha família, em especial à minha mãe Alcione, à minha tia Cilene e à minha avó Ivonice, pelo apoio incondicional ao longo de toda minha vida.

À minha orientadora, a professora Martine, agradeço pela confiança que depositou em mim, por sua paciência e, sobretudo, pelos ensinamentos e conselhos que carregarei sempre comigo. À minha coorientadora, a professora Amanda de Paula, agradeço por sua dedicação constante e pelos valiosos ensinamentos.

Aos professores do PPENGFIS, em especial ao professor Marcos Oriá, agradeço pelas inúmeras orientações, conselhos e conhecimentos transmitidos. Muito obrigado por tudo o que fez por mim, serei eternamente grato.

Aos professores Sérgio e Weliton, agradeço pelas experiências enriquecedoras em suas disciplinas e pelos conhecimentos de laboratório que me ajudaram a evoluir.

Aos meus amigos, que desempenharam um papel essencial ao longo do meu mestrado. À minha querida amiga Larissa Silva, que compartilho doze anos de amizade, agradeço pelo incentivo constante e pela troca de conhecimento. Aos amigos que fiz no mestrado, Brenda, Silvana, Wibson, Stephany, Luana e Lucas. E em especial, Flávia por sua amizade e pelas vivências no laboratório, aos companheiros Gleison, Alexandro, Gilderlan e Andielisson, expresso minha profunda gratidão por tudo o que fizeram por mim ao longo desses anos de convivência. A todos vocês agradeço pelas discussões enriquecedoras, pela irmandade que compartilhamos e, acima de tudo, pelas risadas que tornaram essa jornada mais leve.

E não posso deixar de expressar minha gratidão ao meu amigo Mateus Felipe que sempre acreditou em mim e sempre esteve ao meu lado em todos os momentos e aos amigos do Clube de Astronomia Vega por todo aprendizado ao longo dos anos.

Por fim, agradeço a FACEPE pelo apoio financeiro.

RESUMO

A busca constante por processamento rápido de informações é um objetivo fundamental, sobretudo no âmbito das áreas científicas, com destaque para o campo da computação. Com o avanço tecnológico e o desenvolvimento de unidades de processamento mais sofisticadas, como as GPUs (Unidades de Processamento Gráfico), tornou-se cada vez mais viável lidar com tarefas computacionais intensivas de aprendizado de máquina, impulsionadas por grandes conjuntos de dados. No entanto, esses processadores ainda enfrentam desafios, como consumo de energia e transferência de dados para a memória, que limitam seu desempenho nessas tarefas. Por outro lado, a óptica tem emergido como uma ferramenta poderosa para comunicação e processamento de informações. Há um interesse crescente em utilizar a óptica para realizar cálculos em alta velocidade. Nesta dissertação, exploramos o desenvolvimento de sistemas ópticos experimentais capazes de executar tarefas de classificação de imagens com altas taxas de acurácia. Para isso, investigamos o algoritmo de treinamento de Máquinas de Aprendizado Extremo (Extreme Learning Machine - ELM), um algoritmo de aprendizado de máquina que tem sido pouco explorado na literatura em aplicações ópticas. Através desse algoritmo, obtivemos parâmetros otimizados que foram aplicados em nossos sistemas experimentais. Inicialmente, examinamos um sistema óptico que utiliza um meio espalhador como parte do processamento óptico, resultando na geração de padrões de speckles com estatísticas de intensidade de Rayleigh. Posteriormente, exploramos outra configuração óptica, onde utilizamos técnicas de holografia gerada por computador para criar um sistema de classificação de imagens que realiza a mesma etapa de processamento, porém gera speckles com estatísticas de intensidade não Rayleigh. Comparamos os resultados dos dois sistemas e analisamos o impacto dessa modificação na precisão da classificação de imagens. Observamos que, para esses sistemas, ainda é necessária a utilização de um computador na etapa de classificação. Com base no mesmo algoritmo de treinamento, desenvolvemos duas técnicas de classificação, nas quais obtivemos um discriminador que pode ser facilmente aplicado em sistemas de classificação totalmente ópticos. Avaliamos o desempenho dos sistemas, classificando imagens de dígitos manuscritos dos bancos de dados MNIST e imagens de moda do Fashion- MNIST.

Palavras-chaves: Classificação de imagens, Sistemas ópticos, Máquina de Aprendizado Extremo.

ABSTRACT

The constant pursuit of fast information processing is a fundamental objective, particularly in the field of scientific areas, with a focus on the field of computing. With technological advancement and the development of more sophisticated processing units, such as GPUs (Graphics Processing Units), it has become increasingly feasible to handle computationally intensive machine learning tasks, driven by large datasets. However, these processors still face challenges, such as energy consumption and data transfer to memory, which limit their performance in these tasks. On the other hand, optics has emerged as a powerful tool for communication and information processing. There is a growing interest in using optics to achieve high-speed calculations. In this dissertation, we explore the development of experimental optical systems capable of performing image classification tasks with high accuracy rates. To do this, we investigate the training algorithm of Extreme Learning Machines (ELM), a machine learning algorithm that has been underexplored in the literature in optical applications. Through this algorithm, we obtained optimized parameters that were applied to our experimental systems. Initially, we examine an optical system that uses a scattering medium as part of the optical processing, resulting in the generation of speckle patterns with Rayleigh statistics. Subsequently, we explore another optical configuration, where we use computer-generated holography techniques to create an image classification system that performs the same processing step but generates speckles with non-Rayleigh statistics. We compare the results of the two systems and analyze the impact of this modification on image classification accuracy. We observed that, for these systems, the use of a computer in the classification step is still necessary. Based on the same training algorithm, we developed two classification techniques in which we obtained a discriminator that can be easily applied to fully optical classification systems. We evaluated the performance of the systems by classifying handwritten digit images from the MNIST database and fashion images from Fashion MNIST.

Keywords: Image classification, Optical systems, Extreme Learning Machine.

LISTA DE FIGURAS

1	Fluxograma do aprendizado supervisionado	26
2	Representação de um problema de a) Classificação, b) Regressão .	26
3	Neurônios a) biológicos b) artificial	28
4	(a) Classes de padrões linearmente separáveis. (b) Classes de padrões não linearmente separáveis	29
5	Arquitetura de uma rede neural profunda com duas camadas ocultas.	30
6	Gradiente descendente	32
7	Arquitetura típica de um <i>Reservoir Computing</i>	33
8	Arquitetura de uma Máquina de Aprendizado Extremo	34
9	Multiplicação óptica de matrizes, a) Multiplicação de amplitude, b) multiplicação vetor-matriz c) multiplicação matriz-matriz utilizando um sistema óptico 4f.	39
10	Redes Neurais difrativa. a) multicamadas difrativas b) camada única.	41
11	Transmissão da luz através de uma abertura	42
12	Máquina óptica de aprendizado extremo	45
13	Processo de formação de padrões de speckles no espaço livre a) por superfícies rugosas transmissivas, b) por superfícies rugosas reflexivas.	47
14	a) Imagem da distribuição de intensidades de um padrão de speckle criado por uma superfície rugosa, b) Perfil de intensidade ao longo de uma linha horizontal no centro da imagem.	48
15	Processo de formação de um padrão de speckle.	50
16	Gráfico da distribuição de densidade de probabilidade da intensidade de um padrão de speckle.	54

17	Resultados experimentais para speckle super Rayleigh a) Simplificação do setup experimental, b) Speckle Rayleigh com $C = 0,98$, c) Speckle Super Rayleigh com $C = 1,28$, d) Speckle Super Rayleigh com $C = 2,79$, seguido da curva de densidade de probabilidade de cada padrão, os triangulos azuis referem-se aos speckle Rayleigh, losangos verdes aos speckles super Rayleigh de contraste $C = 1,28$ e os quadrados vermelhos aos speckle super Rayleigh de contraste $C = 2,79$	56
18	Resultados experimentais para speckle sub-Rayleigh a) Padrão de intensidade de speckle sub-Rayleigh com $C = 0,77$, b) Gráfico da densidade de probabilidade da intensidade, os losangos verde referem-se ao speckles sub-Rayleigh e os triângulos azuis aos speckles de Rayleigh.	57
19	Setup Experimental de um classificador óptico utilizando um difusor de luz como elemento espalhador.	61
20	Perfil espacial de um feixe gaussiano a) imagem capturada por uma câmera CCD b) perfil de intensidade do feixe.	63
21	Perfil espacial de um feixe Top hat a) imagem capturada por uma câmera CCD b) perfil de intensidade do feixe.	63
22	Dispositivo LCOS-SLM somente de fase, modelo X13267, desenvolvido pela Hamamatsu.	64
23	Exemplos do conjunto de dados do MNIST	65
24	Esquema do processamento óptico do sistema.	67
25	Esquema do processamento óptico do sistema.	68
26	Caracterização do padrão de speckle gerado pelo difusor. a) Imagem capturada pela CCD onde o padrão de luz no plano da imagem representa o dígito 0, b) Gráfico da densidade probabilidade da intensidade deste padrão.	69
27	Caracterização do padrão de speckle gerado pelo difusor. a) Imagem capturada pela CCD onde o padrão de luz no plano da imagem representa o dígito 6, b) Gráfico da densidade probabilidade da intensidade deste padrão.	70

28	Matriz de Confusão da classificação do sistema óptico que utiliza um meio espalhador.	71
29	Matrizes utilizada na simulação de padrões de speckle. a) Matriz de fase totalmente aleatória, b) Transformada de Fourier do campo de entrada, c) Filtro espacial numérico, d) Campo resultante após o filtro.	73
30	Padrões de speckles gerados numericamente. a) Módulo do campo complexo, b) Padrão de intensidade.	74
31	Caracterização dos padrões de speckle gerados numericamente. a) Distribuição de amplitude, b) Distribuição de intensidade em escala linear-linear, c) Distribuição de intensidade em escala mono-log. . .	74
32	Speckles modificados estatisticamente, a) $\alpha = 0, 2$, b) $\alpha = 0, 5$, c) $\alpha = 1$ d) $\alpha = 1, 5$, gráficos da densidade de probabilidade da intensidade e) $\alpha = 0, 2$, f) $\alpha = 0, 5$, g) $\alpha = 1$ h) $\alpha = 1, 5$	75
33	Setup experimental para classificar imagens a partir de speckles com estatísticas de intensidade não Rayleigh.	76
34	Speckles com estatística de intensidade modificada gerados experimentalmente. Imagens obtidas na CCD para speckle gerado com a) $\alpha = 0, 2$, b) $\alpha = 0, 5$, c) $\alpha = 1$, d) $\alpha = 1, 5$.. Gráficos da densidade de probabilidade e) $\alpha = 0, 2$, f) $\alpha = 0, 5$, g) $\alpha = 1$, h) $\alpha = 1, 5$	77
35	Matriz de confusão da classificação para projeções aleatórias geradas com $\alpha = 1$. Resultados a) Numérico b) Experimental.	78
36	Gráfico normalizado da função de densidade de probabilidade da intensidade para speckles super Rayleigh.	79
37	Gráfico normalizado da função de densidade de probabilidade da intensidade para speckles sub Rayleigh.	79
38	Gráfico da acurácia em função do contraste.	80
39	Matriz de confusão para speckles sub-Rayleigh com $C= 0.25$. Resultados a) Numérico b) Experimental	81
40	Matriz de alvos	82
41	Imagens associadas aos seus respectivos alvos. a) Dígitos manuscritos do MNIST b) Alvos relativo a cada dígito.	83

42	Esquema de predição óptica: após a obtenção da matriz β , ela é codificada em um holograma de fase junto com o dígito de teste por meio de uma multiplicação matricial. Em seguida, o holograma é enviado ao SLM, e através da modulação de fase, obtemos a predição óptica visualizada no monitor de um computador.	83
43	Predição óptica das imagens obtidas na CCD.	84
44	Matriz de confusão da predição óptica.	84
45	Exemplos de matrizes alvo relativa a problemas de duas classes. . .	87
46	Representação do processo de obtenção das matrizes $K(x_i)$ para a etapa de treinamento.	87
47	Diagrama do treinamento com as matrizes de entrada e alvos como matrizes quadradas, onde o objetivo em questão é obter os pesos β	88
48	Diagrama do treinamento com as matrizes de entrada e alvos como matrizes quadradas para resolver problemas de 10 classes.	89
49	Esquema da etapa de classificação, a) Descrição numérica b) Setup experimental.	89
50	Esquema da configuração experimental.	90
51	Resultado óptico experimental da classificação em duas classes. . .	91
52	Resultado óptico experimental da classificação em duas classes das 30 primeiras imagens.	91
53	Matriz de confusão do resultado óptico numérico e experimental da classificação em duas classes	92
54	Máscara de transmissão para classificação óptica de imagens em 10 classes	92
55	a) Imagem experimental obtida pela CCD para classificação óptica de imagens em 10 classes, b) Imagem da CCD com regiões de classificação definidas pelas linhas em vermelho, c) Soma óptica das intensidades.	93
56	Matriz de confusão do resultado óptico numérico e experimental da classificação para dez classes	93
57	Conjunto de dados de dígitos manuscritos do MNIST	95
58	Conjunto de dados de moda do Fashion MNIST	96
59	Descrição da nossa etapa de classificação	97

60	Dígitos MNIST duplicados.	98
61	Etapa de classificação com dígitos duplicados. A) Classificação da o dígito 1 e B) classificação do dígito 0.	99
62	Etapa de classificação com dígitos decuplicados	99
63	Etapa de classificação com imagem de moda decuplicadas	100
64	Reconstrução a partir do holograma de fase	103
65	Reconstrução a partir do holograma de fase	103
66	Erro quadrático médio em função do número de iterações do algo- ritmo IFTA	105
67	Setup Experimental	106
68	Máscaras, A) Dígitos Mnist 2 Classes, B) Dígitos Mnist 10 Classes, C) Imagens de Moda Fashion Mnist	107
69	Resultado do produto de Hadamard dígito 0- 2 Classes	108
70	Resultado do produto de Hadamard dígito 7- 10 Classes	108
71	Matriz de confusão numérica da classificação dos dígitos manuscri- tos do MNIST- 2 Classes	109
72	Matriz de confusão numérica da classificação dos dígitos manuscri- tos do MNIST- 10 Classes	109
73	Matriz de confusão numérica da classificação das imagens de moda do Fashion MNIST	110
74	Gráfico da Acurácia em função no Número de imagens de Teste do dígitos do MNIST	111
75	Imagem experimental seguida do gráfico de intensidade em função no numero de pixels	111
76	Matriz de confusão Experimental MNIST	112
77	Matriz de confusão Experimental FASHION MNIST	113
78	a) Esquema representativo para processo de registro do holograma como um feixe de referência fora do eixo, b) processo reconstrução do holograma com suas diferentes ordens de difração.	128
79	Fases dos Cristais Líquidos	134
80	Aquitetura de um LCOS-SLM	135
81	Representação de uma molécula de cristal líquido	135

82	Representação da modulação de frente de onda feito por um SLM de cristais líquidos.	136
----	------------------------------------------------------------------------------------------------	-----

LISTA DE TABELAS

1	Características do LCOS-SLM	64
2	Características da Câmera CCD	105

LISTA DE EQUAÇÕES

Equação (1)	30
Equação (2)	31
Equação (3)	31
Equação (4)	35
Equação (5)	35
Equação (6)	35
Equação (7)	36
Equação (8)	36
Equação (9)	36
Equação (10)	36
Equação (11)	37
Equação (12)	37
Equação (13)	37
Equação (14)	40
Equação (15)	40
Equação (16)	40
Equação (17)	42
Equação (18)	42
Equação (19)	42
Equação (20)	43
Equação (21)	43
Equação (22)	43
Equação (23)	43
Equação (24)	43
Equação (25)	44
Equação (26)	48
Equação (27)	49
Equação (28)	49
Equação (29)	49
Equação (30)	50
Equação (31)	50

Equação (32)	50
Equação (33)	51
Equação (34)	51
Equação (35)	51
Equação (36)	52
Equação (37)	52
Equação (38)	52
Equação (39)	52
Equação (40)	52
Equação (41)	52
Equação (42)	53
Equação (43)	53
Equação (44)	53
Equação (45)	53
Equação (46)	53
Equação (47)	53
Equação (48)	53
Equação (49)	54
Equação (50)	54
Equação (51)	57
Equação (52)	58
Equação (53)	58
Equação (54)	58
Equação (55)	58
Equação (56)	58
Equação (57)	59
Equação (58)	59
Equação (59)	59
Equação (60)	59
Equação (61)	65
Equação (62)	66
Equação (63)	66
Equação (64)	67

Equação (65)	67
Equação (66)	67
Equação (67)	70
Equação (68)	72
Equação (69)	72
Equação (70)	72
Equação (71)	75
Equação (72)	86
Equação (73)	86
Equação (74)	86
Equação (75)	87
Equação (76)	94
Equação (77)	94
Equação (78)	96
Equação (79)	102
Equação (80)	103
Equação (81)	129
Equação (82)	129
Equação (83)	129
Equação (84)	129
Equação (85)	130
Equação (86)	130
Equação (87)	130
Equação (88)	130
Equação (89)	130
Equação (90)	131
Equação (91)	131
Equação (92)	132
Equação (93)	135
Equação (94)	136

SUMÁRIO

1	INTRODUÇÃO GERAL	21
2	FUNDAMENTAÇÃO TEÓRICA	24
2.1	INTELIGENCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA	24
2.1.1	Aprendizado de máquina	25
2.1.1.1	Aprendizado supervisionado	25
2.1.2	Redes neurais Artificiais	27
2.1.3	Aprendizado profundo	30
2.1.4	Reservoir Computing	32
2.2	MÁQUINA DE APRENDIZADO EXTREMO	34
2.2.1	ELM Regularizado	36
2.3	SISTEMAS ÓPTICOS DE APRENDIZADO DE MÁQUINA	37
2.3.1	Sistemas ópticos difrativos para classificação	41
2.3.2	Sistemas ópticos dispersivos para classificação	45
2.4	SPECKLE	47
2.4.1	Origem dos padrões de speckle	47
2.4.2	Estatística de primeira ordem	48
2.4.2.1	Estatística de amplitude complexa	50
2.4.2.2	Estatística de Intensidade	53
2.4.3	Speckle com estatística de intensidade não- Rayleigh	54
2.4.3.1	Geração de speckle super e sub Rayleigh.	55
2.4.3.2	Speckle com estatística de intensidade de Weibull	58
3	ANÁLISE DA INFLUÊNCIA DA MODIFICAÇÃO DAS PROPRIEDADES ESTADÍSTICAS DA LUZ NA ACURÁCIA DA CLASSIFICAÇÃO DE IMAGENS EM SISTEMAS ÓPTICOS DISPERSIVOS	60
3.1	INTRODUÇÃO	60
3.2	SISTEMA ÓPTICO UTILIZANDO MEIO ESPALHADOR	60
3.2.1	Arranjo experimental e Método	60
3.2.1.1	Codificação do conjunto de dados no SLM	64
3.2.1.2	Descrição do processamento óptico dos dados e treinamento	66
3.3	RESULTADOS	68

3.3.1	Caracterização dos padrões de speckles	68
3.3.2	Avaliação da classificação	69
3.4	INFLUÊNCIA DE SPECKLES COM ESTATÍSTICA DE INTENSIDADE MODIFICADA NA ACURÁCIA DA CLASSIFICAÇÃO	71
3.4.1	Padrões de speckles gerados numericamente	71
3.4.1.1	Codificação do conjunto de dados	75
3.4.2	<i>Setup</i> experimental	76
3.5	RESULTADOS	76
3.5.1	Sistema com predição óptica	81
4	SISTEMA TOTALMENTE ÓPTICO DE CLASSIFICAÇÃO DE IMA- GENS BASEADOS NO ALGORITMO DE MÁQUINA DE APRENDI- ZADO EXTREMO	85
4.1	INTRODUÇÃO	85
4.2	MÉTODO DE CLASSIFICAÇÃO ÓPTICA 1	85
4.2.1	Descrição do treinamento	85
4.2.2	Generalização para problemas de 10 classes	88
4.2.3	Classificação das imagens	88
4.3	RESULTADOS	90
4.4	MÉTODO DE CLASSIFICAÇÃO ÓPTICA 2	94
4.4.1	Descrição	95
4.4.2	Descrição da etapa de classificação	96
4.4.3	Generalização para dez classes	98
4.5	MÉTODOLOGIA EXPERIMENTAL UTILIZANDO SOMENTE UM MO- DULADOR ESPACIAL DE LUZ (SLM) SOMENTE DE FASE.	101
4.5.1	Algoritmo IFTA	101
4.5.2	Projetando as máscaras de fase	104
4.5.3	Setup Experimental e Método	105
4.6	RESULTADOS	107
4.6.1	Resultados numéricos	107
4.6.2	Resultados Experimentais	110
5	CONCLUSÃO	114
	REFERÊNCIAS	116
	APÊNDICE A – HOLOGRAFIA	126

APÊNDICE B – MODULADOR ESPACIAL DE LUZ	133
---------------------------------------------------------	------------

1 INTRODUÇÃO GERAL

Ao longo da história, a relação intrínseca entre o progresso tecnológico e o desenvolvimento da sociedade sempre desempenhou um papel fundamental. Desde os primórdios da civilização, a busca contínua por aprimorar as ferramentas e técnicas disponíveis tem sido um objetivo constante. Esse desejo por inovação tecnológica não só influenciou, mas também impulsionou significativamente o progresso nas áreas científicas, especialmente no campo da computação. O avanço tecnológico se manifestou nas ideias apresentadas por John von Neumann na década de 1940, que deram origem ao conceito de máquina de armazenamento-programa (NEUMANN, 1993). Essas ideias também estabeleceram o princípio fundamental de que a CPU é o “cérebro” do processamento de informações em um computador (NEUMANN, 1993). Isso deu origem a um progresso constante nas Unidades de Processamento Central (CPUs) e nas Unidades de Processamento Gráfico (GPUs) (MIŠIĆ; ĐURĐEVIĆ; TOMAŠEVIĆ, 2012). As GPUs, em particular, não apenas habilitaram avanços em gráficos computacionais, mas também proporcionaram um desempenho de cálculo de alta performance, marcando o início da era da computação moderna.

O desenvolvimento das GPUs não apenas acelerou a evolução da computação, mas também impulsionou um rápido crescimento no campo do Aprendizado de Máquina. Este campo, com raízes que remontam à década de 1950, enfrentou enormes desafios devido às limitações tecnológicas para processar grandes volumes de dados (EUROPE, 2023). Por muito tempo, o Aprendizado de Máquina permaneceu em um estado de estagnação devido à incapacidade de exploração de todo o seu potencial (MINSKY; PAPER, 1969). No entanto, com a introdução das GPUs, esse cenário mudou de maneira radical. Tarefas que costumavam levar dias para serem concluídas podiam agora ser realizadas em questão de horas, graças ao poder de processamento das GPUs.

Desde então, o Aprendizado de Máquina passou a estar presente em uma ampla gama de áreas, incluindo o reconhecimento de fala (RADFORD et al., 2023), (AMODEI et al., 2016), (RIVENSON et al., 2019), (SRINIDHI; CIGA; MARTEL, 2021), diagnóstico médico (YU; BEAM; KOHANE, 2018), (BRIGANTI; MOINE, 2020), análise de imagens (RIVENSON et al., 2019), (SRINIDHI; CIGA; MARTEL, 2021), (CHEN et al., 2017), (MINAEE et al., 2021),

(BARBASTATHIS; OZCAN; SITU, 2019), e pesquisa científica (WANG et al., 2023), (XU et al., 2021). No entanto, a crescente demanda por processamento de grandes volumes de dados em tais áreas, tem vindo acompanhada de desafios, como altos consumos de energia e tempos de processamento relativamente longos. Isso tem motivado cada vez mais a busca por novas soluções de hardware. Recentemente, a utilização da luz para computação e processamento de informações tem ganhado destaque, levando ao desenvolvimento de diversas abordagens fotônicas para atender às demandas de processamento em aprendizado de máquina.

Dentre as abordagens de dispositivos fotônicos, observamos uma gama de inovações que buscam superar as limitações das arquiteturas tradicionais dependentes de unidades centrais de processamento. Esses dispositivos visam alcançar alta velocidade e capacidade de processamento distribuído em larga escala. Suas aplicações incluem tarefas como a multiplicação de matrizes, demonstradas através de técnicas como interferometria, ressonadores de anel e fibras ópticas (SHEN et al., 2017), (ZANG et al., 2019), (FELDMANN et al., 2019), (CHENG et al., 2019). No entanto, eles ainda enfrentam desafios, como a escolha de materiais adequados para a integração óptica com componentes eletrônicos e a compatibilidade com as infraestruturas de computação e comunicação existentes.

Outra abordagem envolve a imitação da estrutura neural do aprendizado profundo na faixa óptica, utilizando arranjos sequenciais de placas difrativas definidas por treinamento computacional (LIN et al., 2018). Essa técnica, embora óptica, processa dados a alta velocidade e alcançou sucesso notável em aplicações de classificação óptica (CHEN et al., 2021). No entanto, requer cuidadosa fabricação de máscaras difrativas para cada aplicação específica. No geral, muitos sistemas ópticos mencionados não são reprogramáveis.

Embora os avanços tenham sido significativos, a pesquisa continua a explorar novas abordagens para atender às crescentes demandas de processamento de dados em aprendizado de máquina e inteligência artificial. A busca por sistemas ópticos inovadores, que sejam eficientes em termos de consumo de energia e processamento de alta velocidade, é fundamental para moldar o futuro do campo. Nesse sentido, esta dissertação se concentra na exploração do potencial da óptica em sistemas experimentais, com foco na classificação precisa de imagens.

Diante disso, a dissertação está estruturada da seguinte forma: No Capítulo 2,

apresentamos uma revisão bibliográfica que aborda conceitos fundamentais de inteligência artificial e aprendizado de máquina. Destacamos algoritmos de aprendizagem, com ênfase nas redes neurais artificiais, que servem como base para o algoritmo de máquinas de aprendizado extremo explorado com detalhes neste capítulo. Além disso, abordamos sistemas ópticos de classificação de imagens, incluindo sistemas difrativos multicamadas e de camada única, bem como sistemas ópticos de classificação baseados na dispersão da luz. Exploramos conceitos relacionados a speckles, especialmente suas propriedades estatísticas, e técnicas para manipular essas propriedades, desviando-se da estatística de Rayleigh. No Capítulo 3, descrevemos a construção do sistema experimental usando speckles no processo de aprendizado de máquina óptico. Começamos explorando o uso de um difusor como meio dispersivo para introduzir complexidade nas imagens de entrada e investigamos seu desempenho na classificação de imagens. Em seguida, utilizamos um modulador espacial de luz de cristais líquidos de modulação somente de fase para criar uma mídia de dispersão holográfica e gerar speckles com estatísticas de intensidade não-Rayleigh. Apresentamos os resultados dessa manipulação e seu impacto na acurácia da classificação de imagens. No Capítulo 4, detalhamos dois sistemas de classificação de imagens puramente ópticos, ambos baseados no algoritmo de treinamento de máquinas de aprendizado extremo. Por fim, no Capítulo 5, apresentamos as conclusões gerais e as perspectivas deste trabalho de pesquisa.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 INTELIGENCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA

A inteligência artificial (IA) é uma área da computação que se dedica ao desenvolvimento de algoritmos que buscam imitar a inteligência humana (RUSSELL, 2010). A ideia de criar máquinas capazes de processar informações de forma semelhante ao cérebro humano remonta ao menos ao século XIX, quando matemáticos como Ada Lovelace e Charles Babbage começaram a explorar a ideia de uma "máquina analítica". Essa "máquina analítica", proposta por Babbage, foi concebida como uma máquina universal programável capaz de executar uma ampla gama de cálculos matemáticos (LUDGATE, 1982). Embora a máquina analítica nunca tenha sido completamente construída durante a vida de Babbage, seu projeto foi considerado um precursor dos computadores modernos.

Um marco importante no desenvolvimento da IA foi o trabalho do matemático Alan Turing na década de 1950. Turing propôs o "teste de Turing", um experimento conceitual que avaliaria se uma máquina poderia exibir um comportamento indistinguível de um ser humano (RUSSELL, 2010). Ao longo das décadas seguintes, a IA passou por altos e baixos. Na década de 1960, houve um grande entusiasmo pelo campo, mas logo se tornou evidente que as abordagens disponíveis na época não seriam capazes de lidar com problemas complexos do mundo real. Isso levou a um período de desencanto e pouco interesse em pesquisas na área (MINSKY; PAPERT, 1969).

No entanto, a partir da década de 1990, a IA começou a se recuperar com o surgimento de novas técnicas de aprendizado de máquina, como as redes neurais artificiais e as árvores de decisão (RUSSELL, 2010). Essas técnicas permitiram que os pesquisadores desenvolvessem sistemas capazes de aprender a partir de dados e realizar tarefas mais complexas do que era possível anteriormente. O aprendizado de máquina, em particular, tornou-se uma área fundamental da IA, pois permite que os computadores aprendam a partir de exemplos passados, identificando padrões e relações entre variáveis (MAHESH, 2020). Diante desse contexto, nesta seção, nosso objetivo é apresentar ao leitor as principais técnicas de aprendizado de máquina e redes neurais artificiais.

2.1.1 Aprendizado de máquina

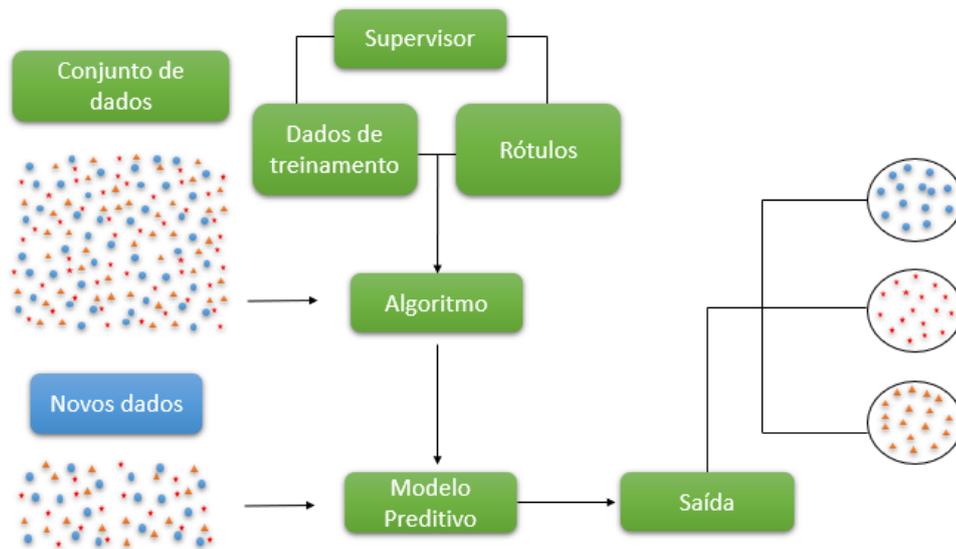
O aprendizado de máquina é uma área de grande importância dentro da inteligência artificial, pois ela possibilita que computadores aprendam a realizar tarefas complexas sem serem explicitamente programados para isso (RASCHKA; MIRJALILI, 2019). Essa subárea da inteligência artificial busca criar algoritmos que possam analisar dados, identificar padrões e, a partir desses padrões, fazer previsões e tomar decisões. Os algoritmos de aprendizado de máquina são capazes de aprender com base em exemplos passados, identificando padrões e relações entre variáveis, e aplicar esse conhecimento para fazer previsões em novos dados (MAHESH, 2020). Além disso, eles são capazes de melhorar sua precisão e desempenho à medida que são expostos a mais dados. Essa capacidade de aprender a partir de dados é fundamental para a realização de tarefas complexas de reconhecimento (ABHISHEK, 2022). O aprendizado de máquina é dividido em várias categorias, como o aprendizado supervisionado, não supervisionado, por reforço e semi-supervisionado. Cada uma dessas categorias é adequada para diferentes tipos de tarefas. No entanto, neste trabalho, nos concentraremos em apresentar os principais conceitos sobre o aprendizado supervisionado, já que é o que faz parte do escopo deste trabalho.

2.1.1.1 Aprendizado supervisionado

O aprendizado supervisionado é uma das categorias mais comuns do aprendizado de máquina. Os algoritmos são treinados com exemplos rotulados, como mostrado na Figura 1, ou seja, dados que já possuem uma resposta correta. Um algoritmo de aprendizado de máquina é treinado com um conjunto de dados rotulados, a partir do qual ele pode aprender a prever rótulos para dados não rotulados. Esses dados não rotulados podem ser dados de teste ou dados coletados no futuro. O objetivo é fazer com que o algoritmo aprenda a mapear entradas para saídas corretas.

Existem duas áreas principais em que o aprendizado de máquina supervisionado é útil: problemas de classificação e problemas de regressão. A Classificação é um processo de atribuir uma categoria ou classe específica a um valor de entrada. É como colocar rótulos em diferentes tipos de objetos. Por exemplo, se tivermos uma imagem de um objeto, a classificação pode envolver identificar se o objeto na imagem

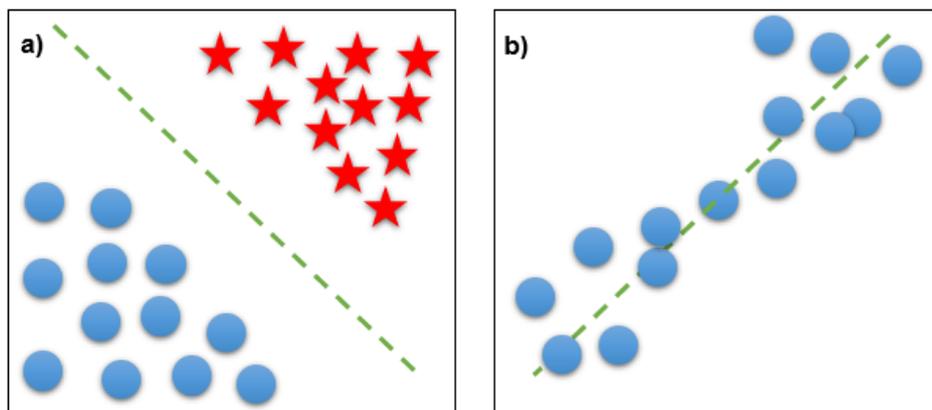
Figura 1 – Fluxograma do aprendizado supervisionado



Fonte: Modificado de Baheti (2023)

é um círculo ou uma estrela, Figura 2 a). A saída da classificação é uma das classes pré-definidas. Essas classes podem ser coisas como "sim" ou "não", "verdadeiro" ou "falso", "azul" ou "vermelho", ou qualquer outro conjunto de categorias que sejam relevantes para o problema em questão.

Figura 2 – Representação de um problema de a) Classificação, b) Regressão



Fonte: Modificado de Baheti (2023)

Nos problemas de regressão, o objetivo não é classificar em categorias, mas sim prever um valor numérico específico. Nesses casos, lidamos com a previsão de valores contínuos, como funções, preços ou quantidades. O propósito é descobrir um padrão matemático que relacione as características de entrada a um valor de saída desejado, assim como exemplificado na Figura 2 b). Dessa forma, podemos construir

um modelo que estime esses valores com base nas características fornecidas. Essa abordagem nos permite entender tendências, fazer previsões e tomar decisões com base em informações numéricas (RASCHKA; MIRJALILI, 2019), (BAHETI, 2023).

Agora que estabelecemos a base do aprendizado supervisionado e suas principais categorias, vamos explorar como as redes neurais artificiais desempenham um papel fundamental no aprendizado de máquina supervisionado, evoluindo ao longo do tempo para permitir a modelagem de relações complexas entre as entradas e as saídas desejadas.

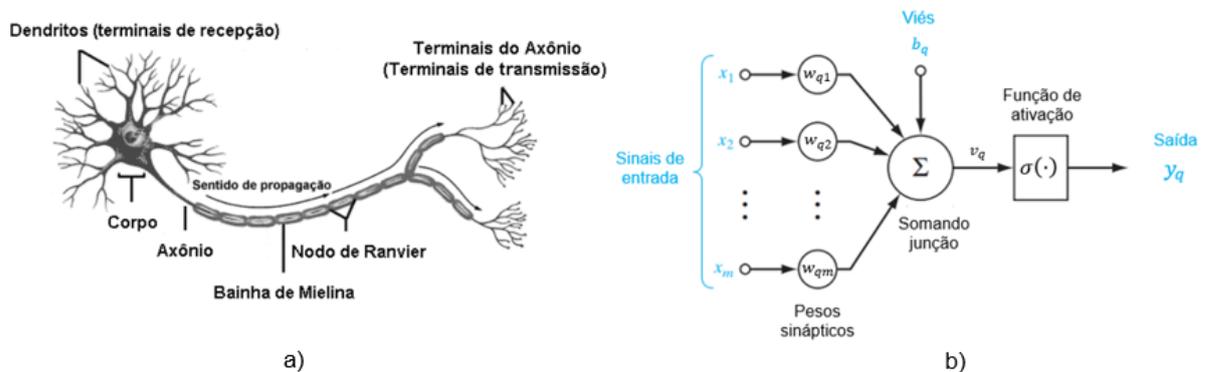
2.1.2 Redes neurais Artificiais

As redes neurais artificiais (RNAs) desempenham um papel fundamental no campo do aprendizado de máquina, sendo inspiradas pela complexa interconexão de neurônios nos sistemas biológicos de aprendizado. No cérebro humano, os neurônios são células especializadas que recebem, processam e transmitem informações por meio de impulsos elétricos e neurotransmissores. Essas células possuem uma estrutura composta por diferentes componentes (Ver Figura 3 a). O corpo do neurônio, também conhecido como soma, é a parte principal do neurônio. É nessa região que se encontra o núcleo, que desempenha um papel fundamental na regulação das atividades celulares e na síntese de proteínas necessárias para o funcionamento e a manutenção do neurônio.

Os dendritos são ramificações que se estendem a partir do corpo da célula. Sua função é receber sinais elétricos e químicos de outros neurônios ou de células sensoriais. Esses sinais são transmitidos por meio de correntes elétricas até o corpo da célula, onde são integrados. O axônio é uma extensão do corpo da célula responsável pela transmissão dos sinais elétricos gerados pelo neurônio. Esses sinais, conhecidos como potenciais de ação, viajam ao longo do axônio em direção aos terminais do axônio. Os terminais do axônio são ramificações localizadas no final do axônio e estabelecem conexões com dendritos de outros neurônios por meio de sinapses. Quando um impulso elétrico alcança os terminais do axônio, neurotransmissores são liberados nas sinapses, transmitindo o sinal para o neurônio seguinte. (MITCHELL et al., 2007),(HAYKIN, 2009), (BEAR et al., 2002).

Em 1943, Warren McCulloch e Walter Pitts desenvolveram um modelo matemá-

Figura 3 – Neurônios a) biológicos b) artificial



Fonte: Santana et al. (2010) e Haykin (2009)

tico de neurônio biológicos. Esse modelo representou um dos primeiros esforços em descrever o funcionamento dos neurônios por meio de formulações matemáticas. O trabalho de McCulloch e Pitts consistiu em simplificar e abstrair os aspectos essenciais dos neurônios biológicos para criar um modelo teórico. Eles se basearam na estrutura e no comportamento observados nos neurônios do cérebro humano, buscando entender as principais características envolvidas na recepção, processamento e transmissão de informações.

O modelo proposto por McCulloch e Pitts capturou a ideia de que os neurônios biológicos podem ser ativados ou inibidos dependendo dos sinais recebidos de outros neurônios. Essa ativação ou inibição foi representada matematicamente por uma função de ativação. Podemos entender esse modelo por meio da Figura 3b, onde os sinais de entrada x_q , que representam os impulsos elétricos de outros neurônios, são considerados como os dados que alimentam uma certa rede neural artificial. Esses estímulos podem ter diferentes graus de excitação para o neurônio receptor, representados pelos pesos sinápticos. Quanto maior o valor do peso, maior a excitação do estímulo. Os pesos sinápticos são representados por w_{qm} , onde q é o índice do neurônio em questão e m refere-se ao terminal de entrada da sinapse correspondente. (HAYKIN, 2009), (MCCULLOCH; PITTS, 1943).

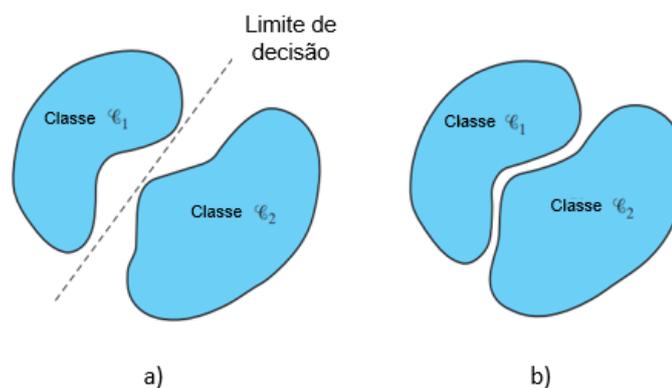
Inspirado pelos trabalhos de Warren McCulloch e Walter Pitts, Frank Rosenblatt propôs um modelo fundamental de neurônio artificial em 1957, conhecido como perceptron. O perceptron, que segue uma abordagem semelhante ao modelo McCulloch-Pitts, opera por meio de uma combinação linear das entradas ponderadas, seguida pela aplicação de uma função de ativação. Essa função determina se o perceptron irá

disparar um sinal de saída, representando a resposta do neurônio, ou se permanecerá inativo.

O modelo do perceptron trouxe um avanço significativo ao introduzir a capacidade de ajuste dos pesos sinápticos durante o processo de aprendizado. Os pesos sinápticos são atribuídos às conexões entre o perceptron e suas entradas, e seu ajuste é crucial para o neurônio aprender e adaptar-se a diferentes padrões de entrada. Essa característica permitiu que o perceptron fosse capaz de aprender e realizar classificações binárias simples.

No entanto, é importante destacar que o perceptron tem limitações quando se trata de representar relações mais complexas e não lineares. Essa limitação foi revelada em um estudo realizado por Marvin Minsky e Seymour Papert em 1969. Essa limitação na capacidade do perceptron de representar relações não lineares se deve à sua estrutura simplificada, que consiste em uma única camada de neurônios com função de ativação baseada em um limiar. O perceptron utiliza um limiar para decidir se o neurônio é ativado ou não, dependendo do valor calculado pela combinação linear das entradas ponderadas (HAYKIN, 2009).

Figura 4 – (a) Classes de padrões linearmente separáveis. (b) Classes de padrões não linearmente separáveis



Fonte: Haykin (2009)

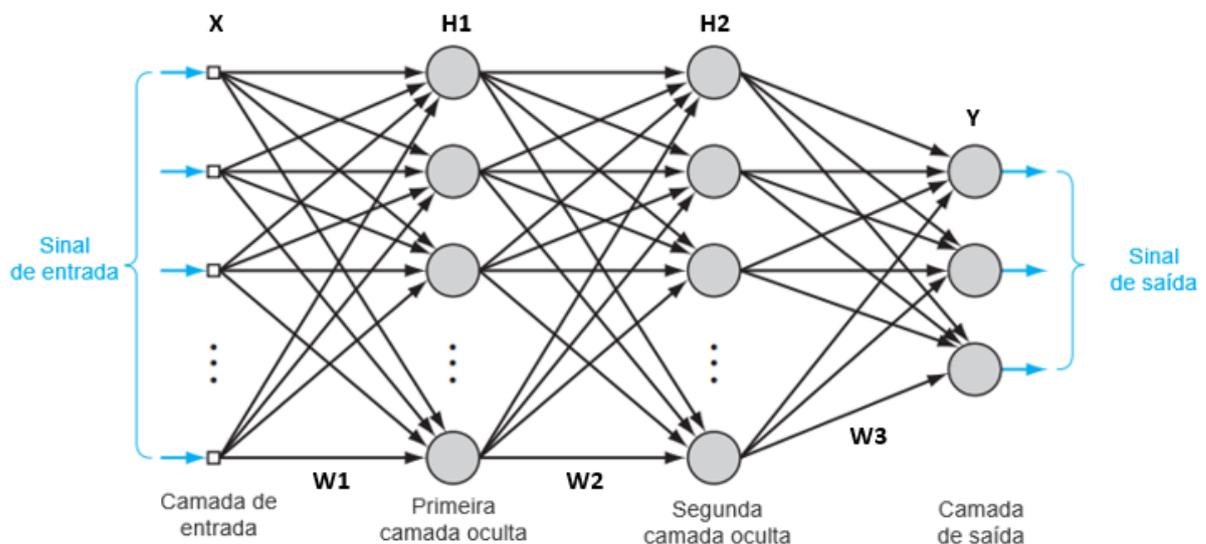
Em outras palavras, o perceptron é capaz de classificar corretamente apenas os dados que podem ser divididos por uma linha reta ou um hiperplano, como visto na Figura 4. No entanto, muitos problemas do mundo real possuem padrões não lineares ou fronteiras de decisão mais complexas. O perceptron não é capaz de aprender e representar esses padrões de forma adequada, o que resulta em baixo desempenho

e precisão. Essa limitação levou ao desenvolvimento de modelos mais avançados, como as redes neurais profundas.

2.1.3 Aprendizado profundo

As redes neurais multicamadas, também conhecidas como redes neurais profundas, foram desenvolvidas para superar a limitação do perceptron na representação de relações não lineares. Essas redes são chamadas de “profundas” devido ao número de camadas intermediárias entre a camada de entrada e a camada de saída. Essas redes consistem em várias camadas de neurônios interconectados, como pode ser visto na Figura 5.

Figura 5 – Arquitetura de uma rede neural profunda com duas camadas ocultas.



Fonte: Haykin (2009)

Para cada neurônio na camada $H1$, calcula-se uma combinação linear das entradas X , ponderadas por um conjunto de pesos específico para cada neurônio. Isso é seguido pela aplicação de uma função de ativação não linear σ . A operação geral para a primeira camada oculta pode ser expressa como:

$$H1 = \sigma(W1X + b1) \quad (1)$$

Onde $W1$ e $b1$ são os pesos associados a cada neurônio da primeira camada, e $b1$ é o vetor de vies para a primeira camada. Seguindo a mesma lógica, para a segunda

camada ($H2$), temos que:

$$H2 = \sigma(W2H1 + b2) \quad (2)$$

E finalmente, para a camada de saída, temos que

$$Y = \sigma(W3H2 + b3) \quad (3)$$

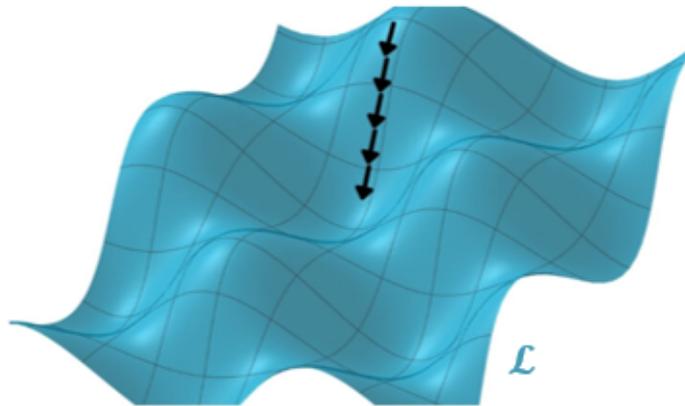
Assim, o treinamento de uma rede neural com esse tipo de arquitetura envolve a otimização dos pesos ($W1$, $W2$, $W3$) e dos vetores de viés ($b1$, $b2$, $b3$) com o objetivo de minimizar uma função conhecida como função de perda. Essa função de perda é uma medida do erro entre as previsões da rede e os rótulos de treinamento (HAYKIN, 2009).

O algoritmo de treinamento que minimiza a função de perda nas redes neurais profundas, é conhecido como retropropagação do erro ou *backpropagation* em inglês, este algoritmo foi um marco no desenvolvimento do campo de aprendizado de máquina. Ele foi concebido por David Rumelhart, Geoffrey Hinton e Ronald Williams com o objetivo de superar as limitações dos métodos de treinamento anteriores, como o perceptron, que tinham dificuldades em lidar com relações não lineares mais complexas (RUMELHART; HINTON; WILLIAMS, 1986).

O *backpropagation* proporcionou um avanço significativo ao introduzir um método sistemático para calcular as contribuições individuais de cada peso na rede para o erro total. Ele funciona através da propagação do erro da saída da rede em direção à camada de entrada, ajustando gradualmente os pesos sinápticos e os vieses ao longo do caminho. Esse processo de retropropagação do erro permite que a rede aprenda a partir dos erros cometidos e refine suas conexões sinápticas para melhorar seu desempenho.

O algoritmo de *backpropagation* é fundamentado no cálculo do gradiente da função de perda em relação aos pesos da rede. Por meio da regra da cadeia da diferenciação, é possível determinar o gradiente descendente e, assim, atualizar os pesos sinápticos de forma a minimizar o erro de predição e zerar o gradiente. Esse processo é iterativamente repetido em um conjunto de dados de treinamento, permitindo que a rede se ajuste aos padrões e melhore seu desempenho ao longo do tempo (NIELSEN, 2015), (RASCHKA; MIRJALILI, 2019).

Figura 6 – Gradiente descendente



Fonte: Science (2022)

Na Figura 6, é possível visualizar uma representação tridimensional da função de perda (\mathcal{L}), na qual são identificados pontos altos e baixos. O objetivo do algoritmo é encontrar os pontos de mínimo da função de perda, ou seja, chegar ao ponto mais baixo da superfície representada. Para atingir essa meta, utiliza-se o gradiente descendente para atualizar os pesos da rede neural. O gradiente descendente orienta o ajuste dos pesos na direção correta, aproximando-nos cada vez mais do mínimo da função de perda (RASCHKA; MIRJALILI, 2019).

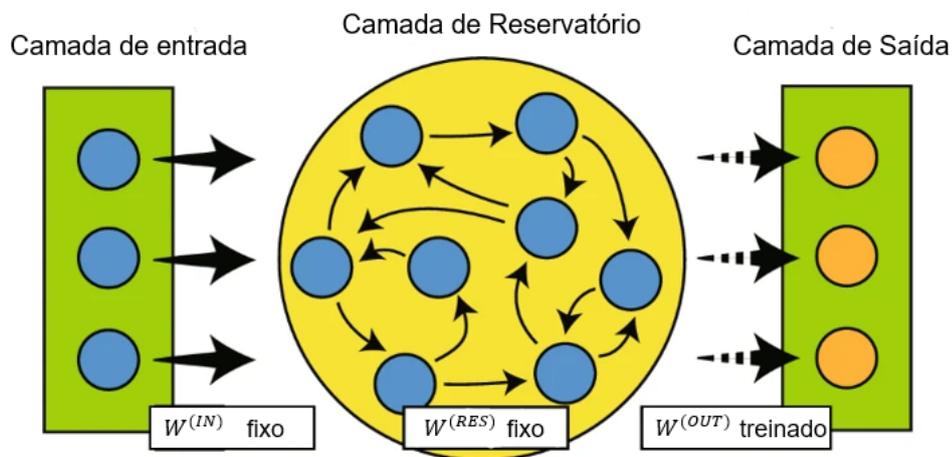
Além disso, é interessante mencionar que, ao longo do tempo, surgiram diversas abordagens alternativas ao *backpropagation*. Isso se deve, em parte, a grande quantidade de parâmetros treináveis, que tende a aumentar com o número de camadas na rede neural, resultando em tempos prolongados de treinamento. Essas abordagens exploram uma variedade de técnicas e algoritmos, com ênfase na eficiência, velocidade de treinamento e na capacidade de lidar com problemas complexos. Duas dessas abordagens notáveis que discutiremos a seguir são o *Reservoir Computing* e o algoritmo de Máquinas de Aprendizado Extremo.

2.1.4 Reservoir Computing

O *Reservoir Computing* é uma abordagem de aprendizado de máquina proposta por Wolfgang Maass e Thomas Natschläger em 2002, com o objetivo de encontrar

uma maneira mais eficiente de treinar redes neurais para lidar com tarefas complexas de processamento de sinais e sequências, especialmente aquelas envolvendo dados temporais (MAASS; NATSCHLÄGER; MARKRAM, 2002). Ao contrário da técnica discutida anteriormente, o *Reservoir Computing* é baseado em redes recorrentes conhecidas como “*reservoirs*” ou reservatórios. Esses reservatórios são compostos por neurônios de estados dinâmicos, nomeados dessa forma devido à sua capacidade de atualizar seus estados internos ao longo do tempo, permitindo assim o armazenamento e processamento dinâmico de informações. Esses neurônios seguem uma modelagem matemática com equações que descrevem a evolução dos estados internos dos neurônios ao longo do tempo em resposta aos sinais de entrada e aos estados internos dos neurônios vizinhos. Além disso, eles são conectados aleatoriamente entre si, formando uma rede de alta dimensionalidade (LUKOŠEVIČIUS; JAEGER, 2009), (MELANDRI, 2014) .

Figura 7 – Arquitetura típica de um *Reservoir Computing*



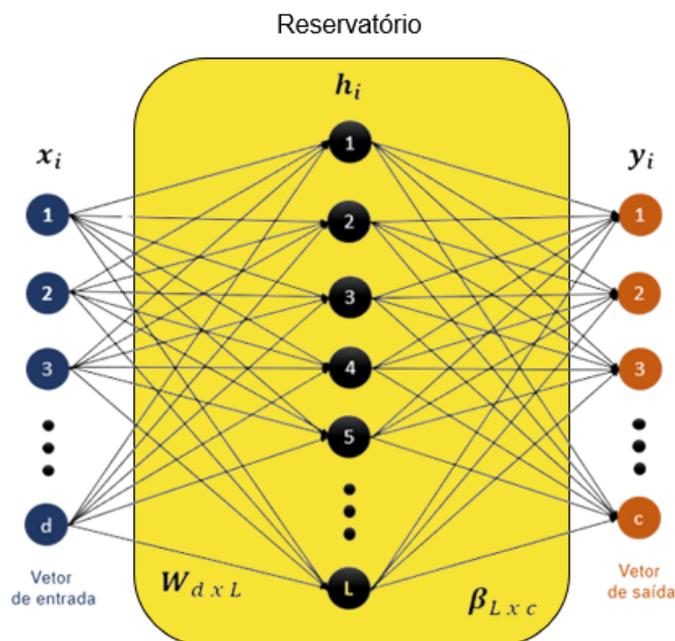
Fonte: Modificado de Sakemi et al. (2020)

A principal característica do *Reservoir Computing* é que o treinamento da rede é realizado apenas na camada de saída, deixando o *reservoir* inalterado. Isso permite que a rede explore a dinâmica complexa do *reservoir* para mapear as entradas para as saídas desejadas. Essa abordagem simplifica significativamente o processo de treinamento e reduz a necessidade de ajustar os parâmetros da rede, tornando o *Reservoir Computing* uma técnica mais fácil de implementar e eficiente em termos computacionais (SAKEMI et al., 2020) .

2.2 MÁQUINA DE APRENDIZADO EXTREMO

Uma abordagem adicional para aprendizado de máquina que se destaca por sua rapidez de treinamento é a Máquina de Aprendizado Extremo (*Extreme Learning Machine- ELM*). Este sistema compartilha algumas semelhanças com o *Reservoir Computing*, uma vez que ambos enfatizam um treinamento que se concentra exclusivamente nos pesos de saída. Além disso, o ELM também possui uma espécie de "reservatório", com a diferença crucial de que no ELM esse reservatório não envolve neurônios em estados dinâmicos, ao contrário do *Reservoir Computing*. Enquanto no *Reservoir Computing*, o reservatório consiste em neurônios que mantêm estados dinâmicos que influenciam o processamento subsequente, no ELM, o reservatório é composto principalmente por uma camada única de neurônios que não possuem estados dinâmicos, simplificando significativamente o treinamento e o uso do modelo. A arquitetura típica de uma Máquina de Aprendizado Extremo está representada na Figura 8.

Figura 8 – Arquitetura de uma Máquina de Aprendizado Extremo



Fonte: Modificado de Akusok et al. (2015)

Este modelo de aprendizado de máquina foi introduzido em 2006 por Haung et. al (HUANG; ZHU; SIEW, 2006). Diferentemente do *backpropagation*, discutido na seção 2.2.1, o ELM concentra-se em treinar uma única camada oculta de uma rede neural ar-

tificial (RNA) *feedforward* (WHITFIELD, 2022) com pesos aleatórios. Assim, no contexto de um treinamento supervisionado, temos N amostras distintas, onde cada amostra \mathbf{x}_i está associada a um rótulo \mathbf{y}_i . Cada \mathbf{x}_i é um vetor $[x_{i1} \dots x_{id}]^T$ de recursos representados por números reais entre 0 e 1, enquanto \mathbf{y}_i é um vetor $[y_{i1} \dots y_{ic}]^T$ que corresponde à saída desejada, d e c são a dimensão das amostras e o número de classes, respectivamente. Definimos \mathbf{X} e \mathbf{T} como matrizes que correspondem a concatenação de todos os \mathbf{x}_i e \mathbf{y}_i , respectivamente, de forma que temos:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_N^T \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} \mathbf{y}_1^T \\ \vdots \\ \mathbf{y}_N^T \end{bmatrix} \quad (4)$$

Para mapear o vetor de entrada \mathbf{x}_i em d combinações lineares diferentes dos seus elementos. Dessa forma, a função de saída de um ELM pode ser escrita como (AKUSOK et al., 2015):

$$\mathbf{y}_i = \sum_{j=1}^L \beta_j \sigma(\mathbf{w}_j \mathbf{x}_i + \mathbf{b}_j) \quad i = 1, \dots, N \quad (5)$$

onde $\sigma(\cdot)$ é uma função de ativação não linear, \mathbf{w}_j são os pesos da entrada e \mathbf{b}_j os vies e β_j são os pesos da saída. A equação 5 pode ser escrita na forma matricial como:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \quad (6)$$

onde,

$$\begin{aligned} \mathbf{H} &= \sigma(\mathbf{W}\mathbf{X} + \mathbf{B}) \\ &= \begin{bmatrix} \sigma(\mathbf{x}_1 \mathbf{w}_1 + \mathbf{b}_1) & \cdots & \sigma(\mathbf{x}_1 \mathbf{w}_L + \mathbf{b}_L) \\ \vdots & \ddots & \vdots \\ \sigma(\mathbf{x}_N \mathbf{w}_1 + \mathbf{b}_1) & \cdots & \sigma(\mathbf{x}_N \mathbf{w}_L + \mathbf{b}_L) \end{bmatrix}_{N \times L}, \\ \boldsymbol{\beta} &= \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times c} \end{aligned}$$

Assim, a solução da equação 6 para determinar os pesos de saída β pode ser obtida por meio do método de mínimos quadrados, alcançada através da minimização da equação 6. Logo, o objetivo é resolver

$$\min_{\beta} \|\mathbf{T} - \mathbf{H}\beta\|^2 \quad (7)$$

Definimos, então $S(\beta) = \|\mathbf{T} - \mathbf{H}\beta\|^2$ logo, a minimização pode ser obtida derivando $S(\beta)$ em relação a β e igualando a zero. Assim, expandindo a expressão de $S(\beta)$ e derivando, temos:

$$\begin{aligned} S(\beta) &= (\mathbf{H}\beta - \mathbf{T})^T (\mathbf{H}\beta - \mathbf{T}) = \mathbf{T}^T \mathbf{T} - \mathbf{T}^T \mathbf{H}\beta + \beta^T \mathbf{H}^T \mathbf{H}\beta \\ \frac{\partial S(\beta)}{\partial \beta} &= -2\mathbf{H}^T \mathbf{T} + 2\mathbf{H}^T \mathbf{H}\beta = 0 \\ \mathbf{H}^T \mathbf{H}\beta &= \mathbf{H}^T \mathbf{T} \end{aligned}$$

Huang, Zhu e Siew (2006) demonstraram que, essa solução pode ser obtida por meio de um pseudo inverso, em particular, a “Inversa generalizada de Moore-Penrose”.

$$\beta = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{T} \quad (8)$$

$$\beta = (\sigma(\mathbf{W}\mathbf{X} + \mathbf{B})^T \sigma(\mathbf{W}\mathbf{X} + \mathbf{B}))^{-1} \sigma(\mathbf{W}\mathbf{X} + \mathbf{B})^T \mathbf{T} \quad (9)$$

$$\beta = \mathbf{H}^\dagger \mathbf{T} \quad (10)$$

onde \dagger denota o inverso, é importante ressaltar também que β tem a menor norma em todas as soluções de mínimos quadrados, e esta solução é única.

2.2.1 ELM Regularizado

A máquina de aprendizado extremo é um tipo de modelo que desperta interesse, principalmente em sua versão restrita da regressão de Ridge, que é conhecida por apresentar maior estabilidade e melhor desempenho de generalização (HUANG et al., 2011). Idealmente, a variável de saída β em que estamos interessados depende de todas as variáveis de entrada, enquanto as próprias variáveis de entrada são independentes. No entanto, para muitos conjuntos de dados, podemos assumir (ou saber)

que as características dos dados sofrem de multicolinearidade, ou seja, podem possuir erros-padrão elevados, o que impacta em muitos modelos de solução de forma negativa. A regressão de Ridge visa melhorar a condição do problema adicionando um termo de regularização da seguinte maneira:

$$\min \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\|^2 + \lambda \|\boldsymbol{\beta}\|^2 \quad (11)$$

onde λ denota o termo de regularização. Podemos observar que esta equação é uma ampliação da equação 2.4

Dependendo do conjunto de amostras de treinamento em questão, pode-se ter mais amostras do que nós ocultos (ou neurônios L), ou seja, ($N > L$), sendo assim, isso leva a uma solução fechada para $\boldsymbol{\beta}$ como:

$$\boldsymbol{\beta} = \mathbf{H}^T(\mathbf{H}\mathbf{H}^T + \lambda\mathbf{I})^{-1} + \mathbf{H}^T\mathbf{T} \quad (12)$$

onde \mathbf{I} é a matriz de identidade de dimensão L . De forma similar, temos uma outra solução para o caso onde tenhamos mais nós ocultos do que amostras de treinamento, ou seja, ($N < L$)

$$\boldsymbol{\beta} = \mathbf{H}^T(\mathbf{H}^T\mathbf{H} + \lambda\mathbf{I})^{-1}\mathbf{T} \quad (13)$$

Embora tenhamos formulado o problema como uma regressão linear regularizada com solução de forma fechada, a versatilidade da estrutura ELM permite que seja facilmente adaptada para tarefas de classificação (HUANG et al., 2011).

O algoritmo da Máquina de Aprendizado Extremo servirá como a base para o desenvolvimento dos sistemas ópticos de classificação abordados neste trabalho de dissertação. A escolha desse algoritmo se justifica pela sua notável eficiência no treinamento e pela ainda limitada exploração de sua aplicação em sistemas ópticos. Assim, na seção seguinte, exploraremos abordagens ópticas em sistemas de classificação que trazem conceitos fundamentais para o desenvolvimento desta pesquisa.

2.3 SISTEMAS ÓPTICOS DE APRENDIZADO DE MÁQUINA

Como já discutido, a inteligência artificial e o aprendizado de máquina têm avançado significativamente em diversas áreas, impulsionados pelo seu grande poder de

aprender a partir de dados. Grande parte desse progresso tem sido impulsionado pela pesquisa em algoritmos de aprendizado, com destaque para as redes neurais. No entanto, o funcionamento e a eficácia de uma rede neural envolve sua capacidade de realizar operações, tanto lineares quanto não lineares.

Uma operação linear refere-se a uma transformação linear aplicada a um conjunto de dados, na qual cada característica ou variável de entrada (definida como X) é ponderada por um coeficiente específico (representado por W) assim como visto na seção 2.1 e, em seguida, somada para gerar uma saída. Isso significa que as operações lineares são transformações simples e proporcionais dos dados, como multiplicação e soma (LUENBERGER; YE et al., 1984).

Por outro lado, as operações não lineares são aquelas em que a relação entre a entrada e a saída não é uma simples multiplicação ou soma. Em vez disso, as operações não lineares introduzem complexidade nos cálculos, permitindo que a rede neural aprenda e modele relações mais complexas nos dados. Essas operações são essenciais para que as redes neurais possam capturar padrões e informações mais sofisticadas que não podem ser representadas apenas por operações lineares (LUENBERGER; YE et al., 1984).

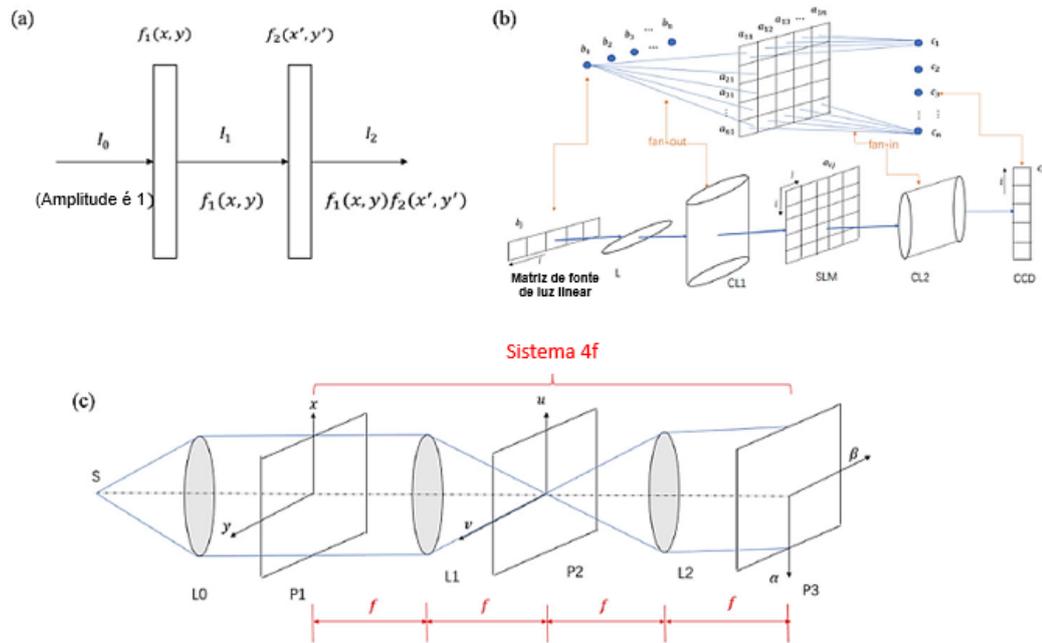
Essas operações frequentemente envolvem uma grande quantidade de cálculos, sobretudo operações de multiplicação e soma, o que, em uma implementação convencional que usa o computador, acaba resultando em um tempo elevado de processamento e custos energéticos substanciais, especialmente quando lidamos com problemas complexos de classificação onde as matrizes possuem dimensões consideráveis.

Frente a essas limitações inerentes aos sistemas eletrônicos convencionais, a óptica tem emergido como uma opção promissora para o processamento rápido de informações. Através da óptica, é possível realizar operações de multiplicação e soma essenciais para o funcionamento das redes neurais de forma extremamente rápida. A Figura 9 ilustra três configurações ópticas distintas para a multiplicação de matrizes.

Na Figura 9a, observamos uma multiplicação óptica direta na qual um feixe luminoso, representado por I_0 com amplitude unitária, incide sobre um padrão específico codificado em uma máscara definida por $f_1(x, y)$. A luz, em relação às informações da máscara, representado agora por I_1 , propaga-se até outra máscara $f_2(x', y')$, resultando em um padrão de saída I_2 que corresponde à multiplicação elemento a elemento

das informações das máscaras $f_1(x, y)$ e $f_2(x', y')$.

Figura 9 – Multiplicação óptica de matrizes, a) Multiplicação de amplitude, b) multiplicação vetor-matriz c) multiplicação matriz-matriz utilizando um sistema óptico 4f.



Fonte: Modificado de Liu et al. (2021)

Na Figura 9b, observamos uma configuração óptica para realizar uma multiplicação óptica vetor-matriz. O processo começa com a utilização de uma barra de VECSEL (*Vertical External Cavity Surface Emitting Laser*). Essa fonte de luz é responsável por simular um vetor, fornecendo um feixe de luz coerente e monocromático. O uso desse laser é fundamental para a criação do vetor de entrada. Para garantir que todos os pontos desse vetor sejam direcionados para uma matriz codificada em um modulador espacial de luz- SLM, a luz proveniente do laser passa por uma série de etapas ópticas. Primeiro, ela é focalizada por uma lente esférica, após a focalização inicial, uma lente cilíndrica é empregada para moldar o feixe de luz e direcioná-lo para um Modulador Espacial de Luz (SLM). O SLM desempenha um papel essencial nesse processo, controlando tanto a fase quanto a amplitude da luz com base nos dados de entrada. Os dados são codificados no SLM na forma de padrões, onde cada ponto do vetor está relacionado a um elemento da matriz. Uma vez modulada, a luz é coletada por outra lente cilíndrica. Posteriormente, o feixe é encaminhado para uma câmera CCD (*Charge-Coupled Device*). A câmera CCD é responsável por capturar a imagem resultante após a multiplicação óptica (GOODMAN; DIAS; WOODY, 1978).

Na Figura 9c, é visto como multiplicar matrizes utilizando uma configuração óptica 4f (CHEN, 1993) (HEINZ; ARTMAN; LEE, 1970). Este sistema é uma técnica óptica que envolve a transformação de uma imagem bidimensional em seu espaço de frequência, também conhecido como domínio de Fourier. Isso é feito usando uma lente que realiza uma transformada de Fourier da imagem. A expressão matemática bidimensional da transformada de Fourier $F(u, v)$ de uma função $f(x, y)$ é dada por:

$$F(u, v) = \iint_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy \quad (14)$$

Onde, u e v representam as coordenadas das frequências espaciais na imagem, e j é a unidade imaginária. A transformada inversa de Fourier é expressa como:

$$f(x, y) = \iint_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} du dv \quad (15)$$

Nesse contexto, um sistema 4f consiste em duas lentes focalizadas (referidas como $L1$ e $L2$ na Figura 9c) separadas por uma distância f , equivalente à sua distância focal. A multiplicação de matrizes é realizada codificando as informações das matrizes em máscaras de amplitude, (referidas na Figura 9c como $P1$ e $P2$) aproveitando as propriedades da transformada de Fourier (GOODMAN; SUTTON, 1996).

A abordagem matemática começa ao representar as matrizes como padrões de amplitude espacial. Suponha que tenhamos duas matrizes $P1$ e $P2$ que desejamos multiplicar. As representações de Fourier dessas matrizes são obtidas pela transformada de Fourier bidimensional, equação 14. Vamos chamar essas representações de $P1(u, v)$ e $P2(u, v)$.

A multiplicação óptica ocorre no plano focal da segunda lente ($L2$), onde as informações da multiplicação das matrizes são codificadas na distribuição de intensidade da luz. O padrão resultante, após a transformada inversa de Fourier, fornece a matriz resultante $P3$. A abordagem matemática para isso é expressa pela convolução no domínio de Fourier:

$$P3(u, v) = P1(u, v) * P2(u, v) \quad (16)$$

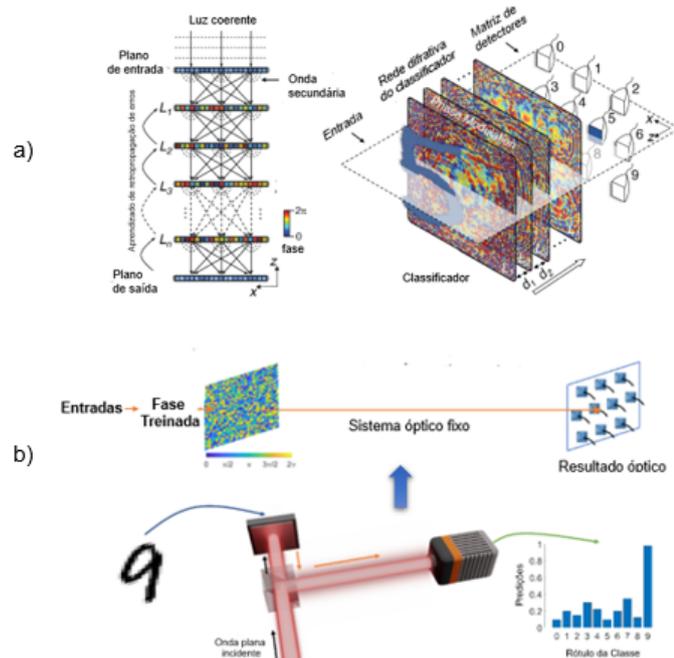
Essa convolução no domínio de Fourier é fundamental para entender como as informações são combinadas. O Teorema da Convolução estabelece que a multiplicação no domínio espacial é equivalente à convolução no domínio de Fourier (GOODMAN,

2005). O sistema óptico 4f, ao realizar essa convolução no plano focal, aproveita a relação entre a multiplicação de matrizes e a convolução no domínio de Fourier.

2.3.1 Sistemas ópticos difrativos para classificação

Com base nessas técnicas ópticas de realização de operações lineares, é possível desenvolver sistemas ópticos que empregam redes neurais para desempenhar tarefas de aprendizado de máquina, com destaque para tarefas de classificação. Isso é especialmente relevante no contexto do trabalho de Lin et al. (2018), que desenvolveu uma rede neural capaz de executar operações lineares com base no princípio da difração da luz. A difração é um fenômeno óptico no qual a luz se propaga e se espalha em diferentes direções ao encontrar obstáculos ou variações em um meio (HECHT, 2023).

Figura 10 – Redes Neurais difrativa. a) multicamadas difrativas b) camada única.



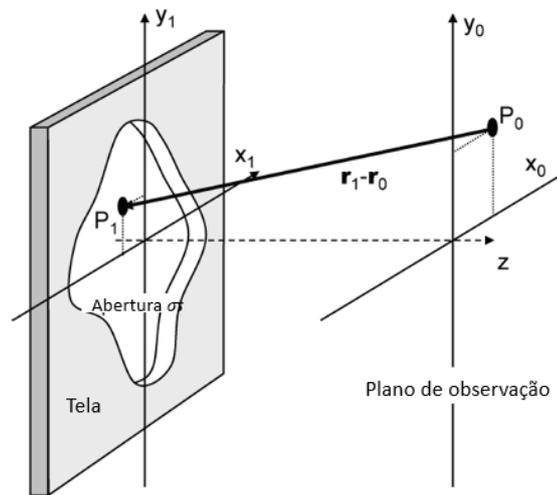
Fonte: Modificado de Lin et al. (2018) e Wu e Zhang (2022)

Para compreender a propagação da luz e sua interação com obstáculos, aberturas ou variação de caminho óptico, é crucial ter conhecimento sobre a teoria escalar da difração. Essa teoria se concentra na compreensão da difração ao considerar apenas uma única componente do vetor campo elétrico ou magnético. Em outras palavras, nosso foco aqui é entender e calcular o campo elétrico de uma onda luminosa em

um ponto P específico, após essa onda interagir com o obstáculo (GOODMAN, 2005) (BAETS; ROELKENS, 2004).

Supondo que o obstáculo está a uma distância finita da fonte luminosa e do plano de observação, a teoria escalar da difração utiliza as aproximações de Fresnel para modelar essa interação de maneira mais precisa. Assim, supondo uma onda monocromática U , e uma abertura que está no plano (x_1, y_1) como mostrado na Figura 11.

Figura 11 – Transmissão da luz através de uma abertura



Fonte: Modificado de Baets e Roelkens (2004)

Assim, o campo é calculado no plano de observação P , que está a uma distância z , como:

$$U(P_0) = \iint_{\sigma} h(x_0, y_0, x_1, y_1) U(x_1, y_1) dx_1 dy_1 \quad (17)$$

Onde,

$$h(x_0, y_0, x_1, y_1) = \frac{-1}{j\lambda} \cos(\mathbf{n}, \mathbf{r}_1 - \mathbf{r}_0) \frac{e^{-jk|\mathbf{r}_1 - \mathbf{r}_0|}}{|\mathbf{r}_1 - \mathbf{r}_0|} \quad (18)$$

onde, λ é o comprimento de onda da luz, $k = 2\pi/\lambda$ é o numero de onda, e

$$|\mathbf{r}_1 - \mathbf{r}_0| = \sqrt{z^2 + (x_0 - x_1)^2 + (y_0 - y_1)^2} \quad (19)$$

Quando a distância axial z é muito maior que as dimensões perpendiculares à direção ao longo do eixo axial. Ou seja, quando estamos lidando com um sistema onde as

dimensões ao longo do eixo z (axial) são significativamente maiores do que as dimensões nas direções perpendicular a esse eixo (transversais). A expressão $\cos(\mathbf{n}, \mathbf{r}_1 - \mathbf{r}_0)$ pode ser aproximada como 1. Essa aproximação é válida para ângulos menores que 18 graus, mantendo um erro inferior a 5%.

Logo, substituindo $|\mathbf{r}_1 - \mathbf{r}_0|$ por z no denominador da equação 18, obtemos a aproximação

$$h(x_0, y_0, x_1, y_1) \approx -\frac{1}{j\lambda z} e^{-jk|\mathbf{r}_1 - \mathbf{r}_0|} \quad (20)$$

A expressão $|\mathbf{r}_1 - \mathbf{r}_0|$ ainda pode ser desenvolvida através de uma expansão binomial, mantendo apenas os dois primeiros termos. Após as manipulações algébricas e substituindo na equação 20, obtemos:

$$h(x_0, y_0, x_1, y_1) \cong \frac{-e^{jkz}}{j\lambda z} e^{-jk/2z[(x_1-x_0)^2+(y_1-y_0)^2]} \quad (21)$$

Diante desta expressão e realizando a integral agora em sobre todo o plano, fazendo $U(x_1, y_1) = 0$ fora da abertura σ , temos:

$$U(x_0, y_0) = \frac{-e^{-jkz}}{j\lambda z} \iint_{-\infty}^{\infty} U(x_1, y_1) e^{-\frac{jk}{2z}[(x_1-x_0)^2+(y_1-y_0)^2]} dx_1 dy_1 \quad (22)$$

Expandindo esta expressão, temos:

$$U(x_0, y_0) = \frac{-e^{-jkz}}{j\lambda z} e^{-\frac{jk}{2z}[x_0^2+y_0^2]} \iint_{-\infty}^{\infty} U(x_1, y_1) e^{-\frac{jk}{2z}[x_1^2+y_1^2]} e^{j\frac{2\pi}{\lambda z}[x_0x_1+y_0y_1]} dx_1 dy_1 \quad (23)$$

Esta integral é comumente chamada de integral de Fresnel. Se definirmos as frequências espaciais como $u = -\frac{x_0}{\lambda z}$ e $v = -\frac{y_0}{\lambda z}$, então temos:

$$U(x_0, y_0) = \frac{-e^{-jkz}}{j\lambda z} e^{-\frac{jk}{2z}[x_0^2+y_0^2]} \iint_{-\infty}^{\infty} U(x_1, y_1) e^{-\frac{jk}{2z}[x_1^2+y_1^2]} e^{-j2\pi[u x_1 + v y_1]} dx_1 dy_1 \quad (24)$$

Este resultado define que o campo $U(x_0, y_0)$ no plano de observação P é a transformada de Fourier bidimensional da distribuição de campos no plano do objeto.

Em um caso em que a distância entre os plano são consideravelmente maiores, ou seja, quando $z \gg \frac{k(x_1^2+y_1^2)_{max}}{2}$ é possível utilizar uma simplificação, denominada de aproximação de campo distante, ou regime de difração de Fraunhofer, onde o termo de fase quadrático pode ser desprezado. Temos, portanto:

$$U(x_0, y_0) = \frac{-e^{-jkz}}{j\lambda z} e^{\frac{-jk}{2z}[x_0^2+y_0^2]} \iint_{-\infty}^{\infty} U(x_1, y_1) e^{j\frac{2\pi}{\lambda z}[x_0x_1+y_0y_1]} dx_1 dy_1 \quad (25)$$

No âmbito do estudo de Lin et al. (2018), é empregado um modelo de difração, denominado de difração de Rayleigh-Sommerfeld (GOODMAN, 2005), uma extensão da teoria de Fresnel, que oferece uma descrição mais precisa da difração em cenários nos quais a aproximação de Fresnel é insuficiente. A equação de Rayleigh-Sommerfeld incorpora termos adicionais que levam em consideração a curvatura da frente de onda, tornando-se especialmente valiosa em situações onde as distâncias são moderadas e a aproximação de Fraunhofer perde sua validade.

No mencionado estudo, os pesquisadores utilizam o modelo de difração para conduzir o treinamento de camadas difrativas responsáveis por modular a fase do campo incidente, assim como mostrado na Figura 9a. Essa modulação é realizada por meio das características difrativas em cada camada, que funcionam como fontes de ondas secundárias moduladas. Ao se combinarem, essas ondas secundárias formam o campo de onda resultante. Nesse contexto, é possível interpretar essas características difrativas como os “neurônios” da rede, conectando-se através da propagação no espaço livre, como descrito em (LIN et al., 2018).

Adotando uma abordagem semelhante, Wu e Zhang (2022) emprega o mesmo modelo de difração para treinar uma rede neural difrativa de apenas uma camada, conforme ilustrado na Figura 9b. O sistema óptico utilizado pelos autores é bastante simples, composto apenas por um laser, um modulador espacial de luz de modulação de fase e uma câmera CCD. Ao realizar o treinamento do modelo com base em alvos predefinidos, ou seja, alvos que definirão os padrões de difração desejados, eles obtêm uma máscara de fase otimizada. Durante o processo de classificação, as imagens a serem analisadas são propagadas até a máscara por meio de computação. Com isso, o padrão resultante é impresso no modulador espacial de luz (SLM), que modula a fase da luz proveniente do laser. Essa modulação resulta no padrão óptico observado, definindo a predição correta da imagem com base na região que possui maior intensidade, conforme ilustrado na Figura 9b.

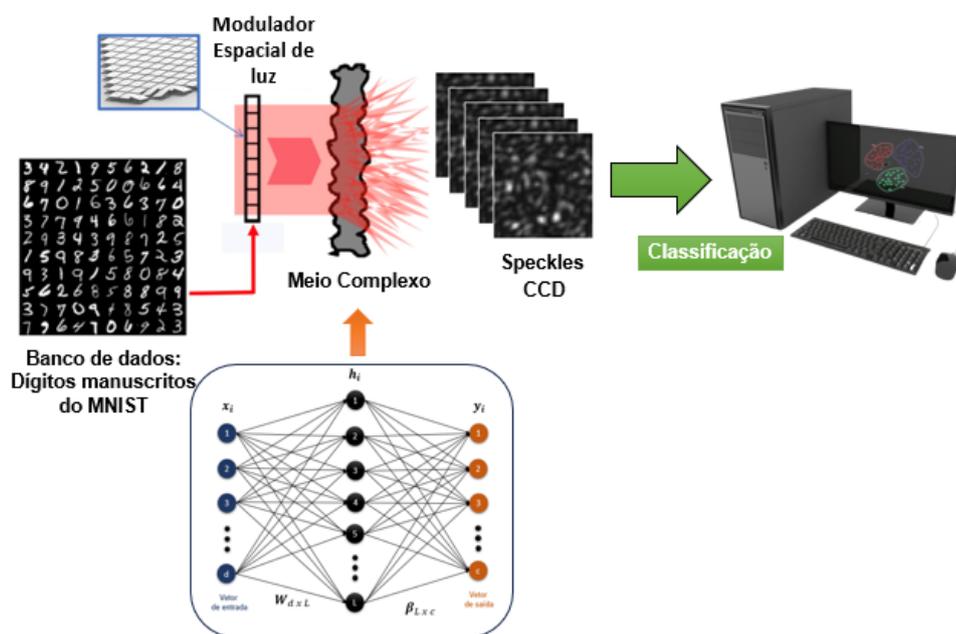
2.3.2 Sistemas ópticos dispersivos para classificação

No contexto da difração, existem outros sistemas ópticos que podem executar operações lineares, e eles estão relacionados ao uso de um meio espalhador como se fossem pesos W , para criar os recursos que compõem as camadas ocultas em sistemas de redes neurais. Isso acontece porque, quando a luz encontra uma superfície rugosa, ela sofre espalhamento, o que significa que a luz se espalha em várias direções e sofre interferência, resultando na formação de padrões não uniformes de luz com pontos claros e escuros, conhecidos como speckles.

Esses speckles, ou padrões de luz, podem ser utilizados de maneira inteligente em sistemas de redes neurais ópticas. As camadas ocultas dessas redes neurais ópticas podem incorporar o meio dispersivo para realizar cálculos e transformações nos dados. A ideia é que, assim como os speckles resultam da interferência da luz, as camadas ocultas podem processar informações de maneira complexa, aproveitando esses padrões para realizar operações lineares e não lineares de forma altamente paralela.

Isso significa que a luz, ao se espalhar no meio dispersivo, não é desperdiçada, mas sim utilizada para processar informações.

Figura 12 – Máquina óptica de aprendizado extremo



Fonte: Modificado de LigthOn (2016)

Essa abordagem é exemplificada no trabalho de Saade et al. (2016), que adota o conceito de espalhamento óptico, onde a luz é dispersa por um meio complexo, para realizar projeções aleatórias nos dados. Essas projeções são integradas ao sistema óptico por meio de um modulador espacial de luz. O padrão de luz resultante, após atravessar o meio difusivo responsável por essas projeções, torna-se componente essencial da camada oculta em uma rede neural, mais especificamente de uma Máquina de Aprendizado Extremo.

A Figura 12 esquematiza uma arquitetura típica desses sistemas. Cada imagem do conjunto de dados é direcionada ao modulador de luz, que ajusta a intensidade da luz proveniente de um laser. Essa luz modulada propaga-se pelo espaço livre até atingir um meio complexo, onde ocorre o espalhamento óptico, gerando padrões de luz distintos que codificam a informação de cada imagem do conjunto de dados. A intensidade do campo desses padrões é, então, capturada por uma câmera CCD. Dai em diante, cada padrão capturado é utilizado para treinar a rede.

Durante o treinamento, esses padrões de luz que contém as informações das imagens do conjunto de dados de treinamento são empregados para calcular a matriz de pesos de saída β , conforme a equação 10.

Na fase de teste, novas imagens (ou imagens de testes, imagens que não foram utilizadas na fase de treinamento) são introduzidas no sistema criando novos padrões de luz. A obtenção da matriz beta otimizada permite que a rede realize a classificação destes novos padrões. Essa classificação é realizada pela multiplicação matricial dessas distribuições de luz com a matriz beta. Esse processo, executado via computador, proporciona uma previsão precisa com base nas características destes padrões aprendidas durante a fase de treinamento.

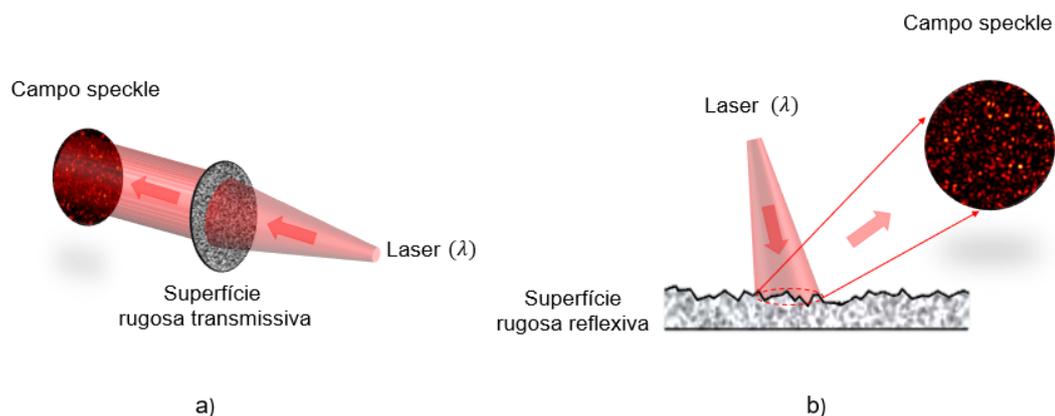
Um dos sistemas ópticos de classificação de imagens desenvolvidos neste trabalho fundamenta-se nesse tipo de abordagem. Na seção seguinte, iremos apresentar as características e propriedades essenciais dessas distribuições de luz, analisando de maneira mais aprofundada a estatística de primeira ordem desses padrões.

2.4 SPECKLE

2.4.1 Origem dos padrões de speckle

Os padrões de speckle são gerados como resultado do fenômeno de interferência que ocorre quando uma luz coerente, como de um laser, interage com uma superfície que apresenta rugosidade. Essa rugosidade pode estar presente tanto em superfícies transmissivas quanto reflexivas, como ilustrado nas Figuras 13a) e 13b), respectivamente. A escala de comprimento da rugosidade nessas superfícies geralmente é da mesma ordem de grandeza do comprimento de onda da luz incidente.

Figura 13 – Processo de formação de padrões de speckles no espaço livre a) por superfícies rugosas transmissivas, b) por superfícies rugosas reflexivas.

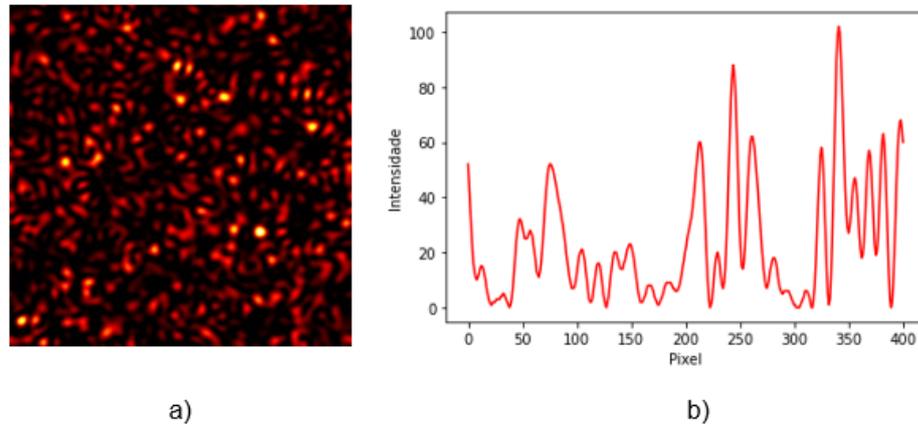


Fonte: Autor, 2023.

Quando a luz incide sobre essa superfície, cada pequeno elemento que a compõe produz uma onda difratada que se propaga em várias direções. À medida que essas ondas difratadas se superpõem em um ponto específico do espaço, elas formam um padrão de interferência granular com intensidades aleatórias, conhecido como padrão de speckle. Nas regiões de interferência construtiva ou destrutiva, ocorrem a formação de grãos claros e escuros, enquanto nos pontos onde a interferência é parcialmente construtiva, temos pontos de luminosidade intermediária.

A Figura 14 ilustra um campo de intensidade típico de um padrão de speckle resultante da iluminação de uma superfície rugosa. Essa imagem foi capturada por uma câmera CCD, permitindo observar o aspecto granular característico do fenômeno de speckle. A distribuição aleatória de intensidade é claramente visível no perfil de intensidade obtido ao longo de uma linha horizontal no centro da imagem.

Figura 14 – a) Imagem da distribuição de intensidades de um padrão de speckle criado por uma superfície rugosa, b) Perfil de intensidade ao longo de uma linha horizontal no centro da imagem.



Fonte: Autor, 2023.

As características aleatórias dos padrões de speckle levam à necessidade de descrevê-los estatisticamente. Nesse sentido, é comum adotar uma abordagem semelhante à utilizada em problemas de caminhada aleatória no plano complexo (GOODMAN, 2007). A análise estatística de primeira ordem desempenha um papel fundamental nesse contexto, permitindo estudar as propriedades individuais de cada grão de speckle. Através dessa análise, é possível obter informações sobre a média, variância e distribuição de probabilidade das intensidades dos grãos de speckle, contribuindo para uma compreensão mais aprofundada do fenômeno.

Nesta seção, examinaremos as propriedades estatísticas dos padrões de speckles, incluindo intensidade, polarização e espectro. Discutiremos como essas propriedades influenciam a forma como a luz interage com os sistemas ópticos e como a informação é capturada.

2.4.2 Estatística de primeira ordem

Considerando o campo elétrico de uma onda luminosa como $E(x, y, z, t)$, no ponto (x, y, z) e no instante t , com polarização linear, a equação de onda para uma onda monocromática pode ser expressa como:

$$E(x, y, z, t) = \mathbf{A}(x, y, z) \exp [j2\pi\nu t] \quad (26)$$

Onde ν representa a frequência óptica e A é a amplitude do fasor do campo, que

é uma função complexa descrita por:

$$\mathbf{A}(x, y, z) = |\mathbf{A}(x, y, z)| \exp [j\theta(x, y, z)] \quad (27)$$

Aqui, $\theta(x, y, z)$ é a fase do campo. Essa equação representa a amplitude do campo elétrico de uma onda luminosa. A intensidade ou irradiação da onda é calculada a partir da amplitude do campo e é dada por:

$$I(x, y, z) = \lim_{T \rightarrow +\infty} \int_{T/2}^{-T/2} |E(x, y, z)|^2 dt = |\mathbf{A}(x, y, z)|^2 \quad (28)$$

Quando um padrão de speckle é formado pela propagação livre no espaço ou através de um sistema formador de imagem, a amplitude do campo elétrico em um ponto específico do plano de observação (x, y) é uma combinação de várias contribuições desfasadas provenientes de diferentes regiões de dispersão na superfície rugosa. Essas contribuições elementares são representadas como fasores, e a amplitude do campo elétrico pode ser expressa como a soma dessas contribuições, conforme a equação (29).

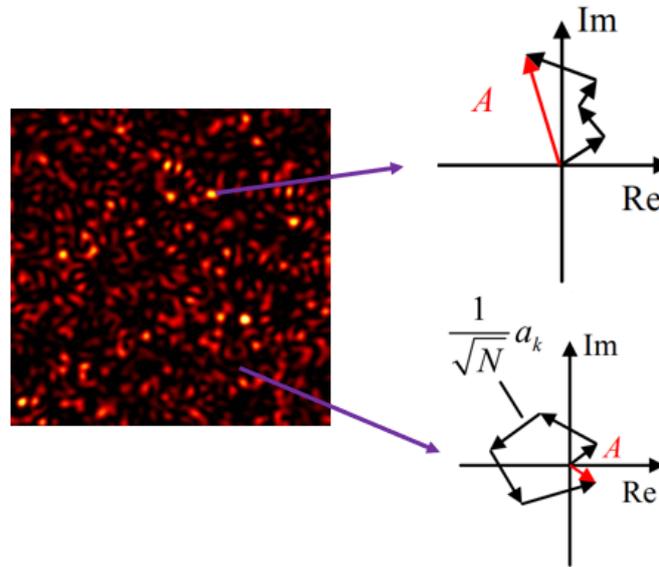
$$\mathbf{A}(x, y, z) = \sum_{q=1}^N \frac{1}{N} \mathbf{a}_q(x, y, z) = \frac{1}{N} \sum_{q=1}^N |a_q| \exp(j\psi_q) \quad (29)$$

Nessa equação, N representa o número de componentes na soma, \mathbf{a}_q é a q -ésima componente com módulo a_q e fase ψ_q , e o fator $\frac{1}{\sqrt{N}}$ é introduzido para preservar o segundo momento da soma quando o número de componentes se aproxima do infinito (GOODMAN, 2007).

Assim, ao analisar a formação do padrão de speckle, consideramos que a contribuição do campo de luz espalhada em um ponto específico $P(x, y, z)$ é o resultado da soma de um grande número de componentes, que representam as contribuições de todos os pontos na superfície de dispersão. Essa abordagem nos permite descrever o padrão de speckle como o resultado da interação dessas contribuições, que estão desfasadas entre si.

Através dessa perspectiva, podemos entender o padrão de speckle como um fenômeno que surge devido à combinação e interferência dessas contribuições espalhadas. Cada componente contribui para a formação do padrão com sua própria magnitude e fase. A Figura 15 ilustra claramente um padrão de speckle, visto do ponto de vista da caminhada aleatória.

Figura 15 – Processo de formação de um padrão de speckle.



Fonte: Autor, 2023.

Além disso, a amplitude complexa $A(x, y, z)$ pode ser decomposta em suas partes real e imaginária, o que possibilita uma análise separada das contribuições dessas partes para as propriedades do campo elétrico.

$$A^{(r)} = \text{Re}\{\mathbf{A}\} = \frac{1}{N} \sum_{q=1}^N |a_q| \cos(\psi_q) \quad (30)$$

$$A^{(i)} = \text{Im}\{\mathbf{A}\} = \frac{1}{N} \sum_{q=1}^N |a_q| \sin(\psi_q) \quad (31)$$

2.4.2.1 Estatística de amplitude complexa

A partir da análise da parte real e imaginária da amplitude complexa do campo, podemos realizar uma análise qualitativa das propriedades estatísticas desse campo. Para isso, são feitas três suposições básicas sobre a estatística dos elementos fasoriais que compõem a soma resultante $\mathbf{A}(x, y, z)$:

1. A primeira suposição é que a_q e ψ_q são estatisticamente independentes de a_p e ψ_p . Isso significa que ter o conhecimento da amplitude e da fase de um fasor componente não implica ter conhecimento da amplitude e da fase de qualquer outro componente.

2. A segunda suposição estabelece que a_q e ψ_q são estatisticamente independentes. Isso implica que a informação sobre a amplitude de um componente fasorial específico não fornece informação sobre a fase do mesmo componente, e vice-versa.

3. A terceira suposição é que a fase ψ_p segue uma distribuição uniforme entre $[-\pi, \pi]$. Isso significa que a fase de cada componente tem a mesma probabilidade de ocorrer em qualquer ponto dentro desse intervalo.

Dessa forma, se considerarmos o caso em que a quantidade N de contribuições de fasores componentes é muito grande, de modo que as partes real e imaginária do campo sejam expressas por uma soma de muitas variáveis aleatórias, podemos aplicar o teorema do limite central (REICHL, 1998) (GOODMAN, 2007). Isso significa que, fazendo N tender ao infinito, as partes real e imaginária do campo, denotadas por $A^{(r)}$ e $A^{(i)}$, respectivamente, seguem uma distribuição gaussiana assintoticamente. Assim, podemos demonstrar que as partes real e imaginária do campo seguem uma distribuição de probabilidade gaussiana, expressa por:

$$P_{r,i}(A^{(r)}, A^{(i)}) = \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{[A^{(r)}]^2 + [A^{(i)}]^2}{2\sigma^2} \right\} \quad (32)$$

onde σ^2 é dado por:

$$\sigma^2 = \sigma_{(r)}^2 = \sigma_{(i)}^2 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{q=1}^N \frac{\langle |a_q|^2 \rangle}{2} \quad (33)$$

A função de densidade de probabilidade mencionada é conhecida como função de densidade Gaussiana circular, pois possui valores constantes cujos contornos formam círculos no plano complexo. Portanto, a amplitude A do fasor é frequentemente referida como uma variável aleatória complexa gaussiana.

Uma consideração adicional a ser feita diz respeito à estatística da amplitude A e da fase θ do fasor resultante.

$$\begin{aligned} A &= \sqrt{[A^{(r)}]^2 + [A^{(i)}]^2} \\ \theta &= \arctan \left(\frac{A^{(i)}}{A^{(r)}} \right) \end{aligned} \quad (34)$$

$$\begin{aligned} A^{(r)} &= A \cos(\theta) \\ A^{(i)} &= A \sin(\theta) \end{aligned} \quad (35)$$

Com base nas relações mencionadas em 34 e 35, é possível obter a função de densidade de probabilidade conjunta da amplitude e da fase.

Assim, podemos expressar a função de densidade de probabilidade conjunta de A e θ utilizando a relação 36, que é derivada da teoria da probabilidade para transformação de variáveis (GOODMAN, 2007).

$$P_{A,\theta}(A, \theta) = P_{r,i}(A^{(r)}, A^{(i)}) \| J \| \quad (36)$$

onde o termo $\| J \|$ representa o Jacobiano de transformação entre as duas variáveis, A e θ , e é definido como:

$$\| J \| = \begin{vmatrix} \frac{\partial A^{(r)}}{\partial A} & \frac{\partial A^{(r)}}{\partial \theta} \\ \frac{\partial A^{(i)}}{\partial A} & \frac{\partial A^{(i)}}{\partial \theta} \end{vmatrix} \quad (37)$$

Portanto, temos:

$$\| J \| = A \quad (38)$$

Dessa forma, a função de densidade de probabilidade conjunta de A e θ é dada por:

$$\rho_{A,\theta}(A, \theta) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{A}{2\sigma^2}\right) \quad (39)$$

onde A e θ assumem valores ($A \geq 0$) e ($-\pi \leq \theta \leq \pi$), respectivamente. Com base na suposição 2, sabemos que A e θ são estatisticamente independentes. Isso significa que podemos encontrar a estatística de A e θ individualmente. Começando com a estatística da amplitude A , temos:

$$\rho_A(A) = \int_{-\pi}^{\pi} \rho_{A,\theta}(A, \theta) d\theta = \frac{A}{\sigma^2} \exp\left(-\frac{A}{2\sigma^2}\right) \quad (40)$$

Essa função é conhecida como a função de densidade de probabilidade Rayleigh. Da mesma forma, podemos encontrar a função de densidade de probabilidade da fase θ da seguinte maneira:

$$\rho_{\theta}(\theta) = \int_0^{\infty} \rho_{A,\theta}(A, \theta) dA = \frac{1}{2\pi} \quad (41)$$

2.4.2.2 Estatística de Intensidade

Sabemos, a partir da equação 28, que a intensidade de uma onda luminosa polarizada pode ser calculada como:

$$I(x, y, z) = |A(x, y, z)|^2 \quad (42)$$

Consequentemente, sua função de densidade de probabilidade será:

$$\rho_I(I) = \frac{1}{2\sigma^2} \exp\left(-\frac{I}{2\sigma^2}\right) \quad (43)$$

onde, ($I > 0$) e $I = 0$ para valores distintos desse intervalo. Podemos obter os momentos dessa expressão por meio de integração direta:

$$\langle I \rangle = \int_0^{\infty} I \rho_I(I) dI = 2\sigma^2 \quad (44)$$

$$\langle I^2 \rangle = \int_0^{\infty} I^2 \rho_I(I) dI = 2(2\sigma^2)^2 \quad (45)$$

Dessa forma, de maneira geral, temos:

$$\langle I^q \rangle = \int_0^{\infty} I^q \rho_I(I) dI = (2\sigma^2)^q q! \quad (46)$$

Como mostrado na equação 44, $\langle I \rangle = 2\sigma^2$. Portanto, podemos reescrever a equação 43 da seguinte forma:

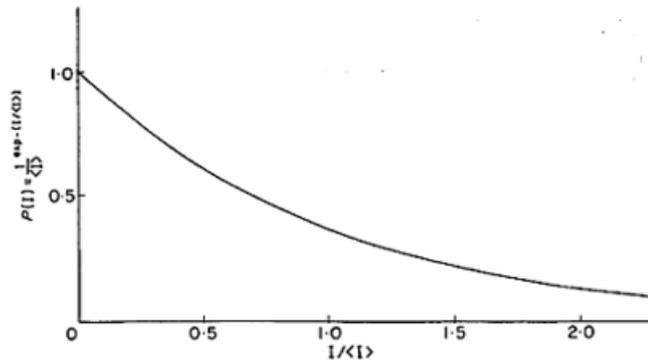
$$\rho_I(I) = \frac{1}{\langle I \rangle} \exp\left(-\frac{I}{\langle I \rangle}\right) \quad (47)$$

Esse tipo de distribuição é conhecido como distribuição de speckles totalmente desenvolvidos e, como ilustrado na Figura 16, a curva característica dessa distribuição é uma exponencial decrescente (GOODMAN, 2007).

Os momentos relacionados à intensidade são descritos na equação 48, e a partir deles, podemos definir outra quantidade importante relacionada ao fenômeno do speckle, que é o contraste C .

$$\begin{aligned} \langle I^2 \rangle &= 2 \langle I \rangle^2 \\ \sigma_I^2 &= \langle I \rangle^2 \\ \sigma_I &= \sqrt{\langle I^2 \rangle - \langle I \rangle^2} = \langle I \rangle \end{aligned} \quad (48)$$

Figura 16 – Gráfico da distribuição de densidade de probabilidade da intensidade de um padrão de speckle.



Fonte: Goodman (2007)

$$C = \frac{\sigma_I}{\langle I \rangle} = \frac{\sqrt{\langle I^2 \rangle - \langle I \rangle^2}}{\langle I \rangle} = \sqrt{\frac{\langle I^2 \rangle}{\langle I \rangle^2} - 1} \quad (49)$$

Podemos observar que o contraste, equação 49 define uma medida da força da flutuação da intensidade do padrão de speckle em relação à média da intensidade. Em speckles de luz que seguem a distribuição de Rayleigh, o contraste assume um valor próximo a 1.

Além disso, podemos definir a função de densidade de probabilidade de a intensidade exceder um determinado limiar (I_t) (*threshold*), que é muito útil em aplicações envolvendo o uso de lasers.

$$P(I \geq I_t) = \int_{I_t}^{\infty} \rho_I(I) dI = \int_{I_t}^{\infty} \frac{1}{\langle I \rangle} \exp\left(-\frac{I}{\langle I \rangle}\right) dI = \exp\left(-\frac{I_t}{\langle I \rangle}\right) \quad (50)$$

2.4.3 Speckle com estatística de intensidade não-Rayleigh

Como vimos, os padrões de speckles, originados na dispersão da luz em um meio espalhador, estão intrinsecamente relacionados com uma distribuição estatística de intensidade conhecida como estatística de Rayleigh. Em outras palavras, essa distribuição segue uma função de densidade de probabilidade caracterizada por um decaimento exponencial. Nesta seção, iremos explorar a capacidade de manipular essa estatística por meio de técnicas documentadas na literatura.

2.4.3.1 Geração de speckle super e sub Rayleigh.

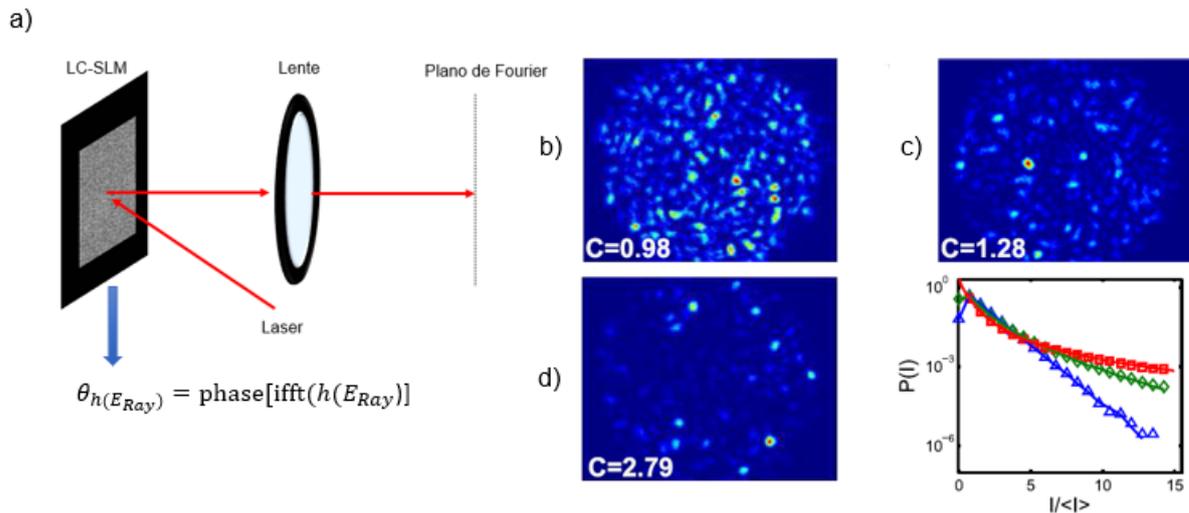
Um marco pioneiro no estudo da manipulação da estatística Rayleigh em padrões de speckles, ocorreu em 2014, graças ao trabalho de Yaron Bromberg e Hui Cao. Eles conduziram uma pesquisa experimental que revelou a possibilidade de criar padrões de speckles com estatísticas que divergem da estatística de Rayleigh, empregando um Modulador Espacial de Luz de cristais líquidos (LC-SLM) (BROMBERG; CAO, 2014). Essa abordagem introduz complexas correlações de alta ordem no campo inicial de luz, resultando na redistribuição da intensidade entre os grãos que compõem o padrão de speckle, enquanto mantém intacta sua estrutura granular característica. Diante disso, tornou-se possível gerar padrões de speckles com contrastes superiores a 1, denominados por Bromberg e Cao como speckles super Rayleigh, bem como contrastes inferiores a 1, categorizados como speckles sub Rayleigh (BROMBERG; CAO, 2014).

Mediante essa técnica, a alteração do contraste nos padrões de speckle é obtida através da geração numérica de padrões de speckle de alto contraste. Isso é realizado ao elevar o campo de speckle Rayleigh, representado como E_{ray} , a uma potência n . Em seguida, é efetuado o cálculo da transformada inversa de Fourier e a fase resultante dessa transformação é aplicada ao (LC-SLM). Em essência, a produção de speckles não Rayleigh é alcançada por meio dessa transformação não linear, expressa como $h(E_{Ray}) = E_{Ray}^n$. Nesse contexto, o parâmetro n , que assume valores inteiros consecutivos, isto é, $n = 1, 2, 3, \dots$, é o que determina a obtenção de speckles resultantes com alto contraste, caracterizando-os como speckles super Rayleigh.

Na Figura 17a, podemos observar o esquema experimental empregado para a geração de padrões de speckle não Rayleigh. Neste experimento, uma fonte de luz coerente, tal como um laser, incide sobre o LC-SLM, no qual uma matriz de fase com pixels correlacionados está incorporada. A luz laser é submetida a modulação de fase pelo LC-SLM e, a partir desse ponto, propaga-se em direção a uma lente que realiza uma transformação de Fourier. No plano focal dessa lente, também denominado plano de Fourier, é onde conseguimos obter o padrão de speckle desejado.

Através da aplicação da equação 49, determina-se o contraste associado a cada padrão gerado, com valores variados de n , como ilustrado na Figura 17, nas subfiguras b), c) e d). As curvas correspondentes ao gráfico da função de densidade de pro-

Figura 17 – Resultados experimentais para speckle super Rayleigh a) Simplificação do setup experimental, b) Speckle Rayleigh com $C = 0,98$, c) Speckle Super Rayleigh com $C = 1,28$, d) Speckle Super Rayleigh com $C = 2,79$, seguido da curva de densidade de probabilidade de cada padrão, os triângulos azuis referem-se aos speckle Rayleigh, losangos verdes aos speckles super Rayleigh de contraste $C = 1,28$ e os quadrados vermelhos aos speckle super Rayleigh de contraste $C = 2,79$.



Fonte: Modificado de Bromberg e Cao (2014)

babilidade $P(I)$, apresentadas em uma escala semi-logarítmica, revelam que, quando o contraste é de $C = 0,98$ (representado pelos triângulos azuis no gráfico), o padrão adere à estatística de Rayleigh. No entanto, as curvas subsequentes retratam padrões de speckles super-Rayleigh, inicialmente com um contraste de $C = 1,28$ na Figura 17 C), e posteriormente com um contraste de $C = 2,79$ na Figura 17 D). Ambas as curvas de $P(I)$ (identificadas pelos losangos verdes e quadrados vermelhos, respectivamente) indicam que a distribuição de densidade de probabilidade exibe um decaimento mais lento em comparação com a exponencial negativa característica da distribuição de Rayleigh.

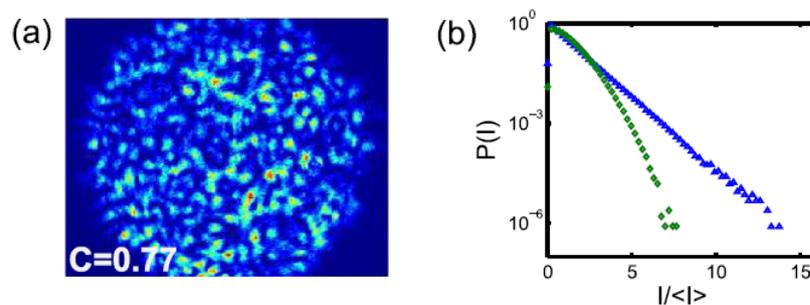
Os autores também investigaram a manipulação do contraste dos speckles Rayleigh para valores inferiores a 1, conhecidos como speckles sub Rayleigh. Nesse cenário, ao reduzir o contraste, a intensidade da luz se distribui de maneira mais homogênea entre os grãos de speckle, ou seja, as variações nos valores de intensidade dos grãos tornam-se insignificantes em relação à média de intensidade desses grãos.

Para reproduzir experimentalmente esse fenômeno, os autores utilizaram uma transformação não linear que satura a amplitude do campo de speckle, mantendo, no entanto, a fase inalterada. Essa transformação pode ser expressa da seguinte maneira:

$$h(E_{Ray}) = \sqrt{1 - e^{-|E_{Ray}|^2}} e^{j\theta_{Ray}} \quad (51)$$

onde θ_{Ray} representa a fase do campo de speckle Rayleigh. Bromberg e Cao observaram que ao aplicar essa transformação não linear e considerar apenas a fase na transformada inversa de Fourier, utilizando o Modulador Espacial de Luz, os resultados obtidos foram bastante semelhantes aos padrões originais de speckle Rayleigh. Para aprimorar isso, adotaram uma abordagem iterativa com base no algoritmo de Gershberg-Saxton (GERCHBERG; SAXTON, 1994) (WU et al., 2021). Este procedimento emprega transformadas de Fourier para propagar o campo luminoso entre o plano do modulador e o plano de Fourier. Durante cada iteração do algoritmo, ajusta-se a amplitude no plano de Fourier de acordo com a transformação desejada, proporcionando correções sequenciais ao campo de luz. Então a fase obtida neste processo de otimização é empregada no modulador para gerar speckles sub-Rayleigh.

Figura 18 – Resultados experimentais para speckle sub-Rayleigh a) Padrão de intensidade de speckle sub-Rayleigh com $C = 0,77$, b) Gráfico da densidade de probabilidade da intensidade, os losangos verde referem-se ao speckles sub-Rayleigh e os triângulos azuis aos speckles de Rayleigh.



Fonte: Bromberg e Cao (2014)

Na Figura 18, observamos o padrão de intensidade de um speckle sub-Rayleigh, caracterizado por um contraste inferior a 1, com valor de $C = 0,77$. Ao analisar a Figura 18b, notamos que a curva verde, que representa a distribuição de intensidade desse padrão, diminui mais rapidamente em comparação com a curva do speckle Rayleigh (curva em azul). Para esse tipo de padrão, é evidente que há uma maior probabilidade de encontrar valores de intensidade mais elevados do que nos speckles de Rayleigh.

2.4.3.2 Speckle com estatística de intensidade de Weibull

Amaral et al. (2015), desenvolveram uma abordagem alternativa para a manipulação da estatística de intensidade em padrões de speckles. Este método envolve a utilização de hologramas de fase (cuja geração esta descrita na subseção A.0.2.1 correspondente ao Apêndice A) em um LC-SLM. Esses hologramas têm a capacidade de codificar tanto a fase quanto a amplitude de um campo de speckle. A amplitude deste campo é ajustada por meio de uma expoente específico, representado como α , que, em contraste com abordagens anteriores, pode assumir valores reais positivos arbitrários (AMARAL; FONSECA; JESUS-SILVA, 2015). A aplicação dessa potência a um campo de speckle Rayleigh modifica o contraste do campo, permitindo a geração de speckles com contrastes tanto superiores quanto inferiores a 1, eliminando assim a necessidade de procedimentos de otimização.

Matematicamente, para o método aplicado pelos autores, supõe que um campo de speckle Rayleigh é dado pela amplitude complexa:

$$E(x, y) = |A| \exp(j\phi(x, y)) \quad (52)$$

Onde a intensidade deste campo pode ser calculada como:

$$I(x, y) = |A(x, y)|^2 \quad (53)$$

A partir do campo complexo, é definida uma transformação apenas na amplitude deste campo dada por:

$$|A| \rightarrow |A|^\alpha \quad (54)$$

Esta transformação eleva a amplitude a uma potencia α , mantendo, então, a fase do campo intacta, assim o campo modificado por essa transformação será dado por:

$$E_T(x, y) = |A|^\alpha \exp(j\phi(x, y)) \quad (55)$$

Logo, a intensidade deste campo após está transformação será definida por:

$$W = |A|^{2\alpha} \quad (56)$$

Onde, W representa uma variável aleatória de intensidade que está relacionada a variável de intensidade de Rayleigh I . Esta relação é definida como:

$$W = f(I) = I^\alpha \quad (57)$$

Como essas duas variáveis aleatórias estão relacionadas por uma função monotônica, as funções densidade de probabilidade estão conectadas pela seguinte relação:

$$\rho_W(W)dW = \rho_I(I)dI \quad (58)$$

$$\rho_W(W) = \rho_I(f^{-1}(W)) \left| \frac{dI}{dW} \right| \quad (59)$$

Nessa equação, $\rho_I(I)$ é a função densidade de probabilidade (FDP) da intensidade dos padrões de speckle de Rayleigh, descrita pela equação 47. Por outro lado, $\rho_W(W)$ representa uma nova função de densidade de probabilidade, dada por:

$$\rho_W(W) = \frac{1}{2\alpha\sigma^2} W^{\frac{1-\alpha}{\alpha}} \exp\left(-\frac{W^{1/\alpha}}{2\sigma^2}\right) \quad (60)$$

Essa nova FDP conhecida como função de densidade de probabilidade de Weibull é obtida a partir da transformação de intensidade mencionada anteriormente, onde W é a variável aleatória resultante da transformação de I com um parâmetro α . É importante destacar que σ^2 representa a variância da distribuição original de intensidade (Distribuição de Rayleigh) (AMARAL; FONSECA; JESUS-SILVA, 2015).

3 ANÁLISE DA INFLUÊNCIA DA MODIFICAÇÃO DAS PROPRIEDADES ESTADÍSTICAS DA LUZ NA ACURÁCIA DA CLASSIFICAÇÃO DE IMAGENS EM SISTEMAS ÓPTICOS DISPERSIVOS

3.1 INTRODUÇÃO

Neste capítulo, apresentaremos uma montagem óptica experimental desenvolvida durante o projeto de mestrado para a classificação de imagens. Esta montagem é fundamentada na representação óptica de um sistema de máquina de aprendizado extremo. Também serão discutidos os resultados obtidos. É importante destacar que nosso trabalho com esse sistema foi dividido em duas etapas distintas. Na primeira etapa, concentramo-nos na construção e avaliação do desempenho do sistema. Utilizamos um meio espalhador para criar projeções aleatórias das imagens contidas nos conjuntos de dados de treinamento e teste em um espaço de alta dimensão. Essa fase nos permitiu explorar a capacidade da nossa configuração em processar dados rapidamente, ao mesmo tempo em que avaliamos sua capacidade de classificar corretamente as imagens de teste. Na segunda etapa, empregamos a técnica da holografia para gerar essas projeções aleatórias com estatísticas de intensidade modificadas. O objetivo principal foi analisar o impacto dessa modificação em nosso sistema e entender como ela influencia o desempenho de classificação.

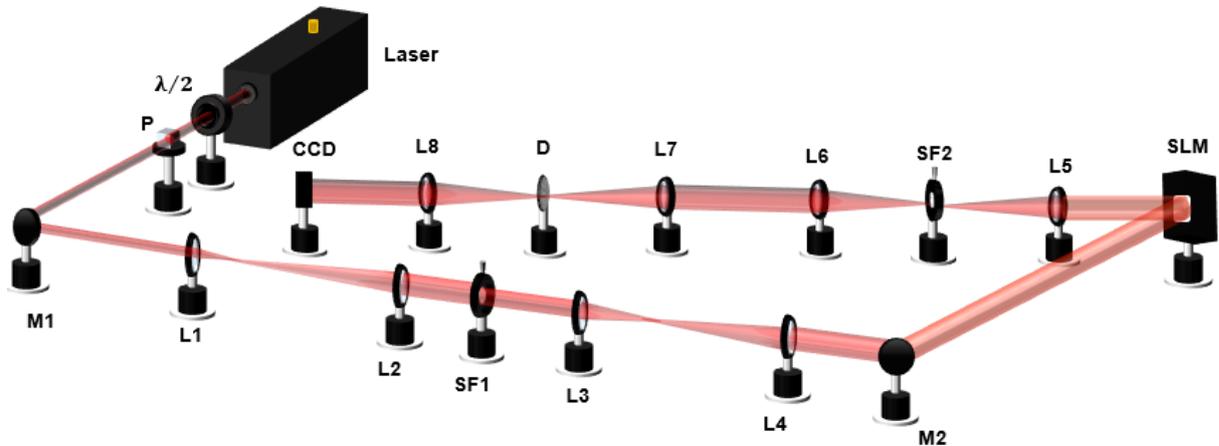
3.2 SISTEMA ÓPTICO UTILIZANDO MEIO ESPALHADOR

3.2.1 Arranjo experimental e Método

Em nosso estudo, iniciamos nossa investigação explorando o uso de um meio espalhador como parte da etapa de processamento óptico dos dados para resolver um problema de classificação de imagens. Desenvolvemos um sistema experimental, conforme ilustrado na Figura 19, projetado especificamente para essa tarefa.

Para construir o sistema, utilizamos um laser de hélio-neônio (He-Ne) que emite luz monocromática a um comprimento de onda de 633 nm e uma potência de 1 mW . O feixe do laser inicialmente possui uma polarização linear vertical devido às características intrínsecas de seu meio ativo e à geometria da cavidade ressonante do laser.

Figura 19 – Setup Experimental de um classificador óptico utilizando um difusor de luz como elemento espalhador.



Fonte: Autor, 2023

(SVELTO; HANNA et al., 2010). No entanto, em nosso sistema, empregamos um modulador espacial de luz do tipo LC-SLM. Esse modulador, que opera apenas com modulação de fase, é responsável por projetar as imagens que desejamos treinar e classificar no sistema óptico.

A modulação de fase ocorre porque as moléculas dos cristais líquidos que compõem o LC-SLM sofrem rotação em resposta a uma diferença de potencial gerada por um sinal enviado por meio de um holograma que codifica as imagens do conjunto de dados. Devido à natureza birefringente dos cristais líquidos, essa rotação altera o índice de refração do meio, permitindo, assim, a modulação de fase do feixe laser. No entanto, é fundamental que a orientação da polarização linear da luz esteja alinhada com a direção do vetor diretor dos cristais líquidos (consulte o Apêndice B para obter mais detalhes).

Para lidar com isso, incorporamos uma lâmina de meia onda ($\lambda/2$) no caminho do feixe. Essa lâmina tem a função de girar a orientação da polarização da luz. Em seguida, o feixe passa por um polarizador (P) que permite a transmissão apenas da componente elétrica na direção do eixo de polarização. Portanto, ao girarmos a lâmina de meia onda, estamos alterando a polarização linear da luz que emerge e, conseqüentemente, modificando a amplitude da componente paralela ao eixo do polarizador. Além de controlar a polarização, o sistema composto pela lâmina de meia onda ($\lambda/2$) e pelo polarizador (P) também permite o controle da intensidade do feixe de luz. Isso ocorre porque a irradiância, que é a potência irradiada por unidade de área, é proporcional ao

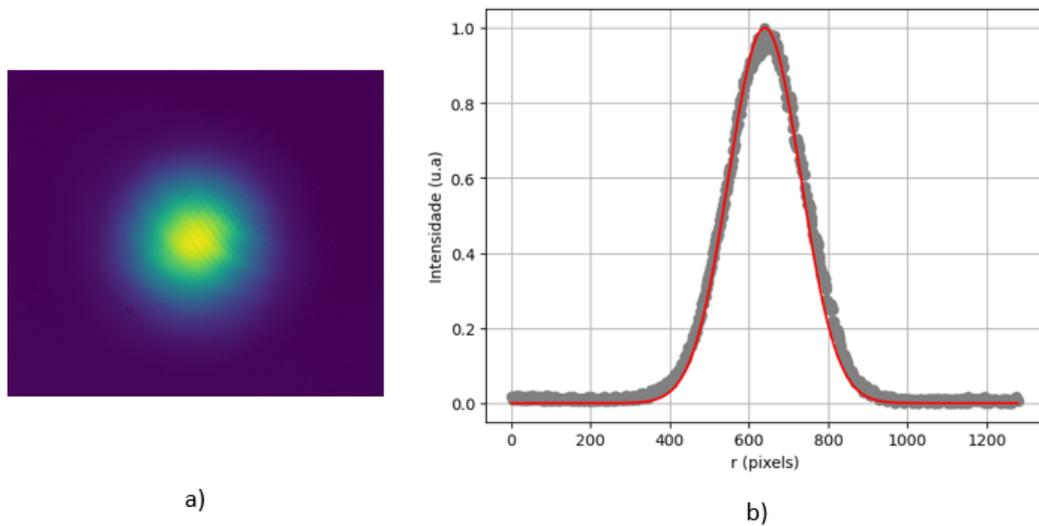
quadrado da amplitude do campo elétrico. Essa relação pode ser facilmente verificada ao analisarmos a expressão matemática da Lei de Malus, que descreve a intensidade transmitida (I_t) em função da intensidade da luz incidente (I_0) e do ângulo (θ) entre a polarização linear da luz e o eixo de transmissão do polarizador (BE; TEICH, 2007).

A equação $I_t = I_0 \cos^2 \theta$, indica que a intensidade transmitida pelo polarizador é máxima quando a polarização linear da luz está alinhada com o eixo de transmissão do polarizador (ângulo $\theta = 0$), resultando em um valor de $\cos^2(0)$ igual a 1. Nesse caso, a intensidade transmitida é igual à intensidade da luz incidente. Por outro lado, quando a polarização linear da luz é perpendicular ao eixo de transmissão do polarizador (ângulo $\theta = 90$ graus), o valor do cosseno é 0, o que resulta em uma intensidade transmitida de zero.

A luz, agora polarizada linearmente na direção vertical (direção do vetor diretor dos cristais líquidos), foi direcionada por um espelho (M1) em direção à lente (L1). Essa lente, juntamente com a lente L2, que possui distâncias focais de $f_1 = 50$ mm e $f_2 = 300$ mm, respectivamente, foi configurada como um telescópio para expandir o diâmetro do feixe laser. Inicialmente, o feixe tinha aproximadamente 1 mm de diâmetro e, após o redimensionamento, passou a ter 6 mm de diâmetro. No entanto, o feixe laser utilizado possui um perfil de intensidade Gaussiano. Esse tipo de feixe, como descrito na Figura 20 segue uma distribuição de intensidade que é descrita por uma função Gaussiana.

Para o nosso experimento, esse tipo de feixe não é ideal, uma vez que não proporciona uma intensidade uniforme em todos os pontos da área ativa do LC-SLM. Isso ocorre porque no perfil gaussiano a intensidade do feixe diminui à medida que nos afastamos do centro (SVELTO; HANNA et al., 2010). Para solucionar esse problema, utilizamos um filtro espacial (SF1), no qual permitimos que apenas a parte mais intensa do feixe passe através da abertura, cortando as bordas. Em seguida, expandimos novamente o feixe utilizando as lentes L3 e L4 ($f_3 = 50$ mm, $f_4 = 400$ mm), obtendo assim um feixe com um perfil espacial aproximadamente uniforme, que se assemelha a um feixe de perfil "Top-Hat". Esse tipo de feixe possui uma distribuição de intensidade plana e uniforme em sua região central, que lembra a forma de um chapéu com a parte superior plana. A intensidade da luz nesse tipo de feixe é praticamente constante na sua região central e cai abruptamente para zero fora dessa região (BELAY et al., 2022). Todavia, para que isso fosse possível, seria necessário um segundo filtro

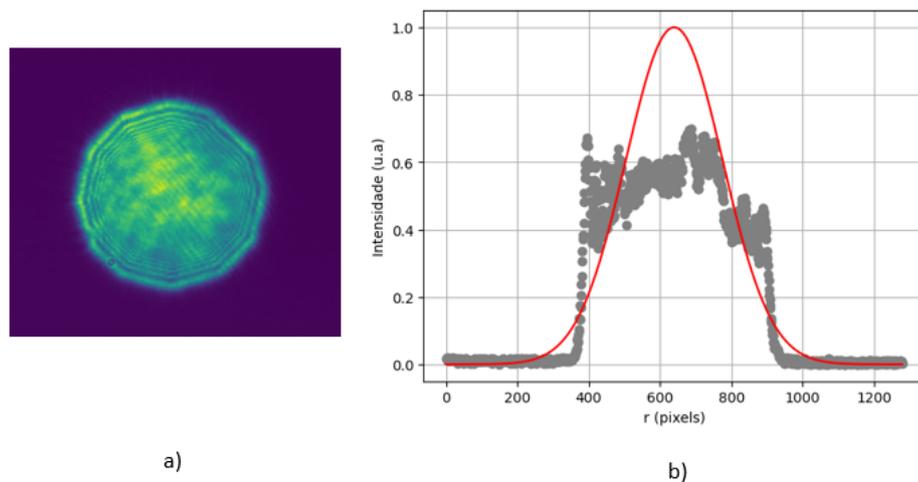
Figura 20 – Perfil espacial de um feixe gaussiano a) imagem capturada por uma câmera CCD b) perfil de intensidade do feixe.



Fonte: Autor, 2023

especial no caminho do feixe, no entanto como nosso laser possui baixa potência, optamos por utilizar apenas um filtro, assim obtemos o feixe como pode ser observado na Figura 21.

Figura 21 – Perfil espacial de um feixe Top hat a) imagem capturada por uma câmera CCD b) perfil de intensidade do feixe.



Fonte: Autor, 2023

Em seguida, a luz é direcionada pelo espelho M2 até um Modulador Espacial de Luz (LC-SLM) (ver o apêndice B) da Hamamatsu, com a tecnologia Lcos-SLM (modelo X13267) da Figura 22 . Na Tabela 1, estão listadas algumas especificações do

modulador utilizado. O SLM empregado em nosso sistema tem a capacidade de modular a fase do feixe incidente. Quando o feixe, viajando na direção z, incide sobre o modulador, ele adquire a fase de uma das imagens do conjunto de dados utilizadas para o treinamento. Essa fase é impressa no feixe de luz por meio de um padrão holográfico criado em um computador (HGC) (ver apêndice A). Em outras palavras, esse processo permite modular uma imagem digital no feixe de luz.

Figura 22 – Dispositivo LCOS-SLM somente de fase, modelo X13267, desenvolvido pela Hamamatsu.



Fonte: Autor, 2023

Tabela 1 – Características do LCOS-SLM

Número de pixels	792 x 600
Pixel pitch	12.5 μ m
Área efetiva	9.9 x 7.5 mm
Orientação do cristal líquido	Alinhamento paralelo
Níveis de sinais de entrada	256 para 8 bit
Níveis de modulação de fase	2 π (ou mais)

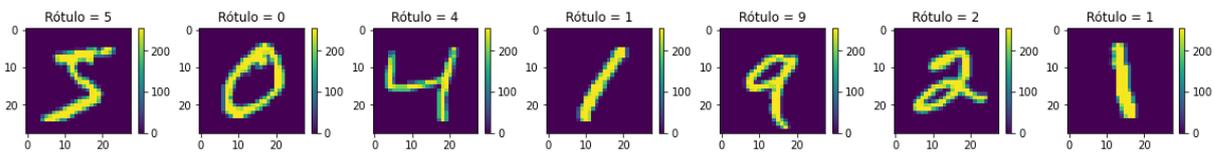
Fonte: Hamamatsu (2017)

3.2.1.1 Codificação do conjunto de dados no SLM

Utilizamos nosso sistema para realizar a classificação de imagens de dígitos manuscritos da base de dados MNIST (*Modified National Institute of Standards and Technology*). Essa base de dados é amplamente utilizada como referência ao implementar técnicas de aprendizado de máquina e métodos de reconhecimento de padrões, pois permite economizar tempo e esforço no pré-processamento e na formatação dos

dados. O MNIST consiste em um extenso conjunto de dados contendo imagens de dígitos manuscritos, cada uma com uma resolução de 28x28 pixels em tons de cinza. As imagens abrangem dígitos de 0 a 9, totalizando 70.000 exemplos. O conjunto de treinamento contém 60.000 imagens, enquanto o conjunto de teste possui 10.000 imagens. Cada imagem está devidamente rotulada com o dígito correspondente que representa. Portanto, temos um total de 10 classes de dígitos, variando de 0 a 9 (TEAM, 2020). Dessa forma, a tarefa realizada é a classificação de uma determinada imagem de um dígito manuscrito em uma das 10 classes que representam os valores inteiros de 0 a 9. Na Figura 23, apresentamos a representação visual dos dígitos manuscritos da base de dados MNIST, cada um deles devidamente rotulado.

Figura 23 – Exemplos do conjunto de dados do MNIST



Fonte: Autor, 2023

Para utilizar a base de dados no nosso sistema experimental, realizamos uma filtragem, selecionando apenas 2000 imagens para o treinamento e 400 imagens para o teste. Em seguida, realizamos uma etapa de pré-processamento utilizando a linguagem de programação *Python*. Durante esse processo, cada uma das imagens foi ampliada para aproximadamente sete vezes o tamanho original. Após a ampliação, foi realizada a normalização de todas as imagens. Isso significa que ajustamos os valores de intensidade de cada pixel para um intervalo entre 0 e 1, garantindo uma escala consistente em todas as imagens. Após o pré-processamento, cada uma das imagens selecionadas foi codificada em um holograma de fase, que será carregado no LC-SLM (Modulador Espacial de Luz). Para realizar essa codificação, definimos uma função de amplitude complexa para cada imagem, que é descrita pela equação:

$$\mathbf{E}_q(x, y) = \sqrt{\mathbf{x}_q(x, y)} \exp(j\pi \mathbf{x}_q(x, y)) \quad (61)$$

Aqui, $\mathbf{x}_q(x, y)$ representa a matriz de intensidade que constitui a imagem do MNIST, onde q varia de 1 a $N = 1000$, correspondendo ao número total de imagens selecionadas para o treinamento e para o teste temos que q varia de 1 a $N = 200$. A amplitude e a fase desse campo complexo são então registradas a partir da função de

transmitância $H_q(x, y) = \exp[j\psi(a_q, \phi_q)]$, que define um holograma de fase gerado por computador. Com isso, $a_q = \sqrt{\mathbf{x}_q(x, y)}$ e $\phi_q = \exp[j\pi\mathbf{x}_q(x, y)]$. Detalhes sobre essa função de transmitância podem ser encontrados na subseção A.0.2 .

3.2.1.2 Descrição do processamento óptico dos dados e treinamento

Após a construção dos hologramas, cada um deles é carregado individualmente no SLM. O SLM é iluminado pela luz laser, e a imagem modulada em fase é propagada até a lente L5. Nesse ponto, ocorre a transformada de Fourier do sinal, levando-o para o domínio das frequências espaciais. Em seguida, a luz passa pelo filtro espacial SF2, que permite apenas a ordem de difração +1 passar, bloqueando as demais.

Após a filtragem espacial, a imagem é reconstruída e coletada utilizando a lente L6, onde ocorre uma nova transformada para o domínio espacial. Nesse momento, consideramos o plano de foco de L6 como o plano de entrada para um sistema de processamento óptico, mostrado em detalhes na Figura 24. Esse sistema é composto por duas lentes, L7 e L8, ambas com a mesma distância focal, além de um difusor D. O objetivo desse sistema é projetar os dados em um espaço dimensionalmente maior, ou seja, realizar uma transição do domínio 2D para o 3D. Isso é alcançado por meio de uma projeção aleatória gerada pelo difusor D. O difusor introduz uma dispersão na luz, criando uma mistura de diferentes ângulos de propagação e, conseqüentemente, uma expansão do espaço no qual os dados são representados (SAADE et al., 2016).

Matematicamente, o campo no plano de entrada é representado por:

$$E_q(x, y) = E_{0q}(x, y) \exp[j\phi_q(x, y)] \quad (62)$$

Após passar pela lente L7, o campo de entrada é transformado para o domínio de Fourier e é representado da seguinte forma:

$$F_q(u, v) = G_{0q}(u, v) \exp[j\phi_q(u, v)] \quad (63)$$

Onde u e v são as respectivas frequências espaciais. Como mencionado anteriormente, o difusor D atua como um meio de dispersão, introduzindo um termo de fase aleatória dado por:

$$d(x, y) = \exp[j\delta(x, y)] \quad (64)$$

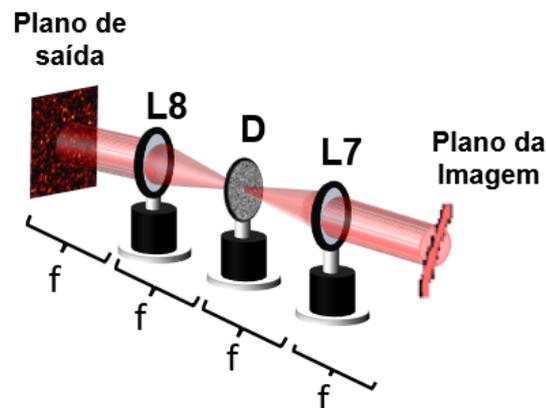
onde $\delta(x, y)$ representa a fase aleatória. Dessa forma, o campo após o difusor é dado por:

$$F'q(u, v) = G_{0q}(u, v) \exp[j(\phi_q(u, v) + \delta(x, y))] \quad (65)$$

A lente L8 realiza novamente uma transformada de Fourier, trazendo o campo de volta para o domínio espacial, onde no foco temos a imagem que exibe um padrão de speckle que corresponde à projeção aleatória da imagem que está em seu respectivo plano, como mostrado na Figura 24. A intensidade desse padrão é então registrada por uma câmera CCD de alta resolução (1280x1024 pixels).

$$I_q(x, y) = |FFT^{-1}(F'_q(u, v))|^2 \quad (66)$$

Figura 24 – Esquema do processamento óptico do sistema.



Fonte: Autor, 2023

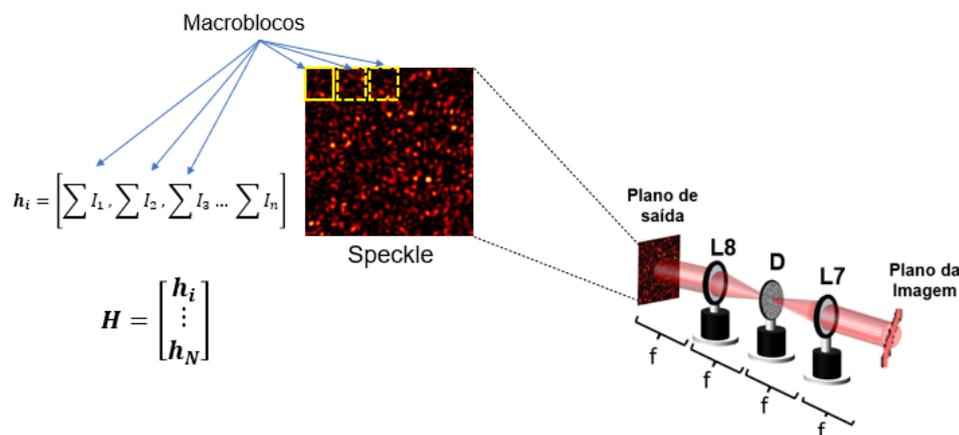
Para reduzir os efeitos de *aliasing* que são efeitos que ocorrem quando as frequências de um sinal não são corretamente representadas devido a uma taxa de amostragem inadequada. Ou seja, se a frequência espacial da imagem obtida é alta e a taxa de amostragem não é suficientemente elevada, as informações de alta frequência podem não ser corretamente representadas na saída do sistema. Assim, ao realizar a transformada de Fourier óptica, as frequências espaciais da imagem são mapeadas em frequências no plano focal. Se a taxa de amostragem não for adequada,

algumas dessas frequências espaciais podem se sobrepor, levando a uma perda de informações e, conseqüentemente, ao *aliasing* (GOODMAN, 2005).

Para contornar esses efeitos nas imagens adquiridas, optamos por selecionar uma região de 600 x 600 pixels em cada imagem. Em seguida, mapeamos cada pixel em macroblocos de tamanho $M \times M$, onde M varia de 2 a 6. O valor de intensidade atribuído a cada macrobloco corresponde à média das intensidades dos pixels contidos dentro dele.

No processo de treinamento do sistema, realizamos a vetorização de cada imagem, ou seja, mapeamos a imagem em um vetor linha. A dimensão do vetor depende do tamanho do macrobloco escolhido. Em seguida, armazenamos esses vetores em um computador. Cada vetor é empilhado para formar uma matriz, denotada por H , assim como mostrado na Figura 25. Com a matriz H formada, prosseguimos com o treinamento do sistema. Esse processo envolve o cálculo dos pesos de saída β , conforme descrito na equação 10.

Figura 25 – Esquema do processamento óptico do sistema.



Fonte: Autor, 2023

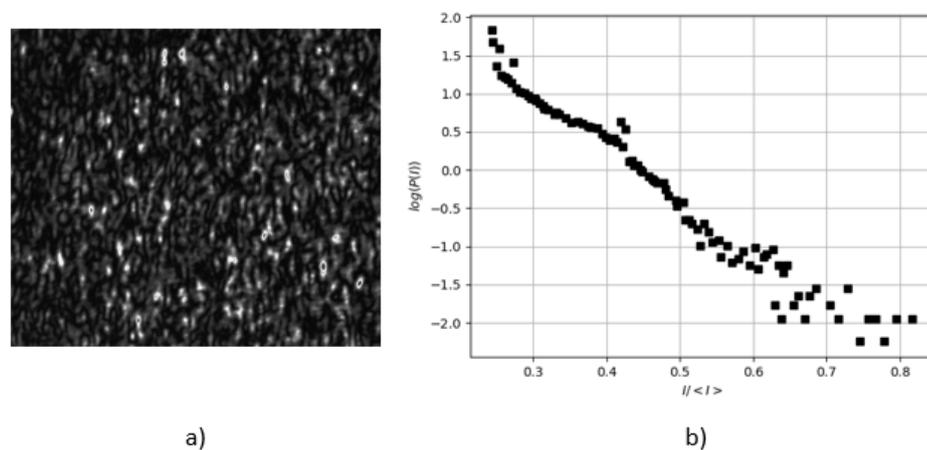
3.3 RESULTADOS

3.3.1 Caracterização dos padrões de speckles

Os padrões de speckles foram gerados utilizando um difusor de luz transparente com um ângulo de difusão de 1 grau. Esse difusor foi iluminado pela luz focalizada

proveniente da lente (L7) de um padrão de luz que representa um dígito manuscrito do MNIST. O padrão de speckle, formado no plano focal da lente (L8), transporta a informação do padrão de luz gerado no plano da imagem, conforme ilustrado na Figura 24. A caracterização do padrão de speckle envolveu uma análise estatística de primeira ordem, conforme descrito na seção 2.4.2. A seguir, apresentamos as funções densidade de probabilidade (FDP) relativas aos speckles formados a partir de diferentes padrões de luz no plano da imagem.

Figura 26 – Caracterização do padrão de speckle gerado pelo difusor. a) Imagem capturada pela CCD onde o padrão de luz no plano da imagem representa o dígito 0, b) Gráfico da densidade probabilidade da intensidade deste padrão.



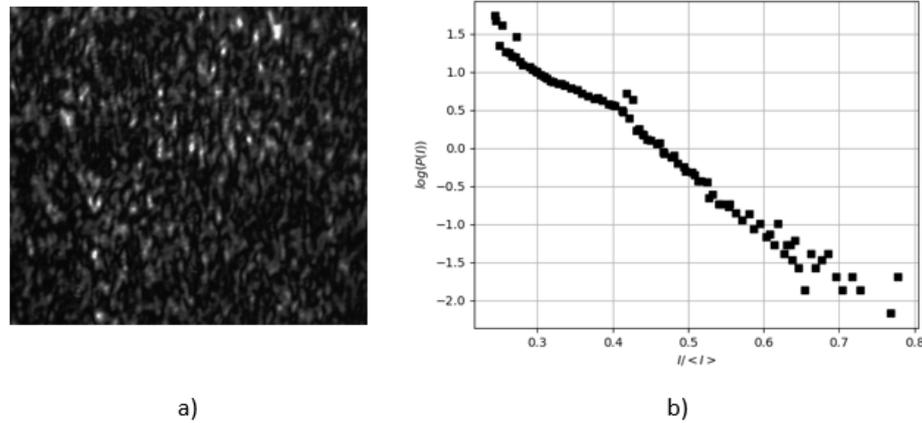
Fonte: Autor, 2023

Ao analisarmos os gráficos das figuras acima, notamos que, independentemente dos padrões de luz no plano imagem do sistema de processamento 4f, a curva de distribuição de intensidade mantém a mesma estatística de Rayleigh, caracterizada por uma exponencial decrescente.

3.3.2 Avaliação da classificação

Para realizar a classificação, empregamos as imagens de teste provenientes do conjunto de dados do MNIST. Estas são imagens que não foram utilizadas durante a etapa de treinamento, representando dados desconhecidos para os quais desejamos que o sistema determine a classe ou rótulo correspondente. O processo óptico segue o mesmo procedimento adotado no treinamento, envolvendo a vetorização de cada imagem de padrão de speckle capturada pela CCD. Em seguida, empregamos

Figura 27 – Caracterização do padrão de speckle gerado pelo difusor. a) Imagem capturada pela CCD onde o padrão de luz no plano da imagem representa o dígito 6, b) Gráfico da densidade probabilidade da intensidade deste padrão.



Fonte: Autor, 2023

o computador para executar a classificação, onde realizamos a multiplicação matricial da matriz de pesos otimizados, obtida durante o treinamento, com cada vetor correspondente a um padrão de speckle cuja imagem a ser classificada está codificada.

Como métrica de avaliação do sistema, utilizamos a acurácia que representa a taxa de previsões corretas feitas por um modelo em relação ao número total de exemplos avaliados, ou seja:

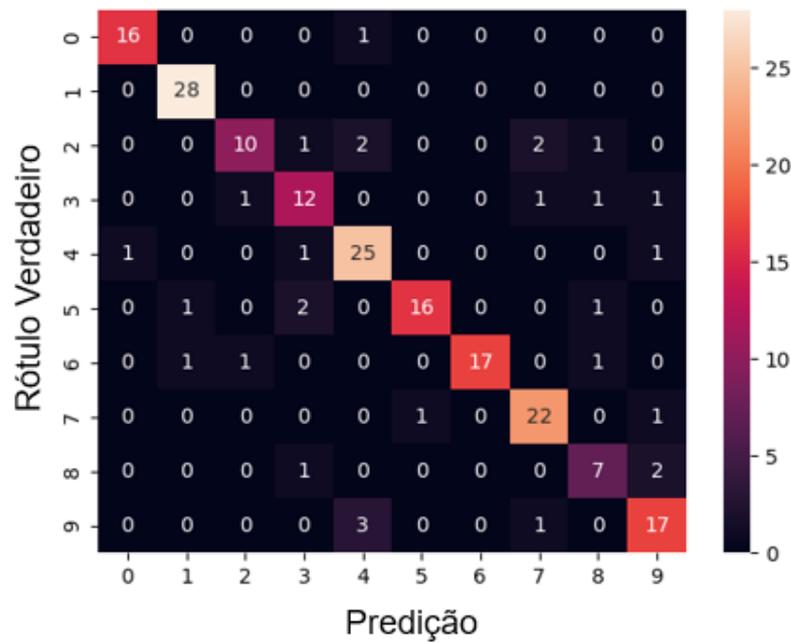
$$\text{Acurácia} = \left(\frac{\text{Número de previsões corretas}}{\text{total de previsões}} \right) \times 100\% \quad (67)$$

Outra métrica que utilizamos para avaliar o desempenho da classificação foi uma tabela denominada de Matriz de Confusão. Ela mostra quantas vezes o modelo acertou ou errou para cada categoria.

Na Figura 28, está representada a Matriz de Confusão da classificação realizada pelo sistema óptico. Nessa tabela, as colunas indicam os rótulos esperados, enquanto as linhas mostram as previsões feitas pelo sistema. Portanto, os números que foram corretamente previstos estão dispostos na diagonal da tabela. Nesta tabela foram classificadas 200 imagens de teste. A partir daí calculamos a acurácia desta classificação, cujo resultado foi de 85%.

Consideramos esse resultado como excelente, dada a quantidade de imagens utilizadas no treinamento (1000 imagens) e as imagens usadas na classificação (200

Figura 28 – Matriz de Confusão da classificação do sistema óptico que utiliza um meio espalhador.



Fonte: Autor, 2023

imagens). No entanto, desejamos investigar se manipulando as propriedades estatísticas dos padrões de speckles gerados durante o processamento, esse resultado seria impactado positiva ou negativamente. Para isso, exploraremos outro sistema óptico capaz de realizar essas modificações nos padrões, e faremos uso de técnicas de holografia, essenciais nesse processo. Para isso é necessário simular numericamente padrões de speckles de estatística não-Rayleigh e codificar as informações do conjunto de dados nestes speckles.

3.4 INFLUÊNCIA DE SPECKLES COM ESTATÍSTICA DE INTENSIDADE MODIFICADA NA ACURÁCIA DA CLASSIFICAÇÃO

3.4.1 Padrões de speckles gerados numericamente

Para simular padrões de speckle numericamente, nos baseamos nos conceitos descrito na seção 2.4. O processo começa pela criação de uma matriz com dimensões $M \times M$, que representa a distribuição de fase aleatória em um intervalo de $[-\pi, \pi]$. Essa distribuição ocorre quando uma frente de onda incide sobre uma su-

perfície rugosa, que pode ser tanto reflexiva quanto transmissiva (que foi discutido em mais detalhes na subseção 2.4.1). Para representar o campo após a reflexão ou transmissão, usamos a equação

$$E_{in} = E_0(x, y)e^{j\phi(x,y)} \quad (68)$$

onde $E_0(x, y)$ é a amplitude do campo assumida como sendo 1 e $\phi(x, y)$ é a matriz de fase gerada. Em seguida, o campo é colocado no plano focal de uma lente, onde é realizada uma transformada de Fourier. Para realizar a simulação numérica, utilizamos a função de transformada rápida de Fourier (FFT) sobre o campo, representada pelo operador \mathcal{F} (SPILSBURY; EUCEDA, 2016). Após a transformada, um filtro espacial circular passa-baixa é aplicado ao campo, com um diâmetro D , dada por $D(x, y)$. Esse filtro além de filtrar as frequências espaciais altas, fornece um pouco de correlação à fase que é totalmente aleatória. O campo resultante após o filtro é dado por:

$$E'_{in} = \mathcal{F}\{e^{j\phi(x,y)}\}D(x, y) \quad (69)$$

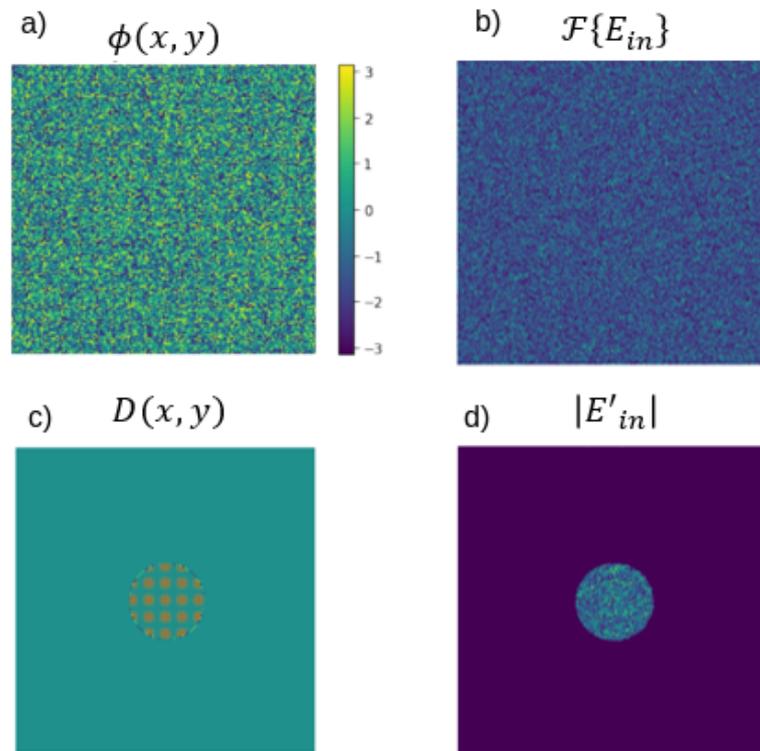
Em seguida, o campo de saída é calculado pela transformada inversa do campo de entrada, resultando em:

$$E_{out} = \mathcal{F}^{-1}\{E'_{in}\} \quad (70)$$

Esse processo é equivalente a colocar outra lente no plano confocal ao filtro espacial, gerando o campo de saída no foco da mesma.

Na Figura 29, é possível observar com clareza a forma de cada matriz utilizada na simulação de padrões de speckle. É importante mencionar que o tamanho de cada speckle é determinado pela razão entre as dimensões da matriz e o diâmetro do filtro passa-baixa utilizado, ou seja, M/D . No entanto, é necessário ter em mente o critério de Nyquist, que estabelece uma relação entre a taxa de amostragem e a frequência máxima que pode ser detectada sem perda de informação (GOODMAN; SUTTON, 1996). Em outras palavras, é preciso garantir que a taxa de amostragem seja suficientemente alta para que todas as informações relevantes sejam capturadas. Caso contrário, pode haver perda de detalhes e distorções na imagem gerada. Isso pode ser assegurado ao seguir o teorema de Nyquist, que estabelece que a taxa de amostragem deve ser

Figura 29 – Matrizes utilizada na simulação de padrões de speckle. a) Matriz de fase totalmente aleatória, b) Transformada de Fourier do campo de entrada, c) Filtro espacial numérico, d) Campo resultante após o filtro.



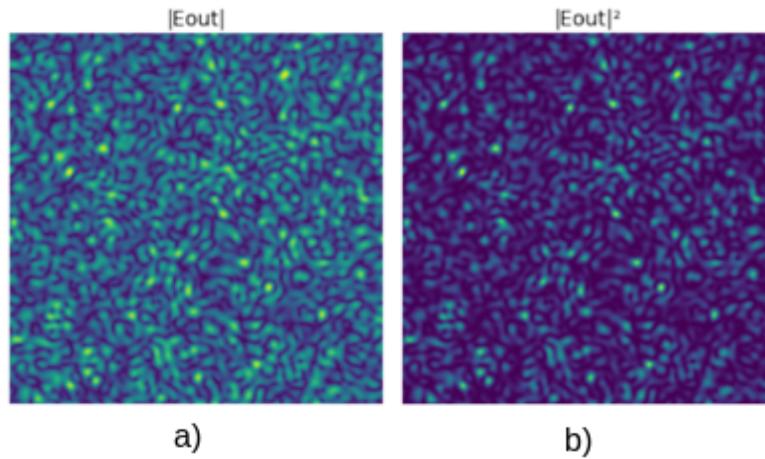
Fonte: Autor, 2023

pelo menos o dobro da maior frequência presente no sinal (GOODMAN, 2005). Logo, atendemos esse critério se a relação M/D for maior ou igual a 2.

A imagem apresentada na Figura 30 exhibe o módulo do padrão de speckle, bem como sua intensidade $I = |E_{out}|^2$. É notável que nossas simulações numéricas conseguem produzir padrões de speckle completamente desenvolvidos. Essa afirmação é reforçada pela análise do padrão de intensidade apresentado na Figura 31. Observa-se que o padrão gerado numericamente segue as propriedades estatísticas descritas por (GOODMAN, 2007).

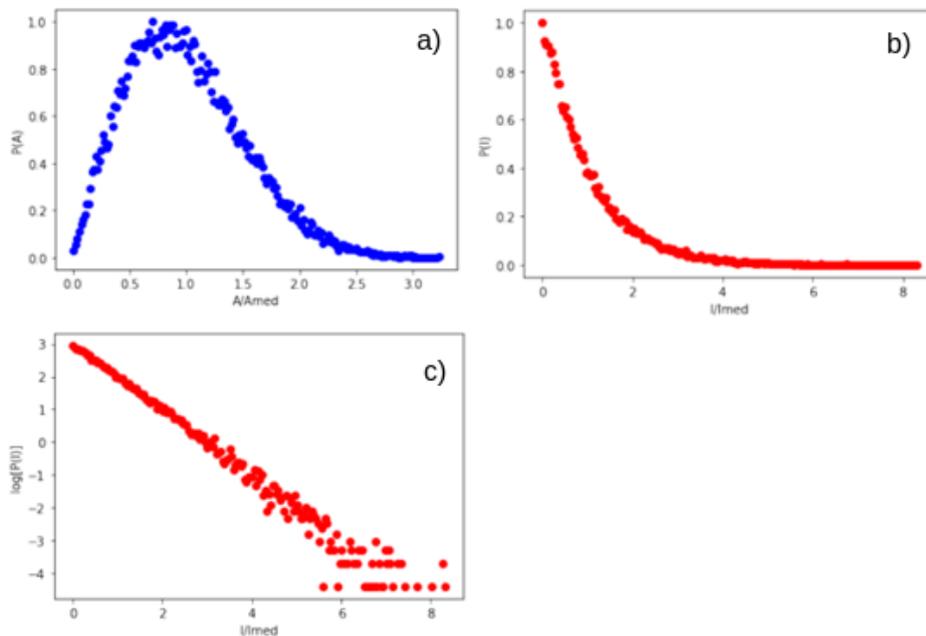
A análise dos gráficos revela que os speckles gerados numericamente seguem a estatística de Rayleigh, cuja função de distribuição da intensidade possui característica de uma exponencial negativa. Para manipular esta estatística utilizaremos o método descrito na seção 2.4.3.2, onde a partir da utilização da função de densidade de probabilidade de Weibull nos desviaremos da estatística de Rayleigh e criaremos speckles super e sub Rayleigh. Essa manipulação é feita elevando o módulo da am-

Figura 30 – Padrões de speckles gerados numericamente. a) Módulo do campo complexo, b) Padrão de intensidade.



Fonte: Autor, 2023

Figura 31 – Caracterização dos padrões de speckle gerados numericamente. a) Distribuição de amplitude, b) Distribuição de intensidade em escala linear-linear, c) Distribuição de intensidade em escala mono-log.

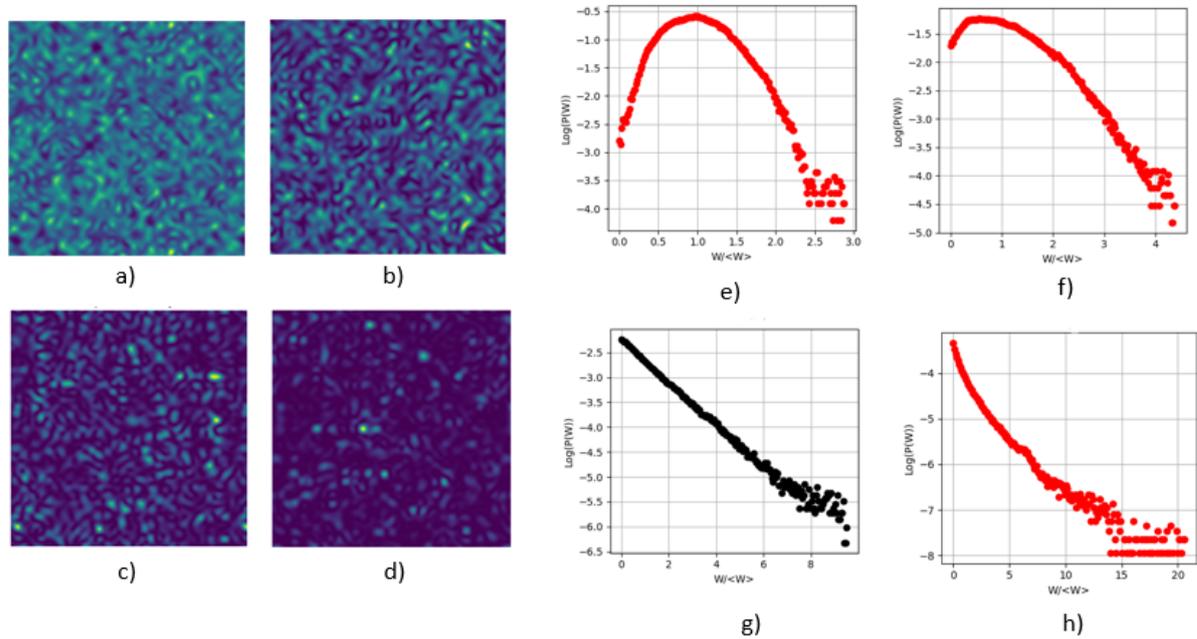


Fonte: Autor, 2023

plitude do campo de saída a potencia 2α . Onde, α corresponde a qualquer valor real positivo. Ou seja, $E_W = |E_{out}|^{2\alpha}$. A partir dessa manipulação na amplitude do campo nos desviamos da estatística de Rayleigh.

Na Figura 32 vemos os padrões de speckles gerados numericamente modificados a partir desta relação. Nas Figuras 32a e 32b, observamos padrões de speckles

Figura 32 – Speckles modificados estatisticamente, a) $\alpha = 0,2$, b) $\alpha = 0,5$, c) $\alpha = 1$ d) $\alpha = 1,5$, gráficos da densidade de probabilidade da intensidade e) $\alpha = 0,2$, f) $\alpha = 0,5$, g) $\alpha = 1$ h) $\alpha = 1,5$



Fonte: Autor, 2023

gerados com $\alpha = 0,2$ e $\alpha = 0,5$, respectivamente. Os gráficos nas Figuras 32e e 32f mostram as funções de densidade de probabilidade desses padrões, indicando um afastamento da estatística de Rayleigh. Conforme ilustrado nas imagens, esses padrões exibem uma distribuição de intensidade mais homogênea e um contraste inferior a 1 quando comparados ao padrão da Figura 32c, com $\alpha = 1$, como evidenciado pelo gráfico correspondente na Figura 32g, que mantém a estatística de Rayleigh. Na Figura 32d, apresenta-se um padrão de speckle gerado com $\alpha = 1,5$, caracterizado pelo gráfico na Figura 32h. Nesse caso, o padrão possui um contraste superior a 1.

3.4.1.1 Codificação do conjunto de dados

Para codificar cada imagem do conjunto de dados nos speckles gerados numericamente, definimos uma função de amplitude complexa para cada imagem, conforme a Equação 61. Posteriormente, a amplitude e a fase relacionadas a essa função são codificadas no campo de speckle, assumindo a forma:

$$E_{in} = E_0(x, y) \sqrt{\mathbf{x}_q(x, y)} e^{j(\phi(x, y) - \pi \mathbf{x}_q(x, y))} \quad (71)$$

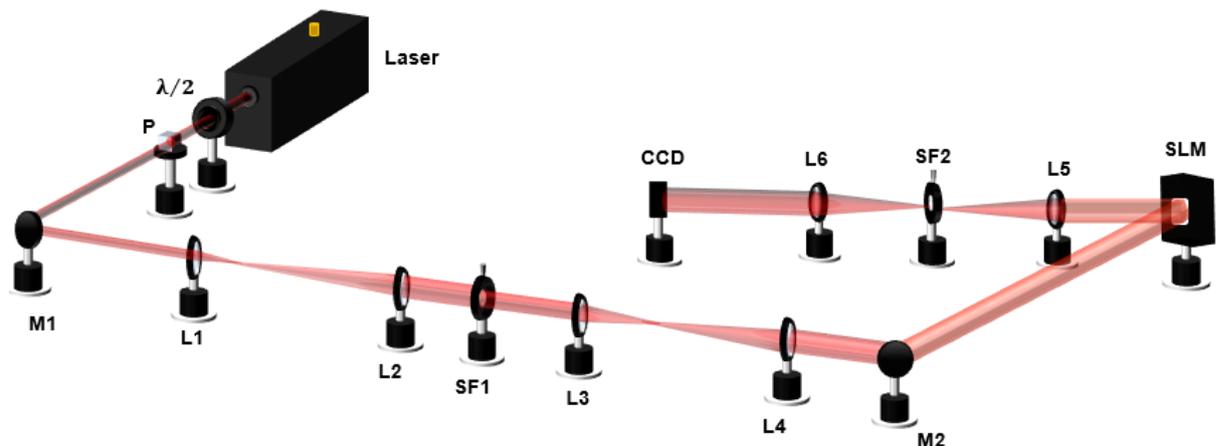
Dessa forma, a amplitude e a fase do campo de saída resultante, conforme des-

crita na Equação 70, são registradas através da função de transmitância $H_q(x, y) = \exp[j\psi(a_q, \phi_q)]$, que define um holograma de fase gerado por computador (ver apêndice A). Esse holograma é então “impresso” no Modulador Espacial de Luz por meio de um computador, sendo responsável por gerar as projeções aleatórias dos dados com estatísticas de intensidade modificadas no sistema óptico.

3.4.2 Setup experimental

Na Figura 33 mostramos o *setup* experimental utilizado para criar uma projeção aleatória dos dados com estatística de intensidade modificada. Este setup se assemelha ao anterior, com a diferença de que as projeções são geradas numericamente e codificadas em hologramas de fase que são enviados para o Modulador SLM.

Figura 33 – Setup experimental para classificar imagens a partir de speckles com estatísticas de intensidade não Rayleigh.



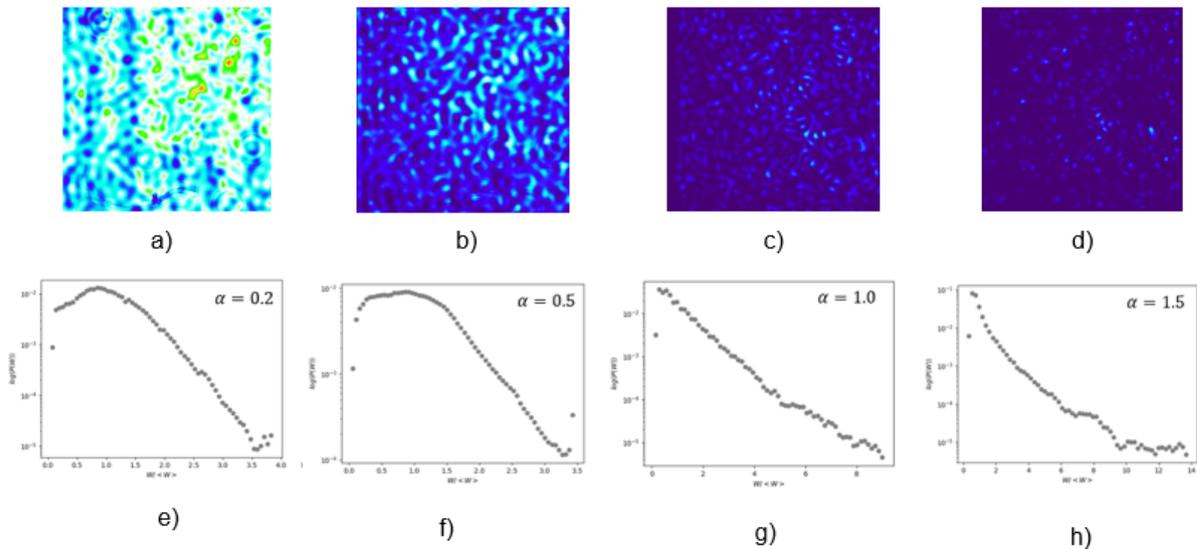
Fonte: Autor, 2023

3.5 RESULTADOS

Os resultados apresentados na Figura 34 demonstram os speckles com estatísticas de intensidade modificadas, gerados experimentalmente. As imagens obtidas na CCD para os speckles gerados com diferentes valores de α são representadas nas subfiguras a), b), c) e d), onde α assume os valores 0,2, 0,5, 1 e 1,5, respectivamente. Os gráficos de densidade de probabilidade correspondentes estão ilustrados nas subfiguras e), f), g) e h), associadas aos mesmos valores de α . Esses resulta-

dos evidenciam a influência direta do parâmetro α na distribuição de intensidade dos speckles, mostrando o uma visão clara das modificações introduzidas nas estatísticas desses padrões.

Figura 34 – Speckles com estatística de intensidade modificada gerados experimentalmente. Imagens obtidas na CCD para speckle gerado com a) $\alpha = 0,2$, b) $\alpha = 0,5$, c) $\alpha = 1$, d) $\alpha = 1,5$.. Gráficos da densidade de probabilidade e) $\alpha = 0,2$, f) $\alpha = 0,5$, g) $\alpha = 1$, h) $\alpha = 1,5$.



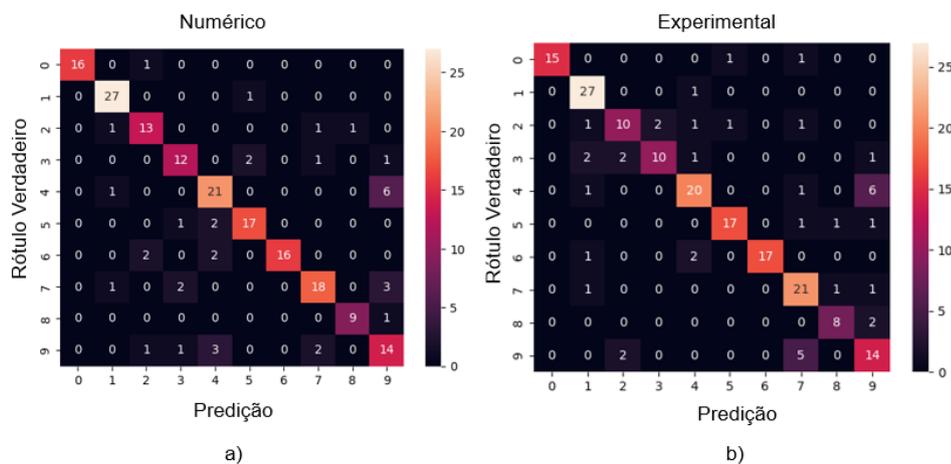
Fonte: Autor, 2023

Agora, iremos avaliar como o valor da acurácia se comporta ao manipular a estatística dos speckles Rayleigh. Para essa análise, utilizamos 1000 imagens durante o treinamento do sistema, seguindo o mesmo conjunto de imagens empregado no treinamento do experimento anterior (seção 3.2). Além disso, realizamos a classificação de 200 imagens de teste, referentes ao mesmo conjunto. Inicialmente, realizamos projeções aleatórias correspondentes às imagens de teste com $\alpha = 1$, o que equivale a speckles de Rayleigh. As imagens foram capturadas por uma câmera CCD. Tanto na etapa de treinamento quanto na etapa de teste, adotamos a mesma metodologia, que compreende o mapeamento de cada pixel das imagens da CCD em macroblocos de tamanho $M \times M$, vetorização, cálculo da matriz de pesos β e classificação, conforme descrito na seção 3.2.

A Figura 35 apresenta as matrizes de confusão numérica e experimental dessa classificação. Os resultados revelam que, na abordagem numérica, alcançamos uma acurácia de 80%. Já na classificação realizada experimentalmente, obtemos uma acurácia de 79.5%.

Acreditamos que a divergência entre o resultado experimental de 79.5% e o anterior de 85% pode ser atribuída ao fato de que o difusor espalha uma quantidade significativamente maior de luz do que o holograma que projetamos. Isso resulta em uma maior quantidade de grãos de speckle (pontos de luz), favorecendo assim o processo de aprendizado.

Figura 35 – Matriz de confusão da classificação para projeções aleatórias geradas com $\alpha = 1$. Resultados a) Numérico b) Experimental.



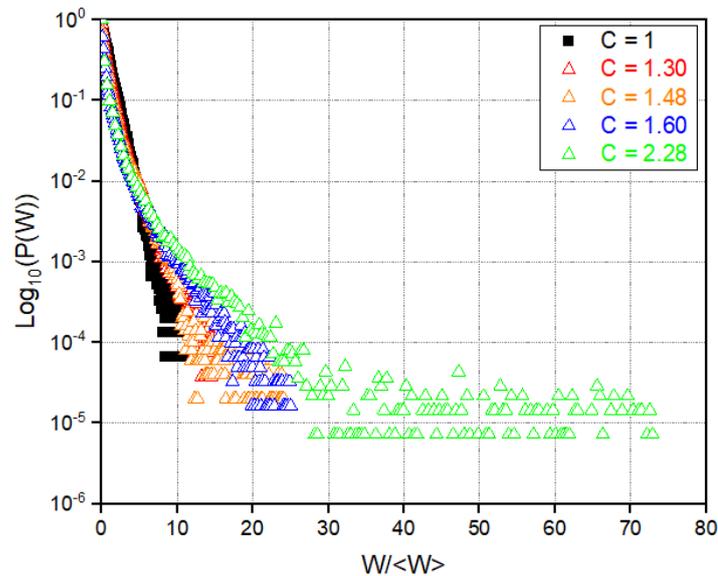
Fonte: Autor, 2023

A partir deste ponto, utilizaremos esse resultado como referência para comparar com as projeções aleatórias que possuem estatísticas não-Rayleigh. Em vez do parâmetro α como referência para a modificação dos padrões, faremos uso do cálculo do contraste. Essa escolha se dá pelo fato de o contraste ser um parâmetro mais direto e intuitivo para avaliar a diferença na distribuição de intensidade nos padrões de speckle. Na Figura 36, apresentamos o gráfico normalizado da distribuição de densidade de probabilidade de speckles super-Rayleigh ($C > 1$), utilizados na experiência. Cada curva exibida corresponde a um padrão com seu respectivo valor de contraste. Já na Figura 37, são apresentadas as curvas normalizadas da densidade de probabilidade de speckles sub-Rayleigh, com cada curva representando um valor de contraste menor que 1.

Como observado, à medida que o contraste aumenta, o decaimento exponencial torna-se mais lento em comparação com a curva de Rayleigh ($C = 1$). Em outras palavras, a partir do gráfico, percebemos que quanto maior o contraste, maior é a probabilidade de termos grãos de speckles com baixa intensidade nos padrões. Quando

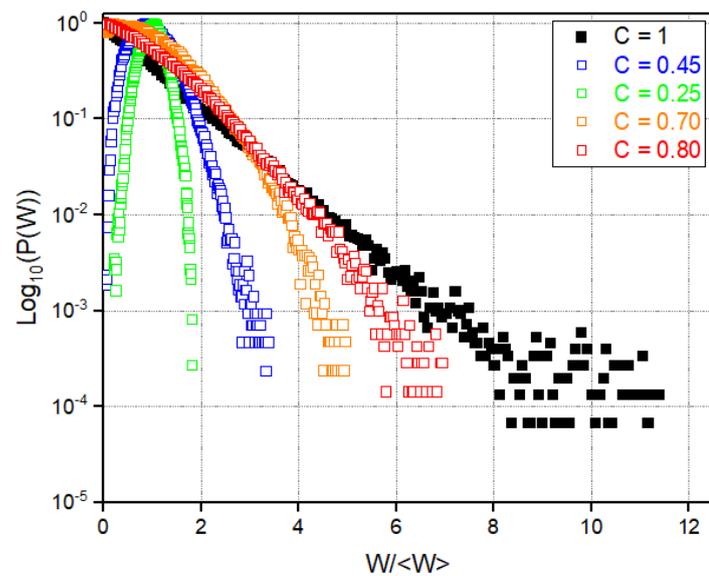
o contraste diminui, o decaimento exponencial se torna mais rápido. Portanto, quanto menor o contraste, maior é a probabilidade de termos grãos de speckles com altas intensidades.

Figura 36 – Gráfico normalizado da função de densidade de probabilidade da intensidade para speckles super Rayleigh.



Fonte: Autor, 2023

Figura 37 – Gráfico normalizado da função de densidade de probabilidade da intensidade para speckles sub Rayleigh.



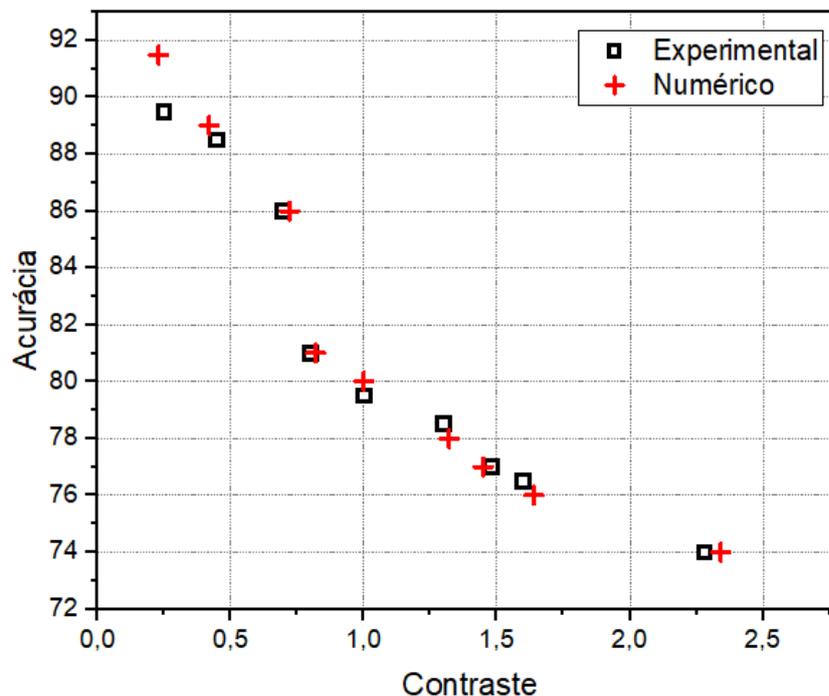
Fonte: Autor, 2023

Na Figura 38, apresentamos os resultados numéricos e experimentais da acurácia na classificação das imagens do conjunto de dados do MNIST em 10 classes.

Demonstramos como esse valor é afetado pelo contraste dos speckles gerados na projeção aleatória dos dados. No gráfico, observamos que à medida que o contraste das imagens diminui, o valor da acurácia tende a aumentar em relação ao resultado obtido com speckles de contraste igual a 1. Além disso, notamos que o valor da acurácia diminui em relação aos speckles de Rayleigh conforme o contraste das imagens aumenta.

Acreditamos que, nas imagens dos speckles sub-Rayleigh, o contraste, que é a relação entre o desvio padrão da intensidade dividido pelo valor médio da intensidade, é crucial. Para baixo contraste, a variação da intensidade em relação ao valor médio é pequena. Portanto, esse tipo de padrão apresenta uma certa saturação nos valores de intensidade de cada grão de speckle. Essa saturação introduz uma não linearidade na etapa de aprendizado e teste das imagens, melhorando a taxa de acerto do sistema na classificação. Já nos padrões de alto contraste, observa-se uma maior variabilidade da intensidade em relação ao valor médio. Em outras palavras, é mais provável termos grãos de speckle com baixas intensidades nesse cenário, o que resulta em uma perda de informações que impacta negativamente na acurácia.

Figura 38 – Gráfico da acurácia em função do contraste.

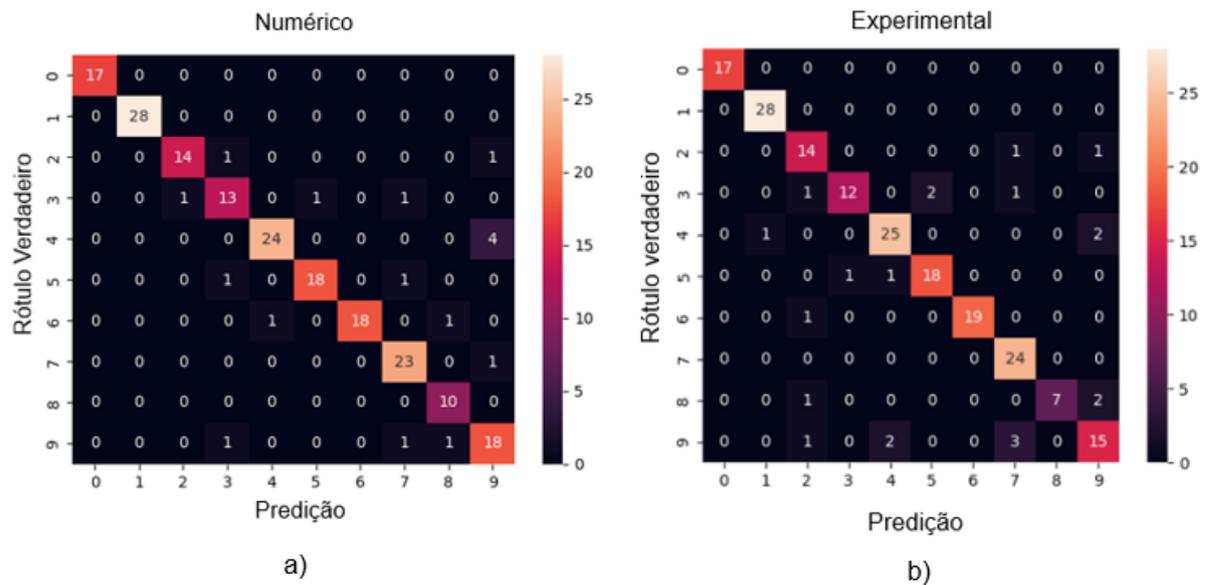


Fonte: Autor, 2023

Dessa forma, na Figura 39, apresentamos a matriz de confusão numérica e experi-

mental referente à classificação das imagens com projeções aleatórias sub-Rayleigh. Nota-se que para o menor valor de contraste, obtemos um resultado numérico de 91.5% de acurácia e um resultado experimental de 89.5%.

Figura 39 – Matriz de confusão para speckles sub-Rayleigh com $C= 0.25$. Resultados a) Numérico b) Experimental .



Fonte: Autor, 2023

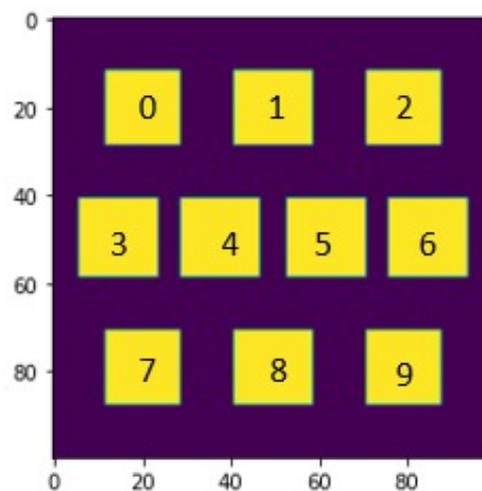
3.5.1 Sistema com predição óptica

Apesar dos bons resultados obtidos, nosso interesse é eliminar a dependência desse sistema do uso do computador na etapa de classificação. Atualmente, essa dependência ocorre devido à necessidade do computador realizar a multiplicação matricial entre a matriz de pesos β e as amostras de teste para definir a classificação correta. Nesse contexto, decidimos que toda a etapa de treinamento do sistema seria realizada via computador, enquanto o sistema óptico seria utilizado exclusivamente para a predição das imagens de entrada, de forma similar ao trabalho de (WU; ZHANG, 2022).

Nesse contexto, treinamos o sistema utilizando o método numérico para gerar projeções aleatórias das imagens, descrito na seção 3.4. Após obter essas projeções, vetorizamos as imagens resultantes e construímos a matriz H , que define as informações da camada oculta em um sistema de aprendizado de Máquinas de Aprendizado

Extremo. Durante o treinamento, calculamos os pesos de saída dessa rede, conforme a equação 10. No entanto, diferentemente do método anterior, definimos que os alvos seriam matrizes associadas a regiões específicas, cada uma correspondendo a um rótulo, como ilustrado na Figura 40. Essa abordagem nos permite obter pesos de saída β otimizados para esse tipo de saída específica.

Figura 40 – Matriz de alvos

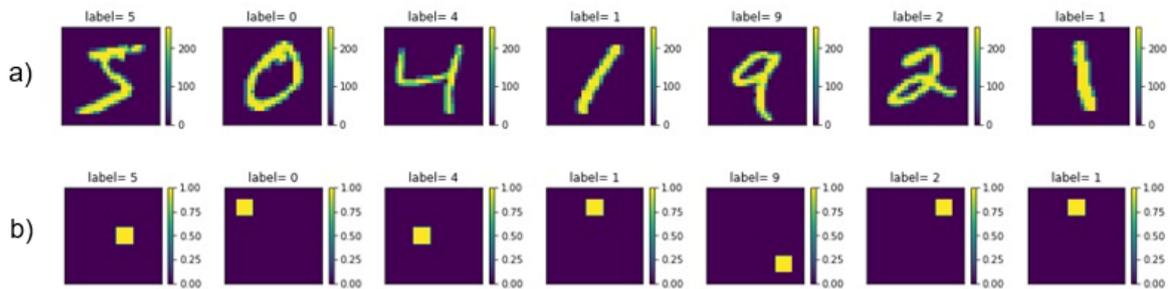


Fonte: Autor, 2023

Observamos na Figura 41 a associação de cada região ao seu dígito correspondente. O processo de classificação envolve a codificação, por meio de um holograma de fase gerado (conforme os métodos já descritos), da multiplicação matricial entre a matriz β é uma das imagens que compõem a matriz H de teste. A Figura 42 ilustra a configuração experimental utilizada e como ocorre essa etapa de predição óptica. Nesse processo, os hologramas de fase são enviados para o Modulador Espacial de Luz (SLM), que modula a fase do laser. A luz modulada é filtrada e capturada pela câmera CCD, proporcionando uma representação óptica do resultado da predição.

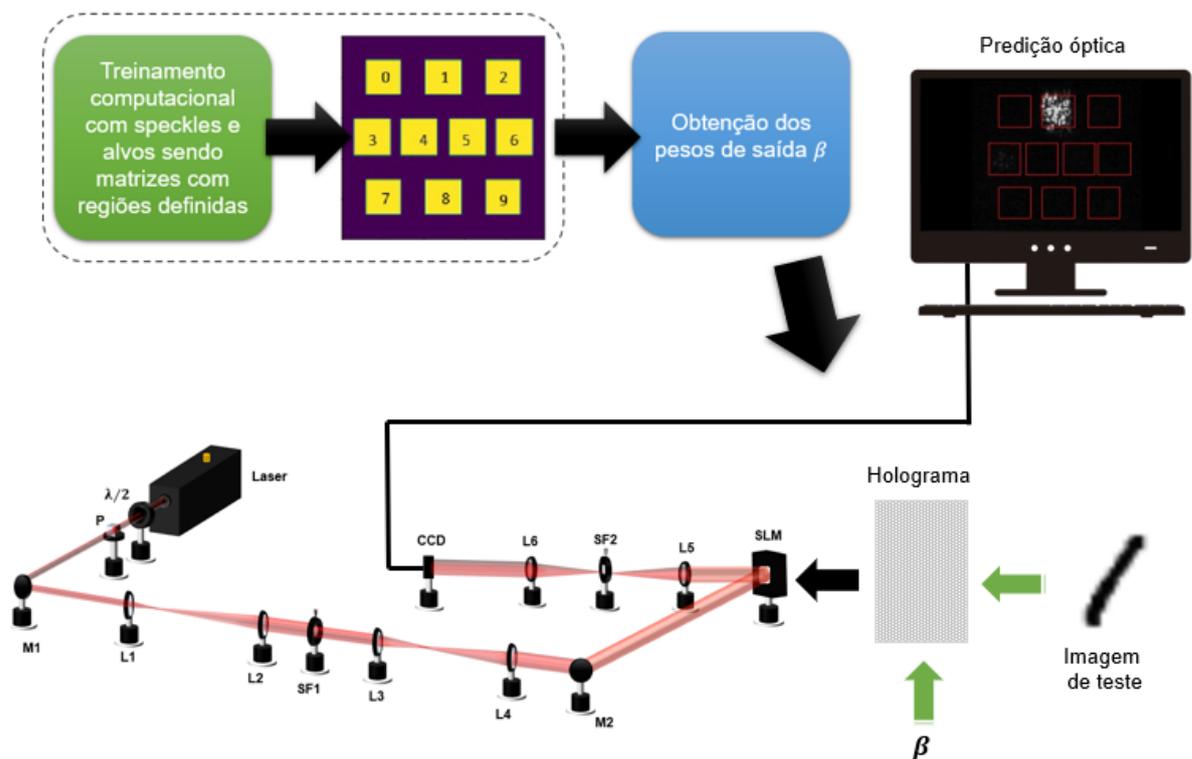
Na Figura 43, são apresentadas algumas imagens capturadas pela câmera CCD relacionadas à predição do sistema. A classificação óptica ocorre quando a região alvo associada ao dígito "1" apresenta maior intensidade. A imagem é correta quando a região de classificação com maior intensidade corresponde a esse dígito. Na Matriz de Confusão apresentada na Figura 44, é possível avaliar o desempenho do sistema

Figura 41 – Imagens associadas aos seus respectivos alvos. a) Dígitos manuscritos do MNIST b) Alvos relativo a cada dígito.



Fonte: Autor, 2023

Figura 42 – Esquema de predição óptica: após a obtenção da matriz β , ela é codificada em um holograma de fase junto com o dígito de teste por meio de uma multiplicação matricial. Em seguida, o holograma é enviado ao SLM, e através da modulação de fase, obtemos a predição óptica visualizada no monitor de um computador.

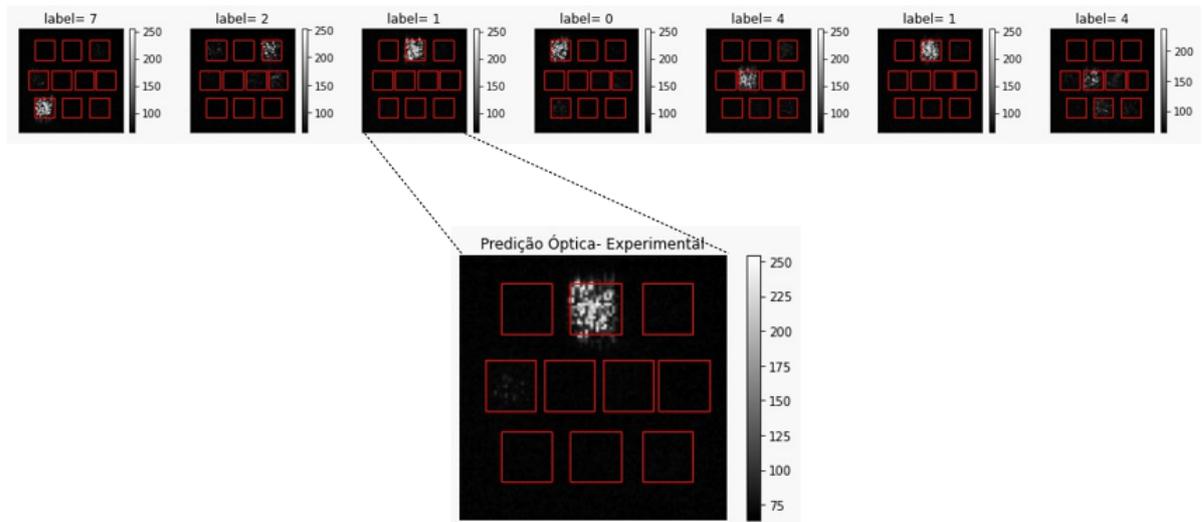


Fonte: Autor, 2023

na classificação de 500 imagens. A acurácia alcançada foi de 80,4%.

Apesar de ser um sistema óptico promissor, é importante destacar que ainda enfrentamos o desafio de codificar uma multiplicação matricial nos hologramas gerados para possibilitar a classificação óptica. Na busca por uma abordagem mais integrada, exploraremos, no próximo capítulo, duas técnicas de classificação óptica. Em vez de

Figura 43 – Predição óptica das imagens obtidas na CCD.



Fonte: Autor, 2023

Figura 44 – Matriz de confusão da predição óptica.



Fonte: Autor, 2023

utilizar uma multiplicação matricial óptica, que requereria vários componentes ópticos, optamos por implementar uma multiplicação óptica elemento por elemento, também conhecida como produto de Hadamard, seguida de uma soma. Essa abordagem visa simplificar o processo óptico de processamento em nosso sistema.

4 SISTEMA TOTALMENTE ÓPTICO DE CLASSIFICAÇÃO DE IMAGENS BASEADOS NO ALGORITMO DE MÁQUINA DE APRENDIZADO EXTREMO

4.1 INTRODUÇÃO

Neste capítulo, apresentaremos algumas técnicas de implementação de sistemas ópticos de classificação de imagens, as quais foram inspiradas no sistema de classificação discutido no capítulo anterior. Embora o sistema em questão execute operações ópticas de processamento, a etapa de classificação ainda depende de um computador. Apesar do custo computacional ser ínfimo após a fase óptica, nosso interesse reside em desenvolver um discriminador óptico que possa realizar a etapa de classificação de forma completamente óptica, eliminando a necessidade de um computador para essa finalidade. Em virtude disso, desenvolvemos dois métodos de classificação que podem ser aplicados em configurações ópticas simples e com poucos componentes.

4.2 MÉTODO DE CLASSIFICAÇÃO ÓPTICA 1

Com o objetivo de criar um sistema óptico eficiente para a classificação de imagens, fizemos a escolha estratégica de adotar também para esse caso o algoritmo de treinamento de Máquinas de Aprendizado Extremo- ELM, descrito na seção 2.2. Ao incorporar o ELM ao nosso sistema óptico, procedemos a adaptações essenciais para atender às nossas necessidades específicas.

Uma das modificações mais significativas consiste na configuração de matrizes de entradas e as matrizes alvo, em formato de matriz quadrada. Essa abordagem se mostrou fundamental para garantir que a matriz de pesos β também seja quadrada. Permitindo uma implementação mais eficiente em um sistema óptico.

4.2.1 Descrição do treinamento

Suponha que tenhamos um conjunto de imagens de treinamento, representado pelo conjunto X , com N exemplos. Cada imagem é representada por um vetor de características x_i de tamanho d , onde cada pixel é um número real entre 0 e 1. Como mencionado anteriormente, para facilitar a etapa de classificação óptica, trataremos os

vetores de entrada como matrizes bidimensionais quadradas. Faremos isso, esticando o vetor x_i por meio do produto de Kronecker com um vetor especial $\mathbf{1}_v$ que é composto por d valores iguais a 1. Esta operação retornará a matriz $\mathbf{K}(x_i)$ de tamanho $d \times d$ que é o vetor x_i esticado.

$$\mathbf{K}(x_i) = \mathbf{x}_i^T \otimes \mathbf{1}_v = \begin{bmatrix} x_1 \cdot 1_{v1} & x_2 \cdot 1_{v1} & x_3 \cdot 1_{v1} & \dots & x_d \cdot 1_{v1} \\ x_1 \cdot 1_{v2} & x_2 \cdot 1_{v2} & x_3 \cdot 1_{v2} & \dots & x_d \cdot 1_{v2} \\ x_1 \cdot 1_{v3} & x_2 \cdot 1_{v3} & x_3 \cdot 1_{v3} & \dots & x_d \cdot 1_{v3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 \cdot 1_{vd} & x_2 \cdot 1_{vd} & x_3 \cdot 1_{vd} & \dots & x_d \cdot 1_{vd} \end{bmatrix} \quad (72)$$

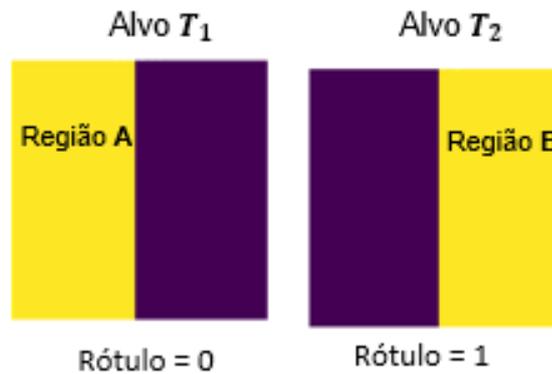
Dessa forma, o conjunto de treinamento \mathbf{X}_c será uma matriz de tamanho $Nd \times d$, formada pela concatenação das matrizes $\mathbf{K}(x_i)$. Assim como mostrado na equação 73.

$$\mathbf{X}_c = \begin{bmatrix} \mathbf{K}(x_1) \\ \mathbf{K}(x_2) \\ \mathbf{K}(x_3) \\ \vdots \\ \mathbf{K}(x_N) \end{bmatrix} \quad (73)$$

Além disso, cada matriz $\mathbf{K}(x_i)$ está associada a uma matriz alvo \mathbf{T}_i de tamanho $d \times d$, que contém diferentes regiões de intensidade relacionadas a uma classe ou rótulo, como mostrado na Figura 45. Essas matrizes alvo são representadas pela matriz \mathbf{T}_c , de tamanho $Nd \times d$, onde cada matriz alvo empilhada representa a classe correspondente a uma imagem de treinamento.

$$\mathbf{T}_c = \begin{bmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \\ \mathbf{T}_2 \\ \vdots \\ \mathbf{T}_N \end{bmatrix} \quad (74)$$

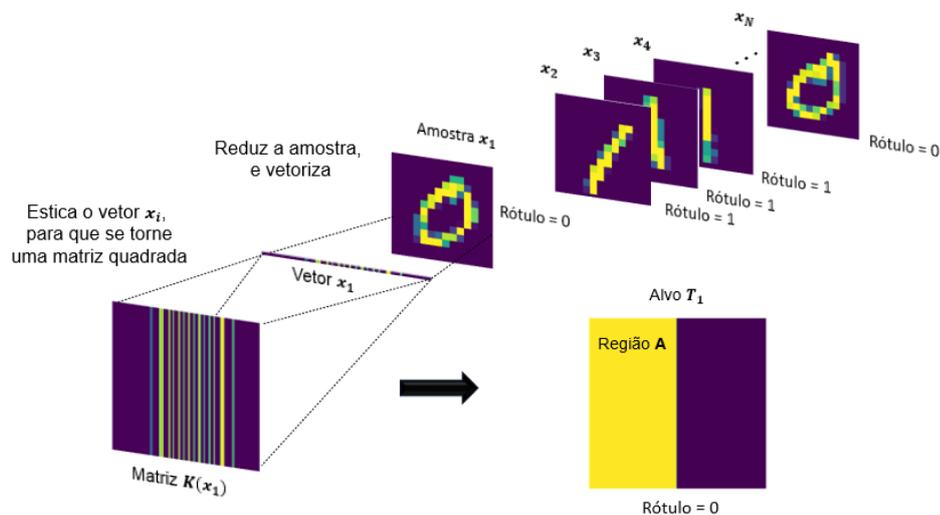
Figura 45 – Exemplos de matrizes alvo relativa a problemas de duas classes.



Fonte: Autor, 2023

O objetivo do treinamento é encontrar uma matriz de pesos que pondere os neurônios das matrizes de entrada, de forma que se aproximem dos valores das matrizes alvo, para que assim possa ser feita a classificação. Isso pode ser melhor compreendido por meio do diagrama das Figuras 46 e 47.

Figura 46 – Representação do processo de obtenção das matrizes $K(x_i)$ para a etapa de treinamento.



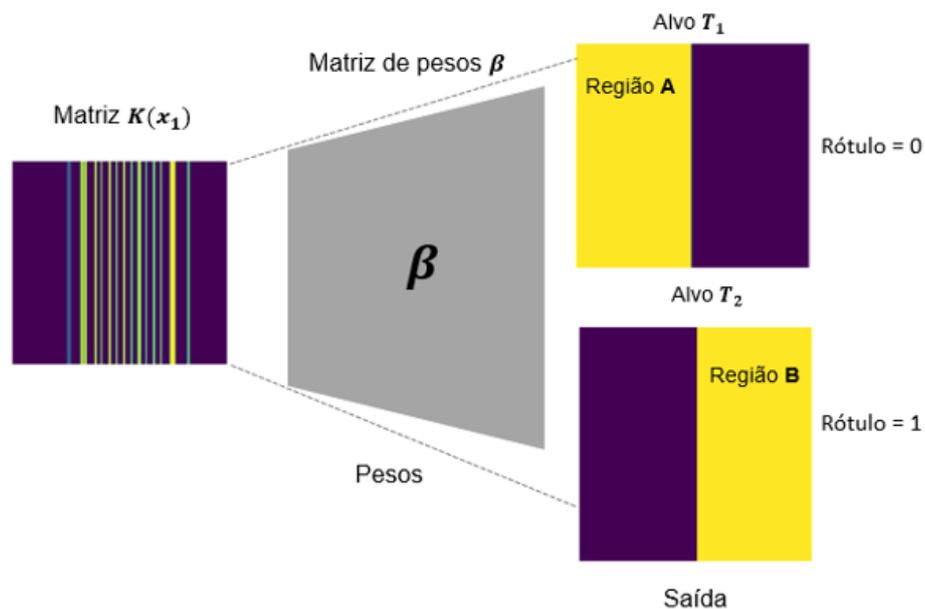
Fonte: Autor, 2023

Para encontrar essa matriz de pesos utilizamos a equação 10, onde neste caso temos:

$$\beta = [Xc]^\dagger Tc \quad (75)$$

Onde, $[Xc]^\dagger$ denota a pseudo-inversa de Moore-Penrose (ZHANG et al., 2019). Neste caso, a dimensão da matriz β é $d \times d$.

Figura 47 – Diagrama do treinamento com as matrizes de entrada e alvos como matrizes quadradas, onde o objetivo em questão é obter os pesos β .



Fonte: Autor, 2023

4.2.2 Generalização para problemas de 10 classes

Anteriormente, foi apresentado um exemplo de como treinamos nosso sistema para solucionar problemas com apenas 2 classes. Agora, vamos explorar como generalizar essa abordagem para problemas com 2 ou mais classes. Para tanto, basta determinarmos uma matriz alvo para cada classe correspondente, onde a área da região alvo pode ser calculada a partir da relação ilustrada na Figura 48, onde d é o número de pixels.

4.2.3 Classificação das imagens

Nossa etapa de classificação segue a arquitetura descrita na Figura 49. Esta arquitetura descreve como classificamos em duas classes os dígitos manuscritos "0" e "1" do conjunto de dados do MNIST. Esta etapa segue, com o processo de esticar um dos vetores de características x'_i do conjunto de imagens de teste X_t , este processo é feito seguindo a equação 72, mostrado na sub-seção 4.2.1.

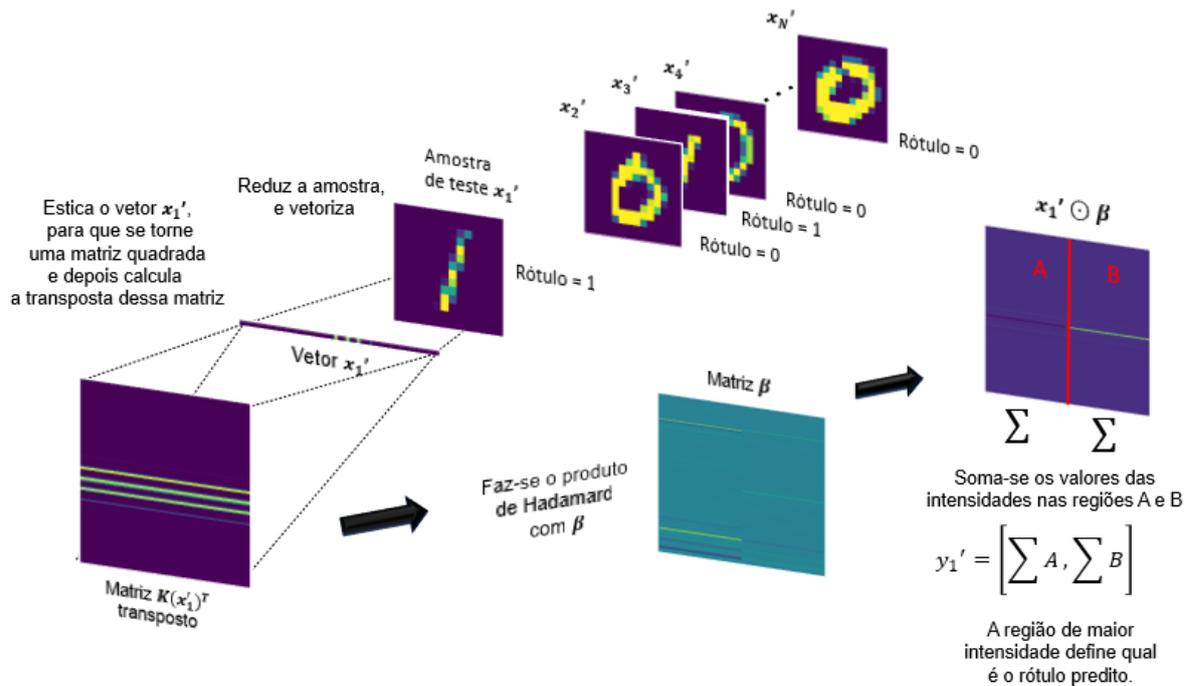
Posteriormente, é calculado a transposta da matriz esticada $K(x'_i)$ para que cada

Figura 48 – Diagrama do treinamento com as matrizes de entrada e alvos como matrizes quadradas para resolver problemas de 10 classes.



Fonte: Autor, 2023

Figura 49 – Esquema da etapa de classificação, a) Descrição numérica b) Setup experimental.

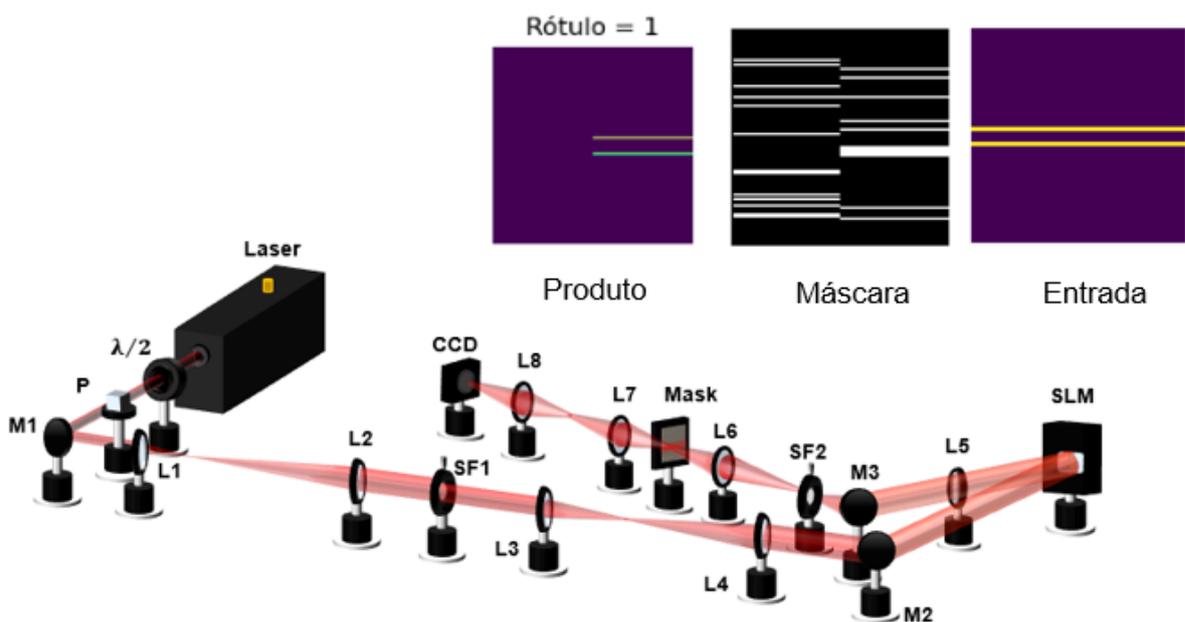


Fonte: Autor, 2023

linha da mesma possa ser ponderada por meio do produto de Hadamard com a matriz β . Em seguida, é feita a soma das intensidades nas regiões A e B desta matriz produto. Estas regiões definem qual a classe que a imagem de entrada corresponde. Então, é comparado qual maior valor de intensidade nestas duas regiões. Se o maior valor corresponder a região A então a imagem é classificada como o dígito 0 e se o maior valor de intensidade for na região B o dígito é classificado como 1.

Inicialmente, procedemos fazendo à discretização das imagens de entrada, assim como da máscara β obtida durante o treinamento. Em seguida, implementamos essa máscara como um discriminador óptico, que modula a amplitude do padrão de luz de entrada gerado por um Modulador Espacial de Luz (SLM). Este processo é visualizado na Figura 50.

Figura 50 – Esquema da configuração experimental.



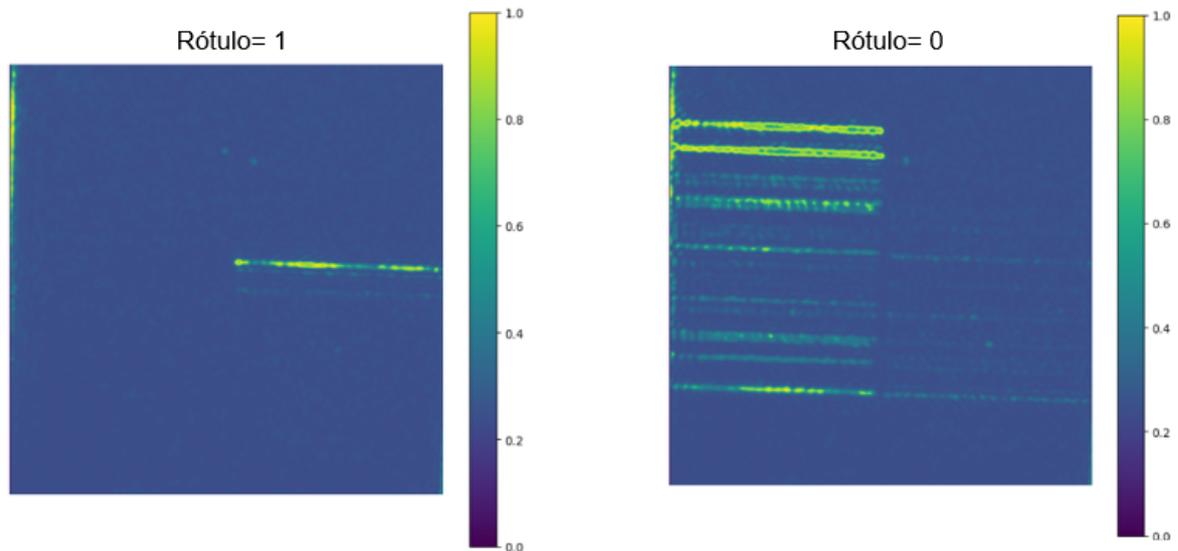
Fonte: Autor, 2023

4.3 RESULTADOS

Um resultado referente à discriminação dos dois números está apresentado na Figura 51. Para o número classificado como 1, a região alvo B exibe uma intensidade mais elevada, enquanto, para o número classificado como 0, a região alvo A apresenta

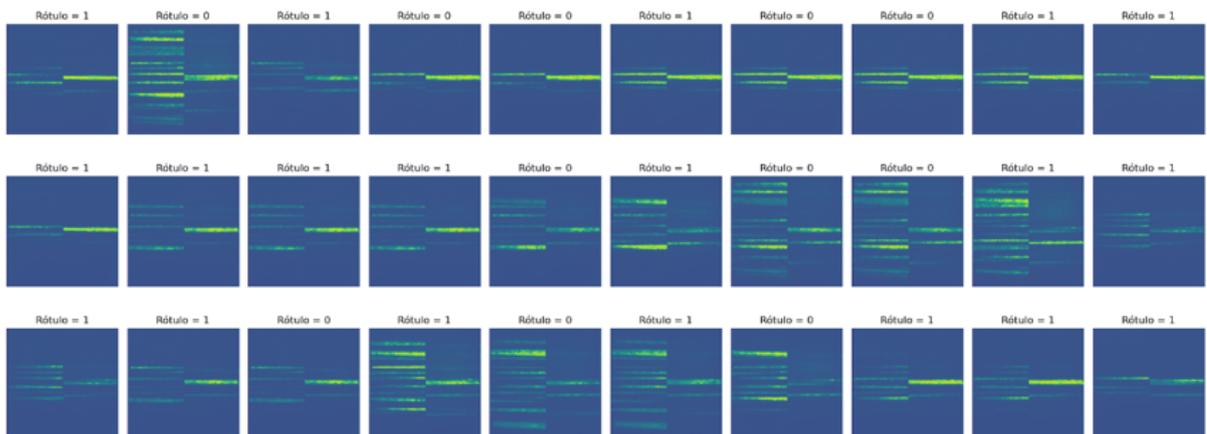
uma intensidade superior. Já na Figura 52, observamos os resultados da discriminação das 30 primeiras imagens de um total de 200 classificadas.

Figura 51 – Resultado óptico experimental da classificação em duas classes.



Fonte: Autor, 2023

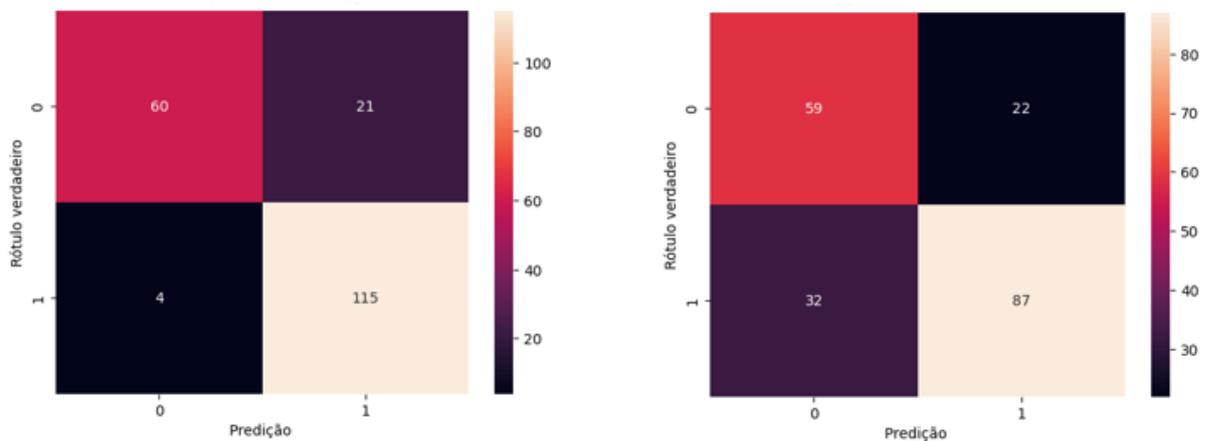
Figura 52 – Resultado óptico experimental da classificação em duas classes das 30 primeiras imagens.



Fonte: Autor, 2023

Na Figura 53, apresentamos a matriz de confusão numérica e experimental referente a essa classificação. A acurácia relativa aos resultados numéricos neste caso foi de 87.5%, enquanto a acurácia experimental foi de 73%. Essa discrepância percentual em relação resultado numérico é em grande parte atribuída à qualidade de impressão da máscara.

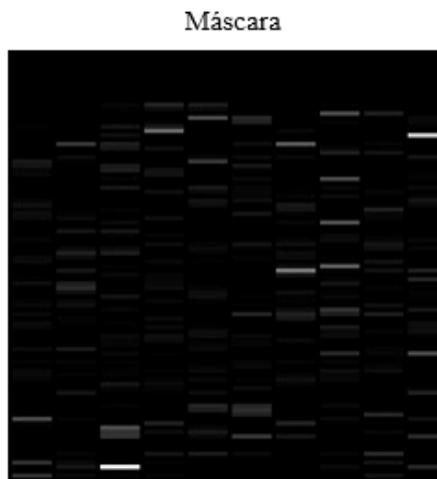
Figura 53 – Matriz de confusão do resultado óptico numérico e experimental da classificação em duas classes



Fonte: Autor, 2023

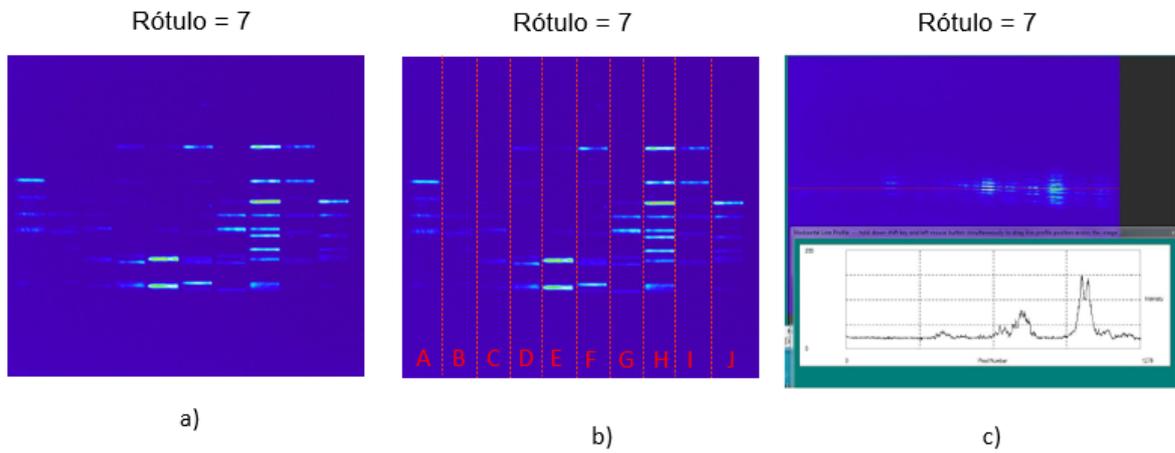
Na sequência de figuras, apresentamos o discriminador (máscara de transmissão, Figura 54), juntamente com o resultado gerado em 10 classes. Observa-se uma classificação precisa, identificando corretamente o número de entrada como 7. Este reconhecimento é evidenciado pela região de maior intensidade, correspondente ao alvo '7' (Figura 55b). A soma das intensidades nessa região pode ser efetuada incorporando uma lente cilíndrica no sistema óptico, localizada na saída da máscara. Essa lente realiza a transformada de Fourier do padrão de luz formado, equivalendo a efetuar uma soma, como ilustrado na Figura 55c.

Figura 54 – Máscara de transmissão para classificação óptica de imagens em 10 classes



Fonte: Autor, 2023

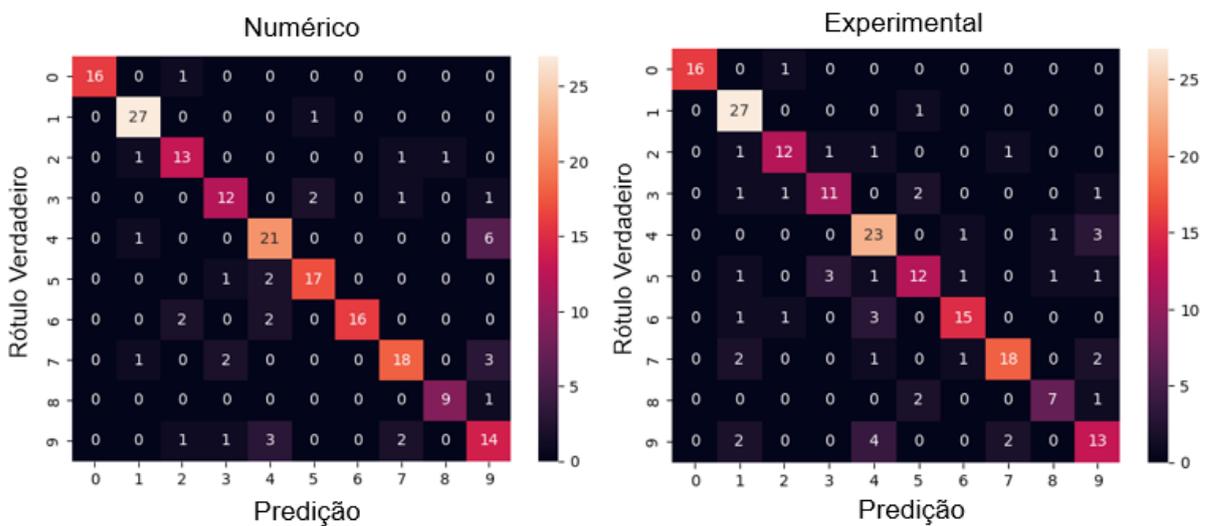
Figura 55 – a) Imagem experimental obtida pela CCD para classificação óptica de imagens em 10 classes, b) Imagem da CCD com regiões de classificação definidas pelas linhas em vermelho, c) Soma óptica das intensidades.



Fonte: Autor, 2023

Na Figura 56, temos a matriz de confusão referente aos resultados numéricos e experimentais para a classificação das imagens em 10 classes utilizando o discriminador óptico. A partir dessas matrizes, observamos que a acurácia para o resultado numérico foi de 81%, considerando um total de 200 imagens classificadas. No caso experimental, a acurácia foi de 77%.

Figura 56 – Matriz de confusão do resultado óptico numérico e experimental da classificação para dez classes



Fonte: Autor, 2023

Com o objetivo de adotar uma abordagem mais direta e precisa na classificação óptica de imagens, sem a necessidade de etapas de transformações das imagens

de entrada como mostrado anteriormente, desenvolvemos um segundo método de classificação.

4.4 MÉTODO DE CLASSIFICAÇÃO ÓPTICA 2

O segundo método de classificação também segue o modelo de treinamento das máquinas de aprendizado extremo. Esta abordagem evolui com a adição de uma simplificação adicional, na qual estabelecemos $W = 1$, e a exclusão de qualquer função não linear na camada oculta. A estrutura básica da ELM, com seu princípio de atribuir pesos aleatórios e fixos W para a camada oculta, atua como a base para o desenvolvimento dessa versão mais simplificada do modelo.

Ao considerar o caso especial em que $W = 1$ na ELM, estamos efetivamente eliminando a ponderação aleatória na camada oculta. Isso resulta em uma camada oculta que não realiza nenhuma transformação não linear nos dados de entrada X . A supressão da função não linear nesse contexto transforma a ELM em um modelo operando linearmente nas características de entrada. À medida que seguimos esse caminho, chegamos à formulação deste modelo linear de treinamento, onde

$$T = \beta X \quad (76)$$

Nessa fórmula, β é determinado utilizando a pseudoinversa de Moore-Penrose aplicada à matriz X , onde

$$\beta = X^\dagger T \quad (77)$$

Dessa forma a matriz β contém os coeficientes lineares que estabelecem uma relação direta entre as características X e os rótulos T . Essa abordagem representa um caso especial e simplificado da ELM, entretanto, ao contrário da ELM convencional, a transformação não linear da camada oculta é substituída por uma relação linear clara.

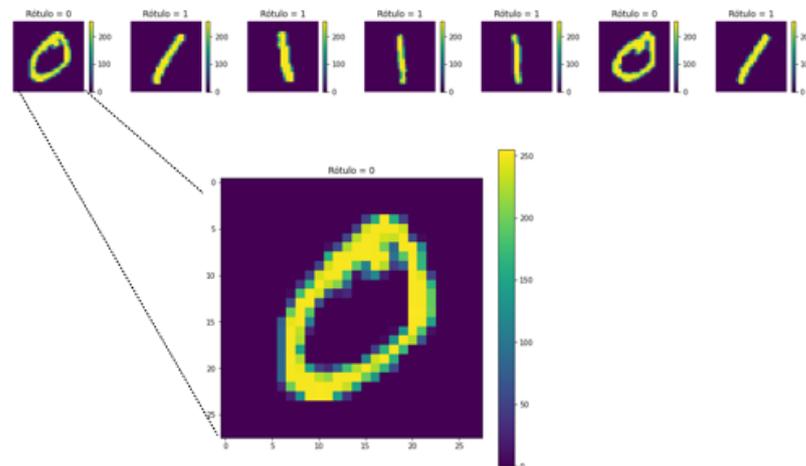
Consequentemente, podemos interpretar este modelo linear como uma simplificação conceitual da ELM, onde a complexidade das transformações não lineares e das ponderações aleatórias é eliminada, resultando em uma estrutura mais direta e linear. Essa estrutura é particularmente aplicável para lidar com problemas de classificação

que podem ser aproximados de maneira eficaz por uma relação linear entre as características e os rótulos.

4.4.1 Descrição

Em nosso processo de treinamento, fizemos uso de dois conjuntos de dados para avaliar o desempenho do nosso modelo. O primeiro conjunto de dados é o MNIST já visto no capítulo anterior e ilustrado na Figura 57, que abrange um total de 70.000 imagens de dígitos manuscritos, variando de 0 a 9. Cada imagem possui dimensões 28×28 pixels. No total, 60.000 imagens foram alocadas para a etapa de treinamento, enquanto 10.000 imagens foram reservadas para o estágio de teste. No entanto, para fins de primeiros testes, nos concentramos em trabalhar com apenas duas classes de dígitos: as classes "0" e "1". Dessa forma, o conjunto de dados de treinamento passou a ter 12.665 imagens de treinamento e 2.115 imagens de teste.

Figura 57 – Conjunto de dados de dígitos manuscritos do MNIST

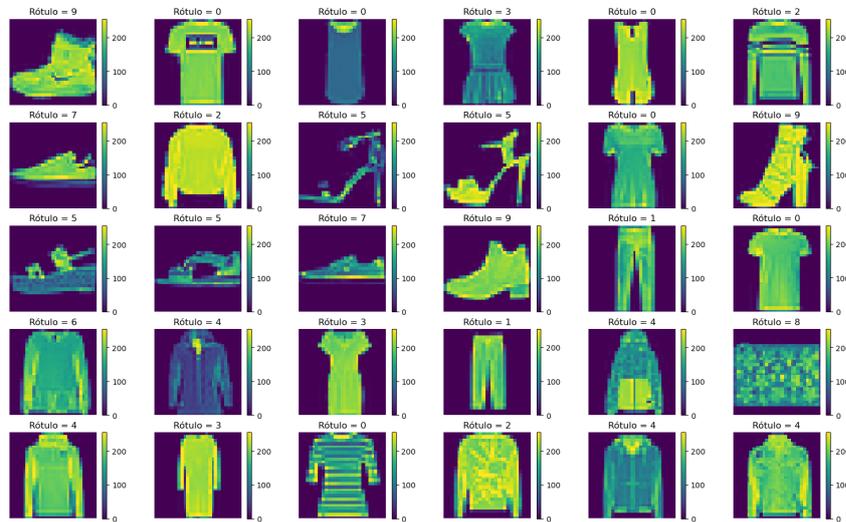


Fonte: Autor, 2023

O segundo conjunto de dados que utilizamos é o Fashion MNIST, que consiste em um conjunto de imagens de moda. Este conjunto de dados é composto por um total de 70.000 imagens de roupas e acessórios de vestuário, distribuídos em 10 classes diferentes, cada uma representando um tipo de vestuário, como : Camiseta/top, calça, pulôver, vestido, casaco, sandália, camisa, tênis, bolsa, bota de tornozelo, rotulados como: 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9, respectivamente, assim como pode ser visto na

Figura 58. Cada imagem também possui dimensões 28×28 pixels, semelhantes ao conjunto de dados MNIST.

Figura 58 – Conjunto de dados de moda do Fashion MNIST



Fonte: Autor, 2023

O processo de treinamento para ambos conjuntos de dados é conduzido com base na equação 77, onde a matriz \mathbf{X} compreende um total de N amostras de treinamento, cada qual consistindo de uma imagem vetorizada $\mathbf{x}_i \in \mathbf{X}$ de dimensão $i \times d$, em que $i \in [1, N]$ e $d = 28^2$.

4.4.2 Descrição da etapa de classificação

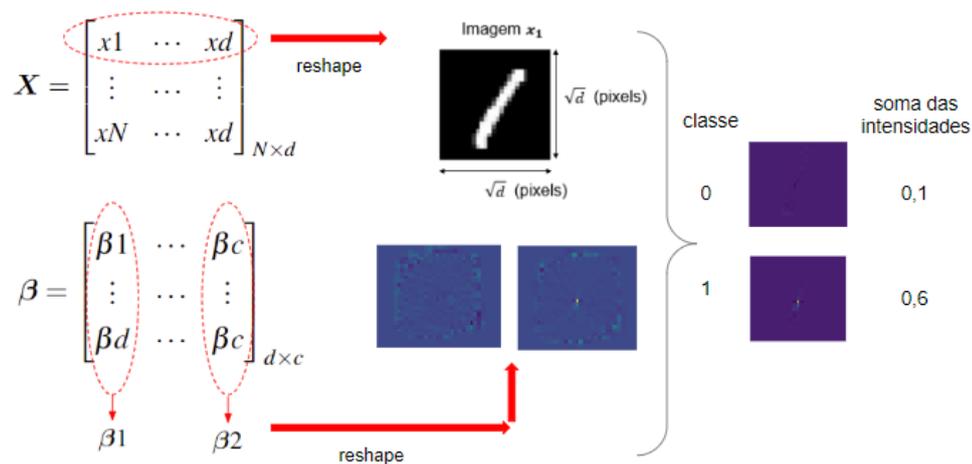
A partir da Equação 77, obtemos a matriz de pesos otimizada β , cuja estrutura é representada por:

$$\beta = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1c} \\ \vdots & \dots & \vdots \\ \beta_{1d} & \cdots & \beta_{dc} \end{bmatrix}_{d \times c} \quad (78)$$

Analisando a Equação 77, percebemos que a matriz β é de dimensão $d \times c$, em que d é a dimensão dos vetores de entrada e c é o número de classes das imagens. Em contextos de sistemas ELM tradicionais, o processo de classificação envolve a multiplicação matricial da matriz de pesos β pelo conjunto de dados de teste. Essa operação possibilita a previsão da classe (ou rótulo) associada à imagem.

Contudo, devido à nossa intenção de transferir essa etapa para um sistema óptico, a abordagem foi adaptada. Em vez de executar uma multiplicação matricial direta, que é mais complicada de aplicar em um sistema físico, optamos por uma técnica alternativa que emprega o produto elemento a elemento seguido de uma soma. Essa modificação permite que o método seja aplicado em diversas configurações de sistemas ópticos, como o uso exclusivo de um modulador espacial de luz de modulação de fase, bem como em sistemas nos quais se utilize tanto um modulador espacial de luz quanto máscaras de transmissão construídas a partir das máscaras de beta. A descrição detalhada desse procedimento é apresentada na Figura 59.

Figura 59 – Descrição da nossa etapa de classificação



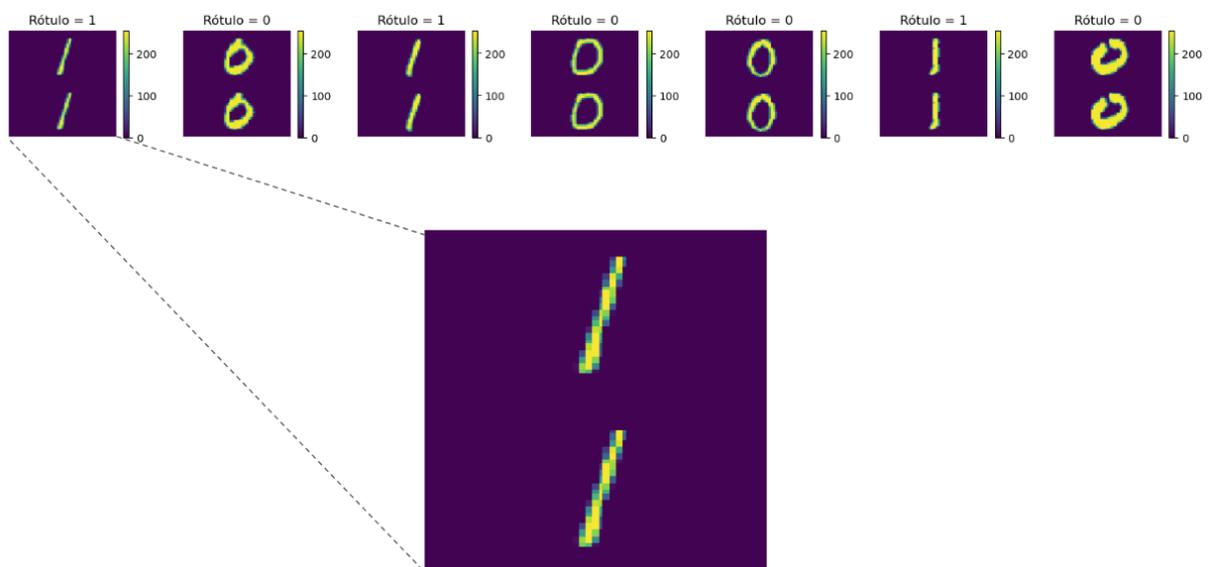
Fonte: Autor, 2023

Em linhas gerais, esse método envolve o cálculo do produto interno dos vetores x_i com β . No entanto, há uma transformação adicional realizada nos vetores, convertendo-os novamente em imagens com dimensões $\sqrt{d} \times \sqrt{d}$ pixels. Similarmente, as colunas da matriz β são transformadas em "máscaras"(imagens) com dimensões $\sqrt{d} \times \sqrt{d}$ pixels. Uma vez que estamos trabalhando com um problema envolvendo apenas duas classes, preparamos duas "máscaras": β_1 (correspondente à classe 0) e β_2 (correspondente à classe 1).

A próxima etapa envolve a multiplicação ponto a ponto da imagem x_i com ambas as máscaras. Em sequência, somamos todos os valores de intensidade resultantes de cada produto. A comparação dos valores acumulados nos permite identificar o maior valor, que por sua vez é associado à classe predita da imagem.

Com o intuito de tornar esse processo mais prático, adotamos uma abordagem adicional: a duplicação das imagens de entrada, conforme representado na Figura 60. Esse método permite a criação de uma única máscara englobando as máscaras β_1 e β_2 . Essa otimização pode ser visualizada de maneira mais clara na Figura 61. Onde, as imagens duplicadas são multiplicadas pela máscara, resultando em uma matriz de produtos. A soma das intensidades nessa matriz e a subsequente comparação das somas constituem a base para a classificação.

Figura 60 – Dígitos MNIST duplicados.

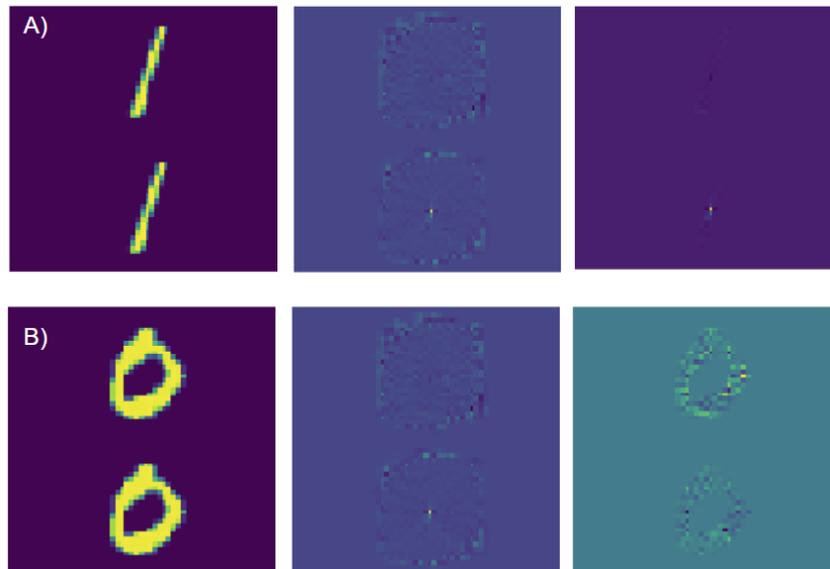


Fonte: Autor, 2023

4.4.3 Generalização para dez classes

Esse método de classificação pode ser estendido para lidar com até dez classes distintas. Após o treinamento ter sido conduzido para todas as dez classes, o procedimento segue uma estrutura semelhante àquela descrita na seção anterior. Inicialmente, as imagens de entrada são preparadas. Em seguida, é criada uma máscara unificada com as correspondentes máscaras $\beta_1, \beta_2, \beta_3, \dots, \beta_{10}$. A etapa subsequente compreende a multiplicação elemento a elemento das entradas decuplicadas com a máscara construída, seguindo o procedimento ilustrado nas Figuras 62 e 63. Posteriormente, as intensidades resultantes de cada produto são somadas e comparadas para determinar a classe predita.

Figura 61 – Etapa de classificação com dígitos duplicados. A) Classificação da o dígito 1 e B) classificação do dígito 0.



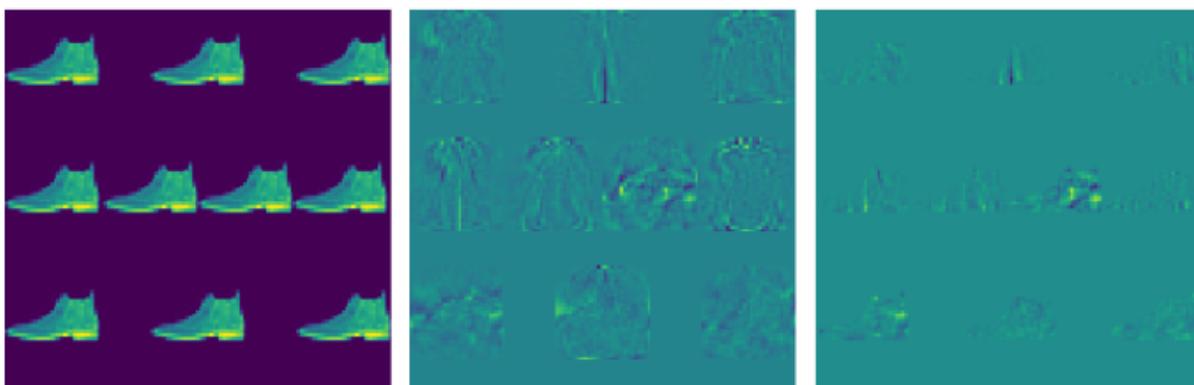
Fonte: Autor, 2023

Figura 62 – Etapa de classificação com dígitos decuplicados



Fonte: Autor, 2023

Figura 63 – Etapa de classificação com imagem de moda decupicadas



Fonte: Autor, 2023

4.5 MÉTODOLOGIA EXPERIMENTAL UTILIZANDO SOMENTE UM MODULADOR ESPACIAL DE LUZ (SLM) SOMENTE DE FASE.

Como mencionado previamente, o treinamento do nosso sistema envolve a busca por matrizes otimizadas (que denominamos de máscaras β), que ponderam os valores de intensidade das imagens de entrada. Com estas máscaras em mãos, o processo de classificação é alcançado ao realizar o produto elemento por elemento das máscaras com as imagens de entrada, seguido pela etapa de soma das intensidades.

Um configuração óptica simples e econômica em termos de componentes ópticos envolve o uso de um único modulador espacial de luz baseado em cristais líquidos (LC-SLM) que modula exclusivamente a fase da luz. Neste caso, as máscaras de fase são projetadas para o modulador, que opera somente na modulação de fase. Como tanto as entradas quanto as máscaras são imagens de intensidade, faz-se necessário o uso do algoritmo de recuperação de fase IFTA (Algoritmo Iterativo de Transformada de Fourier) para codificar a matriz resultante da multiplicação entre as imagens de entrada e as máscaras em matrizes de fase. Essas matrizes serão então carregadas individualmente no SLM de forma que a distribuição de luz resultante da propagação no espaço livre contenha as informações das imagens de entrada multiplicadas elemento por elemento com as máscaras. A etapa de soma, neste caso, será realizada pela soma das intensidades de cada região correspondente a cada classe por meio de uma câmera CCD e um computador.

Neste contexto, esta seção se dedica a detalhar as etapas experimentais, bem como as simulações numéricas que realizamos para implementar a etapa descrita.

4.5.1 Algoritmo IFTA

Como mencionado anteriormente, para codificar a matriz resultante da multiplicação entre as imagens de entrada e as máscaras em matrizes de fase, empregamos o Algoritmo IFTA, também conhecido como Algoritmo de Recuperação de Fase de Gershberg-Saxton. Este algoritmo é uma técnica iterativa amplamente utilizada para reconstruir a fase de duas distribuições de intensidade correlacionadas por uma função de propagação, como a transformada de Fourier.

Esse método é aplicado quando as intensidades correspondentes são conheci-

das nos planos ópticos apropriados: no plano de imagem, a intensidade é dada por $|f(x, y)|^2$, enquanto no plano de Fourier, é expressa como:

$$|F(u, v)|^2 = \left| \iint_{-\infty}^{+\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy \right|^2 \quad (79)$$

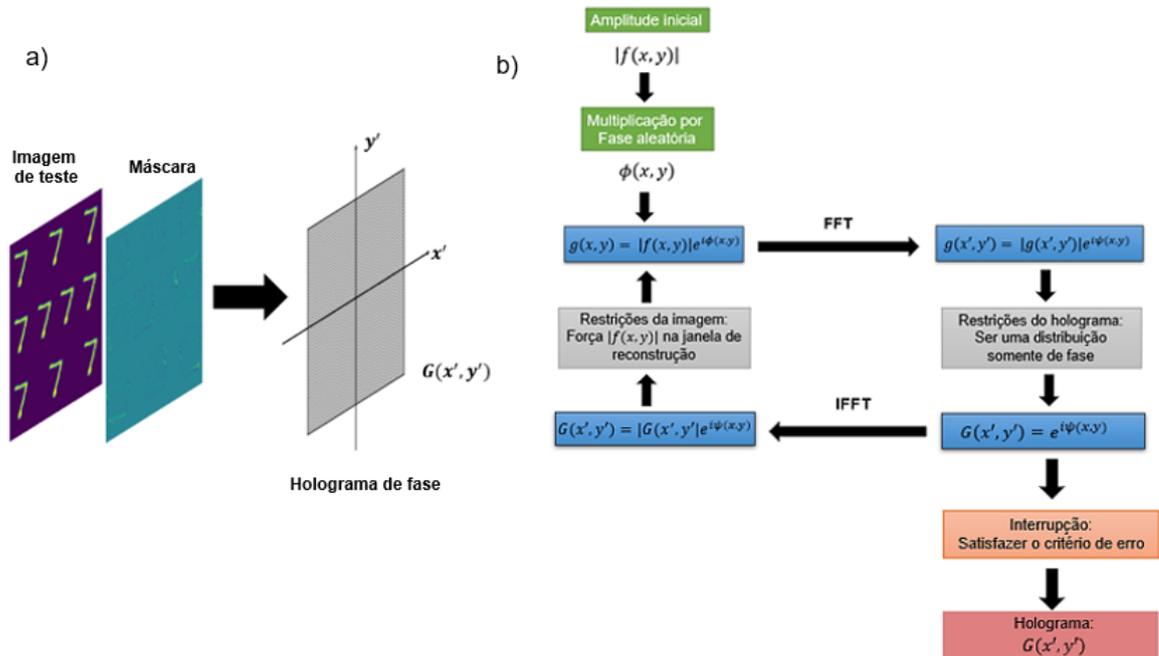
O foco reside em determinar a fase $\psi(x, y)$, sendo muitas vezes suficiente ter acesso à distribuição de fase de um dos planos, já que a fase do outro plano pode ser obtida por meio da transformada de Fourier do plano com fase conhecida.

O procedimento iterativo do algoritmo começa com uma estimativa inicial da fase $\phi(x, y)$, gerada aleatoriamente ou derivada de informações parciais da imagem original. A partir dessa estimativa, a amplitude complexa do campo é calculada utilizando a intensidade medida no plano apropriado, seguindo a fórmula $g(x, y) = |f(x, y)| e^{j\phi(x, y)}$, onde $|f(x, y)|$ é a amplitude inicial que para uma imagem e intensidade é dada por $|f(x, y)| = \sqrt{I}$. Em seguida, essa amplitude complexa é propagada para o plano de Fourier por meio da transformada de Fourier, resultando em $g(x', y') = |G(x', y')| e^{j\psi(x, y)}$, onde $g(x', y')$ é a transformada de Fourier de $g(x, y)$ e x' e y' são as frequências espaciais. A amplitude de $g(x', y')$ é ajustada conforme uma restrição específica que em nosso caso, trata-se forçar com que esse campo seja somente uma distribuição de fase, fazendo que a amplitude $g(x', y')$ seja igual a 1. Além disso, normalizamos a essa distribuição no plano de Fourier para controlar o seu comportamento ao longo do processo iterativo. Isso é feito para evitar instabilidades numéricas ou amplificações excessivas durante as iterações do algoritmo. Em seguida, o espectro é transformado de volta para o domínio espacial, onde aplicamos outra restrição, fazendo com que a a imagem a amplitude da imagem reconstruída esteja dentro do domínio espacial.

Assim, a nova estimativa da fase $\psi(x, y)$ é obtida por meio da transformada inversa de Fourier, utilizando a estimativa da amplitude da imagem e a fase atualizada para gerar uma nova estimativa de fase. Essa fase atualizada é então incorporada à próxima iteração, gerando uma nova estimativa de amplitude e assim por diante, até que o erro entre a imagem reconstruída e a imagem original seja minimizado. A Figura 64 b) ilustra de forma esquemática o processo iterativo do algoritmo.

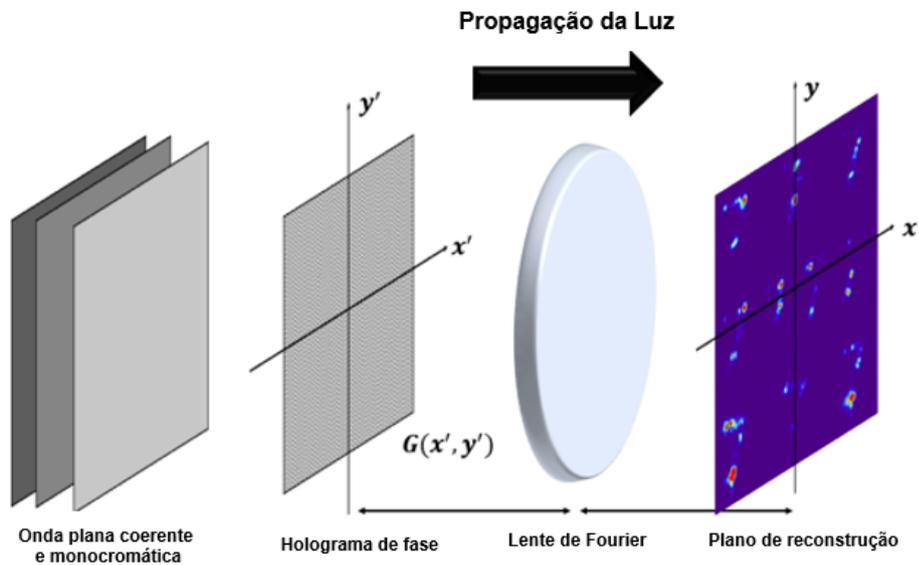
Para avaliar o progresso da reconstrução da imagem ao longo de cada etapa de atualização da fase, medimos o erro quadrático médio (RMSE), conforme definido em (POON; LIU, 2014), como:

Figura 64 – Reconstrução a partir do holograma de fase



Fonte: Autor, 2023

Figura 65 – Reconstrução a partir do holograma de fase



Fonte: Autor, 2023

$$RMSE = \left\{ \frac{1}{MN} \sum_{m,n} [|\Gamma(m, n)| - f(m, n)]^2 \right\}^{1/2} \quad (80)$$

Na equação 80, $f(m, n)$ representa a imagem desejada que se pretende obter. O campo avaliado é representado por $\Gamma(m, n)$, com (m, n) denotando os índices de amostragem ao longo dos eixos x e y , respectivamente. M e N representam o número

de amostras realizadas em cada eixo.

Na Figura 64 a), é apresentado um esquema que descreve a etapa de construção do holograma de fase por meio do algoritmo IFTA, onde ocorre a codificação do produto entre a imagem de entrada e a máscara de fase β .

O processo de reconstrução do holograma é detalhado na Figura 65, em que o holograma é exposto à uma frente de onda monocromática advinda de uma fonte de luz coerente, como um laser. Nesse ponto, a luz incidente interage com o holograma, gerando um padrão de difração. Posteriormente, esse padrão de difração é transformado por Fourier através de uma lente, direcionando as informações para o domínio espacial. Consequentemente, as informações originalmente codificadas no holograma são reconstruídas no plano focal dessa lente, permitindo a visualização ou captura da imagem reconstruída.

4.5.2 Projetando as máscaras de fase

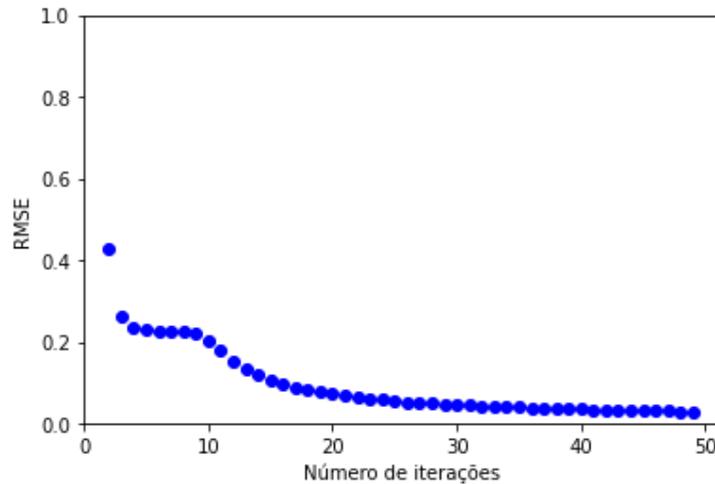
Todas as imagens de entrada, tanto as duplicadas quanto as decuplicadas, bem como as máscaras de β , foram projetadas em um *background* composto por pixels de valor zero, formando uma matrizes de tamanho 204 x 204 pixels.

No entanto, um desafio importante surgiu ao reconstruir essas imagens usando o algoritmo IFTA. O processo padrão de reconstrução resultava na sobreposição das imagens reconstruídas com a ordem de difração zero, que consiste no feixe luminoso refletido pelo LC-SLM. Isso gerava ruído e interferências indesejados nas imagens reconstruídas. Para mitigar esse problema, adotamos uma estratégia de deslocamento. A imagem original foi colocada sobre um fundo expandido de 400 x 400 pixels, deslocado 100 pixels a partir do centro desse plano de fundo. Esse deslocamento permitiu que a imagem reconstruída fosse filtrada espacialmente, assegurando, assim, uma alta qualidade na imagem resultante, livre de interferências da ordem de difração zero.

Feito isso, os hologramas de fase são produzidos após 25 iterações utilizando o algoritmo IFTA. Essa quantidade de iterações foi escolhida por apresentar um erro quadrático médio (RMSE) de 0,12, que é considerado um valor ótimo para a reconstrução. A Figura 66 ilustra como o RMSE diminui ao longo de 50 iterações do algoritmo. Observamos que, com 50 iterações, o valor do RMSE se aproxima ainda mais de zero. No entanto, com 25 iterações, já alcançamos um nível de erro suficientemente baixo

para garantir a qualidade da reconstrução da imagem a partir do holograma.

Figura 66 – Erro quadrático médio em função do número de iterações do algoritmo IFTA



Fonte: Autor, 2023

4.5.3 Setup Experimental e Método

Para a realização deste experimento, foram empregados os seguintes materiais: uma fonte de luz monocromática (Laser HeNe- $\lambda = 633 \text{ nm}$, $P = 1 \text{ mW}$), um modulador espacial de luz de cristais líquidos da tecnologia LCOS-SLM (ZHANG; YOU; CHU, 2014) de modulação de fase, modelo **X13267** da Hamamatsu, uma câmera CCD (charge-coupled device) modelo **DCC3240C** da Thorlabs, além de lentes ópticas e um computador. Algumas das principais especificações técnicas sobre o LCOS-SLM e a câmera CCD são mostradas nas tabelas 1 e 2.

Tabela 2 – Características da Câmera CCD

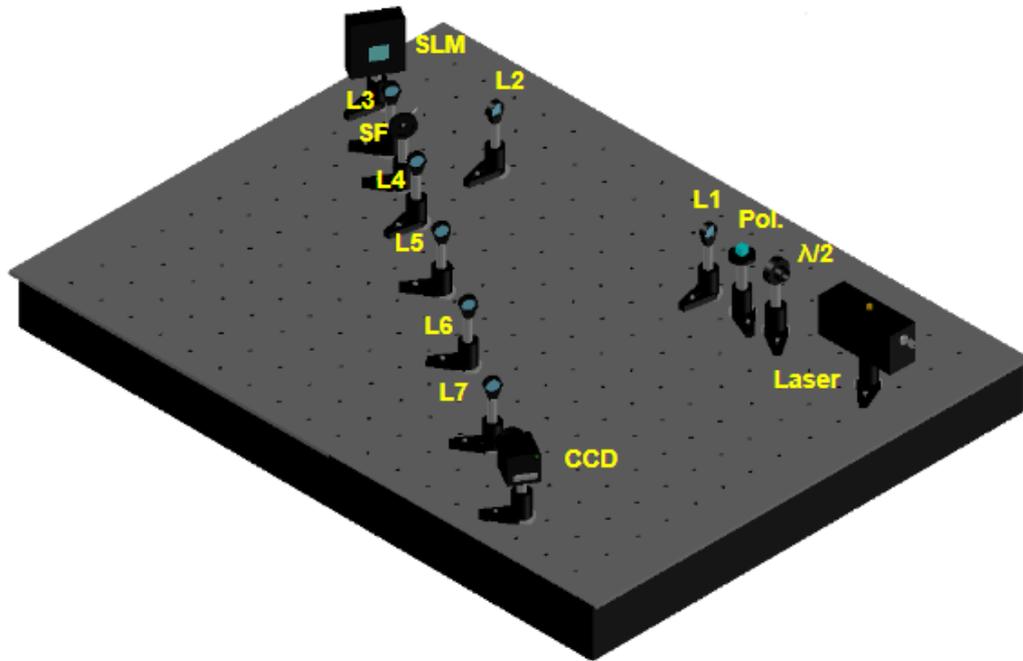
Número de pixels	1280 x 1024
Tamanho do pixel	$5.5 \mu\text{m}$ (quadrado)
Máx. Frame Rate	60.0 fps (Freerun Mode)/ 56.9 fps (Trigger Mode)

Fonte: THORLABS (2020)

Na Figura 67, apresentamos o setup experimental proposto para a realização da classificação das imagens. Nesse arranjo, a fonte de luz monocromática foi linearmente polarizada na direção horizontal, paralela ao vetor diretor dos cristais líquidos, devido à posição do SLM. Essa polarização é obtida utilizando um cristal $\lambda/2$ e um

cubo polarizador, que além de polarizar o feixe também fornece um controle da intensidade do laser. Em seguida, o feixe é expandido pelas lentes L1 e L2 e direcionado, de forma colimada, para o modulador espacial de luz (LCOS-SLM) de fase.

Figura 67 – Setup Experimental



Fonte: Autor, 2023

O LCOS-SLM é então iluminado pela luz laser, e a imagem modulada em fase é propagada até a lente L3. Nesse ponto, ocorre a transformada de Fourier do sinal, levando-o para o domínio das frequências espaciais. Em seguida, a luz passa pelo filtro espacial SF, que permite apenas a ordem de difração +1 passar, bloqueando as demais.

Após a filtragem espacial, a imagem é reconstruída e coletada utilizando as lentes L4 e L5 que estão em uma configuração 4f, a imagem resultante é então gerada no plano focal da lente L4 de forma invertida. Assim, para que a imagem não seja invertida utilizamos novamente um sistema 4f formado pelas lentes L6 e L7. Dessa forma, a imagem resultante, estará no plano focal da lente L7 que é então coletada pela câmera CCD.

4.6 RESULTADOS

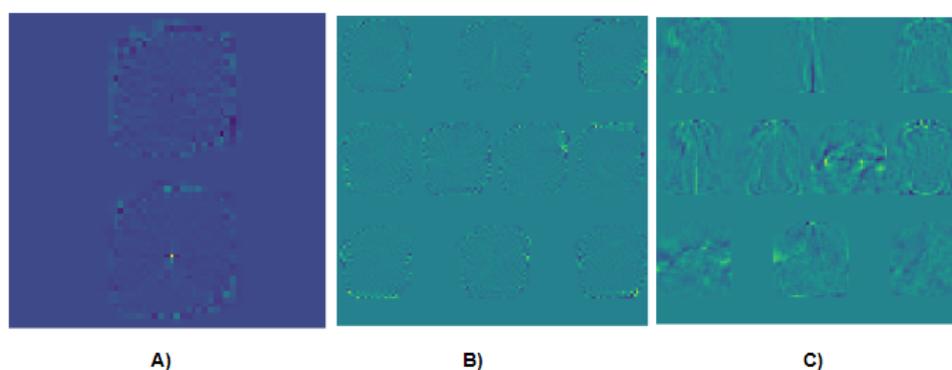
As seguir, apresentaremos os resultados numéricos e experimentais obtidos por meio do nosso sistema experimental utilizando apenas um SLM de fase para validar opticamente tarefas de classificação de imagens. Analisaremos o desempenho do modelo em detalhes, avaliando sua capacidade de distinguir entre as classes de interesse. Essa análise será enriquecida por meio das métricas de desempenho derivadas das matrizes de confusão, que nos fornecerão informações valiosas sobre a precisão e a confiabilidade do nosso sistema óptico de classificação.

4.6.1 Resultados numéricos

No âmbito de nossa investigação numérica, conduzimos uma análise a respeito do desempenho de nosso sistema de classificação óptica, fazendo uso exclusivamente de um modulador espacial de luz que opera unicamente com fase. Além disso, realizamos uma investigação das características das máscaras de β e como isso pode influenciar na métrica de classificação.

Durante a fase de treinamento, executamos o processo de determinação das máscaras otimizadas (máscaras β) por meio da equação 77. Como resultado desse processo, obtivemos máscaras com valores tanto positivos quanto negativos, conforme ilustrado na Figura 68

Figura 68 – Máscaras, A) Dígitos Mnist 2 Classes, B) Dígitos Mnist 10 Classes, C) Imagens de Moda Fashion Mnist

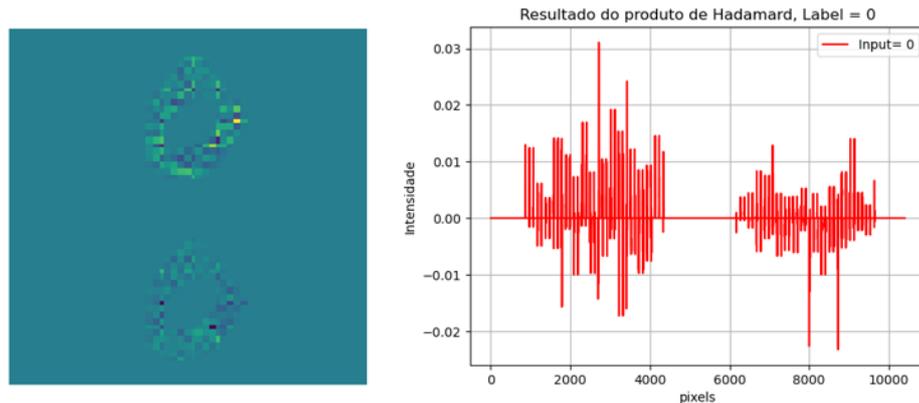


Fonte: Autor, 2023

Essa propriedade das máscaras faz com que, ao multiplicá-las pela imagem, pon-

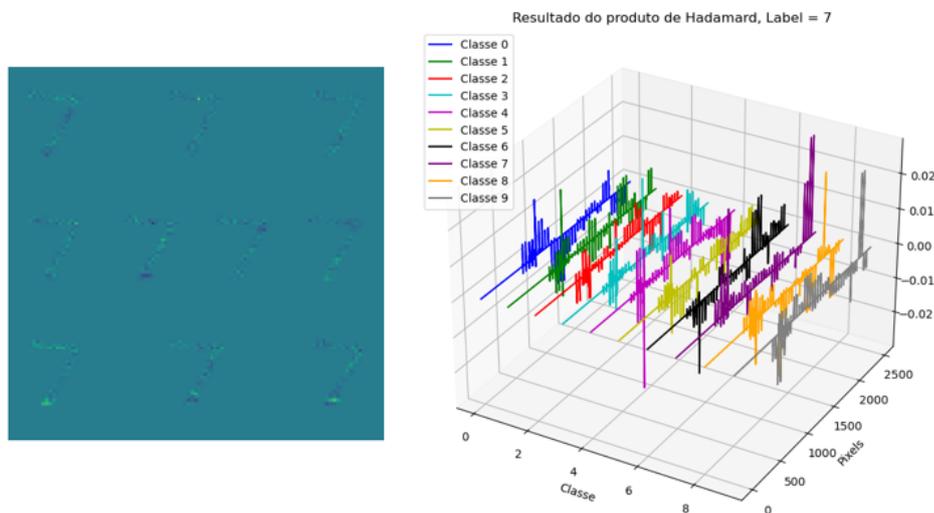
deremos os resultados da multiplicação tanto para valores positivos quanto negativos de intensidade, como pode ser observado nos gráficos apresentados nas Figuras 69 e 70.

Figura 69 – Resultado do produto de Hadamard dígito 0- 2 Classes



Fonte: Autor, 2023

Figura 70 – Resultado do produto de Hadamard dígito 7- 10 Classes

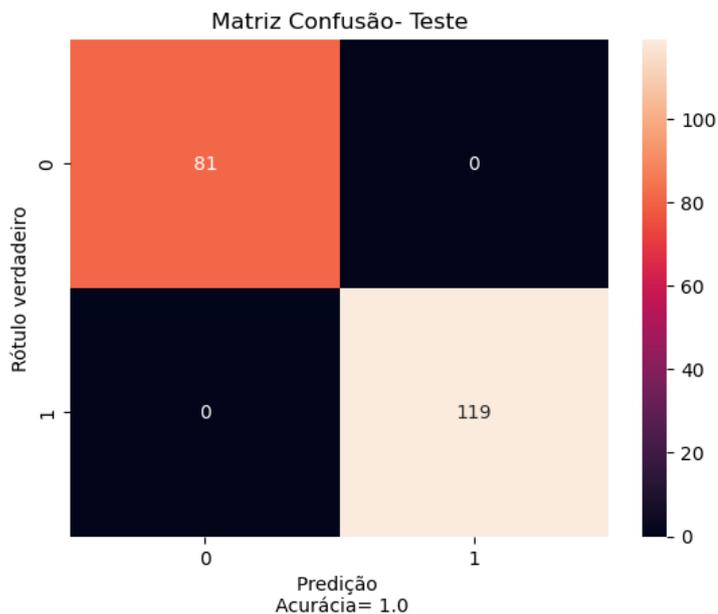


Fonte: Autor, 2023

Os resultados da classificação com essas máscaras, que denominaremos β_{+-} , são representados pelas matrizes de confusão nas Figuras 71, 72 e 73.

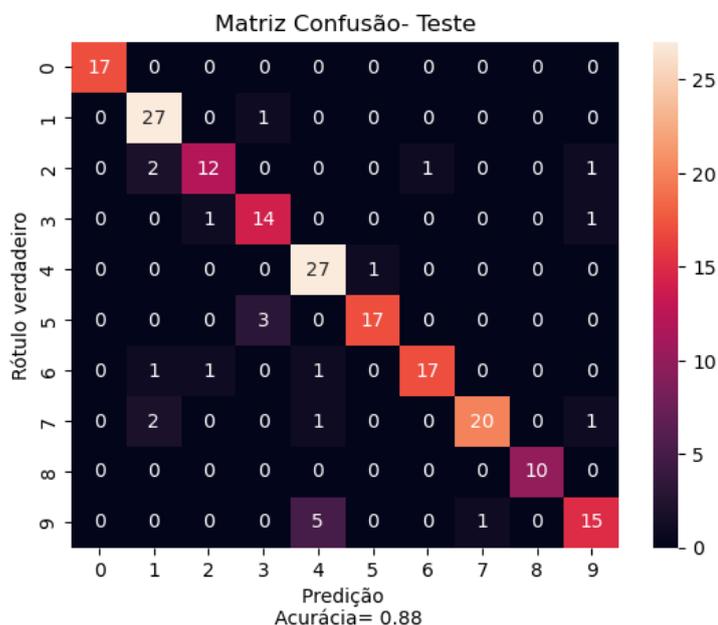
Observamos que, para um total de $N=200$ imagens de teste, obtivemos uma taxa de acerto de 100% na classificação dos dígitos em duas classes (MNIST - 2 Classes). Para a classificação dos dígitos em 10 classes (MNIST - 10 Classes), alcançamos uma taxa de acerto de 88%. Quanto às imagens do Fashion MNIST, obtivemos uma

Figura 71 – Matriz de confusão numérica da classificação dos dígitos manuscritos do MNIST- 2 Classes



Fonte: Autor, 2023

Figura 72 – Matriz de confusão numérica da classificação dos dígitos manuscritos do MNIST- 10 Classes

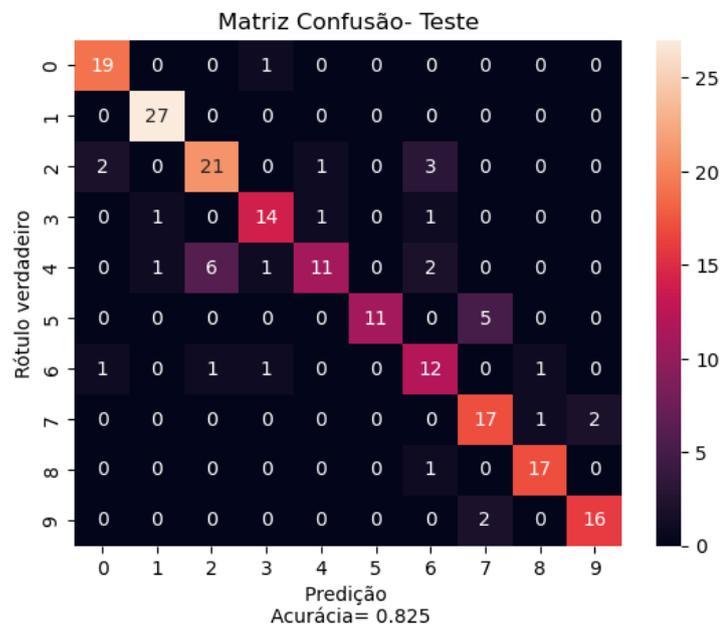


Fonte: Autor, 2023

taxa de acerto de 82.5%. Esses resultados demonstram a eficácia de nosso método de classificação, mesmo quando a etapa de treinamento é totalmente linear.

Na Figura 74, podemos analisar a variação da acurácia em relação à quantidade de imagens de teste. O gráfico revela um comportamento interessante: observamos

Figura 73 – Matriz de confusão numérica da classificação das imagens de moda do Fashion MNIST



Fonte: Autor, 2023

um pico na acurácia quando $N=200$ imagens de teste são utilizadas, atingindo uma taxa de acerto de 88%. Entretanto, à medida que aumentamos o número de imagens de teste, a acurácia começa a oscilar. Quando chegamos a 10.000 imagens de teste, a acurácia é de aproximadamente 85,30%.

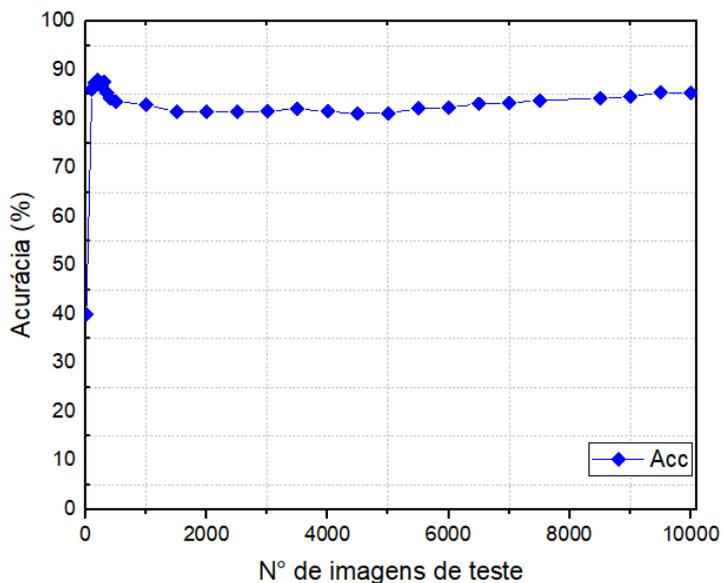
Essa oscilação nos valores de acurácia à medida que aumentamos o número de imagens de teste pode ser atribuída à variabilidade nos lotes de dados. Ao coletar diferentes conjuntos de dados de teste e avaliar a acurácia, é possível que alguns lotes tenham características que favoreçam uma acurácia melhor ou pior em relação aos outros. Essa variação natural nos dados pode influenciar o desempenho do sistema de classificação e resultar na oscilação observada nos resultados de acurácia.

4.6.2 Resultados Experimentais

Na Figura 75, podemos examinar a imagem experimental de intensidade em relação à ponderação realizada pela máscara β . É notável que a região que corresponde à classe 0 apresenta uma intensidade mais significativa, o que fica ainda mais evidente no gráfico ao lado da figura.

O processo de classificação óptica também envolve a soma das intensidades das

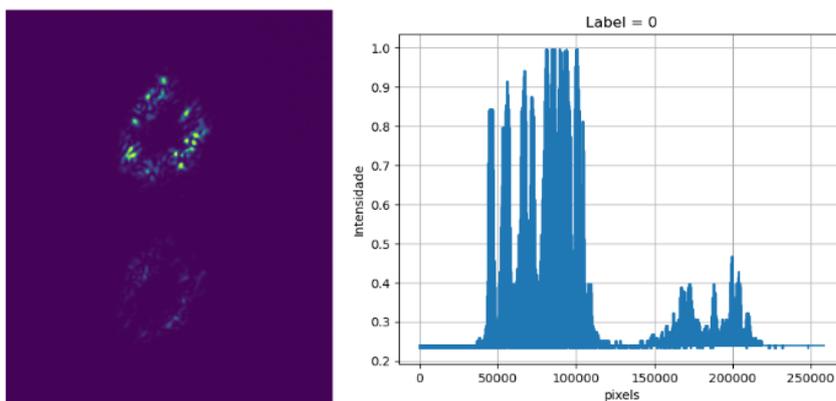
Figura 74 – Gráfico da Acurácia em função no Número de imagens de Teste do dígitos do MNIST



Fonte: Autor, 2023

duas regiões e a comparação entre elas para determinar a classe à qual o dígito em questão pertence. Esse método permite identificar a classe com base na intensidade das regiões específicas da imagem, o que é fundamental para o sistema de classificação óptica.

Figura 75 – Imagem experimental seguida do gráfico de intensidade em função no numero de pixels



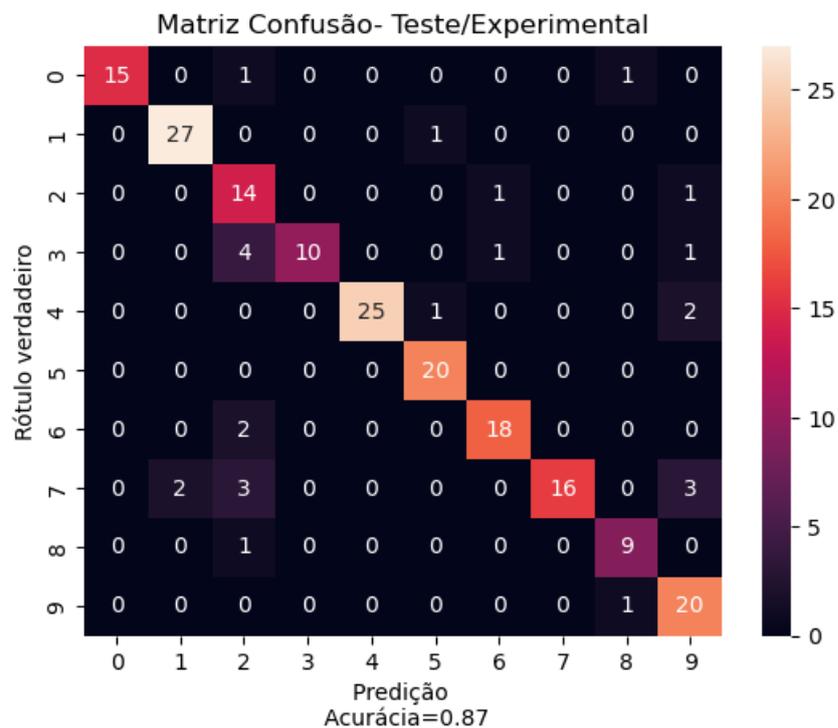
Fonte: Autor, 2023

Nesse contexto, apresentamos as matrizes de confusão referentes à classificação óptica realizada por meio de nosso experimento, utilizando exclusivamente um modulador espacial de luz com modulação unicamente de fase. Na Figura 76, é exibido o número de acertos obtidos pelo sistema experimental ao classificar dígitos manuscri-

tos do conjunto de dados MNIST em 10 classes distintas. Nesse cenário, o sistema alcançou uma taxa de acerto de 87% em relação às 200 imagens de teste utilizadas.

Comparando esse desempenho experimental com o resultado numérico, observamos uma diferença percentual de apenas 1%. Isso demonstra que nosso sistema experimental apresenta um ótimo grau de concordância com as previsões numéricas, validando assim a precisão de nossa abordagem de classificação óptica.

Figura 76 – Matriz de confusão Experimental MNIST



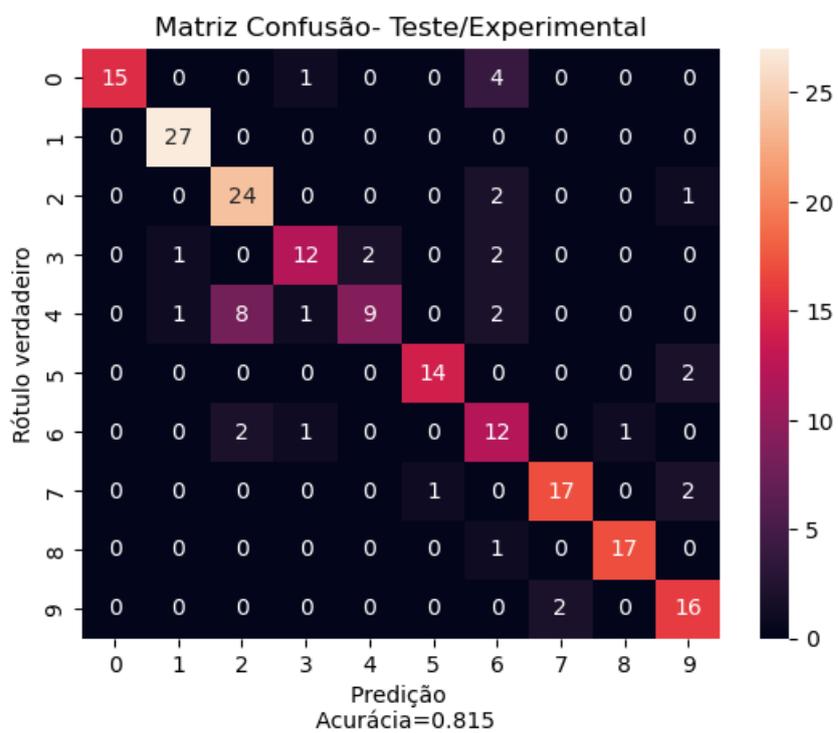
Fonte: Autor, 2023

Além disso, realizamos a tarefa de classificação das imagens de moda do conjunto de dados Fashion MNIST e, nesse cenário, o sistema experimental alcançou uma taxa de acerto de 81.5%. Ao compararmos esses resultados experimentais com as previsões numéricas, mais uma vez, observamos uma diferença percentual de apenas 1%. Essa consistência entre os resultados experimentais e numéricos é um indicativo positivo da robustez e precisão de nosso sistema de classificação óptica.

É importante destacar que a acurácia alcançada na classificação das imagens de moda do Fashion MNIST tende a ser ligeiramente menor devido à maior complexidade dessa base de dados, que inclui roupas e acessórios em várias classes. Mesmo assim, a capacidade do sistema de manter um erro percentual de apenas 1% em relação

aos resultados numéricos é uma boa conquista e valida a eficácia dessa abordagem óptica.

Figura 77 – Matriz de confusão Experimental FASHION MNIST



Fonte: Autor, 2023

5 CONCLUSÃO

O objetivo deste trabalho foi construir e explorar sistemas ópticos experimentais capazes de realizar operações com foco na classificação precisa de imagens. Para desenvolver esses sistemas, partimos do algoritmo de treinamento de máquinas de aprendizado extremo. A fim de compreender a capacidade desse algoritmo, realizamos simulações numéricas e, em seguida, nos dedicamos à construção de um sistema óptico experimental que utiliza um meio espalhador para criar projeções aleatórias dos dados que desejamos classificar. Essas projeções geram complexidade nas imagens, as quais são direcionadas para o sistema óptico por meio de um Modulador Espacial de Luz de Cristais Líquidos (LC-SLM).

Assim, por meio dessas projeções, criamos os recursos necessários para treinar e, posteriormente, classificar as imagens em 10 classes distintas. Observamos que essas projeções geram padrões de speckles que são caracterizados estatisticamente. Em seguida, estudamos a estatística de primeira ordem desses padrões e os caracterizamos como speckles de estatística de intensidade de Rayleigh.

Diante disso, exploramos algumas técnicas de manipulação dessa estatística e, utilizando técnicas de holografia gerada por computador, criamos projeções aleatórias que se desviam da estatística de Rayleigh. Realizamos a classificação de um conjunto de dados de dígitos manuscritos do banco de dados do MNIST e, em seguida, avaliamos o impacto da modificação da estatística de intensidade dessas projeções na acurácia da classificação.

Com base nos resultados obtidos, concluímos que as projeções que possuem o menor contraste, ou seja, aquelas que geram speckles em regime Sub-Rayleigh, melhoram a acurácia do sistema em comparação com os speckles de Rayleigh.

Embora o sistema óptico em questão seja capaz de classificar imagens com uma boa taxa de acerto, ainda é necessário o uso de um computador nesta etapa. Isso ocorre porque, após o treinamento, obtemos uma matriz de pesos otimizada para o conjunto de dados inteiro, e a classificação consiste na multiplicação matricial dessa matriz com os recursos das imagens de teste gerados pelo sistema óptico. Essa multiplicação é realizada pelo computador. Desejávamos, no entanto, que essa etapa de classificação fosse totalmente óptica, ou seja, queríamos ser capazes de prever qual

era a imagem de entrada apenas visualmente, sem a necessidade de um computador para isso.

Portanto, desenvolvemos dois métodos de classificação óptica, ambos baseados no algoritmo de máquinas de aprendizado extremo. No primeiro método, a partir de matrizes de alvos preestabelecidas e de uma transformação nas imagens de treino, criamos um discriminador que pode ser facilmente aplicado em um sistema óptico. Esse discriminador foi então convertido em uma máscara de transmissão capaz de modular a amplitude das imagens que desejamos classificar. Devido a limitações na construção dessa máscara, optamos também por validar essa técnica utilizando o algoritmo IFTA (Inverse Fourier Transform Algorithm) para criar uma máscara difrativa capaz de levar as informações do produto elemento a elemento das imagens de entrada com o discriminador para o domínio óptico e, em seguida, realizar a classificação de forma óptica.

O segundo método de classificação segue uma abordagem mais direta, uma vez que não requer as transformações realizadas no primeiro método, nem alvos predefinidos. Mesmo assim, conseguimos obter discriminadores com base na matriz de pesos obtida no treinamento, que podem ser transformados em máscaras de amplitude, como também podem ser aplicados no sistema de validação usando o algoritmo IFTA. Avaliamos o desempenho dos dois métodos nos sistemas ópticos, classificando imagens de dígitos manuscritos do MNIST, bem como imagens de moda do banco de dados do Fashion-MNIST. Vale destacar que, neste método, a operação realizada utiliza uma multiplicação elemento a elemento, seguida de uma soma para a classificação.

Assim, diante dos resultados numéricos e experimentais, demonstramos a eficácia dessas abordagens de classificação óptica e concluímos que esses sistemas são promissores para a realização da classificação óptica de imagens. Como perspectiva futura, almejamos tornar os dois métodos de classificação desenvolvidos igualmente aplicáveis utilizando um sistema óptico com luz incoerente.

REFERÊNCIAS

- ABHISHEK, K. Introduction to artificial intelligence. [S.l.]: **Redgate**, 2022. Disponível em: <https://www.red-gate.com/simple-talk/development/data-science-development/introduction-to-artificial-intelligence/>. Acesso em: 13 de abril de 2023.
- AKUSOK, A. et al. High-performance extreme learning machines: a complete toolbox for big data applications. [S.l.]: **IEEE Access**, 2015. 1011–1025 p. DOI: 10.1109/ACCESS.2015.2450498. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7140733>. Acesso em 29 de setembro de 2023.
- AMARAL, J. P.; FONSECA, E. J.; JESUS-SILVA, A. J. Tailoring speckles with Weibull intensity statistics. [S.l.]: **Physical Review A**, 2015. 063851 p. DOI: <https://doi.org/10.1103/PhysRevA.92.063851>. Disponível em: <https://journals.aps.org/pr/abstract/10.1103/PhysRevA.92.063851>. Acesso em 29 de setembro de 2023.
- AMODEI, D. et al. Deep speech 2: End-to-end speech recognition in English and Mandarin. In: PMLR. **International conference on machine learning**. [S.l.], 2016. p. 173–182.
- ARRIZÓN, V.; RUIZ, U.; CARRADA, R.; GONZÁLEZ, L. A. Pixelated phase computer holograms for the accurate encoding of scalar complex fields. [S.l.]: **Optical Society of America**, 2007. 3500–3507 p. DOI: <https://doi.org/10.1364/JOSAA.24.003500>. Disponível em: <https://opg.optica.org/josaa/abstract.cfm?uri=josaa-24-11-3500>. Acesso em 10 de setembro de 2023.
- BAETS, R.; ROELKENS, G. Fotonica. **UGent**, February, 2004.
- BAHETI, P. Supervised and Unsupervised Learning [Differences Examples]. [S.l.]: **V7**, 2023. Disponível em: <https://www.v7labs.com/blog/supervised-vs-unsupervised-learning>. Acesso em: 14 de abril de 2023.

BARBASTATHIS, G.; OZCAN, A.; SITU, G. On the use of deep learning for computational imaging. [S.l.]: **Optica Publishing Group**, 2019. 921–943 p. DOI: <https://doi.org/10.1364/OPTICA.6.000921>. Disponível em: <https://opg.optica.org/optica/fulltext.cfm?uri=optica-6-8-921id=416103>. Acesso em 14 de abril de 2023.

BE, A. S.; TEICH, M. Wiley series in pure and applied optics. **Fundamentals of Photonics (second Edition)**, 2007.

BEAR, M. F.; CONNORS, B. W.; PARADISO, M. A. et al. Neurociencia: explorando el cerebro. **Barcelona**, 2002.

BELAY, G. Y.; KINDS, Y.; GOOSSENS, L.; GURUNG, K.; BOSMANS, N.; DILTOER, R.; ERALY, J.; VERVAEKE, M.; THIENPONT, H.; HOOREWEDER, B. V. et al. Dynamic optical beam shaping system to generate gaussian and top-hat laser beams of various sizes with circular and square footprint for additive manufacturing applications. [S.l.]: **Procedia CIRP**, Elsevier, 2022. 75–80 p. DOI: <https://doi.org/10.1016/j.procir.2022.08.134>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2212827122010174>. Acesso em 20 de agosto de 2023.

BRIGANTI, G.; MOINE, O. L. Artificial intelligence in medicine: today and tomorrow. [S.l.]: Frontiers in medicine. **Frontiers Media SA**, 2020. 27 p. Disponível em: <https://www.frontiersin.org/articles/10.3389/fmed.2020.00027/full>. Acesso em 21 de abril de 2023.

BROMBERG, Y.; CAO, H. Generating non-Rayleigh speckles with tailored intensity statistics. [S.l.]: **Physical Review Letters**. APS, 2014. 213904 p. DOI: <https://doi.org/10.1103/PhysRevLett.112.213904>. Disponível em: <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.112.213904>. Acesso em 15 de junho de 2023.

CHEN, J.; PENG, J.; YANG, C.; DUN, X.; CHENG, X.; CHEN, H.; REN, J. Hybrid optical-electronic neural network with pseudoinverse learning for classification inference. [S.l.]: **Applied Physics Letters**. AIP Publishing, 2021. DOI:

<https://doi.org/10.1063/5.0059010>. Disponível em:

<https://pubs.aip.org/aip/apl/article-abstract/119/11/114102/39953/Hybrid-optical-electronic-neural-network-with?redirectedFrom=fulltext>. Acesso em 20 de setembro de 2023.

CHEN, L.-C.; PAPANDREOU, G.; KOKKINOS, I.; MURPHY, K.; YUILLE, A. L.

DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. [S.l.]: **IEEE transactions on pattern analysis and machine intelligence**, 2017. 834–848 p. DOI: 10.1109/TPAMI.2017.2699184.

CHEN, Y. 4f-type optical system for matrix multiplication. [S.l.]: **Optical Engineering**

.1993. 77–79 p. DOI: <https://doi.org/10.1117/12.60078>. Disponível em:

<https://www.spiedigitallibrary.org/journals/optical-engineering/volume-32/issue-1/0000/4f-type-optical-system-for-matrix-multiplication/10.1117/12.60078.short>. Acesso em 19 de junho de 2023.

CHENG, T.-Y.; CHOU, D.-Y.; LIU, C.-C.; CHANG, Y.-J.; CHEN, C.-C. Optical neural

networks based on optical fiber-communication system. [S.l.]: **Neurocomputing**. Elsevier, 2019. 239–244 p. DOI: <https://doi.org/10.1016/j.neucom.2019.07.051>.

EUROPE, C. of. History of Artificial Intelligence. 2023. Disponível em:

<https://www.coe.int/en/web/artificial-intelligence/history-of-ai>. Acesso em: 26 de setembro de 2023.

FELDMANN, J.; YOUNGBLOOD, N.; WRIGHT, C. D.; BHASKARAN, H.; PERNICE,

W. H. All-optical spiking neurosynaptic networks with self-learning capabilities. [S.l.]: **Nature**. 2019. 208–214 p. DOI: <https://doi.org/10.1038/s41586-019-1157-8>.

Disponível em: <https://www.nature.com/articles/s41586-019-1157-8>. Acesso em 14 de junho de 2023.

GABOR, D. A New Microscopic Principle. [S.l.]: **nature**, 1948. 777–778 p. DOI:

<https://doi.org/10.1038/161777a0>. Disponível em:

<https://www.nature.com/articles/161777a0>. Acesso em 19 de agosto de 2023.

GERCHBERG, R.; SAXTON, W. A practical algorithm for the determination of phase

from image and diffraction plane pictures. **SPIE milestone series MS**, SPIE-THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING, v. 93, p. 306–306, 1994.

GESUALDI, M. R. Simplified algorithms of computational holograms: I—for generation of the 3D images. [S.I.]: **Revista Brasileira de Ensino de Física**. SciELO Brasil, 2023. DOI: <https://doi.org/10.1590/1806-9126-RBEF-2022-0248>. Disponível em: <https://www.scielo.br/j/rbef/a/NfYc6tTJ6rQKbHkc4HFKBkJ/abstract/?lang=en>. Acesso em 19 de agosto de 2023.

GOODMAN, J. W. Introduction to Fourier optics. [S.I.]: **Roberts and Company publishers**, 2005.

GOODMAN, J. W. Speckle phenomena in optics: theory and applications. [S.I.]: **Roberts and Company Publishers**, 2007.

GOODMAN, J. W.; DIAS, A.; WOODY, L. Fully parallel, high-speed incoherent optical method for performing discrete Fourier transforms. [S.I.]: **Optics Letters**. Optica Publishing Group, 1978. 1–3 p. DOI: <https://doi.org/10.1364/OL.2.000001>. Disponível em: <https://opg.optica.org/ol/abstract.cfm?uri=ol-2-1-1>. Acesso em 20 de agosto de 2023.

GOODMAN, J. W.; SUTTON, P. Introduction to Fourier optics. **Quantum and Semiclassical Optics-Journal of the European Optical Society Part B**, Springer, v. 8, n. 5, p. 1095, 1996.

HAMAMATSU. LCOS-SLM (Liquid Crystal on Silicon - Spatial Light Modulator). 2017. Disponível em: <http://www.scbeibin.com/dda/htmm/uploadfile/20170106151127877.pdf>. Acesso em: 13 de maio de 2023.

HAYKIN, S. **Neural networks and learning machines**, 3/E. [S.I.]: Pearson Education India, 2009.

HECHT, E. **Optik**. [S.I.]: Walter de Gruyter GmbH & Co KG, 2023.

HEINZ, R.; ARTMAN, J.; LEE, S. Matrix multiplication by optical methods. [S.I.]: **Applied Optic**, 1970. 2161–2168 p. DOI: <https://doi.org/10.1364/AO.9.002161>. Disponível em: <https://opg.optica.org/ao/abstract.cfm?uri=ao-9-9-2161>. Acesso em 19 de junho de 2023.

HUANG, G.-B.; ZHOU, H.; DING, X.; ZHANG, R. Extreme learning machine for regression and multiclass classification. [S.l.]: **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, 2011. 513–529 p. DOI: 10.1109/TSMCB.2011.2168604.

HUANG, G.-B.; ZHU, Q.-Y.; SIEW, C.-K. Extreme learning machine: theory and applications. [S.l.]: **Neurocomputing**, 2006. 489–501 p. DOI: <https://doi.org/10.1016/j.neucom.2005.12.126>.

JANG, S.-W.; CHOI, W.; KIM, S.; LEE, J.; NA, S.; HAM, S.; PARK, J.; KANG, H.; JU, B.-K.; KIM, H. Complex spatial light modulation capability of a dual layer in-plane switching liquid crystal panel. [S.l.]: **Scientific reports**, 2022. 1–14 p. DOI: <https://doi.org/10.1364/JOSAA.24.003500>. Disponível em: <https://doi.org/10.1038/s41598-022-12292-4>. Acesso em 10 de setembro de 2023.

LIGHTON. IMAGE CLASSIFICATION THROUGH A SCATTERING MEDIA. 2016. Disponível em: <https://www.lkb.upmc.fr/opticalimaging/opticalcomputing/>. Acesso em: 11 de outubro de 2023.

LIN, X.; RIVENSON, Y.; YARDIMCI, N. T.; VELI, M.; LUO, Y.; JARRAHI, M.; OZCAN, A. All-optical machine learning using diffractive deep neural networks. [S.l.]: **Science**, 2018. 1004–1008 p. DOI: <https://doi.org/10.1126/science.aat8084>. Disponível em: <https://www.science.org/doi/full/10.1126/science.aat8084>. Acesso em 14 de agosto de 2023.

LIU, J.; WU, Q.; SUI, X.; CHEN, Q.; GU, G.; WANG, L.; LI, S. Research progress in optical neural networks: theory, applications and developments. [S.l.]: **Photonixn**, 2021. 1–39 p. DOI: <https://doi.org/10.1186/s43074-021-00026-0>. Disponível em: <https://photonix.springeropen.com/articles/10.1186/s43074-021-00026-0>. Acesso em 14 de agosto de 2023.

LUDGATE, P. E. On a proposed analytical machine. **The Origins of Digital Computers**, Springer, p. 73–87, 1982.

LUENBERGER, D. G.; YE, Y. et al. **Linear and nonlinear programming**. [S.l.]: Springer, 1984. v. 2.

LUKOŠEVIČIUS, M.; JAEGER, H. Reservoir computing approaches to recurrent neural network training. [S.l.]: **Computer science review**, 2009. 127–149 p. DOI: <https://doi.org/10.1016/j.cosrev.2009.03.005>.

MAASS, W.; NATSCHLÄGER, T.; MARKRAM, H. Real-time computing without stable states: A new framework for neural computation based on perturbations. [S.l.]: **Neural computation**, 2002. 2531–2560 p. DOI: 10.1162/089976602760407955. Disponível em: <https://ieeexplore.ieee.org/abstract/document/6789852>. Acesso em: 13 de agosto de 2023.

MAHESH, B. Machine learning algorithms-a review. [S.l.]: **International Journal of Science and Research (IJSR)**. [Internet], 2020. 381–386 p. DOI: 10.21275/ART20203995.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. [S.l.]: **The bulletin of mathematical biophysics**, 1943. 115–133 p. DOI: <https://doi.org/10.1007/BF02478259>. Disponível em: <https://link.springer.com/article/10.1007/BF02478259>. Acesso em: 15 de abril de 2023.

MELANDRI, L. Introduction to reservoir computing methods. 2014.

MINAEE, S.; BOYKOV, Y.; PORIKLI, F.; PLAZA, A.; KEHTARNAVAZ, N.; TERZOPOULOS, D. Image segmentation using deep learning: A survey. [S.l.]: **IEEE transactions on pattern analysis and machine intelligence**, 2021. 3523–3542 p. DOI: <https://doi.org/10.1109/TPAMI.2021.3059968>. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9356353>. Acesso em 14 de junho de 2023.

MINSKY, M.; PAPERT, S. An introduction to computational geometry. **Cambridge tiass.**, HIT, v. 479, p. 480, 1969.

MIŠIĆ, M. J.; ĐURČEVIĆ, Đ. M.; TOMAŠEVIĆ, M. V. Evolution and trends in GPU computing. [S.l.]: **2012 Proceedings of the 35th International Convention MIPRO**, 2012. 289–294 p. Disponível em: <https://ieeexplore.ieee.org/abstract/document/6240658>. Acesso em 14 de agosto de 2023.

MITCHELL, T. M. et al. Machine learning. [S.l.]: **McGraw-hill New York**, 2007. v. 1.

NEMETALLAH, G.; AYLO, R.; WILLIAMS, L. A. Analog and digital holography with MATLAB. 2015.

NEUMANN, J. V. First Draft of a Report on the EDVAC. [S.l.]: **IEEE Annals of the History of Computing**, 1993. 27–75 p. DOI: <https://doi.org/10.1109/85.238389>. Disponível em: <https://ieeexplore.ieee.org/abstract/document/238389>. Acesso em 15 de maio de 2023.

NIELSEN, M. A. **Neural networks and deep learning**. [S.l.]: Determination press San Francisco, CA, USA, 2015. v. 25.

RADFORD, A.; KIM, J. W.; XU, T.; BROCKMAN, G.; MCLEAVEY, C.; SUTSKEVER, I. Robust speech recognition via large-scale weak supervision. In: PMLR. **International Conference on Machine Learning**. [S.l.], 2023. p. 28492–28518.

RASCHKA, S.; MIRJALILI, V. **Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2**. [S.l.]: Packt Publishing Ltd, 2019.

REICHL, L. **A modern course in statistical physics**. [S.l.]: University of Texas Press, Austin, 1998.

RIVENSON, Y.; WANG, H.; WEI, Z.; HAAN, K. de; ZHANG, Y.; WU, Y.; GÜNAYDIN, H.; ZUCKERMAN, J. E.; CHONG, T.; SISK, A. E. et al. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. [S.l.]: **Nature biomedical engineering**, 2019. 466–477 p. DOI: <https://doi.org/10.1038/s41551-019-0362-y>. Disponível em: <https://www.nature.com/articles/s41551-019-0362-y>. Acesso em 16 de agosto de 2023.

ROSALES-GUZMÁN, C.; FORBES, A. How to Shape Light with Spatial Light Modulators (SPIE, 2017). SL30, p. 1–57.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. [S.l.]: **Nature**, 1986. 533–536 p. DOI:

<https://doi.org/10.1038/323533a0>. Disponível em:

<https://www.nature.com/articles/323533a0>. Acesso em: 15 de abril de 2023.

RUSSELL, S. J. **Artificial intelligence a modern approach**. [S.l.]: Pearson Education, Inc., 2010.

SAADE, A.; CALTAGIRONE, F.; CARRON, I.; DAUDET, L.; DRÉMEAU, A.; GIGAN, S.; KRZAKALA, F. Random projections through multiple optical scattering: Approximating kernels at the speed of light. In: **2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. [S.l.: s.n.], 2016. p. 6215–6219. DOI: <https://doi.org/10.1109/ICASSP.2016.7472872>.

SAKEMI, Y.; MORINO, K.; LELEU, T.; AIHARA, K. Model-size reduction for reservoir computing by concatenating internal states through time. [S.l.]: **Scientific reports**, 2020. 21794 p. DOI: <https://doi.org/10.1038/s41598-020-78725-0>. Disponível em: <https://www.nature.com/articles/s41598-020-78725-0>. Acesso em: 11 de outubro de 2023.

SANTANA, R. A.; BASTOS-FILHO, C. J.; SILVA, D. R.; MARTINS-FILHO, J. F.; CHAVES, D. A. R. Roteamento em redes ópticas transparentes utilizando redes neurais de Hopfield. **Dissertação de Mestrado**, UPE, v. 2010, 2010.

SCIENCE, D. T. I. A. D. Gradiente Descendente Estocástico. 2022. Disponível em: <https://didatica.tech/gradiente-descendente-estocastico>. Acesso em: 24 de maio de 2023.

SHEN, Y.; HARRIS, N. C.; SKIRLO, S.; PRABHU, M.; BAEHR-JONES, T.; HOCHBERG, M.; SUN, X.; ZHAO, S.; LAROCHELLE, H.; ENGLUND, D. et al. Deep learning with coherent nanophotonic circuits. [S.l.]: **Nature photonics**, 2017. 441–446 p. DOI: <https://doi.org/10.1038/nphoton.2017.93>. Disponível em: <https://www.nature.com/articles/nphoton.2017.93>. Acesso em 14 de junho de 2023.

SPILSBURY, M. J.; EUCEDA, A. Transformada rápida de Fourier. [S.l.]: **Revista de la Escuela de Física**, 2016. 45–52 p. DOI: <https://doi.org/10.5377/ref.v4i2.8276>. Disponível em: <https://camjol.info/index.php/fisica/article/view/8276>. Acesso em 20 de agosto de 2023.

SRINIDHI, C. L.; CIGA, O.; MARTEL, A. L. Deep neural network models for computational histopathology: A survey. [S.l.]: **Medical Image Analysis**, 2021. 101813 p. DOI: <https://doi.org/10.1016/j.media.2020.101813>. Acesso em 01 de setembro de 2023.

SVELTO, O.; HANNA, D. C. et al. **Principles of lasers**. [S.l.]: Springer, 2010. v. 1.

TEAM, D. S. Banco de dados MNIST. 2020. Disponível em: <https://datascience.eu/pt/processamento-de-linguagem-natural/banco-de-dados-mnist/>. Acesso em: 13 de maio de 2023.

THORLABS. CMOS-CAMERAS: USB 2.0 and USB 3.0. 2020. Disponível em: <https://www.thorlabs.com/catalogpages/obsolete/2020/DCC3240C.pdf>. Acesso em: 30 de julho de 2023.

WANG, Hanchen et al. Scientific discovery in the age of artificial intelligence. **Nature**, v. 620, n. 7972, p. 47-60, 2023. DOI: <https://doi.org/10.1038/s41586-023-06221-2>. Disponível em: <https://www.nature.com/articles/s41586-023-06221-2>. Acesso em: 29 de junho de 2023.

WHITFIELD, B. Feedforward Neural Networks: A Quick Primer for Deep Learning. 2022. Disponível em: <https://builtin.com/data-science/feedforward-neural-network-intro>. Acesso em: 09 de setembro de 2023.

WU, Lin; ZHANG, Ziyang. Only-train-electrical-to-optical conversion (OTEOC): simple diffractive neural networks with optical readout. **Optics Express**, v. 30, n. 15, p. 28024-28037, 2022. DOI: <https://doi.org/10.1364/OE.462370>. Disponível em: <https://opg.optica.org/oe/fulltext.cfm?uri=oe-30-15-28024&id=479531>. Acessado em: 16 de maio de 2023.

WU, Yang et al. Adaptive weighted Gerchberg-Saxton algorithm for generation of phase-only hologram with artifacts suppression. **Optics express**, v. 29, n. 2, p. 1412-1427, 2021. DOI: <https://doi.org/10.1364/OE.413723>. Disponível em: <https://opg.optica.org/oe/fulltext.cfm?uri=oe-29-2-1412&id=446366>. Acessado em: 16 de maio de 2023.

XU, Yongjun et al. Artificial intelligence: A powerful paradigm for scientific research. **The Innovation**, v. 2, n. 4, 2021. DOI: <https://doi.org/10.1016/j.xinn.2021.100179>. Disponível em: [https://www.cell.com/article/S2666-6758\(21\)00104-1/fulltext](https://www.cell.com/article/S2666-6758(21)00104-1/fulltext). Acessado em: 16 de maio de 2023.

YU, Kun-Hsing; BEAM, Andrew L.; KOHANE, Isaac S. Artificial intelligence in healthcare. **Nature biomedical engineering**, v. 2, n. 10, p. 719-731, 2018. DOI: <https://doi.org/10.1038/s41551-018-0305-z>. Disponível em: <https://www.nature.com/articles/s41551-018-0305-z>. Acessado em: 18 de maio de 2023.

ZANG, Yubin et al. Electro-optical neural networks based on time-stretch method. **IEEE Journal of Selected Topics in Quantum Electronics**, v. 26, n. 1, p. 1-10, 2019. DOI: 10.1109/JSTQE.2019.2957446.

ZHANG, Fan et al. A fast algorithm for multi-class learning from label proportions. **Electronics**, v. 8, n. 6, p. 609, 2019. DOI: 10.3390/electronics8060609. Disponível em: <https://www.mdpi.com/2079-9292/8/6/609>. Acessado em: 18 de maio de 2023.

ZHANG, Zichen; YOU, Zheng; CHU, Daping. Fundamentals of phase-only liquid crystal on silicon (LCOS) devices. **Light: Science & Applications**, v. 3, n. 10, p. e213-e213, 2014. DOI: <https://doi.org/10.1038/lisa.2014.94>. Disponível em: <https://www.nature.com/articles/lisa201494>. Acessado em: 16 de maio de 2023.

APÊNDICE A – HOLOGRAFIA

A holografia é uma técnica óptica-interferométrica utilizada para o registro e reconstrução de imagens tridimensionais. Essa técnica pode ser classificada em dois tipos: holografia clássica (HC) e holografia gerada por computador (HGC). Neste apêndice, apresentaremos os principais conceitos dessas técnicas, além de mostrar definições que nos permitam criar hologramas aplicáveis em nossa experiência.

A.0.1 Holografia clássica

A holografia clássica, também conhecida como holografia analógica, foi desenvolvida pelo físico húngaro Dennis Gabor em 1948. Essa técnica permite registrar não apenas a amplitude, mas também a fase da luz de um objeto. Esse registro é obtido através da interferência de dois feixes de luz, resultando na formação de um padrão holográfico (GABOR, 1948). O processo holográfico compreende duas etapas principais: a gravação e a reconstrução.

Durante a etapa de gravação, consideramos uma onda monocromática representada pela amplitude complexa $U_o(x, y)$, que é chamada de “onda objeto”, e uma segunda onda monocromática plana chamada $U_r(x, y)$, referida como “onda de referência”. Essas duas ondas estão coerentes entre si, o que significa que elas possuem uma relação de fase constante uma em relação a outra. A superposição dessas duas ondas resulta em um padrão de interferência que pode ser registrado em um elemento óptico distribuído num plano (x, y) , como uma transparência ou um meio holográfico. O campo complexo que representa a onda objeto é dado por $U_o(x, y) = |U_o(x, y)| \exp[j\phi(x, y)]$, e a onda de referência, representada pelo campo fora do eixo (*off-axis*), é dada por $U_r(x, y) = |U_r(x, y)| \exp(-2\pi j\zeta_r x)$. Nessa configuração (*off-axis*), a onda de referência é deslocada angularmente em relação à onda objeto, e temos que $\zeta_r = \sin \theta / \lambda$. Ao usar um campo fora do eixo, o holograma resultante pode conter informações adicionais, como a direção e o ângulo da fonte de luz que iluminou o objeto. Essas informações extras podem ser exploradas durante a reconstrução do holograma, proporcionando uma visualização tridimensional mais precisa e realista do objeto gravado (BE; TEICH, 2007). As duas ondas ao interferirem, formam um padrão

de interferência gravado no elemento óptico (holograma). Esse holograma possui uma transmitância complexa que, pode ser relacionada com a intensidade do padrão de interferência (GESUALDI, 2023). A intensidade da luz que chega ao meio de registro (holograma), ou seja, o módulo quadrado da soma dos campos complexos, é registrada durante o tempo de exposição τ , e é representada por $I(x, y)$.

$$\begin{aligned} I(x, y) &= |U_o(x, y) + U_r(x, y)|^2 \\ &= |U_o(x, y)|^2 + |U_r(x, y)|^2 + U_o^*(x, y)U_r(x, y) + U_o(x, y)U_r^*(x, y) \\ &= |U_o(x, y)|^2 + |U_r(x, y)|^2 + 2|U_o(x, y)||U_r(x, y)| \cos \phi(x, y) + 2\pi\zeta_r x \end{aligned}$$

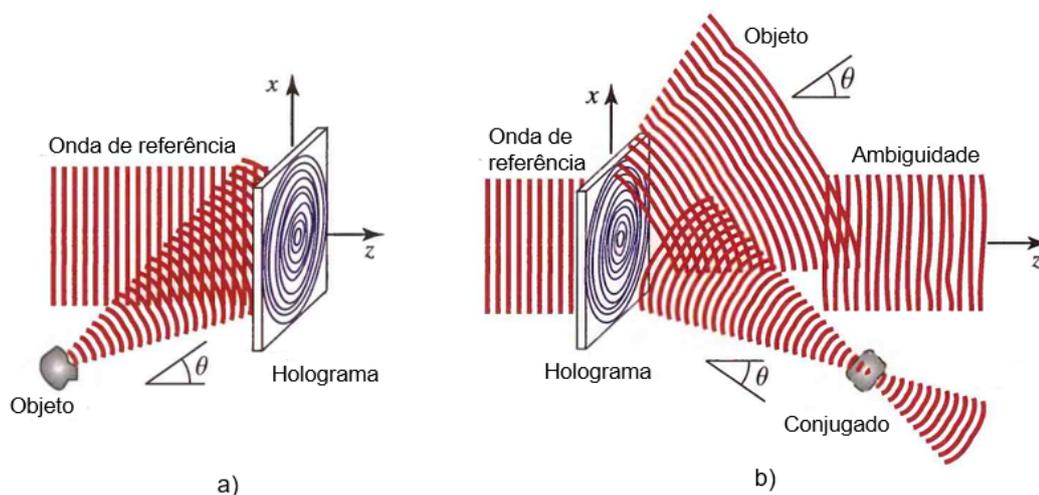
A contribuição dos dois primeiros termos da equação é determinada exclusivamente pelas amplitudes das ondas de forma individual, enquanto o terceiro termo carrega informações tanto da amplitude quanto da fase relativa entre o feixe de referência e o feixe objeto. A figura 78(a) ilustra o processo de registro do holograma. A transmitância complexa do holograma é dada por $t(x, y) = t_0 + \gamma_r I(x, y)$, onde o parâmetro t_0 é definido como o “background” da amplitude de transmissão, que representa uma componente constante ou de referência na transmissão da luz. O parâmetro γ é determinado pelo meio de registro holográfico e está relacionado às propriedades ópticas desse meio. Durante a etapa de reconstrução, que consiste na recuperação da onda objeto, o holograma é iluminado com a mesma onda de referência utilizada na etapa de registro. Isso resulta na formação da amplitude complexa da onda transmitida, representada por $U(x, y)$, que pode ser escrita como:

$$U(x, y) = U_r(x, y)t(x, y)$$

A.0.2 Holografia gerada por computador

Com o avanço da tecnologia computacional, surgiu a holografia gerada por computador (HGC), também conhecida como holografia digital. Diferentemente da holografia clássica, que envolve a interferência de feixes de luz reais, a holografia gerada por computador utiliza modelos matemáticos e simulações computacionais para criar imagens holográficas. Na holografia gerada por computador, o processo de gravação do holograma é realizado de forma computacional. Em vez de utilizar placas fotográficas ou filmes holográficos, os dados do holograma são armazenados digitalmente.

Figura 78 – a) Esquema representativo para processo de registro do holograma como um feixe de referência fora do eixo, b) processo reconstrução do holograma com suas diferentes ordens de difração.



Fonte: BE e Teich (2007)

Esses dados representam as informações de amplitude e fase da onda de luz, permitindo uma maior flexibilidade e controle sobre a imagem holográfica resultante. Uma das principais aplicações da holografia gerada por computador na óptica é a capacidade de modular frentes de onda. Isso é possível através do uso de dispositivos de gravação como os Moduladores Espaciais de Luz-SLM (NEMETALLAH; AYLO; WILLIAMS, 2015). Os SLMs são dispositivos eletrônicos que podem modificar as propriedades da luz, como a amplitude, fase e polarização, em pontos específicos do feixe de luz. Essa capacidade de controlar e modular as frentes de onda permite a criação de padrões complexos e precisos, resultando em imagens holográficas de alta resolução e detalhamento.

A.0.2.1 Holograma de fase

Existem diversas abordagens numéricas para a geração de hologramas por computador, sendo que muitas delas se baseiam nos princípios da óptica de Fourier. Esses princípios permitem a transformação do campo de onda do domínio espacial para o domínio da frequência, facilitando a manipulação e a reconstrução das informações holográficas. Um exemplo de holograma amplamente utilizado é o holograma de fase, proposto por Arrizón et al. (2007), que permite codificar informações tanto sobre a

fase quanto sobre a amplitude da luz. Nesse tipo de holograma, um campo complexo é representado por uma função que descreve a amplitude e a fase da luz, dado por:

$$E(x, y) = A(x, y) \exp [j\phi(x, y)] \quad (81)$$

A amplitude, denotada por $A(x, y)$, varia no intervalo de $[0, 1]$, enquanto a fase, representada por $\phi(x, y)$, varia de $[-\pi, \pi]$. Um aspecto destacado por Arrizón em seu trabalho é a importância de garantir que os valores das amplitudes complexas da função $E(x, y)$ pertençam a um conjunto de números complexos com módulos menores ou iguais a 1, denotado como conjunto Ω_E . O objetivo é codificar o campo $E(x, y)$ através de uma função de transmitância de fase. No entanto, de forma geral, um holograma que codifica uma modulação complexa arbitrária $E(x, y)$ possui uma transmitância complexa restrita a valores contidos em um subconjunto específico. Para o caso de um holograma de fase, este subconjunto é formada por pontos contidos dentro de um círculo unitário traçado num plano complexo (ROSALES-GUZMÁN; FORBES, 2017). A expressão que descreve a função de transmitância, capaz de codificar tanto a fase quanto a amplitude, é dada por:

$$H(x, y) = \exp[j\psi(A, \phi)] \quad (82)$$

Aqui, $\psi(A, \phi)$ representa a função de modulação de fase, que é essencial para a holografia gerada por computador. Através dessa função, é possível obter uma codificação adequada da amplitude e fase, desde que encontremos uma solução que satisfaça a equação mencionada anteriormente. Portanto, a função $\psi(A, \phi)$ desempenha um papel fundamental na obtenção dos resultados desejados na geração de HGCs. Arrizón et al. (2007) propõe uma abordagem eficiente para encontrar as possíveis soluções de $\psi(A, \phi)$, representando a equação de transmitância como uma série de Fourier no domínio da fase ϕ , resultando em:

$$H(x, y) = \sum_{q=-\infty}^{\infty} H_q(x, y) \quad (83)$$

Nessa equação, cada termo $H_q(x, y)$ é definido por:

$$H_q(x, y) = C_q^A \exp(jq\phi) \quad (84)$$

O coeficiente C_q^A é determinado por:

$$C_q^A = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp [j\psi(A, \phi)] \exp (-jq\phi) d\phi \quad (85)$$

A reconstrução do campo original $E(x, y)$ pode ser realizada considerando o primeiro termo da equação (13), desde que a seguinte identidade seja satisfeita:

$$C_1^A = Aa \quad (86)$$

Nessa expressão, C_1^A representa o coeficiente do primeiro termo da série de Fourier e A é a amplitude complexa original do campo $E(x, y)$. A quantidade a é um valor positivo.

A razão pela qual essa identidade é importante para a reconstrução do campo está relacionada à relação entre a amplitude complexa original e a amplitude do primeiro termo da série de Fourier. Ao satisfazer essa identidade, garantimos que a amplitude do primeiro termo seja proporcional à amplitude original do campo.

No trabalho de Arrizón, foram apresentados três tipos de hologramas de fase, cada um com suas próprias características e eficiências distintas. Dentre esses tipos, vamos nos concentrar no terceiro holograma proposto por Arrizón, pois é o que utilizamos em nosso trabalho. Para esse tipo de holograma, consideramos funções de $\psi(A, \phi)$ que possuem simetria ímpar. A expressão para $\psi(A, \phi)$ nesse caso pode ser escrita como:

$$\psi(A, \phi) = f(A) \sin (\phi) \quad (87)$$

Com base nessa modulação de fase, podemos representar a função de transmitância como uma série de Fourier. A partir dessa representação, obtemos os coeficientes necessários para a construção desse tipo específico de holograma. Como a função de modulação possui propriedades especiais, fazemos uso da identidade de Jacobi-Anger, que é dada por:

$$\exp [jx \cos (\theta)] = \sum_{q=-\infty}^{\infty} j^q J_q(x) \exp(jq\theta) \quad (88)$$

$$\exp [jx \sin (\theta)] = \sum_{q=-\infty}^{\infty} J_q(x) \exp(jq\theta) \quad (89)$$

Essa identidade desempenha um papel fundamental ao decompor a função exponencial complexa de senos ou cossenos em uma série de exponencias complexas

ponderadas pelas funções de Bessel de q -ésima ordem. Combinando as equações 87, 82 e 89, temos:

$$H(x, y) = \sum_{q=-\infty}^{\infty} J_q[f(A)] \exp(jq\theta) \quad (90)$$

Ao comparar a equação 90 com a equação 84, podemos ver que os coeficientes da série de Fourier são dados por:

$$C_q^A = J_q[f(A)] \quad (91)$$

Da relação 86, temos que $J_1[f(A)] = Aa$, é resolvida por inversão numérica para obter os valores de $f(A)$. O valor máximo de a para o qual a equação acima é satisfeita é $a = 0.58$, que corresponde ao máximo valor da função de Bessel de primeira ordem $J_1(x)$, ocorrendo quando $x = 1.84$.

É importante ressaltar que, ao codificar um campo óptico em um holograma de fase, o espectro de Fourier $F\{E(x, y)\} = E'(\xi, \eta)$ do campo codificado $E(x, y)$ está centrado no plano de Fourier em $(\xi, \eta) = 0$. Essa propriedade implica que os espectros dos diferentes termos H_q da equação 13, que representam as componentes de difração associadas ao holograma, também se encontram centrados no mesmo plano.

No entanto, essa centralização no plano de Fourier dificulta a recuperação do campo original por meio de técnicas de filtragem espacial convencionais. Isso ocorre porque, ao aplicar um filtro espacial para isolar uma determinada ordem de difração, todas as outras ordens também são afetadas devido à sobreposição dos espectros.

Para contornar esse problema, é necessário modificar o holograma adicionando uma modulação de fase específica. Essa modulação consiste em uma fase linear dada por $2\pi(\xi_0 x + \eta_0 y)$, onde ξ_0 e η_0 são constantes que determinam o deslocamento espacial introduzido.

Ao adicionar essa modulação de fase ao holograma, a função de transmitância é modificada de acordo com essa nova fase. Essa modificação permite um isolamento espacial das ordens de difração, separando-as no plano de Fourier e facilitando a recuperação do campo original.

Dessa forma, a função de transmitância modificada, considerando a adição da modulação de fase $2\pi(\xi_0 x + \eta_0 y)$, é expressa por:

$$H(x, y) = \sum_{q=-\infty}^{\infty} H_q(x, y) \exp[2\pi j(q\xi_0 x + q\eta_0 y)] \quad (92)$$

A inclusão dessa modulação de fase possibilita a recuperação seletiva das ordens de difração desejadas, permitindo assim a reconstrução do campo óptico original com maior precisão e qualidade.

APÊNDICE B – MODULADOR ESPACIAL DE LUZ

Os moduladores espaciais de luz (SLMs) desempenham um papel indispensável no processamento óptico de informações, permitindo a conversão de dados eletrônicos em sinais ópticos que são espacialmente modulados (GOODMAN, 2005). Essa modulação possibilita o controle preciso da fase, amplitude ou polarização da luz, desempenhando um papel crucial em diversas áreas da ciência e tecnologia. As aplicações dos SLMs abrangem campos como óptica adaptativa, microscopia, holografia e comunicações ópticas, contribuindo significativamente para o avanço dessas áreas (JANG et al., 2022). Entre os diversos tipos de SLMs existentes, um dos mais importantes é o modulador espacial de luz de cristal líquido (LC-SLM), que é composto por uma camada de cristal líquido. Esses dispositivos permitem a manipulação controlada das propriedades ópticas por meio da aplicação de um campo elétrico aos cristais líquidos presentes na camada.

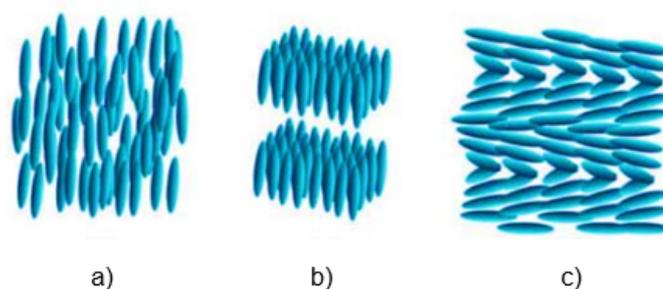
Neste apêndice, iremos aprofundar os princípios e o funcionamento deste dispositivo, que desempenhou um papel fundamental no desenvolvimento do nosso trabalho.

B.0.1 Modulador espacial de luz de Cristal Líquido (LC-SLM)

Os cristais líquidos (LCs) são materiais que apresentam propriedades únicas, situadas entre os estados líquido e sólido. Esses materiais consistem em moléculas ou cadeias moleculares que exibem uma ordem parcial em relação à orientação e/ou posição espacial. Essa organização parcial permite que os cristais líquidos sejam classificados em diferentes fases, que dependem da disposição e do movimento das moléculas. As três principais fases dos cristais líquidos são: nemáticos, esméticos e colestéricos, como ilustrado na Figura 79.

Na fase nemática, as moléculas têm uma orientação preferencial, mas não possuem uma estrutura de longo alcance. Na fase esmética, as moléculas formam camadas ou filamentos com diferentes orientações. Essa estrutura intermediária permite maior mobilidade das moléculas em relação à fase nemática. Já na fase colestérica, as moléculas se organizam em uma estrutura helicoidal tridimensional (GOODMAN, 2005), (ROSALES-GUZMÁN; FORBES, 2017).

Figura 79 – Fases dos Cristais Líquidos



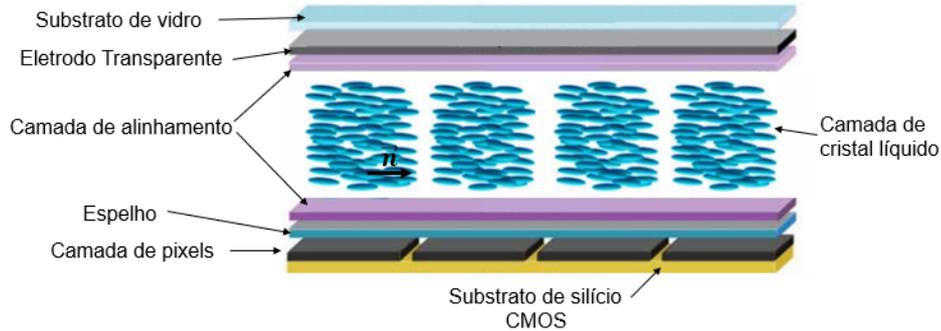
Fonte: Modificado de Rosales-Guzmán e Forbes (2017)

A fase mais amplamente utilizada nos moduladores espaciais de luz de cristais líquidos é a fase nemática. Onde, as moléculas estão organizadas de forma paralela, mas não estão empilhadas em camadas ordenadas. A fase nemática oferece a vantagem de permitir uma resposta rápida e precisa à aplicação de um campo elétrico externo. A capacidade de alinhar e reorientar rapidamente as moléculas nemáticas em resposta a campos elétricos externos permite a modulação eficiente da fase e da amplitude da luz incidente.

Uma das tecnologias mais utilizadas nos moduladores espaciais de luz de cristais líquidos é a LCOS (Liquid Crystal on Silicon, ou Cristal Líquido sobre Silício). A arquitetura de um dispositivo LCOS-SLM é evidenciada na Figura 80. Estes dispositivos combinam as propriedades dos cristais líquidos com a tecnologia de um substrato de silício CMOS (Complementary Metal Oxide Semiconductor) abaixo da camada de pixels. A camada de silício CMOS tem a finalidade de fornecer um substrato robusto e eletronicamente controlável para os pixels do dispositivo, que são formados por eletrodos com uma superfície espelhada de alumínio. Esta tecnologia possibilita que esses dispositivos modulem a fase, a amplitude ou a polarização por meio da transmissão ou reflexão da luz nestes moduladores, com uma resolução espacial altamente precisa (ZHANG; YOU; CHU, 2014).

Os LCOS-SLMs são projetados de forma a permitir a modificação da orientação das moléculas de cristal líquido (LC) de acordo com a tensão do pixel aplicada. Os cristais líquidos são materiais anisotrópicos, o que significa que suas propriedades físicas mudam com a direção, nesse sentido essas moléculas têm um índice de refração diferente ao longo de pelo menos dois dos três eixos como pode ser observado

Figura 80 – Aquitetura de um LCOS-SLM



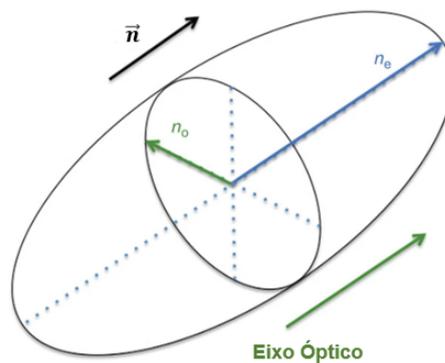
Fonte: Modificado de (ROSALES-GUZMÁN; FORBES, 2017)

na Figura 81, essa propriedade é chamada de birrefringência e o controle dessa propriedade é o que determina o tipo de modulação da luz. A birrefringência é expressa como:

$$\Delta n = n_e - n_o \quad (93)$$

Onde, n_e e n_o são os índices de refração nos eixos extraordinário e ordinário, respectivamente. Assim, devido a anisotropia, se as moléculas de LC rotacionam ocorre uma alteração do índice de refração, mas apenas em um estado de polarização.

Figura 81 – Representação de uma molécula de cristal líquido



Fonte: (ZHANG; YOU; CHU, 2014)

Se o campo de luz de entrada estiverem linearmente polarizada na direção do eixo (n_e) do cristal líquido, o dispositivo opera como um modulador de fase. Isso ocorre porque a alteração do índice de refração afeta o comprimento do percurso óptico e a velocidade da luz. Se a luz estiver linearmente polarizada e alinhada a um ângulo de

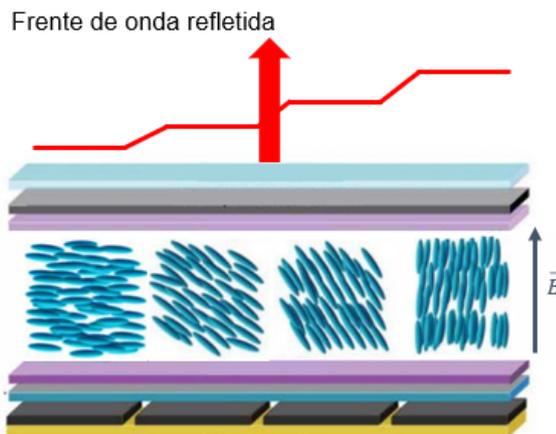
45° em relação ao eixo principal, permite atraso entre os eixos ordinário e extraordinário, resultando em uma mudança na polarização. Isso ocorre porque o estado de polarização é uma combinação vetorial dos dois vetores elétricos ortogonais da onda eletromagnética.

Se a luz do SLM passa por um polarizador linear, a modulação de amplitude ocorre devido ao controle eletrônico da birrefringência e, portanto, do estado de polarização.

Abordando a modulação de fase de maneira específica, como ilustrado na Figura 82, quando o SLM recebe tensão aplicada aos eletrodos, ocorre uma mudança na orientação do vetor diretor, \vec{n} , ou seja, a camada de cristal líquido muda seu índice de refração de forma controlada. Isso resulta em uma mudança na fase da luz que reflete através do dispositivo. A equação a seguir descreve a variação da fase $\Delta\Phi$ da onda óptica em função da tensão aplicada V . Com isso vemos que a aplicação de uma tensão elétrica modifica as propriedades ópticas do cristal líquido, alterando assim a fase da luz que passa por ele. Isso pode ser evidenciado na Figura 82.

$$\Delta\Phi(V) = \frac{2\pi}{\lambda}(n_e - n_o)d \quad (94)$$

Figura 82 – Representação da modulação de frente de onda feito por um SLM de cristais líquidos.



Fonte: Modificado de Rosales-Guzmán e Forbes (2017)

Onde observamos o arranjo físico do SLM, e como a fase da onda luminosa é modificada em resposta à variação da tensão aplicada.