



UNIVERSIDADE FEDERAL DO AMAZONAS  
FACULDADE DE TECNOLOGIA  
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

RECONHECIMENTO DE ÍRIS FAKE UTILIZANDO DESCRITORES DE  
MOMENTOS DE ZERNIKE E CLASSIFICADORES DE FORMAS

Geórgio Sá Colares

Manaus  
Dezembro de 2024



UNIVERSIDADE FEDERAL DO AMAZONAS  
FACULDADE DE TECNOLOGIA  
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

RECONHECIMENTO DE ÍRIS FAKE UTILIZANDO DESCRITORES DE  
MOMENTOS DE ZERNIKE E CLASSIFICADORES DE FORMAS

Geórgio Sá Colares

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

Orientador: Waldir Sabino da Silva Júnior

Manaus  
Dezembro de 2024

## Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

C683r Colares, Georgio Sa  
Reconhecimento de íris fake utilizando descritores de momentos de Zernike e classificadores de formas / Georgio Sa Colares . 2024  
76 f.: il. color; 31 cm.

Orientador: Waldir Sabino da Silva Júnior  
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Amazonas.

1. Reconhecimento de íris. 2. Momentos de Zernike. 3. Classificadores. 4. Descritores circulares. 5. Ataque de spoofing. I. Silva Júnior, Waldir Sabino da. II. Universidade Federal do Amazonas III. Título



Ministério da Educação  
Universidade Federal do Amazonas  
Coordenação do Programa de Pós-Graduação em Engenharia Elétrica

### FOLHA DE APROVAÇÃO

Poder Executivo Ministério da Educação  
Universidade Federal do Amazonas  
Faculdade de Tecnologia  
Programa de Pós-graduação em Engenharia Elétrica

Pós-Graduação em Engenharia Elétrica. Av. General Rodrigo Octávio Jordão Ramos, nº 3.000 - Campus Universitário, Setor Norte - Coroado, Pavilhão do CETELI. Fone/Fax (92) 99271-8954 Ramal:2607. E-mail: ppgee@ufam.edu.br

GEÓRGIO SÁ COLARES

### RECONHECIMENTO DE ÍRIS FAKE UTILIZANDO DESCRITORES DE MOMENTOS DE ZERNIKE E CLASSIFICADORES DE FORMAS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Aprovada em 06 de dezembro de 2024.

#### BANCA EXAMINADORA

Prof. Dr. Waldir Sabino da Silva Júnior - Presidente  
Prof. Dr. Gabriel Matos Araújo - Membro Titular 1 - Externo  
Prof. Dr. Frederico da Silva Pinagé - Membro Titular 2 - Externo

Manaus, 29 de novembro de 2024.



Documento assinado eletronicamente por **Gabriel Matos Araújo, Usuário Externo**, em 10/12/2024, às 16:05, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Frederico da Silva Pinagé, Professor do Magistério Superior**, em 11/12/2024, às 10:55, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

---



Documento assinado eletronicamente por **Waldir Sabino da Silva Júnior, Professor do Magistério Superior**, em 16/12/2024, às 10:55, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

---



A autenticidade deste documento pode ser conferida no site [https://sei.ufam.edu.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufam.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **2348650** e o código CRC **B2C31659**.

---

Av. Octávio Hamilton Botelho Mourão - Bairro Coroado 1 Campus Universitário Senador Arthur Virgílio Filho,  
Setor Norte - Telefone: (92) 3305-1181  
CEP 69080-900 Manaus/AM - [mestrado\\_engelettrica@ufam.edu.br](mailto:mestrado_engelettrica@ufam.edu.br)

Referência: Processo nº 23105.050721/2024-73

SEI nº 2348650

---

Criado por [31183646291](#), versão 3 por [31183646291](#) em 29/11/2024 11:21:37.

*Gostaria de expressar minha  
gratidão ao autor da frase "A  
ciência é uma forma de iluminar  
o mundo, de construir uma  
civilização baseada no  
conhecimento."...*

# Agradecimentos

- Primeiramente à minha mãe Maria Odinaíde, ao meu pai Heraldo, irmãos Áudrea e Geandre e ao principalmente ao meu amor Ana Maewe que nesse tempo deram apoio, tiveram paciência, e muita compreensão, por acreditarem na minha dedicação. Saibam que vocês foram fundamentais nesse período da minha vida.
- Ao meu orientador Waldir Sabino da Silva Júnior, por ser um amigo leal, compreensivo e paciente durante a orientação, me ajudando nos momentos desafiadores durante todo o mestrado, e sendo um exemplo de pesquisador para mim.
- Ao Professor Celso Barbosa Carvalho, por possibilitar minha continuidade nos estudos do mestrado, e compreenderem a importância desta formação para meu desenvolvimento pessoal e profissional.

Resumo da Dissertação apresentada à UFAM como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica

## RECONHECIMENTO DE ÍRIS FAKE UTILIZANDO DESCRITORES DE MOMENTOS DE ZERNIKE E CLASSIFICADORES DE FORMAS

Geórgio Sá Colares

Orientador: Waldir Sabino da Silva Júnior

Programa: Pós-Graduação em Engenharia Elétrica

Sistemas biométricos têm atraído o interesse de pesquisa, devido aos crescentes aperfeiçoamentos de métodos de fraude na autenticação de acesso a estes sistemas. Como uma sugestão de solução à falha de segurança de acesso a estes sistemas, nesta dissertação, desenvolveu-se um sistema de detecção de íris sintéticas utilizando momentos de Zernike como descritor de formas não alinhadas ou circulares. A contribuição deste trabalho está em investigar o desempenho dos momentos de Zernike quando combinados com classificadores de redes neurais artificiais para detecção de imagens de íris sintéticas. Nos experimentos, foram utilizadas bases de dados criadas a partir das combinações realizadas de outras bases de dados disponíveis em estudos na literatura relacionada. Os momentos de Zernike são aplicados localmente como descritores de formas, fornecendo bons resultados quando se utilizam classificadores de redes neurais mais complexos. A avaliação de desempenho do sistema proposto foi realizada para momentos de Zernike combinados com os três tipos diferentes de classificadores: *Multilayer Perceptron*, *Convolutional Neural Network* e *Residual neural network*. Para a combinação Zernike com MLP, teve-se a acurácia de 79%, enquanto que para o caso de combinação de Zernike com CNN e Zernike com ResNet na classificação, obteve-se, respectivamente, valores de 90% e 93% de acurácia.

**Palavras-chave:** Reconhecimento de íris, Momentos de Zernike, Classificadores, Descritores circulares, Ataque de *spoofing*.

Abstract of Dissertation presented to UFAM as a partial fulfillment of the requirements for the degree of Master in Electrical Engineering

FAKE IRIS RECOGNITION USING ZERNIKE MOMENT DESCRIPTORS  
AND SHAPE CLASSIFIERS

Geórgio Sá Colares

Advisor: Waldir Sabino da Silva Júnior

Department: Postgraduate in Electrical Engineering

Biometric systems have attracted research interest due to the increasing improvements in fraud prevention for access authentication to these systems. As a suggested solution to the security holes in accessing these systems, this dissertation developed a synthetic iris detection system using Zernike moments as a descriptor for non-aligned or circular shapes. The contribution of this work is to investigate the performance of Zernike moments when combined with artificial neural network classifiers for detecting synthetic iris images. In the experiments, databases were created from combinations of other databases available in studies on the referenced literature. Zernike moments are applied locally as shape descriptors, providing good results when using more complex neural network classifiers. The performance evaluation of the proposed system was performed for Zernike moments combined with three different types of classifiers: *Multilayer Perceptron*, *Convolutional Neural Network* and *Residual neural network*. For the combination of Zernike with MLP, the accuracy was 79%, while for the combination of Zernike with CNN and Zernike with ResNet in the classification, the accuracy values were 90% and 93%, respectively.

**Keywords:** Iris recognition, Zernike moments, Classifiers, Circular descriptors, Spoofing attack.

# Lista de Tabelas

2.1	Comparação de métodos com algoritmos de ataques de <i>spoofing</i> . . .	10
3.1	Tabela de termos de momentos de Zernike por ordem e número de repetições . . . . .	22
6.1	Sistema de momentos de Zernike com MLP onde os parâmetro de ordem de momentos de Zernike igual à $n = 5$ . . . . .	51
6.2	Sistema de momentos de Zernike com MLP onde os parâmetro de ordem de momentos de Zernike igual à $n = 9$ . . . . .	52
6.3	Sistema de momentos de Zernike com MLP onde os parâmetro de ordem de momentos de Zernike igual à $n = 12$ . . . . .	52
6.4	Sistema proposto onde a ordem de momentos de Zernike varia em números inteiros de $n = 0$ à $n = 12$ , com base de dados pré-processada CASIA, janela <i>non-overlapping</i> de 16x16, e classificador MLP conforme descrito em 6.4.1 . . . . .	53
6.5	Sistema de momentos de Zernike com CNN onde os parâmetro de ordem de momentos de Zernike igual à $n = 5$ . . . . .	54
6.6	Sistema de momentos de Zernike com CNN onde os parâmetro de ordem de momentos de Zernike igual à $n = 9$ . . . . .	55
6.7	Sistema de momentos de Zernike com CNN onde os parâmetro de ordem de momentos de Zernike igual à $n = 12$ . . . . .	55
6.8	Sistema proposto onde a ordem de momentos de Zernike varia em números inteiros de $n = 0$ à $n = 12$ , com base de dados pré-processada CASIA, janela <i>non-overlapping</i> de 16x16, e classificador CNN conforme descrito em 6.4.2 . . . . .	56

6.9	Sistema de momentos de Zernike com ResNet onde os parâmetro de ordem de momentos de Zernike igual à $n = 5$ . . . . .	57
6.10	Sistema de momentos de Zernike com ResNet onde os parâmetro de ordem de momentos de Zernike igual à $n = 9$ . . . . .	57
6.11	Sistema de momentos de Zernike com ResNet onde os parâmetro de ordem de momentos de Zernike igual à $n = 12$ . . . . .	58
6.12	Sistema proposto onde a ordem de momentos de Zernike varia em números inteiros de $n = 0$ à $n = 12$ , com base de dados pré-processada CASIA, janela <i>non-overlapping</i> de 16x16, e classificador ResNet conforme descrito em 6.4.3 . . . . .	58
6.13	Comparação de resultado de métricas entre sistema de referência e sistema proposto. . . . .	60
6.14	Sistema de referência implementado com o algoritmo do sistema proposto e removendo a técnica do LBPV. . . . .	61
6.15	Comparação entre os três sistemas implementados utilizando MLPe base de dados IITD de 6000 imagens, e com ordem do momento de Zernike igual à 10. . . . .	62

# Lista de Figuras

3.1	Imagem de estrutura ocular humana. Fonte: [1]. . . . .	15
3.2	Representação visual dos Momentos de Zernike até 6 <sup>a</sup> ordem. Fonte: [2]. . . . .	21
3.3	(a) Erro geométrico ao transpor a imagem ao disco unitário. (b) Erro numérico devido a limitação na representação do disco usando <i>pixels</i> . Fonte: [3]. . . . .	23
3.4	Exemplo de classificador de rede neural <i>multilayer perceptron</i> . . . . .	26
3.5	Exemplo de classificador de rede neural convolucional. . . . .	27
3.6	Exemplo de classificador de rede neural residual. . . . .	29
4.1	Imagem do sistema referência. Fonte: [4]. . . . .	32
4.2	Imagem do sistema modificado proposto. . . . .	35
4.3	Imagem normalizada em métodos usuais de reconhecimento de iris. . . . .	36
4.4	Exemplo de imagem normalizada usando o método proposto. . . . .	37
4.5	Exemplo de janelamento sem sobreposição de janelas da imagem ( <i>non-overlapping</i> ). . . . .	38
4.6	Exemplo de momento de Zernike aplicado à diferentes tamanhos de janela da imagem. . . . .	38
5.1	Diagrama de pré-processamento realizado antes do início do treinamento do sistema proposto com 2000 imagens rotuladas na entrada. . . . .	40
5.2	Diagrama de pré-processamento realizado antes do início do treinamento do sistema proposto com 6000 imagens rotuladas na entrada. . . . .	41
5.3	Exemplos de imagens da base de dados CASIA Iris Interval. . . . .	41
5.4	Exemplos de imagens da base de dados CASIA Iris <i>Thousand</i> . . . . .	42
5.5	Exemplos de imagens da base de dados CASIA Iris <i>Syn</i> . . . . .	42

5.6	Exemplos de imagens da base de dados IIT Delhi Iris. . . . .	43
5.7	Exemplos de imagens da base de dados IIIT Delhi CLI. . . . .	44
6.1	Imagem do sistema modificado proposto. . . . .	49
6.2	Classificador MLP do sistema proposto. . . . .	51
6.3	Classificador CNN do sistema proposto. . . . .	54
6.4	Classificador ResNet do sistema proposto. . . . .	56
B.1	Representação de uma imagem digital $f(x, y)$ no domínio espacial. . .	76

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Objetivo da Dissertação . . . . .	2
1.1.1	Objetivos geral . . . . .	2
1.1.2	Objetivos Específicos da Dissertação . . . . .	2
1.2	Organização da dissertação . . . . .	3
<b>2</b>	<b>Trabalhos relacionados</b>	<b>5</b>
2.1	Introdução . . . . .	5
2.2	Trabalhos relacionados . . . . .	5
<b>3</b>	<b>Fundamentos Teóricos</b>	<b>11</b>
3.1	Fundamentos sobre imagem sintética . . . . .	11
3.2	Pré-processamento de imagens . . . . .	12
3.3	Conceitos sobre sistema biométrico . . . . .	13
3.4	Sistemas biométricos de imagem de íris . . . . .	14
3.5	Ataques de <i>spoofing</i> e os tipos de problemas em sistemas biométricos	16
3.6	Conceitos sobre momentos . . . . .	17
3.7	Momento de Zernike . . . . .	19
3.8	Conceitos sobre <i>machine learning</i> . . . . .	22
3.9	Classificador <i>multilayer perceptron</i> . . . . .	24
3.10	Classificador de rede neural convolucional . . . . .	25
3.11	Classificador de rede neural residual . . . . .	28
<b>4</b>	<b>Metodologia</b>	<b>30</b>
4.1	Introdução . . . . .	30
4.2	Sistema de referência - <i>baseline</i> . . . . .	30

4.3	Sistema proposto . . . . .	32
4.4	Sistema proposto versus baseline . . . . .	35
<b>5</b>	<b>Base de dados</b>	<b>39</b>
5.1	Introdução . . . . .	39
5.1.1	Base de dados CASIA-Iris . . . . .	40
5.1.2	Base de dados IITD-Iris . . . . .	43
<b>6</b>	<b>Experimentos e Resultados</b>	<b>45</b>
6.1	Introdução . . . . .	45
6.2	Setup . . . . .	45
6.3	Métricas de desempenho . . . . .	46
6.4	Análise de desempenho . . . . .	48
6.4.1	Momento de Zernike com MLP . . . . .	50
6.4.2	Momento de Zernike com CNN . . . . .	53
6.4.3	Momento de Zernike com ResNet . . . . .	55
6.4.4	Análise de resultados . . . . .	58
<b>7</b>	<b>Conclusão</b>	<b>63</b>
7.1	Proposta para Trabalhos Futuros . . . . .	64
	<b>Referências Bibliográficas</b>	<b>66</b>
<b>A</b>	<b>Artigos Publicados</b>	<b>74</b>
A.1	Artigos publicados em colaboração . . . . .	74
<b>B</b>	<b>Fundamentos de imagem</b>	<b>75</b>
B.1	Representações da imagem digital . . . . .	75
B.2	Componentes de imagens digitais . . . . .	75

# Capítulo 1

## Introdução

A área de pesquisa em reconhecimento de padrões visa detectar dados em classes reconhecidas durante a extração de características [5,6]. Alguns dos exemplos de aplicações de métodos da área de reconhecimento de padrões são processamento de sinais de voz, identificação de assinaturas, detecção de comportamento, identificação facial, entre outros [5]. As autenticações de biometria estão cada vez mais presentes em produtos utilizados em nossos cotidianos. Nesse mesmo contexto, existe uma crescente gama de tipos de ataques a estes sistemas, feitos por engenheiros sociais, criminosos cibernéticos e outros mal-intencionados que utilizam ataques de *spoofing* em sistemas biométricos. Neste contexto, identificar e verificar a autenticidade da identidade dos usuários ao acessar um sistema biométrico ganha relevância. Este reconhecimento pode ser através de detecção de características biométricas do rosto, das impressões digitais, da geometria da mão, da íris, da retina, da face, entre outros, e é um fator importante para proteção de dados e controle de acesso [5,7]. A área de *presentation attack detection* ou PAD fez retornar o interesse em pesquisas de reconhecimento de íris.

A biometria das pessoas está sujeita ao envelhecimento e outros fatores de desgaste do corpo, como a face e impressões digitais, mas a biometria de íris dos olhos tem uma relativa regularidade ao longo da vida quando comparada com a face, palma da mão ou mesmo impressão digital. A detecção ou reconhecimento de íris se utiliza de métodos de identificação de características extraídas de uma íris em algum banco de dados de imagens [8]. A biometria através da íris dos olhos tem robustez à ocorrência de correspondência de ambiguidades, ou seja, não deve gerar

confusão de entradas de usuários distintos no sistema como se fossem iguais [9]. A identificação de usuário para acesso pode ser aplicada por técnicas de detecção de íris, quais podem utilizar-se de métodos de classificação [10,11]. Dentre as soluções disponíveis com alta capacidade de detecção e reconhecimento de íris estão as redes neurais artificiais [12]. As redes neurais artificiais são estudadas há longo tempo, mas somente com o aumento do nível da capacidade de processamento computacional é que foi possível aplicar em sistemas mais complexos [13].

A técnica principal aplicada neste trabalho é o momento de Zernike para detecção de padrões de imagens de íris com e sem ataques *spoofing*. Relacionado aos momentos de Zernike, o polinômio de Zernike é bastante difundido na medicina, na área da oftalmologia, para descrever deformidades na íris dos olhos. Os fundamentos dos momentos de Zernike estão no estudo de óptica, pesquisados pelo físico Fritz Zernike. A matemática aplicada aos momentos de Zernike, nesta dissertação, é utilizada como descritor de formas circulares ou não alinhadas [14,15]. Para tal, manipulam-se parâmetros e variáveis do sistema proposto para detecção de imagens de íris utilizadas como ataques de *spoofing*.

## 1.1 Objetivo da Dissertação

### 1.1.1 Objetivos geral

Essa dissertação tem como objetivo avaliar a utilização de técnicas de momentos de Zernike em combinação com outras técnicas de classificação de imagens de íris, separadas por duas classes de imagens: imagens com íris reais e imagens com íris sinteticamente processadas. Em suma, será aplicado um reconhecimento de ataques de *spoofing* nessas imagens de íris.

### 1.1.2 Objetivos Específicos da Dissertação

- (1) Contribuir para o desenvolvimento de sistemas mais robustos à ataques de *spoofing* usando imagens de íris sintéticas;
- (2) Implementar o sistema de detecção de íris utilizando o descritor momento de Zernike com classificadores de formas;

- (3) Desenvolver os classificadores que interpretem as características de formas circulares e não alinhadas presentes nas extrações do método de momento de Zernike das imagens de íris;
- (4) Desenvolver os classificadores de modo que o mesmo consiga determinar se há ou não ataques de *spoofing* através das características dos momentos de Zernike da imagem de íris;
- (5) Verificar o comportamento do sistema proposto, aplicando mudança de parâmetros internos nas técnicas utilizadas;
- (6) Utilizar métricas que avaliem resultados de robustez, de invariância à rotação da imagem, de tolerância à ruídos e à redundância de informação.

## 1.2 Organização da dissertação

Esta dissertação está organizada conforme abaixo.

- No Capítulo 02, apresenta-se trabalhos que embasam a teoria utilizada como fundamentos, e que elucidam a ideia por trás do método proposto nesta dissertação.
- No Capítulo 03, inicia-se os fundamentos teóricos utilizados na dissertação, apresentando os conceitos fundamentais da área de reconhecimento de íris, e contextualizando o problema de ataques *spoofing*. Em seguida, verifica-se os conceitos de momentos de Zernike e de algumas técnicas de classificação de formas. Por fim, apresenta-se os trabalhos relacionados com o sistema proposto nesta dissertação.
- No Capítulo 04, é explanado a metodologia desta dissertação e do artigo de referência. Neste contexto, são explicados os blocos de ambos os sistemas, e discutidos os valores e parâmetros das técnicas utilizadas no método proposto comparando com o sistema de referência.
- No Capítulo 05, são apresentadas as bases de dados que foram utilizadas durante o desenvolvimento da metodologia e experimentos desta dissertação.

- No Capítulo 06, são apresentadas as métricas utilizadas, e os resultados dos experimentos, acrescentando considerações finais sobre cada método aplicado e o sistema proposto.
- No Capítulo 07, tem-se por fim as conclusões encontradas nesta dissertação, discutindo os resultados obtidos nos experimentos, e finalizando com as considerações finais sobre o sistema proposto.

# Capítulo 2

## Trabalhos relacionados

### 2.1 Introdução

Os trabalhos relacionados a esta dissertação visam explicar a ideia antes do desenvolvimento, estudo de viabilidade e de melhorias a serem aplicadas no método de momentos de Zernike combinado com outras técnicas de classificação de imagens de íris reais e sintéticas. Foram avaliadas outras dissertações e artigos com algumas similaridades de classificação, detecção ou métodos similares, aplicados a imagens reais ou com *spoofing*. Focou-se em avaliar a teoria que envolve a detecção de imagens de íris, ou outros processamentos, ou informações relevantes contidas nos trabalhos relacionados.

### 2.2 Trabalhos relacionados

Os pesquisadores Smita Khade, Shilpa Gite e Sudeep Thepade [16] demonstram uma abordagem de detecção de ataques *spoofing* em íris, e usam coeficientes extraídos das transformadas cossenos discretas [17], transformada Haar [18] e transformada híbrida [17]. Os autores utilizam seis bases de dados e nove tipos de classes de ataques spoofing para íris. Como modelos de treinamento de aprendizado de máquina, utilizam os classificadores de máquina de vetor de suporte (SVM do inglês, *support vector machine*), *naive Bayes*, *random forest* e árvore de decisão (algoritmo J48) [19], assim como a combinação desses métodos, e MLP (do inglês, *multilayer perceptron*) para detectar a autenticidade da íris [19]. Em termos de resultados, há

nove variantes de ataques *spoofing* e sete classificadores, todos aplicados a cada uma das transformadas. A transformada híbrida para imagem de íris com algoritmo RF mostrou o melhor resultado de detecção com 99,95% de precisão.

Os pesquisadores Manu Gupta e Sakshi Gupta [20] utilizam momentos de Zernike como descritores de forma e classificadores para determinar tumores cerebrais. Os descritores são extraídos de imagens segmentadas do volume tumoral. A classificação é feita utilizando-se SVM, *naive Bayes*, e *bagged trees* [19] e utilizaram validação cruzada com *k-fold* igual a 5. Um preditor decide o grau do glioma à partir das características classificadas. Como métrica, utiliza-se acurácia, sensibilidade e especificidade. Em termos de resultados, o método obteve precisão de 88% para SVM, 90% para *naive Bayes*, e 88% para *bagged trees* na classificação de gliomas de baixo e alto grau.

Em um outro trabalho, um método de verificação de imagem de íris falsa ou real, usando um descritor de imagem por estatística binarizada denominada BSIF (do inglês, *Binarized Statistical Image Features*) [21], é apresentado por Manar Ramzy Dronky e Wael Khalifa e Mohamed Roushdy. O BSIF é um descritor baseado no descritor LBP (do inglês, *local binary pattern*) [22]. Foram utilizadas 04 bases de dados de imagens, incluindo imagens de íris com lentes de contato, imagens de íris *spoofing* sintéticas, imagens *spoofing* impressas. Inicialmente, o método separa em dois grupos de imagens: segmentadas e normalizadas. Aplica-se dois filtros residuais [23] e um filtro de *Sobel* [24], antes da fase de extração de características, para melhorar a capacidade de discriminação do BSIF. Em paralelo, junto com BSIF na extração de recursos, é feita uma amostragem de 50% e calculado o histograma das imagens filtradas. Na classificação, um SVM é utilizado para determinar se a imagem é real ou falsa. Imagens não segmentadas, com BSIF e os três filtros, tiveram os melhores resultados, com 98,2% de acurácia.

O pesquisador Bineet Kaur propõe um sistema robusto de detecção de falsificação de rosto utilizando recursos com invariância local [25]. As bases de dados utilizadas foram CASIA-FASD, *Replay-Attack* e Oulu-NPU. Recursos e artefatos foram extraídos por momentos de Zernike invariantes à rotação, transformadas harmônicas polares [26], momentos ortogonais discretos como Krawtchouk, Tchebichef e Dual-Hahn [27–29]. As estruturas do sistema são todas as combinações de extratores em

série com os classificadores. Invariância à rotação, escala e translação são recursos locais que viabilizam detectar pequenas variações em uma imagem genuína ou falsa. O método proposto apresenta desempenho superior em comparação às técnicas disponíveis no levantamento da literatura, onde a transformada polar combinada com momento de Zernike obteve para as bases de dados: CASIA-FASD atingiu precisão de 99,98%, *Replay-Attack* alcançou 99,98% e Oulu-NPU atingiu 99,95%.

Os pesquisadores Gunjan G., Aditya R. e Susanta M. [30] abordam um método denominado de DSCE (do inglês, *deep supervised class encoding*) para casos de detecção de três tipos de ataques de apresentação, chamados de PAD [31], para imagens de íris: impressas, com lentes de contato e sintéticas. O DSCE é um *autoencoder* baseado em método de aprendizado supervisionado para classificação de recursos. O objetivo do DSCE é minimizar os erros de reconstrução e classificação simultaneamente durante a fase de treinamento. DSCE é empregado para projetar uma estrutura *iris-PAD* denominada *DeepI*, para perceber o acesso falso a uma íris. Resultados experimentais em diferentes bancos de dados mostram que o *DeepI* baseado em DSCE supera os atuais métodos de detecção de ataques *spoofing* com íris. O desempenho do modelo *DeepI* baseado em DSCE foi comparado com métodos existentes para detecção de ataques de apresentação em biometria de íris. A comparação foi realizada com métodos não supervisionados e supervisionados.

Os pesquisadores Smita K., Shilpa G. e Sudeep T. propõem [32] um método de detecção de ataques de falsificação de íris utilizando uma técnica de codificação denominada TSBTC (do inglês, *thepade's sorted block truncation coding*) e recursos locais de matriz de coocorrência de nível de cinza ou GLCM (do inglês, *gray-level co-occurrence matrix*). Foram utilizadas 04 bases de dados, pré-processadas com redimensionamento da imagem. O TSBTC extrai recursos de textura de cor global, e o GLCM extrai detalhes locais de textura fina. Após a classificação, utiliza-se um algoritmo preditivo para determinar se a imagem de íris é real ou falsa. As métricas utilizadas são acurácia, precisão, *recall* e *f-score*. O combinado do random forest, com algoritmo J48 e MLP e a fusão, atingiu 99,68% de precisão, e o método random forest com fusão atingiu 95,57%.

Os pesquisadores Meiling Fang, Naser Damer [33], entre outros, propuseram o método descritivo A-PBS de fusão de atenção espacial multiescala para reconhe-

cimento de íris que executa supervisão de rótulo binário com redes densas baseadas em CNN (do inglês, *convolutional neural network*) [34]. Uma convolução bidimensional  $7 \times 7$  é realizada em seguida de um bloco de *max pooling* [35]. Um bloco de rede densa é processado, onde o sistema de supervisão em termos de pixel é capaz de capturar os recursos refinados em nível de *pixel/patch*, enquanto o mecanismo de atenção orienta a rede neural a encontrar automaticamente as regiões que mais contribuem para uma melhor precisão. É utilizado um bloco reorganizador, e em seguida repetido o processo de rede densa. O diagrama em bloco finaliza com uma camada FC que é introduzida para gerar uma pontuação de previsão supervisionada por rótulo binário. A classificação supervisionada de rótulo binário reconhece recursos semânticos capturados. Experimentos extensos são realizados formados no LivDet-Iris 2017 e na base de dados IIITD-WVU com eficácia de 72%, Clarkson com 92,35% e Notre Dame com 90%.

Os pesquisadores Akshay Agarwal e Afzel Noore [36] propuseram um algoritmo para ataque *spoofing* unificado baseado em CNN, e utilizam a fusão de uma variedade de imagens de íris e pontuações de classificação calculadas sobre essas informações. O sistema utiliza um conjunto de informações para aumentar a confiabilidade no classificador; no tratamento das variações, ocorre uma avaliação extensiva do algoritmo proposto em diferentes cenários (controlados e não controlados). Nesse trabalho, foram utilizados ataques *spoofing* de íris de lente de contato 3D. Para a avaliação, utilizam-se os protocolos experimentais, que refletem configurações não controladas para ataque de apresentação de íris de lentes de contato 3D, onde as imagens pertencem tanto a condições controladas quanto adversas de imagem. As contribuições do algoritmo baseado na arquitetura CNN generalizada são desenvolvidas pela fusão precoce e tardia de informações complementares; também é feita uma comparação com algoritmos baseados em texturas; também é feita uma comparação com algoritmos baseados em texturas, CNN profundos ajustados e algoritmos de última geração existentes. Os experimentos apresentaram taxas de acerto de 94,7% para teste com 50% de imagens não controladas, e 98,2% para testes com 100% das imagens controladas.

Contudo, o método proposto nesta dissertação acrescentou um bloco de pré-processamento de redimensionamento da imagem e um bloco de normalização previ-

amente à aplicação das técnicas de detecção e classificação de ataques de *spoofing*. O método proposto permite mudanças de parâmetros e variáveis que influenciam nos momentos de Zernike quando combinados com classificadores projetados para classificar formas circulares ou não alinhadas. O trabalho desta dissertação se destaca ao empregar em combinação o momento de Zernike com MLP, momento de Zernike com CNN e momento de Zernike com Resnet. O trabalho utilizou as duas principais bases de dados: CASIA e IITD, em um sistema com imagens mais complexas, com *spoofing* menos perceptíveis.

A comparação direta entre os sistemas apresenta a abordagem deste trabalho em cenários em que o momento de Zernike é de baixa, média e alta ordem do polinômio, sendo o valor máximo utilizado para ordem de  $n$  igual à 12. O método apresenta um aproveitamento das repetições das ordens de momento de Zernike. Outra abordagem foi o uso de classificadores MLP, CNN e ResNet, mais dedicados ao tipo de entrada no processamento.

\* Sistema do artigo de referência

Tabela 2.1: Comparação de métodos com algoritmos de ataques de *spoofing*

Autor	Pré-processamento	Deteção	Tipo de classificador	Base de dados	Métricas
Colares, G.S	- Redimensionamento; - Normal. em escala; e translação; - Janelamento sem sobreposição;	Momentos de Zernike	- MLP - CNN - ResNet	- IITD-Sys - IITD Iris CSD - CASIA Iris Syn - Casia Iris Thousand	- Acurácia - Precisão - Recall - F1-Score
Kohli, N., Yadav, D., Vatsa, M., et al.	- Redimensionamento; - Janelamento sem sobreposição;	- Momentos de Zernike - vLBP	MLP	- IIT-Delhi CLI - IITD IIS - SDB - IIT Delhi Iris - MID - CSD	Acurácia (Desvio Padrão)
Khade, S., Gite, S., Thepade, S., et al.	- Separa por tipo de Spoofing; - Converte em Imagem grey; - Redimensionamento;	- DCT - Filtro Haar - Transformada Híbrida	- NB - J48 - SVM - RF - Métodos combinados	- LivDet-Iris 2015 - IITD Iris CSD - CASIA Iris Syn - IITD contact len iris - Warsaw-Post - Mortem-Iris - MBGC	- Acurácia - APCER - NPCAR - ACER - FPR - TPR
Gupta, M., Gupta, S., et al.	- Ponderação sobre a imagem; - Estimação dos campos de bias; - Correção dos campos de bias;	- Momento de Zernike; com análise estatística usando t-test.	- SVM - Naive Bayes - Bagged trees	- MICCAI 2012 Challenge database	- Acurácia - Sensitividade - Especificidade
Dronky, M. R., Khalifa, W., Roushdy, M.	- Segmentação regional de iris; - Normalização	- 50% amostragem reduzida - BSIF 8 bits - Histograma 8 bits	- SVM	- CASIA-iris-Fake - CASIA-iris-Syn	- Acurácia - Taxa de erro
Kaur, B.	-	- LBP - SIFT - PHT - Momentos de Zernike - Momento Krawtchouk - Momento Dual-Hahn - Momento Tchebichef	- SVM com top-k recursos + Friedman test	- CASIA FASD - OULU-NPU - Replay-Attack	- Acurácia - BPCER - APCER - ACER
Gautam, G., Raj, A., Mukhopadhyay, S.	Adiciona ruído à imagem	Encoder CNN (FCN-DAE) + Rede neural espaço latente	- Decoder + Convolução 2D + Imagem reconstruída - Deep CNN	- IIT-D CLI - IITD Iris CSD - Multi-Sensor - Synthetic - IITD Iris Spoofing - IIT Delhi Iris	- Acurácia - Sensitividade - Especificidade - PPV - NPV - informação - F1-Score
Khade, S., Gite, S., Thepade, S. D., et al.,	- Redimensionamento;	Fusão dos extratores - Global feature extractor TSBTC N-ary - Local feature extractor GLCM	- SVM - J48 - Random Forest - Random Tree - Naive Bayes - Método Ensemble + algoritmo preditivo	- Clarkson LivDet2013 - Clarkson LivDet2015 - IITD CSD - IITD Contact Lens	- Acurácia - F-measures - APCER - NPCER - ACER
Fang, M., Damer, N., Boutros, F., et al.,	-	- DenseNet - PBS - APBS	-	- NDCLD13 - NDCLD15 - IIT-D CLI - LivDet-Iris 2017	- APCER - BPCER - HTER - CCR - TDR - FDR
Agarwal, A., Noore, A., Vatsa, M., et al.,	Soma imagem original com imagem CLAHE pré-processada	Fusão de métodos: IPAD CNN de imagem original + IPAD CNN de imagem de pré-processamento	-	- UnMIPA - MUIPA	- BPCER - APCER - HTER

# Capítulo 3

## Fundamentos Teóricos

### 3.1 Fundamentos sobre imagem sintética

Uma imagem digital é feita a partir de um modelo numérico, que é armazenado digitalmente como uma descrição objetiva. Supondo esta imagem gerada a partir deste modelo, o sistema visual humano faz uma análise subjetiva [37]. Imagens sintéticas têm detalhes de textura, de sombras, entre outras distinções que não exatamente seriam perceptíveis à visão humana, mas são imagens que se aproximam do real o suficiente para fornecer um certo grau de realismo. Normalmente, as imagens digitais sintéticas são geradas a partir de um conjunto de imagens digitais reais, com o objetivo de obter uma representação de imagens sintéticas. O desenvolvimento de algoritmos cada vez mais especializados está diminuindo as diferenças entre imagens reais e sintéticas. Esse desenvolvimento é acompanhado por um custo de processamento no treinamento do método, de falhas para casos de modelagem generalista [38].

O sistema computacional pode ser comparado a um artista que pinta um quadro a partir de sua descrição. Mas o artista pode realizar uma correção se sua pintura não lhe parecer correta em seu sistema visual. O computador que gera uma imagem não recebe esse *feedback* sobre a qualidade da imagem. Para isso, é essencial definir características que tragam um grau de realismo à imagem sintética [38].

Como exemplo, são listadas algumas dessas características a seguir: curvas suaves e superfícies, efeitos de técnicas de amostragem digital, distorções, erros e ruído na aquisição de dados, a aparência visual e a interpretabilidade da imagem,

relação do *pixels* com suas vizinhanças de *pixels*, respostas temporais, espectrais e espaciais, características de cor e espacial na imagem, secções de áreas da imagem, erros de medições quantitativas, erros de redução ou ampliação de imagem, ajustes de contraste linear e não linear, erro de racionamento de banda, erro de filtragem espacial; erros ao restaurar uma imagem, problema de pequenos detalhes suprimidos na imagem, problemas na melhoria de imagem degradada [37, 38].

Um modo de agrupar as características das imagens sintéticas é dividi-las em dois grupos: probabilístico e por classificação [37, 38]. Os métodos probabilísticos têm o objetivo de produzir um modelo para representar alguma característica de imagens sintéticas e reais. Para isso, extraem informações, como p.ex. valor central, média, mediana, desvio padrão, distribuição, curtose, momentos, entre outros [37]. Em métodos por classificação, um classificador é treinado para separar características como marcadores. Esses marcadores são métricas ou valores quantificáveis para descrever as imagens, p.ex., estão relacionados à cor, textura, formas, manchas de cores e cantos contidos nas imagens [37], entre outras. Atualmente, algoritmos se utilizam desses métodos para gerar imagens sintéticas cada vez mais realistas; em contraponto, é ainda mais desafiador projetar sistemas que identifiquem se uma imagem é sintética ou não. Por fim, espera-se que esses sistemas reconheçam padrões nas imagens, consigam observar propriedades na imagem, encontrem pontos de interesse na imagem [38], entre outras informações para detectar se uma imagem é real ou sintética.

## 3.2 Pré-processamento de imagens

Pré-processamento de imagens refere-se à manipulação inicial dos dados para uma padronização ou calibração da imagem. Então, antes de aplicar o método, alguns sistemas optam pela padronização das imagens da base de dados, quando isso permite extrair informações relevantes ao sistema. Em complemento, é necessário adequar as imagens na extração de informações ou de características, aplicando-se algumas técnicas de pré-processamento [39]. Dentre essas técnicas de pré-processamento, o redimensionamento de imagem é comumente aplicado e consiste em modificar as dimensões, a resolução ou o tamanho do arquivo da imagem [39]. Perceba que o

redimensionamento de uma imagem resulta na perda de detalhes da imagem ou pixelização [40].

Outro pré-processamento realizado em imagens é a normalização. Existem diversos tipos de técnicas de normalizações. A mais comum é a normalização em escala, que é um método utilizado para reorganizar os dados [38]. Uma técnica de normalização chamada de Min-Max é uma técnica de escalonamento em que os recursos serão redimensionados para que os dados sejam adequados em um intervalo de valores, p.ex. o intervalo de  $[0, 1]$ . Outra normalização é em translação, qual a coordenada da imagem original  $(x, y, z)$  é transladada para  $(x', y', z')$  em relação às coordenadas de si mesma. A translação da imagem pode ser obtida realizando deslizamento da imagem, cálculo de função densidade, cálculo de momentos, entre outros [38].

Por último, temos o janelamento (do inglês, *Windowing*), que são submatrizes de dimensões menores às dimensões da imagem, e utiliza somente uma área de mesmas dimensões da janela sobre a imagem. Por fim, estas áreas de interseção entre a submatriz e a imagem serão processadas com algum método pré-determinado.

A imagem será dividida em vários blocos com mesmas dimensões da janela. Quando uma janela tem *pixels* que também contém em sua janela vizinha, diz-se que é um janelamento por sobreposição (do inglês, *overlapping window*). Caso não contenham *pixels* semelhantes entre janelas vizinhas, o tipo de janelamento é sem sobreposição (do inglês, *non-overlapping Window*). Esse conjunto de janelas ou blocos de *pixels* irá receber o método a ser aplicado.

### 3.3 Conceitos sobre sistema biométrico

Existem diversos tipos de sistemas biométricos, alguns utilizam a geometria das mãos ou as estruturas oculares ou até mesmo a análise de grafia e escrita, que verifica a autenticidade de um documento assinado pela pressão e modo de escrever [41–43]. Métodos de medidas e cálculos de características fisiológicas estão relacionados a sistemas biométricos. A autenticação biométrica são formas de identificação e controle de acesso utilizadas na área da computação, como p.ex., ao ser utilizada para identificar indivíduos em um sistema de vigilância. Para os seguin-

tes identificadores biométricos específicos: face, impressão digital, íris, retina, mão, orelha, assinatura e dinâmica de teclas, pode-se reafirmar os recursos típicos usados para identificar, listar o hardware de aquisição, explicar técnicas de extração de recursos e avaliar seus pontos fortes e fracos [42, 43]. Conceitos de sistema biométrico têm diferentes técnicas de identificação e autenticação de indivíduos de maneira automatizada, com diferentes níveis de confiabilidade e desempenho. Para um ou mais identificadores biométricos, pode-se construir implementações para treinar um modelo de um sistema biométrico [43–46].

### 3.4 Sistemas biométricos de imagem de íris

Primeiramente, nesta seção, apresentam-se os conceitos de estruturas e funções dos olhos. Tudo inicia com a luz que é introduzida no olho através da córnea, uma camada transparente e curva na frente da íris e da pupila. A córnea serve como uma cobertura protetora da parte frontal do olho e ajuda a focar a luz na retina. Na parte central do olho, após a córnea, a luz passa pela pupila [47]. A íris é a área circular colorida nos olhos, circunda a pupila, e sua função é controlar a quantidade de luz que entra no olho. O tamanho da pupila é controlado pela ação do músculo esfíncter pupilar, que contrai a pupila para a entrada de menos luz, e pelo músculo dilatador, que permite a entrada de mais luz em ambientes escuros. Atrás da íris fica a lente, que focaliza a luz na retina através do controle de pequenos músculos ciliares sobre o cristalino. A retina contém as células fotorreceptoras, que, dentre estas, a mais sensível é uma pequena área chamada mácula, com milhões de fotorreceptores chamados de cones. A imagem visual torna-se rica em detalhes com alta densidade de cones na mácula. Por fim, na seção frontal, é preenchido com um fluido chamado humor aquoso, enquanto que na seção posterior contém um fluido gelatinoso chamado humor vítreo [38, 45, 46]. Uma imagem da estrutura ocular encontra-se na Figura 3.1.

Em sistemas biométricos de íris, para usar essas informações da íris como forma de identificação, é necessário realizar um processo de quatro etapas [48]:

- A etapa de captura consiste em uma extração da imagem do olho esquerdo e do direito, com uma determinada qualidade podendo ou não utilizar luz,

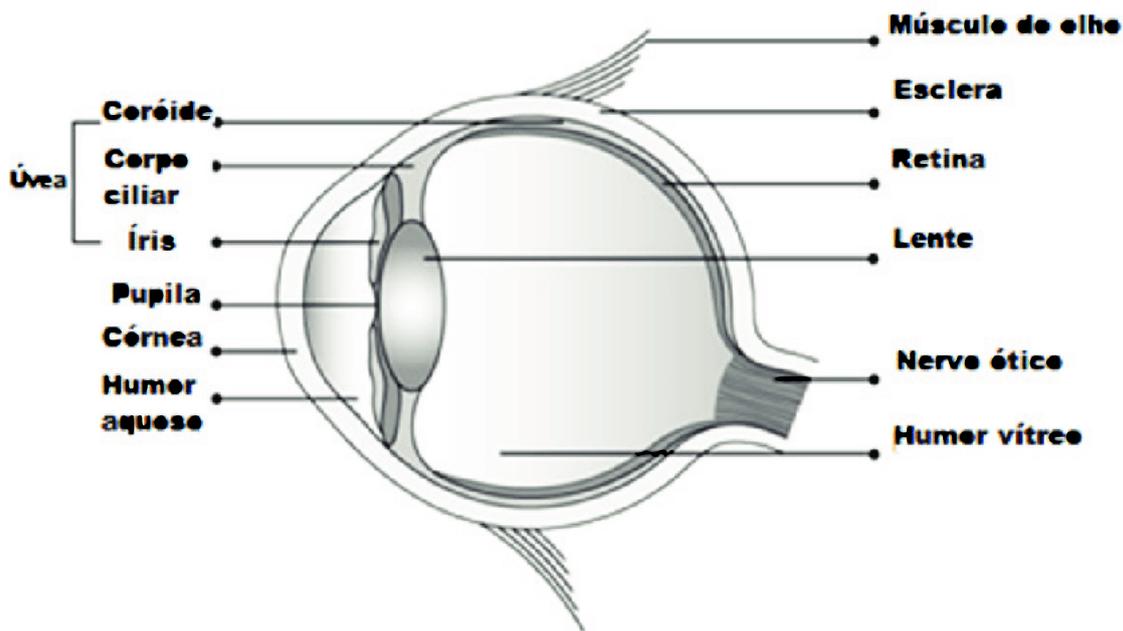


Figura 3.1: Imagem de estrutura ocular humana. Fonte: [1].

dependendo da técnica de aquisição.

- Na etapa de verificação e controle, o método extrai as informações necessárias a partir da imagem. Para tal, o método deve garantir um nível de qualidade das informações tal que sejam úteis no objetivo do sistema biométrico.
- Na etapa de processamento, usando algoritmos de realce, compressão, entre outros pré-processamentos que ajudem a remover distorções de imagem e outros artefatos.
- Por fim, a etapa de criação de modelo, as informações presentes na imagem são então traduzidas em um modelo biométrico. De acordo com o modelo, e a vulnerabilidade à falsificação biometria irá variar [41, 42].

Apesar de ter muito em comum quando comparado a outros métodos de verificação biométrica, a biometria de íris apresenta algumas vantagens como maior dificuldade de falsificação [5]. As informações da íris são exclusivas para cada usuário e requerem tecnologia específica na etapa de aquisição. Outra vantagem são as dimensões e condições da íris que variam pouco à medida que envelhecemos [42]. Os avanços nas câmeras tornaram possível escanear uma íris humana e, dependendo do equipamento, até a alguns metros de distância e ainda com resultados confiáveis.

A tecnologia ainda está sujeita a alguns dos problemas semelhantes em outras autenticações biométricas, sendo vítima de falsificação de identidade. A realidade de qualquer método de autenticação deve tratar tanto com a identidade real do usuário que está tentando acessar, quanto com a identidade falsa tentando acessar o sistema. O foco está no fato de que a segurança de autenticação eficaz e robusta deve ter métodos de identidade para garantir que a pessoa que acessa um sistema é autenticamente um usuário do mesmo [41].

### 3.5 Ataques de *spoofing* e os tipos de problemas em sistemas biométricos

Ataques de *spoofing* são quando sistemas ou indivíduos tentam falsificar ou se passar por um usuário na autenticação, com objetivo de invadir, acessar ou inserir dados restritos em um sistema alvo [5, 14, 49]. Atualmente, a presença de dispositivos portáteis inteligentes com as aplicações biométricas é mais frequente em nosso cotidiano, assim como a ocorrência de ataques de *spoofing*. No ataque de *spoofing*, o sistema invasor tenta falsear as restrições de acesso apresentando informações falsas, com o intuito de enganar o sistema alvo da invasão [14, 49, 50]. Para tal, apresenta características falsas, que o sistema alvo da invasão irá comparar com as características obtidas em seu banco de dados para autenticação.

O tipo de técnica usada no ataque depende muito do tipo de verificação do sistema a ser atacado. Como exemplo, consideremos um sistema biométrico que permite acesso com autenticação facial, e que em sua base de dados tem informações da face do usuário [49–51]. Então, o sistema invasor irá tentar entrar com uma informação falsa do usuário, no caso mais simples uma imagem impressa da face do usuário, ou em caso mais complexo, a invasão é efetuada com uma imagem fabricada contendo características comuns aos usuários do sistema o suficiente para enganar o processo de autenticação. Outro exemplo seria para a voz, onde o sistema alvo da invasão autentica através de sinal contendo as características vocais do usuário. Surpreendentemente, o sinal nem precisa ser compreensível por um humano, desde que exiba os sinais vocais determinísticos com as características da identidade atacada [49, 52].

Baseando-se na norma ISO/IEC 30107-1:2023 [53, 54] tecnologia da informação, os tipos de ataques de *spoofing* de íris são classificados em: *fake printed iris attack*, de modo simples são tentativas de invasão ao sistema alvo via autenticação de imagem impressa da íris; outro tipo de ataque de *spoofing* seria usando *textured contact lens*, ou seja, lentes de contato texturizadas que, por meios de oclusões, modificam as características da íris ao ponto de se tornar semelhante o suficiente para autenticação de um sistema; e por fim o ataque de *spoofing* através de *synthetic iris images*, ou seja, o uso de imagens de íris sintéticas geradas via métodos computacionais, que modificam ou criam imagens de íris similares às reais ao ponto de se tornar o mais verossímil à imagem de íris real em uma autenticação.

A globalização das informações atua a favor de sistemas de ataque de *spoofing*, tornando o acesso via dados biométricos possível de ser fraudável. Fotos dos usuários ou mesmo vídeos podem estar disponíveis em vários sites na internet, a voz dos usuários pode ser facilmente gravada e examinada à distância, moldes de rostos podem ser criados utilizando inteligência artificial. Com isso, na atualidade os sistemas *anti-spoofing* são extremamente necessários nos sistemas em dispositivos com *software* [4, 50].

### 3.6 Conceitos sobre momentos

Antes da definição de momentos, abordaremos alguns conceitos. Primeiramente, uma variável aleatória discreta  $\mathcal{X}$  é definida por descrever um número contável de valores distintos de um conjunto em eventos probabilísticos em estudo [55]. A probabilidade de um evento é definida por  $P(\mathcal{X})$ . Sendo  $\mathcal{X}$  uma variável aleatória contida em  $1 \leq x \leq N$ , a função que define o conjunto de probabilidades é  $p(x_i) = P(\mathcal{X} = x_i)$  onde  $\sum_{i=0}^N p(x_i) = 1$  [55]. A função distribuição de probabilidade  $F(\mathcal{X})$  é nada mais que esse conjunto de informações de probabilidades  $P(\mathcal{X})$  em um intervalo  $x$ , ou seja,  $F(x) = P(\mathcal{X} \leq x)$  [55]. A esperança  $E(\mathcal{X})$  refere-se à média dos valores de  $\mathcal{X}$ , denotada por  $\mu$ , e é também o primeiro momento de uma distribuição de probabilidade [7, 56]. De forma resumida, a esperança  $E(\mathcal{X})$  pode ser entendida como uma média ponderada dos possíveis valores das variáveis aleatórias, e para melhor entender, caso a probabilidade desses valores seja igual,

tem-se a média aritmética [9, 56, 57]. O centroide ou centro de massa é determinado por coordenadas, que aplicando em uma imagem bidimensional serão os eixos  $x$  e  $y$ , e é calculado pela razão entre o momento de primeira ordem ( $\mu_{1,0}$  ou  $\mu_{0,1}$ ) dividido pelo momento de ordem zero  $\mu_{0,0}$  de cada eixo de coordenadas conforme a equação abaixo:

$$X_c = \frac{\mu_{1,0}}{\mu_{0,0}}, \quad Y_c = \frac{\mu_{0,1}}{\mu_{0,0}} \quad (3.1)$$

A variância  $Var(\mathcal{X})$  é o segundo momento de uma função de distribuição de probabilidade  $F(\mathcal{X})$  e fornece uma medida de variabilidade em torno da esperança  $Var(\mathcal{X}) = E[\mathcal{X} - E(\mathcal{X})]^2$  [55]. Resumidamente, seria o quanto as variáveis se desviam em relação à média, ou em casos particulares é o desvio padrão  $\sigma$ . Cada distribuição pode ser associada a parâmetros, como a covariância  $Cov(\mathcal{X}, \mathcal{Y}) = E[(\mathcal{X} - \mu_x)(\mathcal{Y} - \mu_y)]$  que permite comparar duas variáveis aleatórias (neste caso  $\mathcal{X}$  e  $\mathcal{Y}$ ) e relacioná-las entre si, gerando novas informações sobre uma distribuição [56, 58]. O centroide  $(X_c, Y_c)$ , média  $\mu$ , variância  $Var(\mathcal{X})$  e covariância  $Cov(\mathcal{X}, \mathcal{Y})$  compõem as medidas mais comumente usadas nas funções densidade de probabilidade [56, 57]. O terceiro e quarto momentos de uma distribuição são normalmente utilizados no estudo das formas de uma probabilidade. Em momentos de terceira e quarta ordem pode-se obter mais informações das funções densidade de probabilidade. Em casos particulares, os momentos de terceira ordem são a assimetria, e nos momentos de quarta ordem é a curtose, ou seja, elevação ou achatamento da função densidade de probabilidade [56, 58].

Com tais definições, denota-se momentos como o conjunto estatístico de características de uma distribuição de probabilidade. Essas características são certas medidas quantitativas, como exemplo as apresentadas anteriormente, como centroide, média, variância, assimetria, curtose, entre outras. A equação fundamental para momentos discretos é expressa conforme a seguir:

$$M_{p,q}(x, y) = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} f(x, y)x^p y^q \quad (3.2)$$

onde  $p, q = \{0, 1, 2, \dots, \infty\}$  e  $M_{p,q}(x, y)$  é o  $(p + q)$ -ésima ordem do momento da função nas coordenadas  $x$  e  $y$ , e  $f(x, y)$  é a função base ou *kernel* de ponderação. Momentos, quando aplicados ao processamento de imagens, descrevem o conteúdo

da imagem com relação aos eixos da mesma, capturando informações geométricas globais, entre outras características, como as abordadas sobre momento e propriedades analogamente encontradas em momentos na física.

### 3.7 Momento de Zernike

O momento de Zernike é um tipo de momento ortogonal do domínio complexo, pois seu kernel é um conjunto de polinômios de Zernike ortogonais definidos no interior de um disco unitário, usualmente representado em espaço de coordenadas cartesianas ou polares [9, 59]. Os momentos de Zernike bidimensional, definidos pela variável  $ZM$ , de ordem  $n$  e com  $m$  repetições das ordens, e  $N$  e  $M$  são os números de pixels em cada eixo  $ZM$  é expresso por:

$$ZM_{n,m} = \frac{(n+1)}{\pi} \sum_{x=1}^N \sum_{y=1}^M f(x,y) V_{n,m}^*(\rho, \theta) \quad (3.3)$$

A ordem  $n$  tem valores positivos inteiros ou zero, enquanto  $m$  assume valores inteiros positivos e negativos definidos por  $n - |m|$ : par, e  $|m| \leq n$ .  $ZM_{n,m}$  é limitado aos valores dentro de  $x^2 + y^2 \leq 1$ , sendo a magnitude de momentos complexos ortogonais à imagem e capturando informações globais sobre a imagem. O polinômio de Zernike  $V_{n,m}$  é definido como:

$$V_{n,m}(\rho, \theta) = R_{n,m} \exp(jm\theta) \quad (3.4)$$

sendo  $n - |m|$ : par, e  $|m| \leq n$ . A variável  $\rho$  é o tamanho do vetor de origem para  $(x, y)$  no sentido anti-horário e  $\theta$  é o ângulo entre o vetor  $\rho$  e o eixo  $x$  no sentido anti-horário. O  $R_{n,m}$  é a parte real do polinômio de Zernike  $V_{n,m}$ , definido a seguir:

$$R_{n,m}(p) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s!(a-s)!(b-s)!} p^{(n-2s)} \quad (3.5)$$

onde  $a = \frac{(n+|m|)}{2}$  e  $b = \frac{(n-|m|)}{2}$ .

Os momentos de Zernike têm base ortogonal. Uma representação usual, por imagem, é exposta na Figura 3.2. Os momentos de Zernike são invariantes apenas à rotação. Para obter invariância à escala e translação, as imagens devem ser nor-

malizadas usando seus momentos regulares. Os *features* invariantes à rotação são extraídos de imagem normalizada.

Momentos Zernike é um método muito útil no reconhecimento de padrões e análise de imagens, onde são bem utilizados por suas propriedades de ortogonalidade e de invariância de rotação [59, 60]. No entanto, verificando pelas equações anteriores, o cálculo direto destes momentos tem um alto custo computacional, limitando a sua utilização especialmente em ordens maiores. Um dos métodos empregados para redução dos esforços de custo computacional está na quantização das coordenadas utilizadas [59, 60]. A literatura elucida que momentos de Zernike oferecem desempenho superior em comparação com outros métodos baseados em momentos. Momentos Zernike são definidos como descritores de imagens muito eficazes, usados em muitas aplicações de processamento digital de imagens [7]. Por outro lado, esse processo de digitalização afeta a precisão dos momentos; conseqüentemente, diversas de suas propriedades são afetadas. O erro geométrico e o erro de integração numérica são dois erros principais dessa discretização [7].

Devido à sua propriedade de ortogonalidade, os momentos de Zernike caracterizam uma imagem de forma individual por possuírem redundância mínima de informações. A separação distinta da contribuição individual de cada momento de cada ordem é uma das vantagens da propriedade da ortogonalidade. Essa propriedade também adiciona robustez ao ruído para o sistema que a utiliza [9]. Considerando as informações de magnitudes dos momentos de Zernike, estas apresentam invariância à rotação da imagem. A mesma informação pode se tornar invariante à escala e translação somente aplicando algumas simples transformações na imagem [9, 59]. Para garantir invariância à translação, a imagem original é transformada novamente, cujos primeiros momentos sejam iguais a zero ( $m_{01}$  e  $m_{10}$ ),  $\alpha$  e  $\beta$  valores previamente determinados. Resumidamente, uma imagem pode ser normalizada à escala e translação transformando conforme a equação abaixo:

$$g(x, y) = f\left(\frac{x}{\alpha} + \bar{x}, \frac{y}{\alpha} + \bar{y}\right) \quad (3.6)$$

onde  $\alpha = \sqrt{\frac{\beta}{m_{00}}}$ .

Momentos de Zernike trazem consigo o erro de integração numérica e o erro geométrico. O erro de integração numérica é gerado quando a função imagem é

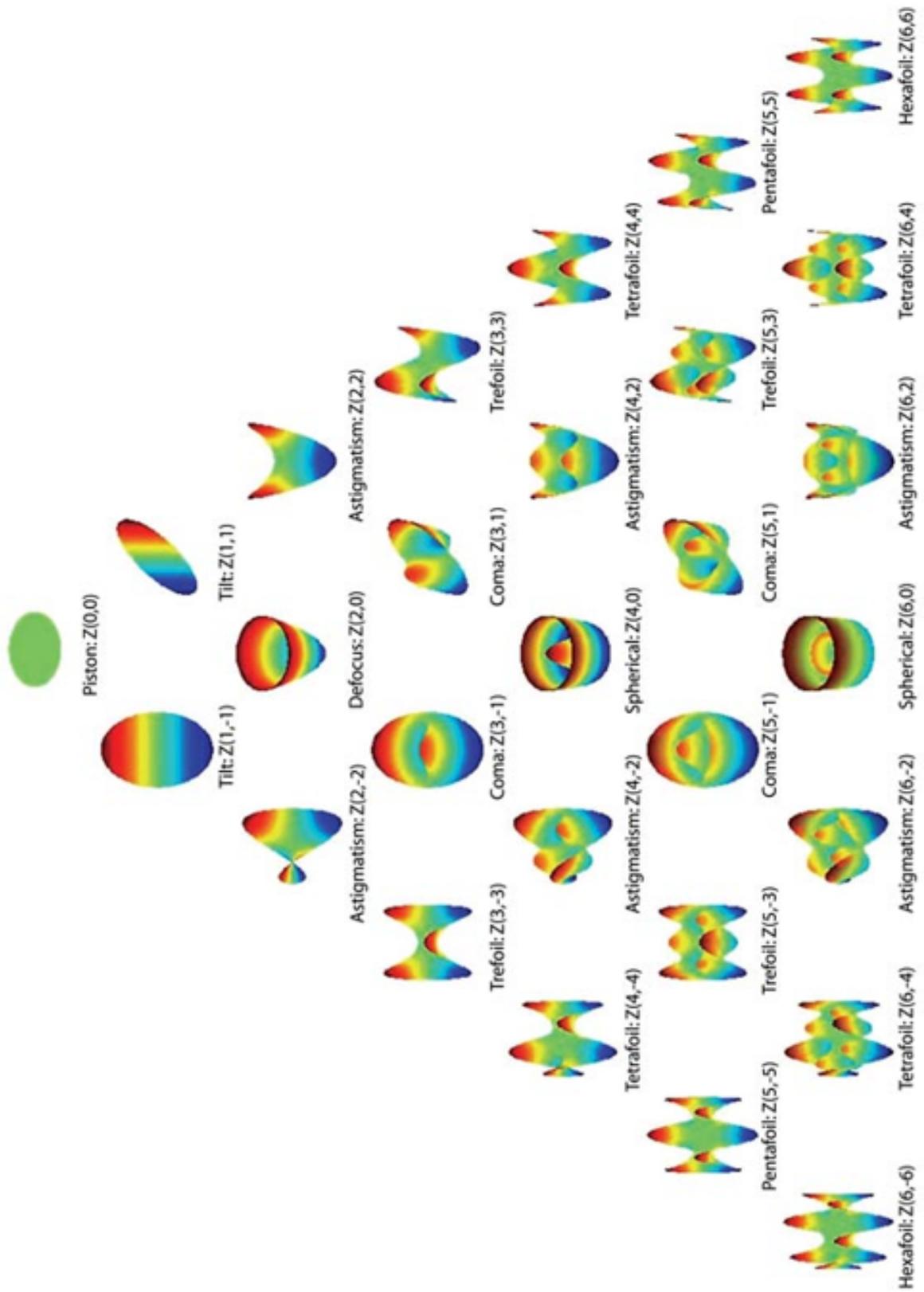


Figura 3.2: Representação visual dos Momentos de Zernike até 6ª ordem. Fonte: [2].

Tabela 3.1: Tabela de termos de momentos de Zernike por ordem e número de repetições

Ordem (n)	Graus com Repetições (m)	Momentos de Zernike
0	0	$ZM_{0,0}$
1	1	$ZM_{1,1}$
2	0   2	$ZM_{2,0}ZM_{2,2}$
3	1   3	$ZM_{1,3}ZM_{3,3}$
4	0   2   4	$ZM_{4,0}ZM_{4,2}ZM_{4,4}$
5	1   3   5	$ZM_{5,1}ZM_{5,3}ZM_{5,5}$
6	0   2   4   6	$ZM_{6,0}ZM_{6,2}ZM_{6,4}, ZM_{6,6}$
7	1   3   5   7	$ZM_{7,1}ZM_{7,3}ZM_{7,5}ZM_{7,7}$
8	0   2   4   6   8	$ZM_{8,0}ZM_{8,2}ZM_{8,4}ZM_{8,6}ZM_{8,8}$
9	1   3   5   7   9	$ZM_{9,1}ZM_{9,3}ZM_{9,5}ZM_{9,7}ZM_{9,9}$
10	0   2   4   6   8   10	$ZM_{10,0}ZM_{10,2}ZM_{10,4}ZM_{10,6}ZM_{10,8}ZM_{10,10}$
11	1   3   5   7   9   11	$ZM_{11,1}ZM_{11,3}ZM_{11,5}ZM_{11,7}ZM_{11,9}ZM_{11,11}$
12	0   2   4   6   8   10   12	$ZM_{12,0}ZM_{12,2}ZM_{12,4}ZM_{12,6}ZM_{12,8}ZM_{12,10}ZM_{12,12}$

discreta e, para simplificar o cálculo do momento, as funções do polinômio de Zernike são ajustadas à posição central dos *pixels*, e é realizada uma aproximação de ordem zero para a integração dupla envolvida no cálculo dos momentos de Zernike [7, 61]. No erro geométrico, o problema é gerado ao transpor o domínio do quadrado que define o domínio da imagem, convertendo-o em um disco unitário, que, deste modo, não há um mapeamento perfeito do domínio quadrado para um domínio circular. Com tal, irão participar do cálculo do momento de Zernike somente aqueles *pixels* cujos centros estão no círculo ou que estejam dentro dele [60]. No cálculo do momento, os *pixels* cujos centros ficam fora do disco unitário, mas que ainda fazem parte do disco unitário, são deixados de fora dos cálculos e geram o erro geométrico [7, 59]. Na Figura 3.3 exemplifica os casos de erros gerados pelos momentos de Zernike.

### 3.8 Conceitos sobre *machine learning*

*Machine learning* é um subcampo da inteligência artificial que tem como objetivo criar modelos que exibam comportamentos mimetizados da capacidade de aprender da inteligência humana. No processo de aprendizado de máquina (do inglês, *machine learning*), o sistema computacional recebe um conjunto de dados e uma tarefa para executar, e com base nesses dados, o sistema computacional encontrará uma forma de realizar a tarefa [62]. *Machine learning* é dividida em dois

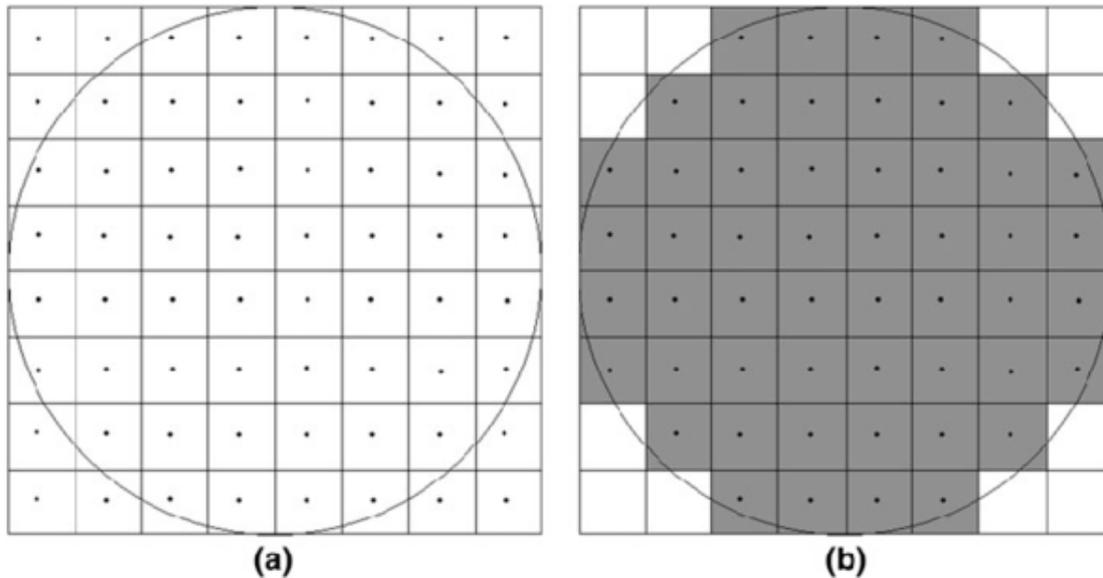


Figura 3.3: (a) Erro geométrico ao transpor a imagem ao disco unitário. (b) Erro numérico devido a limitação na representação do disco usando *pixels*. Fonte: [3].

grupos, o primeiro é o aprendizado de máquina supervisionado, onde um conjunto de dados recebe rótulos utilizados no treinamento do modelo. Esse rótulo ou *label* é a resposta para a tarefa final na saída do sistema ou a resposta correta fornecida com o conjunto de dados. Um exemplo de aprendizado supervisionado seria um sistema previamente treinado para identificar imagens com faces humanas, onde ter ou não as faces seriam os rótulos. O outro grupo de *machine learning* é o aprendizado não supervisionado, onde o algoritmo busca por padrões nos dados sem rotulação, na tentativa de aprender com os erros [12, 62]. Um exemplo desse aprendizado seria um sistema de recomendações de compras online, onde o sistema aprende com o tempo as preferências e melhores recomendações de compra aos clientes. Um meio-termo desses grupos de *machine learning* seria o aprendizado de máquina por reforço, que mescla o treinamento da máquina a tomar decisões corretas a partir de tentativas e erros, usando paradigmas de recompensa e punição ao processar dados. Um algoritmo comumente usado em *machine learning* é o de redes neurais artificiais, que seguem a ideia de rede neural humana, são sistemas de nós interconectados ou células de processamento ou neurônios, ou seja, são estruturas matemáticas que desempenham funções e estão organizados por camadas. A rede neural artificial tem camadas de entrada, ocultas e de saída [12, 62].

Sistemas de *machine learning* com aprendizado supervisionado podem ser divididos em duas categorias: de regressão e de classificação. A *machine learning*

com aprendizado supervisionado de classificação tem como meta aprender uma regra geral que consiga mapear corretamente a entrada para as saídas. A entrada é dividida em grupos ou classes, e as saídas são categorizadas. O *machine learning* com aprendizado supervisionado de regressão se difere por ter resultados numéricos contínuos ou discretos na saída resultado de saída do modelo. A regressão também se utiliza de dados prévios, e avalia o erro entre a saída do modelo e a saída estimada, que o modelo tenta chegar o mais próximo possível do valor estimado, e essa diferença de valores é uma medida de erro do algoritmo.

Outro modo de agrupar *machine learning* é dividindo em dois grupos: *shallow learning* e *deep learning*. A *shallow learning* é um tipo de *machine learning* com aprendizado tradicional dos sistemas computacionais. Essa é uma categoria de algoritmos de aprendizado de máquina que normalmente envolve uma única camada de transformação. Algoritmos *shallow learning* visam encontrar padrões e relações entre os dados de entrada para fazer previsões ou decisões de uma tarefa [12]. Alguns exemplos são: regressão linear, SVM (do inglês, *support vector machines*), *random forest*, *Naive Bayes*, dos k-NN (do inglês, *k-nearest neighbors*) [12, 13]. O *deep learning* é composto por múltiplas camadas de nós interconectados ou neurônios, que trabalham juntos para processar e aprender com os dados de entrada. *Deep learning* é o ramo do *machine learning* baseado na arquitetura de redes neurais artificiais como CNN (do inglês, *convolutional neural network*), ResNet (do inglês, *residual neural network*), GAN (do inglês, *generative adversarial network*) entre outros [19, 34, 63].

### 3.9 Classificador *multilayer perceptron*

Por muito tempo, a teoria de redes neurais para *perceptron* afirmava que o mesmo poderia representar qualquer circuito lógico, mas a literatura tinha ressalvas pois o mesmo não poderia representar a porta XOR (do inglês, *exclusive OR*). O perceptron de múltiplas camadas ou MLP (do inglês, *multilayer perceptron*) foi desenvolvido para resolver essa limitação [19]. O mesmo possui camadas de entrada e saída, e uma ou mais camadas ocultas com muitos neurônios empilhados juntos. No *perceptron*, o neurônio deve ter uma função de ativação que impõe um limite

como os da função ReLU ou função sigmoide, enquanto no *multilayer perceptron*, os neurônios podem usar qualquer função de ativação arbitrária [19,64]. A rede neural MLP varia suas conexões entre os neurônios de sua arquitetura, podendo ser em alguns casos totalmente conectados entre si, e quando essas conexões se estendem densamente da entrada até a saída da rede, essa rede neural é denominada *fully connected*.

O *multilayer Perceptron* pertence à categoria de algoritmos *feed forward*. Nestes, as entradas são combinadas com os pesos iniciais em uma soma ponderada e submetidas à função de ativação, assim como no Perceptron, mas a diferença é que cada combinação linear é propagada para a próxima camada [19,64]. Assim é feito o aprendizado na MLP, e o algoritmo é chamado de *backpropagation*, podendo apresentar algumas variações na regra da cadeia (*chain rule*), ou seja, parâmetros são ajustados durante o processamento nessa diferenciação do gradiente. Resumidamente, *backpropagation* atribui valores iniciais aleatórios para pesos (chamados de *bias*) aos neurônios, e na camada oculta realiza cálculos dos valores dos neurônios, considerando a correlação entre os mesmos, atualizando os pesos dos neurônios nas camadas ocultas e na saída. Por fim, com base no erro calculado, repete o processo até satisfazer a condição de erro estimado [19,64].

### 3.10 Classificador de rede neural convolucional

No ano de 1998, os pesquisadores Yann Lecun e equipe publicaram o artigo *Gradient-based learning applied to document recognition* [65], onde utilizam *backpropagation* em uma rede *feed forward* e várias camadas ocultas, construindo uma rede que consegue tratar de várias características ao mesmo tempo, posteriormente chamada de redes neurais convolucionais ou CNN (*convolutional neural network*). Essas redes funcionam processando grandes quantidades de dados ao mesmo tempo, em formato matricial numérico de várias dimensões, e processam através da convolução com uma matriz filtro chamada *kernel* de dimensões  $x \times y$ , resultando em  $M_i$  mapa de características na entrada, e  $M_o$  na saída. As CNNs normalmente consistem em três tipos de camadas: a camada convolucional, a camada de *pooling* e a camada totalmente conectada [19]. Nas camadas convolucionais aplicam-se filtros

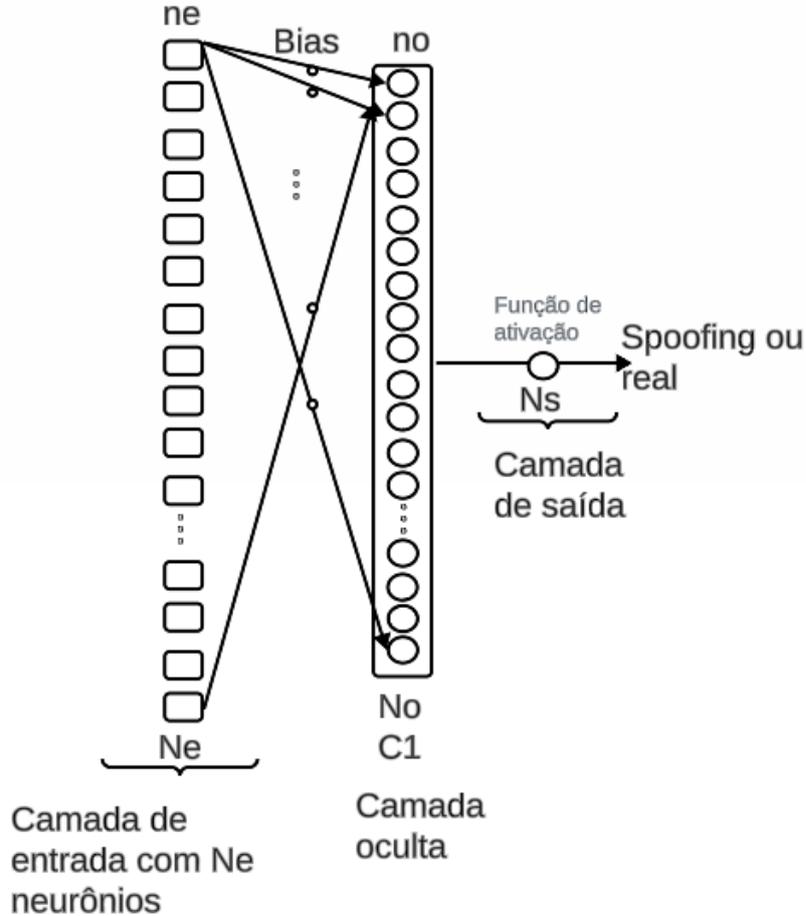


Figura 3.4: Exemplo de classificador de rede neural *multilayer perceptron*.

nos dados de entrada e buscam por recursos ou características, formando um mapa de características nesse processo. Neste tipo de rede, o somatório dos *bias*  $b$  é definido pelo produto de  $\sum_{i=0}^{N-1} b_i = xyM_iM_o$ , sendo  $N$  o número de nós da camada [34]. As dimensões  $x \times y$  do filtro *kernel* definem o campo de percepção da rede neural, e a quantidade de filtros em uma rede neural é proporcional à quantidade de mapas de características utilizados na mesma. O fator de deslocamento ou distância de cada passo do *kernel* sobre a matriz de entrada é denominado *stride*. Durante a convolução, em alguns produtos do filtro pela matriz de entrada, tem-se elementos que ficam de fora da matriz. O ajuste é realizado adicionando zeros nessas partes, e a técnica é chamada de *padding*. O *padding* pode ser *full padding*, onde adiciona-se zeros nas bordas, *same padding*, onde as dimensões da matriz de entradas são mantidas, *No padding*, nesse caso a convolução é descartada.

Normalmente, existem várias camadas convolucionais conectadas por meio de camadas de *pooling* [19]. As primeiras camadas convolucionais extraem recursos

gerais ou de baixo nível, enquanto as camadas posteriores aprendem detalhes mais sutis ou recursos de alto nível. As camadas de *pooling* diminuem o tamanho do mapa de características convolucionais para reduzir os custos computacionais [19]. Com isso, as camadas de *pooling* reduzem a complexidade computacional ao longo das etapas de treinamento. Além disso, tornam a rede neural mais robusta a ruídos e pequenas variações de translação de objetos em imagens. Em contraponto, as camadas de *pooling* intrinsecamente geram perdas de informações, o que pode ser problemático para sistemas com tarefas de alta precisão ou de detalhamento. A facilidade de ocorrer *overfitting* é outra desvantagem que deve-se ter atenção nessa camada.

Camadas totalmente conectadas, são como redes neurais com camada totalmente conectadas ou *fully connected* dentro da CNN, e essas aprendem padrões globais com base na saída de características de alto nível das camadas convolucionais e de *pooling*, gerando padrões globais. Depois que os dados passam pela camada totalmente conectada, a camada de saída emite suas previsões ou probabilidades [19]. Cada camada tem finalidades diferentes, executam tarefas nos dados e acrescentam quantidades crescentes de complexidade. Em suma, em cada camada, as dimensões matriciais mudam enquanto atravessam a rede neural convolucional, e na saída geram um conjunto de possíveis probabilidades de classes. A classe definida pela rede neural será a de maior probabilidade [19].

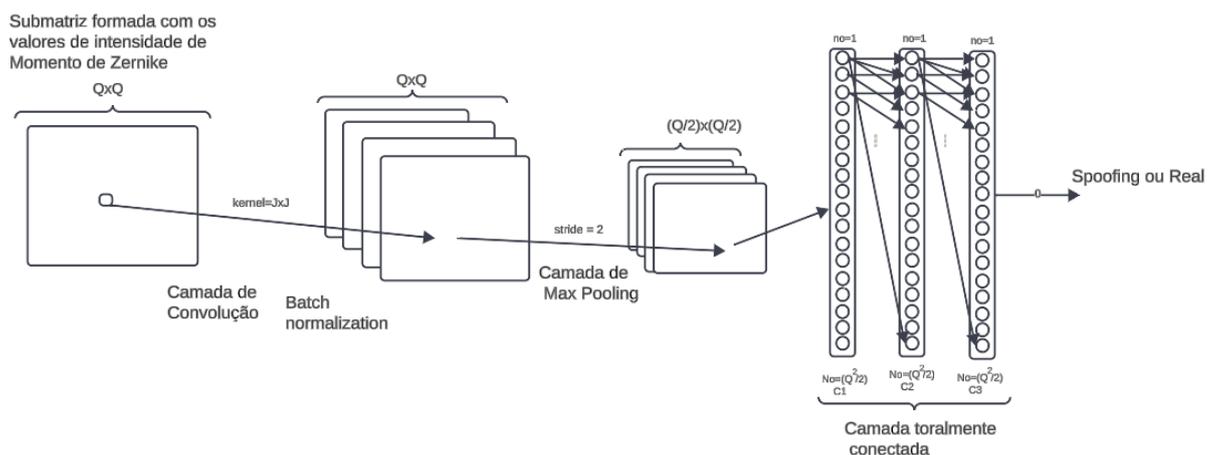


Figura 3.5: Exemplo de classificador de rede neural convolucional.

### 3.11 Classificador de rede neural residual

Com o surgimento de séries de tipos de redes neurais convolucionais profundas, a área de processamento de imagens teve diversos avanços. Nessa arquitetura, construir uma rede neural profunda permite a extração de recursos ou características de diversos níveis. A rede neural residual, chamada de ResNet (do inglês, *Residual Neural Network*), é um tipo específico de rede neural profunda que foi introduzido em 2015 no artigo “*Deep Residual Learning for Image Recognition*” [66] dos autores Kaiming He e Xiangyu Zhang. O diferencial foi que estas redes residuais são constituídas por blocos residuais [19, 66].

A primeira diferença na ResNet é que existem conexões diretas que pulam algumas camadas, podendo variar de modelo para modelo. Esta conexão é chamada de *skip connection* e é o cerne dos blocos residuais. Devido a esta conexão de salto, a saída da camada não é a mesma agora. Sem usar esta conexão de salto, a entrada é multiplicada pelos pesos da camada, seguida pela adição de um termo de polarização [66].

A arquitetura ResNet é geralmente dividida em partes, cada uma contendo vários blocos residuais com diferentes profundidades. A primeira parte da rede compreende a camada convolucional, seguida de *max pooling*, para reduzir as dimensões espaciais da entrada. A segunda até  $N$ -parte da rede contém filtros. A parte final da rede consiste no *pooling* médio global e em uma camada totalmente conectada que produz a saída.

Por outro lado, assim como em outras redes neurais profundas, a ResNet tem problema com o treino, pois depende de muito desempenho do modelo para a inicialização da ResNet. À medida que projetamos redes cada vez mais profundas, torna-se imperativo entender como a adição de camadas pode aumentar a complexidade e a expressividade da rede. Ainda mais importante é a capacidade de projetar redes onde adicionar camadas torna as redes estritamente mais expressivas, em vez de apenas diferentes. Para fazer algum progresso, precisamos de um pouco de matemática.

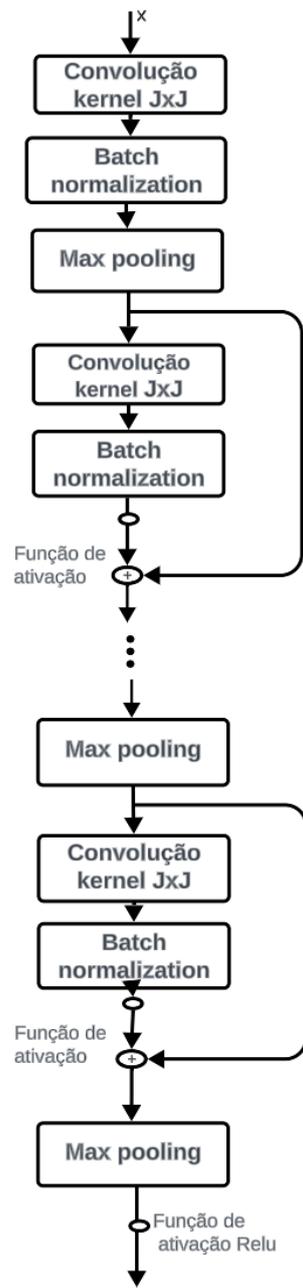


Figura 3.6: Exemplo de classificador de rede neural residual.

# Capítulo 4

## Metodologia

### 4.1 Introdução

Este capítulo apresenta a metodologia para desenvolvimento do algoritmo de momento de Zernike com classificadores, para extração de formas circulares ou não alinhadas. Em seguida, apresenta-se o sistema de referência. O capítulo está organizado em seções, que apresentam os detalhes do sistema de referência (*baseline*), e do método proposto para detecção de ataques de *spoofing*. É explicado como foram utilizados os momentos de Zernike e os classificadores de redes neurais, e por fim, um comparativo entre ambos.

Com este método proposto, aplica-se uma alternativa para a abordagem desta dissertação com o objetivo de aplicar um sistema que realize detecção de *presentation attack detection* propondo previamente o redimensionamento, normalização em escala e translação, dividir a imagem em janelas menores de *non-overlapping*, e assim aplicar momentos de Zernike em baixas, médias e alta ordem junto com classificadores de diferentes complexidades. E por fim, apresentar comparação entre os sistemas propostos neste capítulo, com o sistema de referência.

### 4.2 Sistema de referência - *baseline*

Os pesquisadores Kohli, N., Yadav, D., Vatsa, M., et al. [4] propuseram que o projeto do sistema de referência para detectar os ataques de *spoofing* do tipo *synthetic iris images*, *fake printed iris images* e *contact lens* utilizando bases de

dados de imagens biométricas de íris, como mostrado em 4.1. No início do sistema, tem a imagem de entrada de dimensões de  $N \times N$  *pixels*, com  $N = 256$ , e em escala de cinza, com somente um canal. O sistema entrega um vetor de características a um classificador de rede neural. O vetor de características, através da técnica de fusão, é constituído pela concatenação das características extraídas pela técnica LBPV, a qual extrai informações de variação de iluminação e textura da imagem, e a técnica de Momentos de Zernike de multiordem combinadas com um janelamento sem sobreposição, que extrai características não alinhadas locais da imagem.

O bloco de *non-overlapping* divide a imagem em um número inteiro de janelas de iguais dimensões  $P \times P$ , e sem sobreposição. No [4] são utilizados múltiplos divisores comuns das dimensões  $P \times P$  da imagem de entrada, onde adotou-se os valores de  $P=\{4, 8, 16\}$ . O bloco de janelamento *non-overlapping* está em série com o bloco da técnica do momento de Zernike.

O bloco de momentos de Zernike de multi-ordem, para que capture as mudanças nas formas contidas na imagem de íris, independente de ser sintética ou real. Estas formas extraídas da imagem de íris são características de formas circulares da imagem, p.ex. contornos e curvas. Os cálculos dos momentos de Zernike de multi-ordem são definidos pelo somatório das multiplicações dos conjuntos ortogonais de polinômios por cada uma das janelas da imagem de entrada dividida antes no bloco de *non-overlapping*. No bloco de momentos de Zernike de multi-ordem tem polinômio do momento de Zernike de ordem  $n$ , variando em números inteiros de 0 a 10.

O bloco que utiliza a técnica LBPV, ou seja, *local binary pattern variance*, calcula os valores em janelas da imagem com base em uma função binária a partir do pixel central da janela no sentido horário, e as estrutura em um histograma. No sistema de referência, o LBPV é usado para extrair *features* ou características da imagem da íris, principalmente informações de contraste, acrescentando variações da informação diferenciada da técnica original.

Os resultados das técnicas LBPV e *non-overlapping* com momentos de Zernike são características extraídas da imagem de entrada, que são unificadas em um bloco denominado fusão em nível de características. Este bloco realiza uma concatenação das características extraídas anteriormente nos blocos multiordem de

momento de Zernike e do LBPV, formando um vetor de características.

O vetor de características concatenado (fundido) é usado como entrada para um bloco de rede neural artificial (*multilayer perceptron*) de três camadas, entrada, oculta e saída. Esse classificador tem  $H$  nós na camada oculta, sendo aplicado o  $H = 170$ , e usa algoritmo de *backpropagation* chamado gradiente conjugado escalonado. Na rede neural MLP, foi aplicado *cross-validation* de *k-fold*, com  $k$  igual à 5. Na saída tem-se a decisão da rede neural, determinando se a imagem da íris é *spoofing* ou normal.

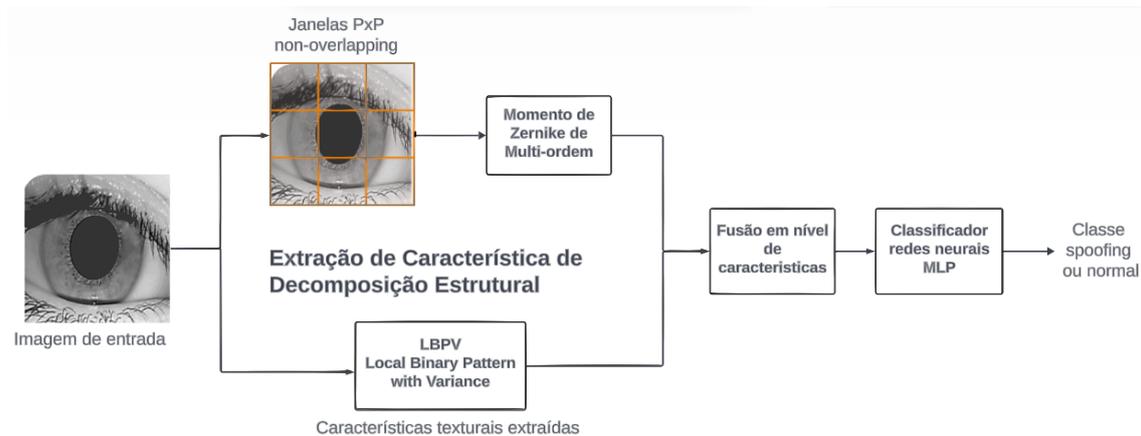


Figura 4.1: Imagem do sistema referência. Fonte: [4].

### 4.3 Sistema proposto

O sistema proposto foi projetado para detecção de *presentation attack detection* ou PAD do tipo *synthetic iris images*. O sistema proposto deste trabalho, inicialmente, tem a entrada de uma imagem de uma base de dados que foi organizada e rotulada previamente de forma automática. E a entrada do sistema foi padronizada, fazendo pré-processamentos antes de aplicar o método proposto. O sistema proposto completo é apresentado na Figura 4.2.

As bases de dados utilizadas têm diferentes dimensões e cenários distintos na aquisição da imagem. No bloco de redimensionamento foi utilizada a técnica de *resize* ou redimensionamento. Nessa técnica é aplicada readequação ao intervalo de cada dimensão da imagem a um novo intervalo a ser processado. Essas diferentes dimensões na imagem de íris são todas padronizadas com dimensões de  $N \times N$ , onde para este caso o  $N = 256$ .

Em seguida, o bloco de normalização da imagem em escala e em translação é aplicado. A saída do bloco de normalização em escala e em translação padroniza as propriedades globais da imagem processada resultante. A normalização em escala, padronizando a imagem para o intervalo de mínimo e máximo, no caso foi aplicado o intervalo de valores de 0 à 1. Ainda no terceiro bloco é realizada a normalização em translação, qual se utiliza do cálculo de centroide usando-se da equação de momentos de imagem 4.1, onde encontra-se o centroide da densidade da imagem, ou seja, a coordenada que a imagem é invariante à translação radialmente.

$$\mu_{i,j}(x, y) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x, y)x^i y^j \quad (4.1)$$

onde  $i, j = \{0, 1, 2, \dots, N - 1\}$ ,  $N$  é o número de *pixels* nas dimensões da imagem  $N \times N$ , e  $\mu_{i,j}(x, y)$  é o  $(i + j)$ -ésimo ordem da função momento da imagem, e  $I(x, y)$  é a função base da imagem referente à intensidade nas coordenadas  $x$  e  $y$ . Com esse momento, calcula-se o centroide da imagem para definir qual é o centro de distribuição da função.

$$Centroide (X_c, Y_c) = \left( \frac{\mu_{1,0}}{\mu_{0,0}}, \frac{\mu_{0,1}}{\mu_{0,0}} \right) \quad (4.2)$$

O bloco de *non-overlapping* divide a imagem de entrada em janelas quadradas de tamanho  $P \times P$  não sobrepostas.  $P$  é um divisor comum às dimensões da imagem. A saída é um conjunto de  $K$  janelas  $P \times P$ , onde no sistema proposto é  $K = \frac{256 \times 256}{P \times P}$ . As janelas são definidas pelo cálculo do fator de divisão da imagem em janelas quadradas sem sobreposição. No trabalho proposto, para esse fator resultante utilizamos três tamanhos de janelas com dimensões, com  $P = \{4, 8, 16\}$ , sendo  $K = \{4096, 1024, 256\}$  respectivamente a cada janela.

O bloco de momentos de Zernike tem como entrada cada uma das  $K$  janelas extraídas no bloco de *non-overlapping*. Para este bloco, é inicialmente calculado o valor do centroide, para cada uma das janelas, que resulta em pares de coordenadas  $(x_i, y_j)$  para cada coordenada central  $(i, j)$  da janela. Uma máscara de disco unitário é aplicada sobre a janela  $P \times P$  de modo que o centroide seja o centro da máscara, e assim definindo a área do disco unitário. Nos *pixels* fora da área do disco unitário de cada janela  $P \times P$  têm seus valores zerados.

Com base na imagem da janela, são definidos parâmetros como ordem de momento de Zernike e o número de repetições de graus no polinômio de Zernike. Estes parâmetros são utilizados para realizar o cálculo do momento de Zernike, efetuando a soma de todos os valores da janela limitada ao disco unitário e multiplicados pelos valores do polinômio de Zernike conjugado. Como o conjunto de polinômios de Zernike é bem determinado, é construída uma tabela armazenando os valores do polinômio de Zernike.

Resumidamente, para o cálculo de momentos de Zernike, é definida a soma do produto da janela do bloco de *non-overlapping* multiplicados pelos valores de polinômios de Zernike da tabela. Cada momento de Zernike  $(n, m)$  descreve os momentos geométricos circulares calculados. O resultado final dos cálculos de todos os momentos de Zernike para todas as janelas é um vetor de  $V$  termos de amplitudes ortogonais ao disco unitário. Neste bloco são extraídas as *features* descritivas das formas da imagem de íris, onde a ordem polinomial do momento de Zernike  $n$  varia de 0 à 12 em números inteiros. Nessa configuração, consegue-se aplicar os momentos de Zernike desde baixa ordem até momentos de alta ordem. A saída do bloco de momento de Zernike é um conjunto de vetores de magnitude descrevendo as características extraídas das imagens.

No método proposto foram utilizados os classificadores MLP, CNN e ResNet. A saída do classificador é o resultado da rede neural determinando se a imagem de íris da entrada do sistema é sintética ou real. Para treinamento do sistema proposto, a imagem de entrada é rotulada em *spoofing* ou real a partir de um pré-processamento realizado anteriormente. Esse conjunto de vetores é reorganizado e aplicado a três classificadores: MLP, CNN e ResNet.

Na rede neural MLP, foi aplicado *cross-validation* de *k-fold*, igual à 4 e 10. MLP foi utilizada como uma rede totalmente conectada, com a entrada reorganizada para receber o vetor de tamanho  $V \times 1$ , com duas camadas ocultas com 256 neurônios e função de ativação Relu, e um neurônio na saída com função sigmoide. No CNN, foi realizada uma reorganização do vetor de características  $V$ , e tem-se como objetivo mapear as *features* ou características, e a entrada reorganizada em uma matriz  $M \times N$ , sendo  $V_k$  o vetor intensidade e  $V_k = M \times N$ . O *kernel* tem dimensão  $3 \times 3$ . Na camada de *max pooling*, o *stride* é 2. A camada totalmente conectada com função

Relu, e na saída 1 neurônio com função *Softmax*. A Resnet foi utilizada com uma arquitetura profunda, a entrada também é reorganizada em  $M \times N$ , com camadas realizando convoluções simples com *kernel*  $3 \times 3$  e *stride* de 2, formando bloco de conexões residuais. O *zero-padding* é igual a 1, e cada camada passa adiante sem modificações. A soma das camadas passa à saída o resultado da classificação.

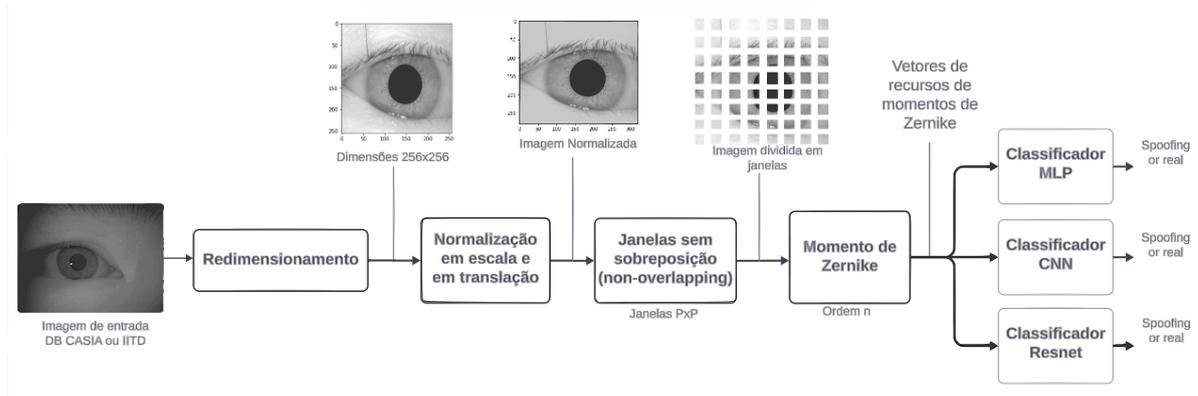


Figura 4.2: Imagem do sistema modificado proposto.

## 4.4 Sistema proposto versus baseline

No trabalho de N. Kohli, D. Yadav, M. Vatsa, et. al., [4] foi utilizado o sistema DESIST como sistema de referência ao sistema proposto neste trabalho. Ambos os sistemas têm o propósito de conseguir identificar a diferença, ou seja, detectar imagens de íris reais e imagens de íris *spoofing*. O sistema de referência parte da metodologia de extrair características locais de momentos de Zernike em virtude da divisão da imagem em janelas não sobrepostas para extrair informações de formas e contornos, e acrescenta a técnica LBPV para extrair informações de textura das imagens. Ambas as extrações são complementares em seu objetivo, e após a fusão em um vetor de características distintas. Este vetor compõe a entrada de uma rede neural *Multilayer perceptron* simples, a qual deve classificar se a imagem contém algum tipo de *spoofing*, ou seja, se a imagem é do tipo íris *printed attack*, íris sintética, ou íris com lentes de contato texturizadas. Outras técnicas de reconhecimento de íris precisam transpor a imagem de modo a mapear a íris em formato plano. Isso acrescenta deformações da informação, transformando formas não lineares da imagem em disformes tentativas de formas retilíneas, conforme visto

na imagem a seguir na Figura 4.3.

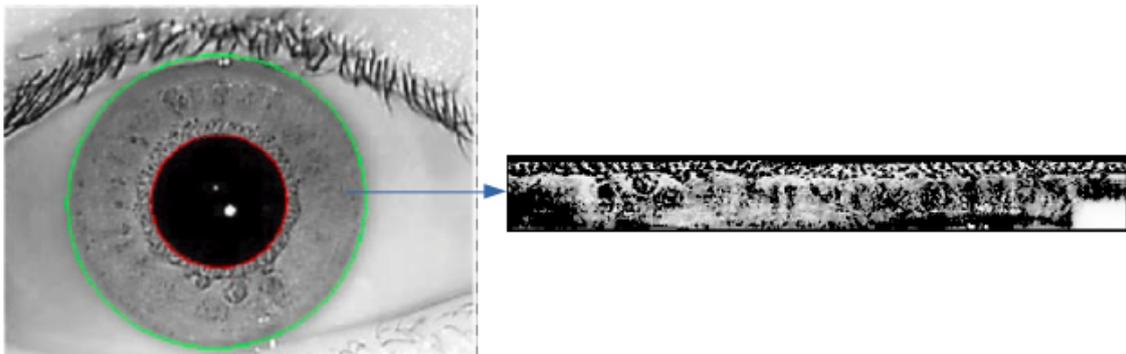


Figura 4.3: Imagem normalizada em métodos usuais de reconhecimento de iris.

No sistema proposto, momento de Zernike agrega como vantagem a possibilidade de processar a imagem sem necessidade de transformar a imagem em formatos que tragam deformação, ou mesmo mais processamentos para adequar a entrada do sistema para a aplicação da técnica. Momento de Zernike extrai as características de formas circulares na imagem, que são formas bastante recorrentes em imagens de íris. Independente do tipo de coordenadas que os cálculos estejam se orientando, momentos de Zernike processam a imagem considerando-a de formas radiais. O sistema proposto utiliza momento de Zernike, onde as propriedades de invariância à escala, invariância à translação e invariância à rotação são também aplicáveis nas características de formas circulares extraídas, garantindo um nível de robustez ao sistema.

O redimensionamento permite que sejam processadas outras bases de dados, e igualmente comparadas no processo de aplicação do método proposto. A normalização de escala e em translação são outros diferenciais para o método proposto, pois permite que garanta à imagem em processamento que tenha invariância em escala e em translação, permitindo que a função de distribuição de densidade para os *pixels* seja radialmente posicionada na imagem. Momentos de Zernike, quando aplicados diretamente em uma imagem, irão extrair as características de formas circulares e não alinhadas globalmente na imagem. O janelamento sem sobreposição permite, após a divisão da imagem em janelas menores, que a técnica de momento de Zernike seja aplicada em partes locais na imagem. Uma consequência da aplicação de janelamento é o aumento de informações de características extraídas da imagem.

As ordens de momento de Zernike para o sistema proposto são calculadas

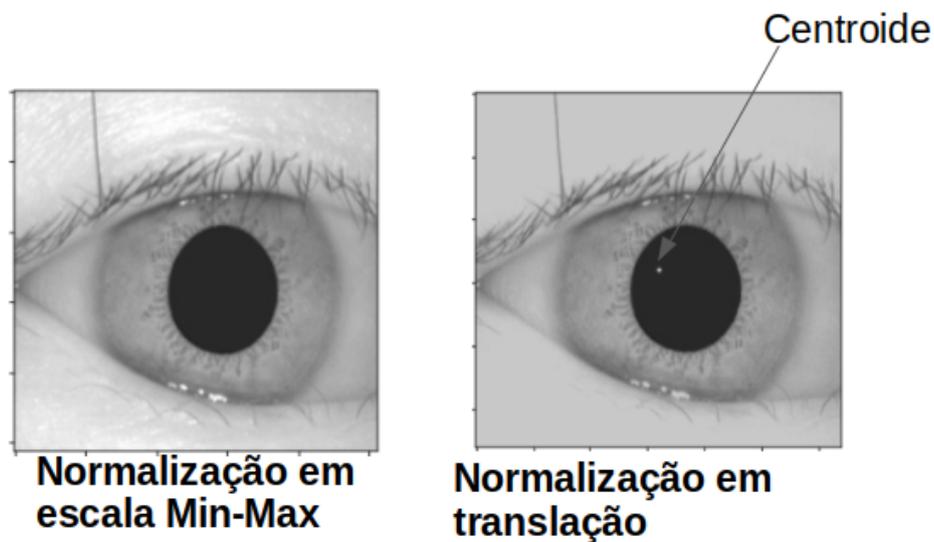


Figura 4.4: Exemplo de imagem normalizada usando o método proposto.

até a décima segunda ordem, o que acrescenta duas ordens a mais em relação ao método de referência. O que se destaca nesse aumento da ordem é que são variáveis exponenciais, de forma que acrescenta muito mais informação ao sistema proposto, e um desafiador processo computacional. De forma simplificada, se processar uma imagem com momento de Zernike de ordem  $n = 10$  teria  $2^{10} = 1024$  processamentos, proporcionalmente, para  $n = 12$  teria-se  $2^{12} = 4096$  processamentos, ou seja, 4 vezes mais processamentos, e, conseqüentemente, muito mais informação extraída.

Para o classificador, o aumento de complexidade estrutural das redes neurais, como exemplo a CNN e a ResNet, agrega na extração de características do sistema. A ideia em plano de fundo é verificar se, aplicando redes neurais de técnicas mais complexas, também teremos resultados melhores. São utilizadas redes neurais CNN e ResNet com estruturas comumente utilizadas na literatura, de modo que adicione similaridade entre os métodos da literatura com o sistema proposto ao realizar comparação adequada aos trabalhos relacionados. A rede neural MLP é construída com mais uma camada, para possibilitar a verificação se a classificação corresponde ao aumento da complexidade da rede neural, também sendo um diferencial para o sistema proposto.

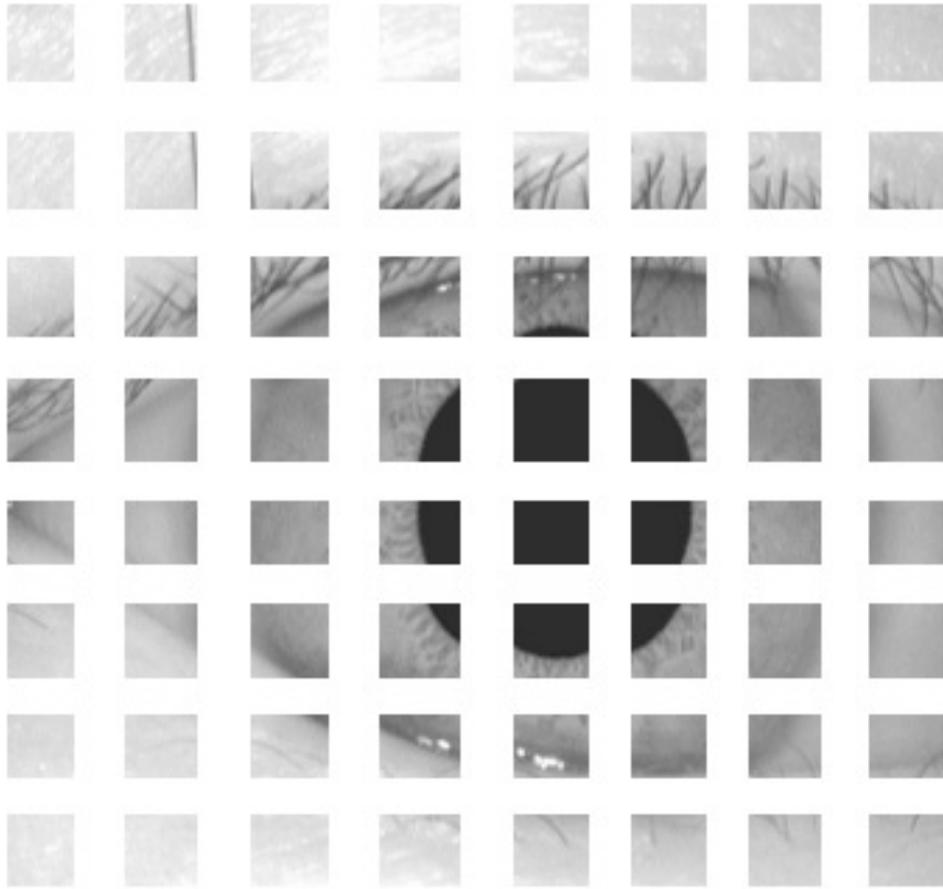


Figura 4.5: Exemplo de janelamento sem sobreposição de janelas da imagem (*non-overlapping*).

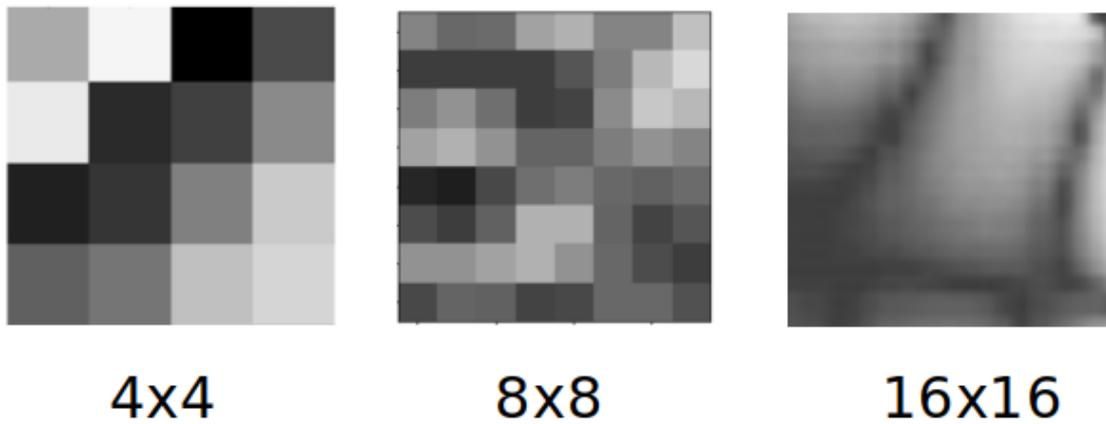


Figura 4.6: Exemplo de momento de Zernike aplicado à diferentes tamanhos de janela da imagem.

# Capítulo 5

## Base de dados

### 5.1 Introdução

No sistema de referência, as bases de dados formam um novo conjunto composto de base de dados IIT Delhi Iris que são IIIT-Delhi CLI, IIIT-Delhi IIS, SDB, MID e o CSD que é uma combinação de bases de dados IIT de imagens de íris, todos com dimensão  $256 \times 256$ . Ao treinar o método apresentado em [4] mostrado na Figura 4.1, as imagens de entrada de dimensões  $256 \times 256$  são organizadas em uma base de dados que é formada por 50% de imagens de íris *spoofing* e 50% de imagens de íris reais.

No trabalho proposto, as bases de dados escolhidas são combinação de IIT Delhi Iris *database* junto com IITD Iris CSD, e a combinação de CASIA Iris *Syn* com CASIA Iris *thousand*. As bases de dados utilizadas formam dois grupos de bases de dados de imagens de íris sintéticas e reais. O primeiro grupo é formado pelas bases de dados IITD-*Syn* e o IITD Iris CSD, um combinado de bases de dados IITD de imagens de íris sintéticas e reais de dimensão  $256 \times 256$ . No segundo grupo, é formada pela base de dados CASIA Iris *Syn* e CASIA Iris *thousand*, formados por imagens de íris sintéticas e reais de dimensão  $320 \times 240$ .

- CASIA - conjunto formado pelas base de dados CASIA-Iris-Synthetic e CASIA-Iris-Thousand.
- IITD - conjunto formado pelas base de dados IITD *database* e *combined spoofing database* (CSD).

Em ambas as bases de dados, e para todos os casos, o sistema foi submetido ao treinamento, validação e teste, totalizando duas possibilidades de quantidade de imagens: 2000 e 6000 imagens, seguindo a Figura 5.1.

Para o sistema proposto, o *redimensionamento* é utilizado devido às variações de dimensões das bases de dados. A partir das bases de dados IITD-*Syn*, IITD Iris CSD, CASIA Iris *Syn* e CASIA Iris *thousand* essas novas bases de dados têm todas as imagens padronizadas com dimensão de  $256 \times 256$ . As imagens de íris redimensionadas que formam uma base de dados são organizadas de duas maneiras: 50% reais e 50% sintéticas. Ainda no treinamento, a imagem de entrada é organizada e rotulada a partir de uma das duas bases de dados CASIA ou IITD, ambas contendo imagens de íris sintéticas (com *spoofing*) e reais rotuladas. O pré-processamento gera uma base de dados que tem dois casos de quantidade de imagens rotuladas: 2000 e 6000 imagens, conforme Figuras 5.1 e 5.2.

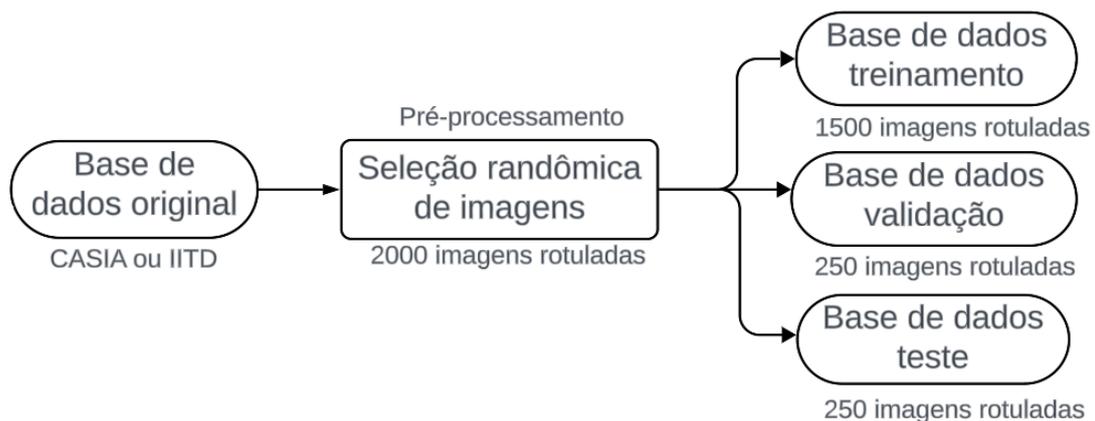


Figura 5.1: Diagrama de pré-processamento realizado antes do início do treinamento do sistema proposto com 2000 imagens rotuladas na entrada.

### 5.1.1 Base de dados CASIA-Iris

Desde a década de 90 tem-se pesquisas em investigação sobre problemas na área de íris biométrica. O avanço dessas investigações se deu pela disponibilidade de bases de dados de imagem de íris projetadas com rigor científico e em quantidade amostral suficiente. Dentre estes conjuntos de dados disponíveis publicamente tem-se a CASIA *Iris Image Database* ou CASIA-Iris para abreviar. O CASIA *Iris Image Database* (CASIA-Iris) foi lançado para a comunidade biométrica internacional e

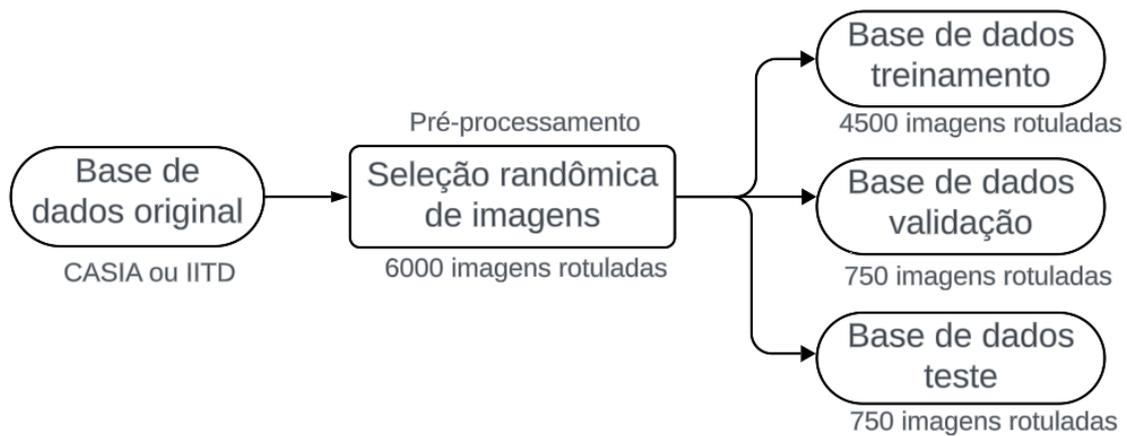


Figura 5.2: Diagrama de pré-processamento realizado antes do início do treinamento do sistema proposto com 6000 imagens rotuladas na entrada.

atualizado do CASIA-IrisV1 até CASIA-IrisV4, que contém 54.601 imagens de íris de mais de 1.800 indivíduos reais e 1.000 indivíduos virtuais. Essas imagens de íris são arquivos JPEG de 8 bits em níveis de cinza. Seis conjuntos de dados foram coletados ou sintetizados: CASIA-Iris-Syn, CASIA-Iris-Interval, CASIA-Iris-*Lamp*, CASIA-Iris-*Distance*, CASIA-Iris-*Thousand*.

A base de dados CASIA-Iris-*Interval* é uma das bases de dados utilizadas para compor a entrada do sistema proposto. A CASIA-Iris-*Interval* apresenta um fluxo luminoso adequado para imagens de íris, podendo capturar imagens de íris muito claras. A CASIA-Iris-*Interval* é bem adequada para o estudo de características detalhadas da textura das imagens de íris.

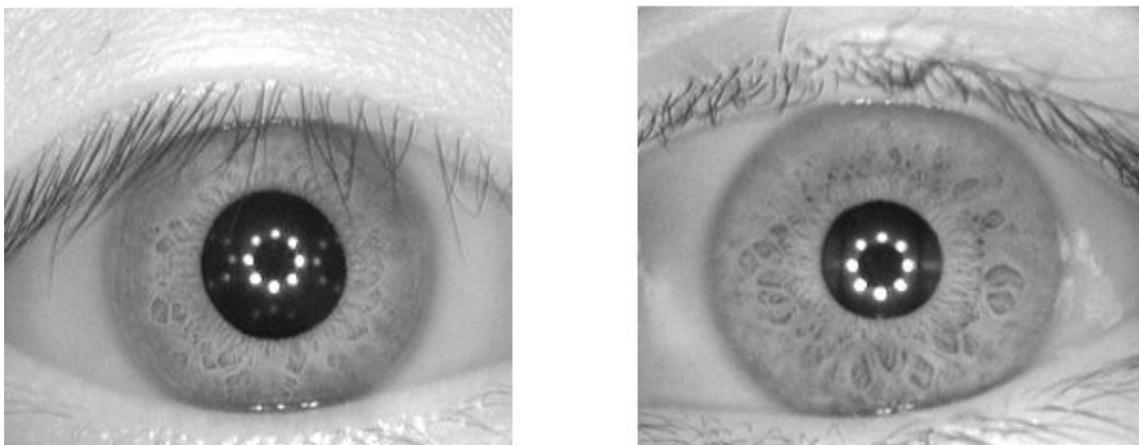


Figura 5.3: Exemplos de imagens da base de dados CASIA Iris Interval.

Outra base de dados utilizada nesta dissertação é a CASIA-Iris-*Thousand*, que é formada por 20.000 imagens de íris de 1.000 indivíduos. As variações de

classe em *CASIA-Iris-Thousand* são óculos e reflexos na imagem. Como *CASIA-Iris-Thousand* é o primeiro conjunto de dados de íris disponível, ele é bem adequado para estudar as características da íris e para testar novos métodos de classificação.

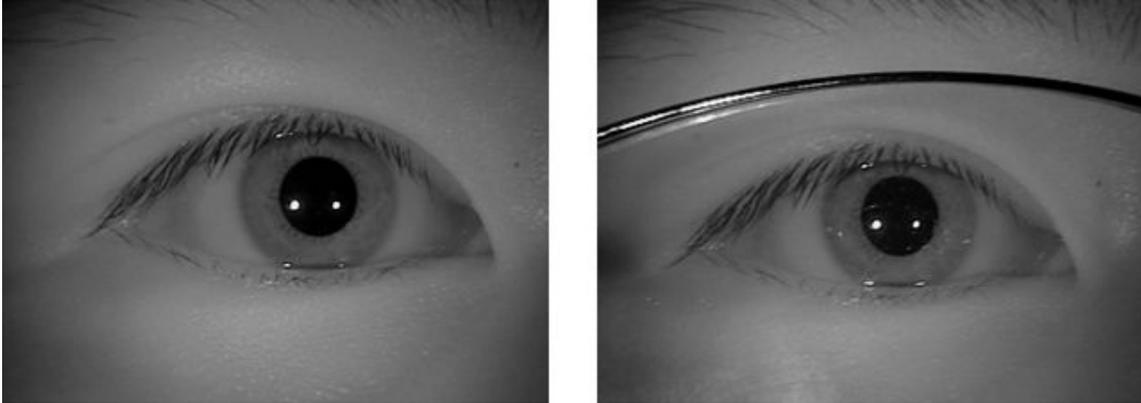


Figura 5.4: Exemplos de imagens da base de dados *CASIA Iris Thousand*.

Uma outra base de dados utilizada na dissertação foi a *CASIA-Iris-Syn*, que tem 10.000 imagens de íris sintetizadas de 1.000 classes. Texturas de íris das imagens sintetizadas são realizadas automaticamente a partir de um subconjunto de *CASIA-Iris*. Em suma, imagens de íris artificiais se tornam mais realistas ao incorporar as características circulares de íris às imagens de íris reais. Em contraponto, no acréscimo dessas características das íris em imagens sintetizadas, também trazem deformação, desfoque e rotação. Esse problema é desafiador na representação e correspondência de características da íris.

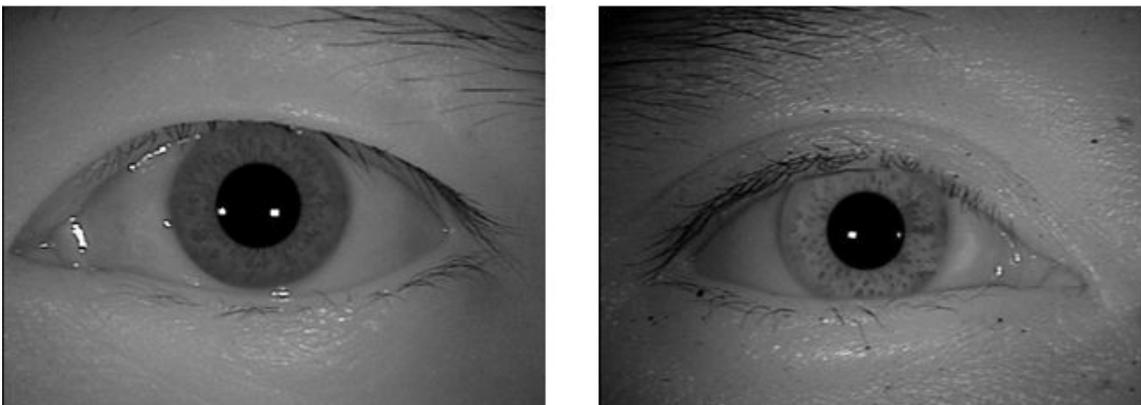


Figura 5.5: Exemplos de imagens da base de dados *CASIA Iris Syn*.

Conforme informado pelo fornecedor da base de dados, os resultados de desempenho testados nessa base de dados de imagens de íris sintetizadas têm características estatísticas semelhantes à base de dados de íris reais. Com tal, o uso de

CASIA-Iris-Syn para pesquisa de reconhecimento de íris é bem atrativo por ter um histórico de ser bem-sucedido na maioria das pesquisas de reconhecimento de íris.

### 5.1.2 Base de dados IITD-Iris

Como na base de dados anterior, a avaliação de desempenho depende muito da disponibilidade de uma base de dados de íris com mais amostras adquiridas. No instituto IIT Delhi, o laboratório de pesquisa em biometria vem coletando imagens de íris desde 2007 para construção de um banco de dados cada vez maior em escala e disponível para domínio público.

O banco de dados de íris do IIT Delhi consiste em imagens de íris coletadas. A aquisição de imagens foi escrita para adquirir e salvar essas imagens em formato bitmap, e de 224 indivíduos. A base de dados de 1120 imagens é organizada em 224 pastas. A resolução dessas imagens é de  $320 \times 240$  *pixels* e todas essas imagens foram adquiridas em ambiente controlado. No sistema proposto, as imagens foram modificadas para o formato PNG.

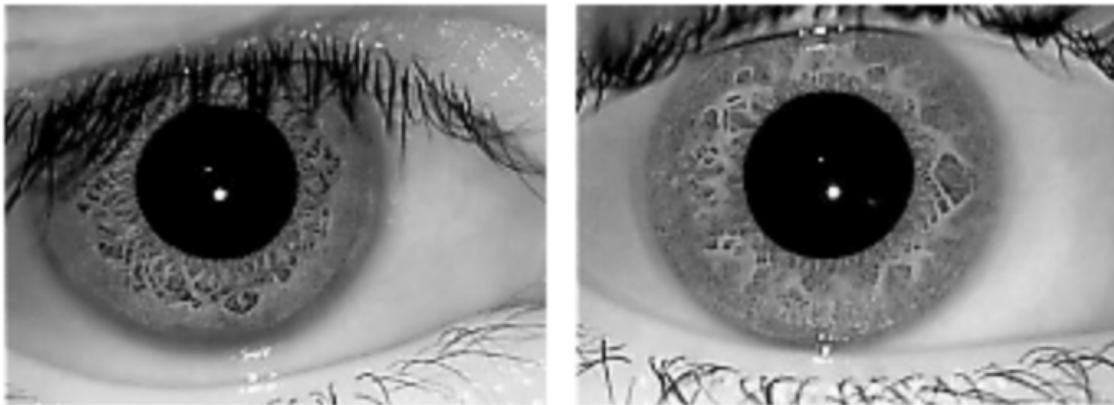


Figura 5.6: Exemplos de imagens da base de dados IIT Delhi Iris.

A base de dados IIT Delhi CSD (do inglês *Combined Spoofing Database*) consiste em imagens de outras bases de dados da IIIT-Delhi Contact Lens Iris (CLI), da IIITD Iris Spoofing (IIS) e da base de dados Multi-sensor Iris. Essa base de dados foi criada pelos pesquisadores N. Kohli, D. Yadav, M. Vatsa, et al., que são autores do artigo [4] e criaram essa base de dados para testar o sistema de referência desta dissertação.

Base de dados IIIT-Delhi CLI (do inglês *Contact Lens Iris*) é composta por

imagens de 101 indivíduos. As imagens foram capturadas com e sem lente, inclusive com lentes texturizadas. A base de dados IIIT-Delhi CLI é utilizada para criar o banco de dados IIS. A base de dados IIITD IIS (do inglês *Iris Spoofing* ) é uma base de dados de íris que tem como ataques de *spoofing* de *fake printed attack image*. No ataque *print+capture*, a entrada é composta por imagens de íris impressas, e no ataque *print+scan*, as imagens de íris impressas são escaneadas. A base de dados SDB sintético, desenvolvida pelo pesquisador Galbally et al. usando a técnica de Markov *Random Field* com recursos de íris, contém imagens de 1000 indivíduos. Base de dados de íris do IIT Delhi contém imagens de íris reais de 224 indivíduos. A base de dados de íris MID (do inglês *multisensor iris database*) é composta por representações da classe real, imagens de íris de 547 indivíduos.

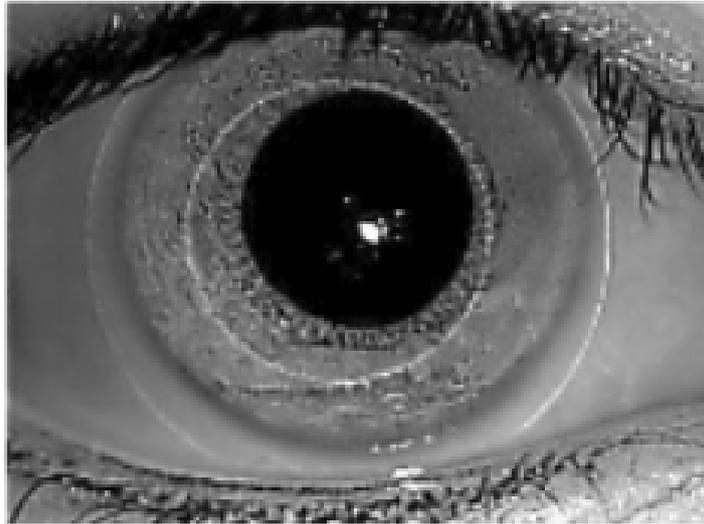


Figura 5.7: Exemplos de imagens da base de dados IIIT Delhi CLI.

# Capítulo 6

## Experimentos e Resultados

### 6.1 Introdução

Neste capítulo serão apresentadas as etapas do procedimento experimental, assim como as análises dos resultados e avaliação da metodologia proposta para detecção de ataque de *spoofing*. No decorrer da execução dos experimentos, objetivase encontrar respostas para questionamentos surgidos durante a pesquisa:

- Quais os resultados encontrados nas métricas de desempenho de detecção de imagens de iris *spoofing* quando aplicados o método proposto?
- Entre os classificadores utilizados, algum obteve melhores resultados quando comparado aos demais?
- Com os resultados das métricas de desempenho nas duas base de dados criadas a partir das bases de dados de referência, é viável generalizar o uso dos método proposto de detecção de imagens de iris *spoofing*?

### 6.2 Setup

O setup dos experimentos realizados nesta dissertação possui a seguinte configuração:

- Ambiente de desenvolvimento: os algoritmos foram implementados utilizando o sistema operacional Linux Ubuntu 22.04.3 LTS, em um processador Intel(R) Core(TM) i7-7500U CPU 2,70 GHz.

- Para o desenvolvimento dos algoritmos, foi utilizado a linguagem de programação Python [67] na versão 3.7.6 em ambiente virtual Conda, com o Anaconda Navigator, e utilizando o compilador Spider 4.0.1, IDE em python.
- Em conjunto, algumas API foram utilizadas a fim de executar o TensorFlow 2.9.0 e Keras 2.9.0, bibliotecas de código aberto para redes neurais artificiais. Foram utilizadas outras bibliotecas auxiliares como NumPy 1.18.1, Pandas 1.0.1, PIL 7.0.0, Mahotas 1.4.12 entre outros.

### 6.3 Métricas de desempenho

As métricas utilizadas nesta dissertação usadas como critérios na avaliação foram: acurácia, precisão, *recall* ou revocação e *F1-score*, todas as métricas normalmente utilizadas na literatura. Aplicada matriz de confusão, para indicar os erros e acertos do método proposto, comparando com o resultado esperado. No contexto de *presentation attack detection*, detectar corretamente os ataques de *spoofing* em uma imagem é primordial. Com tal, dentre os termos abordados nas métricas, tem-se:

- O verdadeiro-positivo (do inglês *true-positive*) ou abreviadamente o TP, onde a imagem é sintética com diagnóstico positivo pelo sistema, ou seja, de forma assertiva;
- O verdadeiro-negativo (do inglês *true-negative*) ou TN, onde o caso é de uma imagem real e o diagnóstico é definido como negativo, ou seja, sistema classifica imagem de iris real assertivamente como uma imagem de iris real;
- O falso-positivo (do inglês *false-positive*) ou FP, onde é o caso da imagem real e o diagnóstico é positivo, ou seja, sistema erra classificando a imagem real como de iris sintética;
- O falso-negativo (do inglês *false-negative*) ou FN, onde o caso é de uma imagem de iris sintética, e o diagnosticado como negativo, ou seja, predita erroneamente como imagem de iris real.
- A acurácia é a classificação correta das imagens, sejam sintéticas ou reais,

dividida pelo total de imagens na entrada do sistema, ou seja, é definida como:

$$Accuracy = \left( \frac{TP + TN}{TP + TN + FP + FN} \right) \quad (6.1)$$

- A precisão é referente ao número acertos da classe que são realmente imagens sintéticas, divididos pelo número total de exemplos previstos na classe, no caso todas as classes previstas como imagens sintéticas:

$$Precision = \left( \frac{TP}{TP + FP} \right) \quad (6.2)$$

- O *Recall* para esta dissertação é quando o número de imagens de iris sintéticas identificadas corretamente dividida pelo número total de imagens que foram acertadas:

$$Recall = \left( \frac{TP}{TP + FN} \right) \quad (6.3)$$

- A métrica *F1-score* é uma média calculada entre a precisão e o *recall*, e é importante quando ambas métricas são consideradas no sistema, ou seja, quando precisa expressar a média de acertos de imagens sintéticas em relação ao conjunto das imagens na base de dados, e o mesmo em relação ao conjunto de predição das imagens:

$$F1-Score = \left( 2 \frac{Precision \times Recall}{Precision + Recall} \right) = \left( \frac{2 TP}{2 TP + FP + FN} \right); \quad (6.4)$$

A base de dados de imagens de íris de classes reais e *spoofing* são classificadas correta e incorretamente pela estrutura proposta. O treinamento é realizado no conjunto todo conforme modificações realizadas no procedimento experimental. Para tal, precisa-se informar as variáveis do sistema proposto, que são base de dados CASIA e IITD formadas pelo pré-processamento na Figura 5.1 e 5.2, o tamanho das bases de dados 2000 e 6000, respectivamente. Além disso, informar qual janela *non-overlapping* foi utilizada:  $4 \times 4$ ,  $8 \times 8$  ou  $16 \times 16$ , informar a ordem do momento de Zernike a ser calculado no processamento em específico, que pode variar entre os números inteiros de  $n = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ , informar a rede a ser aplicada do conjunto de redes MLP, CNN ou ResNet. Com essas informações, pode-

se apresentar e executar o sistema proposto, e obter os resultados das métricas e montar a matriz de confusão.

## 6.4 Análise de desempenho

A Seção 6.4 apresenta o procedimento experimental, com as alterações relacionadas às características específicas para cada algoritmo. Este processo é fundamentado na metodologia descrita no Capítulo 4, especialmente no diagrama geral de blocos do sistema proposto na Figura 4.2. Para melhor entendimento, este diagrama é novamente apresentado com o conjunto de variáveis na Figura 6.1 e com as etapas descritas abaixo.

Bloco de redimensionamento: todas as imagens da entrada serão convertidas para o formato de dimensões  $256 \times 256$ , independente das dimensões anteriores nas imagens na base de dados original IITD e CASIA, conforme Capítulo 5.

Bloco de normalização em escala e translação: para garantir a padronização das imagens de entrada, é aplicada normalização com Min-Max em escala, com valores no intervalo de 0 à 1. A normalização em translação, é calculado o centróide de cada imagem usando momentos de imagens, no intuito de ter simetria para a função distribuição de probabilidade.

Bloco de janela sem sobreposição ou *non-overlapping*: neste bloco do sistema, a imagem é dividida em sub-matrizes ou janelas, que variam em três possibilidades de dimensões  $4 \times 4$ ,  $8 \times 8$  e  $16 \times 16$ . Em cada ciclo de execução do sistema proposto, será calculado para uma dessas dimensões das sub-matrizes.

Bloco de Momentos de Zernike: são descritores robustos e precisos que caracterizam formas circulares de imagem de íris. Para momentos de Zernike, essas formas são extraídas usando a propriedade de ortogonalidade, invariância à translação e escala dos polinômios de Zernike, e conseqüentemente refletidos para os momentos de Zernike. Esta área de descrição está delimitada ao disco com raio unitário, conforme descrito na seção 4.3. Cada ciclo de execução do sistema proposto usa uma ordem do maior polinômio de Zernike durante o cálculo de momentos de Zernike. Estas ordens variam no conjunto de valores de  $n$ , com  $n = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ . O objetivo é explorar os níveis de extração de características que variam com a ordem

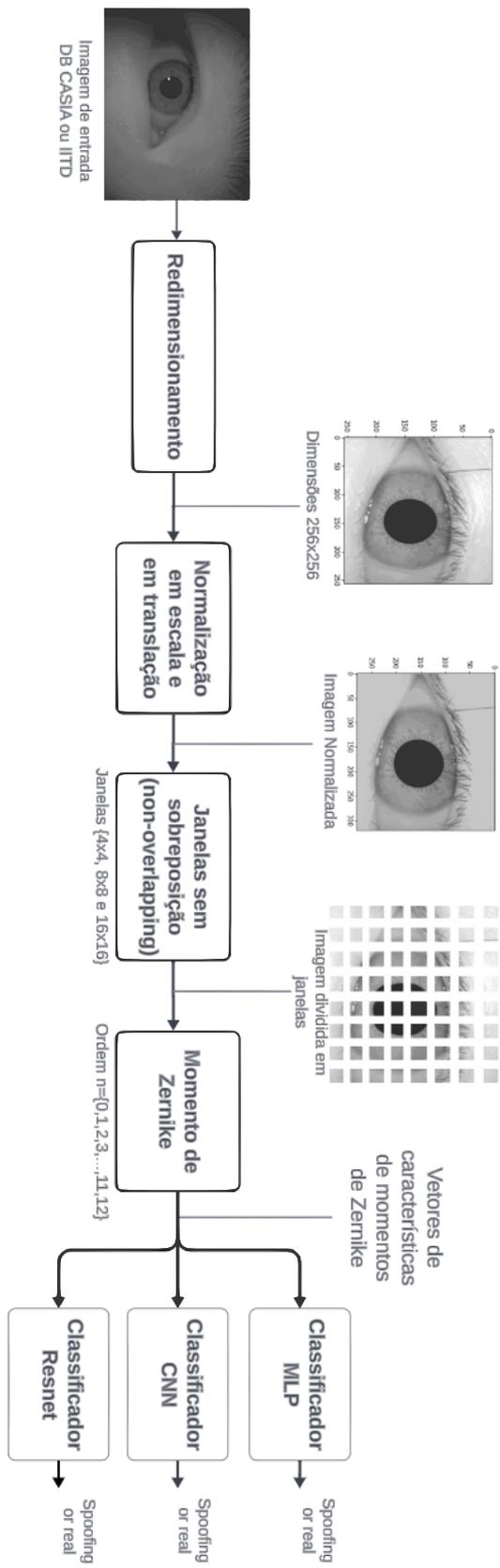


Figura 6.1: Imagem do sistema modificado proposto.

do polinômio de Zernike;

Bloco do classificador: Utiliza-se como classificadores três métodos de redes neurais: MLP, CNN e ResNet. No treino das redes, aplicamos dois casos: *k-fold* igual à 4 e 10. A quantidade de neurônios na camada oculta foi projetada de acordo com o arranjo das informações de entrada e do tipo de rede. Abordaremos mais detalhadamente nas seções a seguir.

Após o pré-processamento já temos definido as imagens da base de dados para validação e teste do sistema proposto, conforme diagrama em blocos descrito na Figura 6.1 e seguindo a quantidade pré-determinada no *k-fold* inicialmente definido. O redimensionamento tem o padrão de  $256 \times 256$ . As normalizações em escala e em translação são aplicadas conforme técnica explicada na Seção 6.4. O bloco de janela sem-sobreposição (*non-overlapping*) será utilizado somente uma das opções de janela por vez, com armazenamento realizado individualmente por imagem. Para o bloco de momento de Zernike, será processada uma ordem por vez, e no armazenamento também será individualmente armazenado por imagem, que será armazenado em um *buffer* de memória. Para o classificador, terá na entrada esse *buffer* de memória com os valores das características extraídas, e o tamanho do mesmo varia de acordo com a ordem do momento de Zernike utilizada no ciclo de execução. Inicialmente, o *buffer* de memória é organizado de acordo com o tipo de rede neural que é definido para o ciclo de execução, entre as redes neurais MLP, CNN e ResNet. Por exemplo, o *buffer* de memória é organizado para um vetor de 256 valores de entrada da rede MLP, mas para a CNN e ResNet é organizado como uma matriz de  $16 \times 16$ . Também são pré-carregados os valores e parâmetros das configurações da rede neural obtidas no treinamento, como, por exemplo, valores de taxa de aprendizado, taxa de decaimento, tamanho do kernel, valor do *padding* e *stride*, valores de *batch* e épocas.

#### 6.4.1 Momento de Zernike com MLP

Sobre o sistema onde tem-se a combinação de momentos de Zernike com MLP, é apresentando as características da rede, onde na entrada tem 256 nós, tem 2 camadas ocultas com 256 nós cada, e na camada de saída com 1 nó e função de ativação ReLU. O algoritmo gradiente de descida e retropropagação (ou em inglês *backpropagation*). A taxa de aprendizado é igual à 0.001 e decaimento de 0.0001.

Com relação ao lote *batch* é de 32, épocas igual à 200, e perda calculada pelo algoritmo *binary crossentropy*, e *k-fold* foram de 4 e 10. Com estes valores de configuração da rede para treino teremos acurácia, precisão, *recall* e *F1-score*.

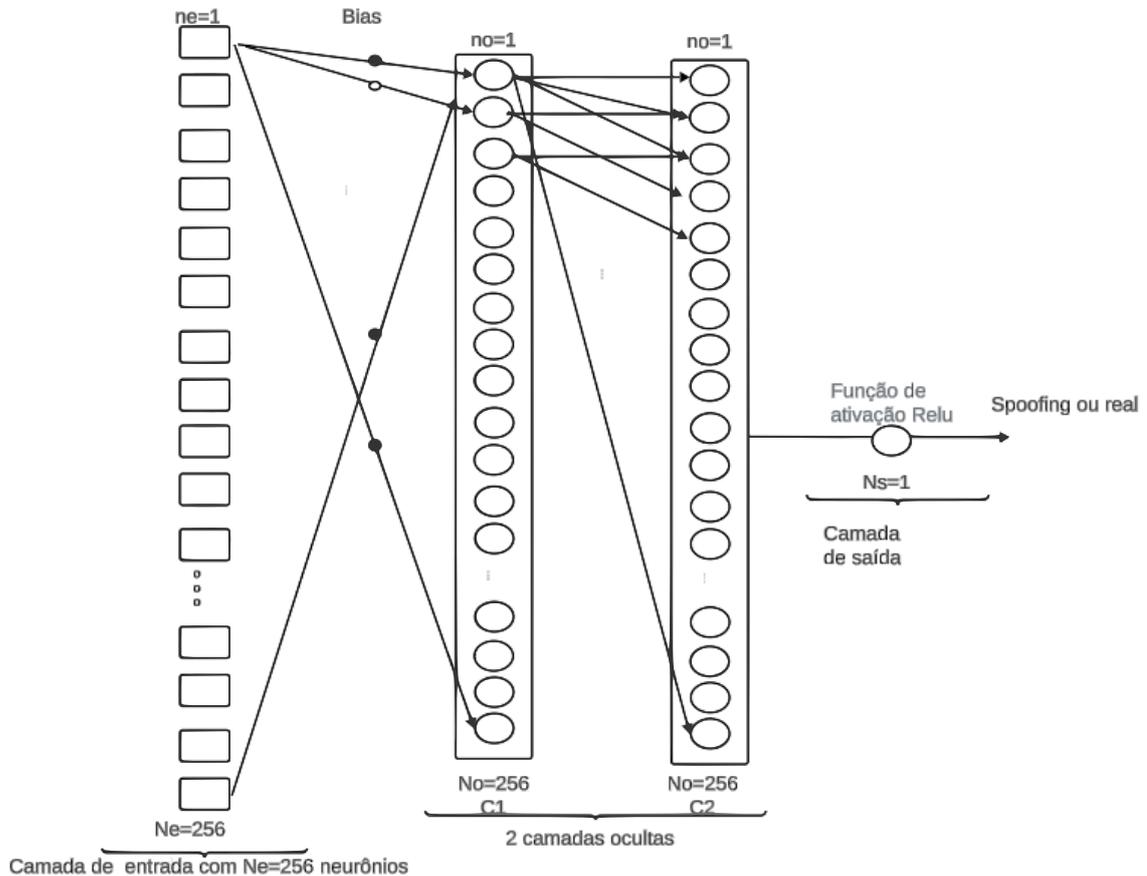


Figura 6.2: Classificador MLP do sistema proposto.

Tabela 6.1: Sistema de momentos de Zernike com MLP onde os parâmetro de ordem de momentos de Zernike igual à  $n = 5$

Base	Qtd	Janela	Acurácia	Precisão	Recall	F1-Score
CASIA	2000	4x4	0,56	0,34	0,83	0,48
		8x8	0,56	0,31	0,63	0,42
		16x16	0,62	0,38	0,86	0,53
	6000	4x4	0,77	0,54	0,65	0,59
		8x8	0,73	0,46	0,66	0,55
		16x16	0,72	0,45	0,57	0,50
IITD	2000	4x4	0,62	0,36	0,67	0,47
		8x8	0,70	0,45	0,84	0,59
		16x16	0,66	0,42	0,86	0,56
	6000	4x4	0,68	0,39	0,45	0,42
		8x8	0,77	0,53	0,82	0,64
		16x16	0,83	0,62	0,77	0,69

Tabela 6.2: Sistema de momentos de Zernike com MLP onde os parâmetro de ordem de momentos de Zernike igual à  $n = 9$

Base	Qtd	Janela	Acurácia	Precisão	Recall	F1-Score
CASIA	2000	4x4	0,56	0,33	0,71	0,45
		8x8	0,68	0,43	0,84	0,57
		16x16	0,66	0,42	0,97	0,59
	6000	4x4	0,81	0,59	0,75	0,66
		8x8	0,84	0,64	0,79	0,71
		16x16	0,85	0,67	0,82	0,74
IITD	2000	4x4	0,56	0,33	0,76	0,46
		8x8	0,72	0,46	0,62	0,53
		16x16	0,56	0,35	0,84	0,49
	6000	4x4	0,71	0,42	0,43	0,43
		8x8	0,75	0,51	0,63	0,56
		16x16	0,82	0,65	0,57	0,61

Tabela 6.3: Sistema de momentos de Zernike com MLP onde os parâmetro de ordem de momentos de Zernike igual à  $n = 12$

Base	Qtd	Janela	Acurácia	Precisão	Recall	F1-Score
CASIA	2000	4x4	0,77	0,53	0,81	0,64
		8x8	0,82	0,59	0,86	0,70
		16x16	0,74	0,49	1,08	0,67
	6000	4x4	0,87	0,71	0,81	0,76
		8x8	0,91	0,77	0,89	0,82
		16x16	0,91	0,80	0,87	0,83
IITD	2000	4x4	0,51	0,32	0,84	0,46
		8x8	0,56	0,35	0,86	0,50
		16x16	0,62	0,40	1,02	0,58
	6000	4x4	0,74	0,48	0,54	0,51
		8x8	0,68	0,38	0,48	0,43
		16x16	0,79	0,55	0,87	0,68

Precisão de detecção varia para detecção de *spoofing* de íris em diferentes bases de dados separadamente usando a estrutura do sistema proposto. Observe que o treinamento é realizado no conjunto para os testes, são relatados resultados relativos a ataques de *spoofing*. Abaixo mostramos os resultados para MLP com momentos de Zernike de ordem 5, 9 e 12 e *non-overlapping*  $4 \times 4$ ,  $8 \times 8$ , e  $16 \times 16$ . O nível de acurácia do sistema apresenta melhorias com uma rede mais adequada ao padrão de imagens com características ou *features* circulares. A seguir, apresentamos o resultado da MLP variando a ordem de momentos de Zernike no conjunto de  $n = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ .

Tabela 6.4: Sistema proposto onde a ordem de momentos de Zernike varia em números inteiros de  $n = 0$  à  $n = 12$ , com base de dados pré-processada CASIA, janela *non-overlapping* de  $16 \times 16$ , e classificador MLP conforme descrito em 6.4.1

Ordem ZM	Classificador	Acuracia	Precisao	Recall	F1-Score
0	MLP	0,62	0,31	0,42	0,35
1		0,64	0,33	0,45	0,38
2		0,65	0,35	0,47	0,40
3		0,67	0,38	0,49	0,43
4		0,70	0,42	0,52	0,46
5		0,72	0,45	0,57	0,50
6		0,75	0,50	0,70	0,58
7		0,77	0,52	0,74	0,61
8		0,79	0,55	0,78	0,64
9		0,85	0,67	0,82	0,74
10		0,88	0,71	0,86	0,78
11		0,89	0,74	0,87	0,80
12		0,91	0,80	0,87	0,83

#### 6.4.2 Momento de Zernike com CNN

Para o sistema de momentos de Zernike com o classificador CNN teremos os seguintes valores de configuração da rede: a entrada é uma matriz  $16 \times 16$ , a camada de convolução tem valor de *kernel* de  $3 \times 3$ , *stride* igual à 1 e *padding* igual à 1, e a função de ativação utilizada é a Relu. Na camada de *pooling* é aplicado o *max pooling* de  $2 \times 2$ , *stride* igual à 2, a função de ativação dessa camada também é Relu. A camada de rede densamente conectada é composta por 3 camadas de 128 neurônios cada camada e a função de ativação é Relu.

O algoritmo de retropropagação ou *backpropagation* é o gradiente descendente estocástico. Na rede de classificação CNN foram utilizados os seguintes valores gerais da rede: usado o algoritmo de *batch normalization* para *batch size* igual à 32, a taxa de aprendizado é de 0.001. Os valores de *k-fold* utilizados foram 4 e 10. A saída do sistema tem a função de ativação *Softmax*.

Os blocos de redimensionamento, normalização, *non-overlapping* e momentos de Zernike não têm mudanças quando comparados ao treinamento. Para o classificador, inicialmente são definidas configurações, ou seja, valores e parâmetros para qual rede neural CNN irá ser utilizada. Com isso, para avaliação e testes são carregadas as configurações da rede neural obtidas no treinamento. São configuradas tanto as características específicas da rede quanto os valores obtidos para as variáveis

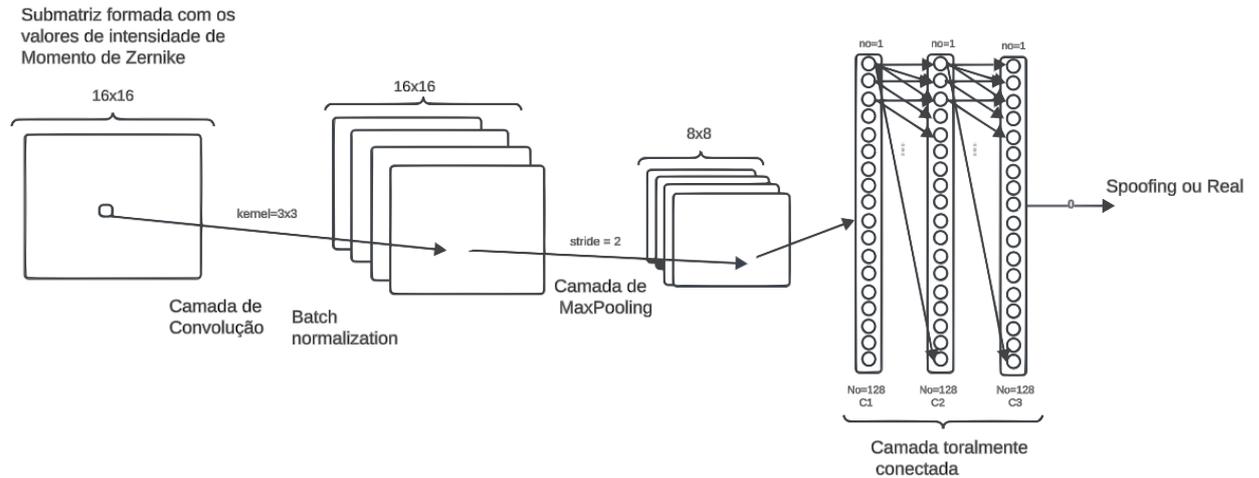


Figura 6.3: Classificador CNN do sistema proposto.

no treinamento. Por fim, as métricas são calculadas e armazenadas para estudo. A seguir, apresentamos os resultados de teste da CNN com base de dados de 2000 e 6000, *non-overlapping* de  $4 \times 4$ ,  $8 \times 8$  e  $16 \times 16$ , e momento de Zernike de ordem de 5,9 e 12. O nível de acurácia, precisão, *recall* e *F1-score* do sistema apresenta melhoras com uma rede neural mais adequada ao padrão de imagens.

Tabela 6.5: Sistema de momentos de Zernike com CNN onde os parâmetro de ordem de momentos de Zernike igual à  $n = 5$

Base	Qtd	Janela	Acurácia	Precisão	Recall	F1-Score
CASIA	2000	4x4	0,73	0,48	0,79	0,60
		8x8	0,64	0,40	0,81	0,53
		16x16	0,67	0,41	0,73	0,53
	6000	4x4	0,84	0,68	0,70	0,69
		8x8	0,81	0,60	0,71	0,65
		16x16	0,83	0,63	0,81	0,71
IITD	2000	4x4	0,86	0,70	0,79	0,75
		8x8	0,59	0,36	0,84	0,51
		16x16	0,57	0,34	0,79	0,48
	6000	4x4	0,94	0,86	0,93	0,89
		8x8	0,87	0,66	0,94	0,78
		16x16	0,82	0,66	0,53	0,59

Apesar da variação da quantidade de elementos na entrada, notou-se que pra baixas ordens de Momento de Zernike (0 à 5) o resultado depende muito da imagem conter detalhes o suficiente para distinguir as formas e contornos, possivelmente por não extrair tantas características adequadamente da imagem. Destaque para a ordem zero, onde o modelo repassa à saída com produto de 1, sem ganho ou perdas,

Tabela 6.6: Sistema de momentos de Zernike com CNN onde os parâmetro de ordem de momentos de Zernike igual à  $n = 9$

Base	Qtd	Janela	Acurácia	Precisão	Recall	F1-Score
CASIA	2000	4x4	0,81	0,59	0,84	0,69
		8x8	0,72	0,47	0,87	0,61
		16x16	0,85	0,65	0,89	0,75
	6000	4x4	0,92	0,82	0,87	0,84
		8x8	0,87	0,72	0,78	0,75
		16x16	0,89	0,73	0,89	0,80
IITD	2000	4x4	0,86	0,69	0,81	0,74
		8x8	0,74	0,49	0,79	0,61
		16x16	0,64	0,39	0,75	0,51
	6000	4x4	0,90	0,83	0,75	0,79
		8x8	0,81	0,59	0,71	0,65
		16x16	0,88	0,70	0,89	0,78

Tabela 6.7: Sistema de momentos de Zernike com CNN onde os parâmetro de ordem de momentos de Zernike igual à  $n = 12$

Base	Qtd	Janela	Acurácia	Precisão	Recall	F1-Score
CASIA	2000	4x4	0,87	0,67	0,92	0,78
		8x8	0,91	0,76	0,92	0,83
		16x16	0,89	0,72	0,92	0,81
	6000	4x4	0,95	0,86	0,95	0,90
		8x8	0,93	0,82	0,93	0,87
		16x16	0,94	0,84	0,94	0,88
IITD	2000	4x4	0,63	0,38	0,75	0,51
		8x8	0,78	0,53	0,86	0,66
		16x16	0,59	0,36	0,78	0,49
	6000	4x4	0,95	0,85	0,95	0,90
		8x8	0,78	0,56	0,64	0,60
		16x16	0,90	0,79	0,81	0,80

somente o classificador após a janela de *non-overlapping*.

### 6.4.3 Momento de Zernike com ResNet

O classificador ResNet tem como configuração os seguintes valores: na entrada tem-se uma matriz de  $16 \times 16$ . Bloco da camada de convolução tem *kernel* de  $3 \times 3$ , *padding* igual à 1 e *strides* igual à 1. Em seguida, utilizou-se *Batch normalization* padrão. No bloco de *max pooling*  $2 \times 2$  e *stride* igual à 2.

A estrutura residual é formada por um bloco de camada de convoluções com *kernel* igual à  $3 \times 3$ , *padding* igual à 1 e *strides* igual à 1; em seguida, um bloco

Tabela 6.8: Sistema proposto onde a ordem de momentos de Zernike varia em números inteiros de  $n = 0$  à  $n = 12$ , com base de dados pré-processada CASIA, janela *non-overlapping* de  $16 \times 16$ , e classificador CNN conforme descrito em 6.4.2

Ordem ZM	Classificador	Acuracia	Precisao	Recall	F1-Score
0	CNN	0,75	0,50	0,57	0,53
1		0,77	0,53	0,61	0,57
2		0,80	0,57	0,72	0,64
3		0,81	0,59	0,76	0,67
4		0,82	0,61	0,76	0,68
5		0,83	0,63	0,81	0,71
6		0,86	0,67	0,85	0,75
7		0,87	0,70	0,87	0,78
8		0,89	0,73	0,88	0,80
9		0,89	0,73	0,89	0,80
10		0,90	0,76	0,89	0,82
11		0,91	0,78	0,91	0,84
12		0,94	0,84	0,94	0,88

*Batch normalization*, um bloco de função de ativação ReLU, e por fim, nesse bloco residual, tem um bloco de *max pooling* ( $2 \times 2$ ), o *stride* de 2. Essa estrutura residual se repete. No bloco de saída, a função de ativação ReLU. As configurações globais da rede ResNet são *k-fold* igual à 4 ou 10, com taxa de aprendizado igual à 0.01, o *batch size* igual à 32.

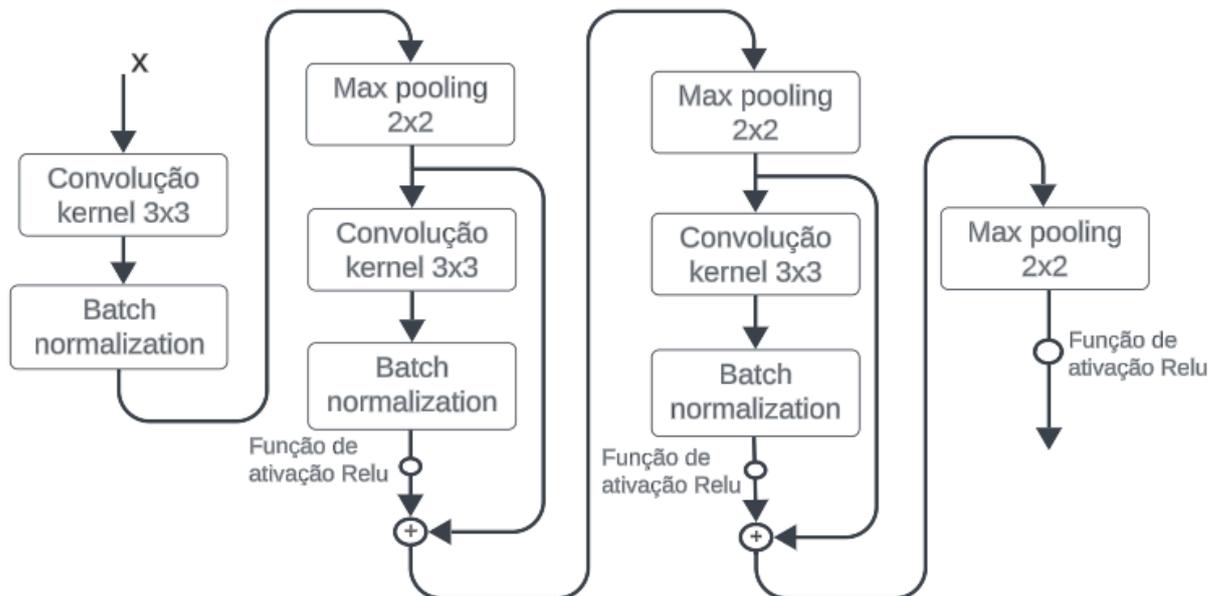


Figura 6.4: Classificador ResNet do sistema proposto.

O resultado dos testes da ResNet com a base de dados 2000 e 6000, janela de

*non-overlapping*  $4 \times 4$ ,  $8 \times 8$  e  $16 \times 16$  e momento de Zernike (ZM) de ordem 5, 9 e 12. Com o acréscimo de mais estruturas convolutivas e as técnicas de redes residuais, resultaram em um aumento da acurácia. O sistema teve dificuldades ao processar com submatrizes  $4 \times 4$ , pelo tamanho da matriz não mostrar um padrão na vizinhança de pixels, tornando menos distinguíveis as características de formas circulares na imagem e, nesse contexto, os melhores resultados foram para as submatrizes de  $16 \times 16$ .

Tabela 6.9: Sistema de momentos de Zernike com ResNet onde os parâmetro de ordem de momentos de Zernike igual à  $n = 5$

Base	Qtd	Janela	Acurácia	Precisão	Recall	F1-Score
CASIA	2000	4x4	0,74	0,49	0,79	0,61
		8x8	0,74	0,49	0,86	0,62
		16x16	0,77	0,53	0,79	0,63
	6000	4x4	0,79	0,58	0,60	0,59
		8x8	0,87	0,72	0,81	0,76
		16x16	0,85	0,67	0,80	0,73
IITD	2000	4x4	0,70	0,45	0,84	0,59
		8x8	0,67	0,43	0,90	0,58
		16x16	0,68	0,43	0,83	0,57
	6000	4x4	0,81	0,60	0,74	0,67
		8x8	0,79	0,57	0,58	0,58
		16x16	0,83	0,70	0,55	0,61

Tabela 6.10: Sistema de momentos de Zernike com ResNet onde os parâmetro de ordem de momentos de Zernike igual à  $n = 9$

Base	Qtd	Janela	Acurácia	Precisão	Recall	F1-Score
CASIA	2000	4x4	0,86	0,66	0,90	0,76
		8x8	0,78	0,53	0,87	0,66
		16x16	0,83	0,62	0,84	0,72
	6000	4x4	0,84	0,66	0,73	0,70
		8x8	0,92	0,81	0,89	0,85
		16x16	0,90	0,75	0,87	0,81
IITD	2000	4x4	0,66	0,41	0,81	0,55
		8x8	0,76	0,52	0,84	0,64
		16x16	0,60	0,37	0,87	0,52
	6000	4x4	0,77	0,55	0,55	0,55
		8x8	0,80	0,58	0,74	0,65
		16x16	0,79	0,57	0,68	0,62

Para casos de ordem de momento de Zernike iguais à 6 à 9, o modelo tem resultados mais assertivos. Mesmo assim, os melhores resultados foram para ordens

Tabela 6.11: Sistema de momentos de Zernike com ResNet onde os parâmetro de ordem de momentos de Zernike igual à  $n = 12$

Base	Qtd	Janela	Acurácia	Precisão	Recall	F1-Score
CASIA	2000	4x4	0,91	0,77	0,92	0,84
		8x8	0,92	0,78	0,92	0,85
		16x16	0,91	0,78	0,90	0,84
	6000	4x4	0,91	0,81	0,82	0,81
		8x8	0,95	0,88	0,94	0,90
		16x16	0,93	0,84	0,89	0,87
IITD	2000	4x4	0,89	0,74	0,89	0,81
		8x8	0,82	0,61	0,76	0,68
		16x16	0,80	0,58	0,78	0,66
	6000	4x4	0,86	0,68	0,80	0,74
		8x8	0,81	0,60	0,75	0,66
		16x16	0,93	0,84	0,89	0,86

maiores, de 9 à 12, onde consegue-se extrair mais características. Outra influência foi o tamanho da janela, que tenha mais características distinguíveis de formas circulares e semelhanças de *pixels* na vizinhança.

Tabela 6.12: Sistema proposto onde a ordem de momentos de Zernike varia em números inteiros de  $n = 0$  à  $n = 12$ , com base de dados pré-processada CASIA, janela *non-overlapping* de 16x16, e classificador ResNet conforme descrito em 6.4.3

Ordem ZM	Classificador	Acuracia	Precisao	Recall	F1-Score
0	ResNet	0,74	0,47	0,51	0,49
1		0,75	0,50	0,52	0,51
2		0,77	0,53	0,60	0,56
3		0,80	0,59	0,71	0,64
4		0,83	0,63	0,78	0,69
5		0,85	0,67	0,80	0,73
6		0,86	0,68	0,82	0,74
7		0,86	0,69	0,83	0,75
8		0,87	0,70	0,86	0,77
9		0,90	0,75	0,87	0,81
10		0,91	0,79	0,89	0,84
11		0,92	0,82	0,89	0,85
12		0,93	0,84	0,89	0,87

#### 6.4.4 Análise de resultados

Nas métricas, os resultados variam na detecção de *spoofing* de íris, nos tamanhos da base de dados, cada mudança processada separadamente, mas usando a

mesma estrutura do sistema proposto. Observe que os testes a seguir são realizados no conjunto de testes, relatando os resultados referentes à detecção de imagem *spoofing*. Abaixo mostram-se os resultados para MLP com base de dados de 2000 e 6000 imagens integrantes, janelas de *non-overlapping* de dimensões  $4 \times 4$ ,  $8 \times 8$  e  $16 \times 16$ , momento de Zernike de ordem 10 e classificador MLP, CNN e ResNet. Na tabela da seção 6.4, apresentamos o resultado da MLP, CNN e ResNet, variando a ordem de momento de Zernike em graus de baixa, média e alta ordem no conjunto de  $n = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ . Por fim, as métricas são calculadas e armazenadas para estudo e apresentadas a seguir.

Abaixo apresenta-se a Tabela 6.13, uma comparação dos melhores resultados de acurácia entre o sistema de referência e o sistema proposto em seus três classificadores. Em todos os casos de comparação, o sistema apresentou resultados similares para ambas as bases de dados compostas por IITD e por CASIA. Com relação à variação da janela de *non-overlapping*, tivemos melhores resultados para valores de janelas  $16 \times 16$ , esperando uma área o suficiente para abranger melhor as formas circulares existentes na imagem quando comparadas à janela de  $4 \times 4$ . Com relação à variação de momentos de Zernike, os valores foram crescentes junto com a ordem do polinômio, com os melhores resultados para momentos de Zernike de ordem  $n = 12$ , isso devido à quantidade de características extraídas ser maior nestes. É observada uma saturação de valores em momentos de alta ordem e janelas de *non-overlapping*  $4 \times 4$ . Na comparação entre os valores do sistema de referência e o sistema de momentos de Zernike com MLP, temos o melhor resultado do sistema de referência, sendo que o mesmo tem mais informação extraída da imagem devido à técnica de LBPV, detalhes de luminância que se mostraram importantes na imagem. Quando comparados o sistema de referência e o sistema de momentos de Zernike com CNN, a técnica do sistema proposto foi mais promissora, em partes pela técnica de classificação CNN extrair mais características das imagens, mesmo esta sendo modificada no processamento de momentos de Zernike, apresentando na entrada do classificador uma imagem com características circulares com valores de *pixel* correlacionados por representar essas características. O sistema de momentos de Zernike com a rede residual ResNet apresenta os melhores resultados quando comparados ao sistema de referência. O sistema proposto usando o classificador

ResNet apresentou mais informações de características da imagem, o que era esperado pelos conceitos da rede ResNet, preservando a informação de entrada da rede e unindo-a com informação extraída pela recorrência de várias redes convolucionais. Abaixo é apresentada a Tabela 6.13 com o melhor valor do sistema de referência e das três variantes do sistema proposto.

Tabela 6.13: Comparação de resultado de métricas entre sistema de referência e sistema proposto.

Autor	Algoritmo de Classificação	Acurácia (%)
Kohli, N., Yadav, D., Vatsa, M., et al.	DESIST Framework Sistema de referência	82.20
Colares, G.S.	ZM + MLP	79
	ZM+ CNN	90
	ZM + Resnet	93

Além disso, foram desenvolvidos classificadores que fazem o reconhecimento de características de formas extraídas das imagens de íris pertencentes aos grupos de base de dados IIT Delhi Iris e CASIA Iris, de modo a determinar características de ataque de *spoofing*, com base na informação observada nas imagens usando técnicas mais apropriadas ao problema. A janela de *non-overlapping*  $P \times P$  tem os valores de  $4 \times 4$ ,  $8 \times 8$  e  $16 \times 16$ . Na proposta, utilizou-se para referência no estudo o cenário com as técnicas combinadas com momentos de Zernike, e também isoladamente sem momentos de Zernike.

No Classificador, a entrada de características será classificada utilizando uma rede neural MLP de quatro camadas, sendo duas camadas ocultas, e usa algoritmo de *backpropagation*. Na rede neural MLP, foi aplicado *cross-validation* de *k-fold* igual à 5 executado duas vezes. Na saída tem-se a decisão da rede neural, determinando se a imagem da íris é *spoofing* ou real. A saída é uma decisão da rede neural determinando se a íris é falsificada ou não. O sistema descrito acima é apresentado em sua simplificação no diagrama em blocos do sistema referência na Figura 4.1 do artigo. Nas métricas, em todas as etapas de treinamento, validação e testes, faz-se a avaliação do desempenho com a acurácia e a matriz de confusão.

Na saída tem-se a decisão da rede neural, determinando se a imagem da íris é *spoofing* ou real. Nas métricas, em todas as etapas de treinamento, validação e

testes, faz-se a avaliação do desempenho com a acurácia, precisão, *recall* e *F1-score* assim avaliados em uma matriz de confusão. Apesar da variação da quantidade de elementos na entrada, notou-se que pra baixas ordens de Momento de Zernike (0 à 5), os resultados foram aleatórios, possivelmente por não extrair tantas características da imagem. Destaque para a ordem zero, onde o modelo repassa à saída com produto de 1, sem ganho ou perda.

Para uma melhor comparação entre os métodos de empregando MLP e o sistema de referência, foi proposto mais o seguinte experimento. Para este experimento, foi utilizado o algoritmo do sistema proposto aplicado aos moldes do sistema de referência, conforme feito no experimento anterior com resultado de 82,2%, e utilizando a base de dados IITD, mas desta vez removendo a técnica do LBPV. O resultado abaixo deste experimento é apresentado na Tabela 6.14 também é organizado na disposição crescente da ordem de momento de Zernike.

Tabela 6.14: Sistema de referência implementado com o algoritmo do sistema proposto e removendo a técnica do LBPV.

<b>Ordem ZM</b>	<b>Classificador</b>	<b>Acuracia</b>	<b>Precisao</b>	<b>Recall</b>	<b>F1-Score</b>
<b>0</b>	<b>MLP</b>	0.64	0.34	0.45	0.39
<b>1</b>		0.67	0.37	0.50	0.43
<b>2</b>		0.70	0.42	0.59	0.49
<b>3</b>		0.71	0.43	0.60	0.51
<b>4</b>		0.73	0.47	0.62	0.54
<b>5</b>		0.77	0.53	0.64	0.58
<b>6</b>		0.78	0.54	0.66	0.60
<b>7</b>		0.78	0.55	0.67	0.60
<b>8</b>		0.79	0.56	0.67	0.61
<b>9</b>		0.79	0.57	0.67	0.62
<b>10</b>		0.80	0.59	0.67	0.63
<b>11</b>		0.81	0.60	0.67	0.63
<b>12</b>		0.81	0.61	0.68	0.64

Destaque para a tabela 6.15, onde as condições distintas entre o sistema de referência e os outros dois sistemas simulados com o algoritmo proposto foram muito atenuadas. Vale ressaltar que o valor extraído no artigo de referência é consequência de um algoritmo com as otimizações empregadas. Outro ponto é o uso do número de repetições de momentos de Zernike, ou seja, uso de menos termos dos momentos de Zernike no sistema de referência. O valor simulado é próximo ao valor obtido no sistema do artigo referência, contudo, pode-se apontar também como motivo as dife-

renças na organização da bases de dados. Com os resultados simulados no sistema de referência e os outros dois simulados, pode-se também destacar a similaridade entre os valores obtidos simulando os sistemas propostos. Comparar o sistema proposto com o sistema de referência sem LBPV é mais justificável para realizar esta paridade. O sistema proposto tem menos informação qualitativa extraída da imagem, sem informação de textura. Em contraponto, o sistema proposto extrair informações quantitativas da imagem, com a utilização de muito mais termos do número de repetições de momentos de Zernike, conseqüentemente há mais características de formas não lineares da imagem .

Tabela 6.15: Comparação entre os três sistemas implementados utilizando MLPe base de dados IITD de 6000 imagens, e com ordem do momento de Zernike igual à 10.

<b>Autor</b>	<b>Algoritmo de Classificação</b>	<b>Acurácia</b>
<b>Kohli, N., Yadav, D., Vatsa, M., et al.</b>	<b>Sistema de Referência</b>	82,2%
<b>Colares, G.S.</b>	<b>Sistema Proposto</b>	79,0%
	<b>Sistema de Referência sem LBPV</b>	80,0%

# Capítulo 7

## Conclusão

Nesta dissertação, investigou-se a influência da criação de um sistema acrescentando blocos de pré-processamento previamente, antes de executar a tarefa de detecção de imagens de íris em ataques de *spoofing* do tipo íris sintética. Entre as etapas de pré-processamento, o redimensionamento da imagem e a normalização não se mostraram primordiais no resultado final. Em complemento, investigou-se os parâmetros e variáveis que têm influência na melhoria de momentos de Zernike quando combinados com classificadores de formas circulares ou não alinhadas. O processo de integração das técnicas foi realizado com o sistema concluído para os métodos de momento de Zernike com MLP, momento de Zernike com CNN e momento de Zernike com ResNet.

Para ambas as bases de dados, o sistema proposto de momentos de Zernike com ResNet é o algoritmo com o melhor desempenho em termos de acurácia e precisão. Esta técnica tem a capacidade de melhorar tanto em acurácia, precisão quanto o *recall* em relação às outras abordagens, sugerindo que pode lidar melhor com as características de formas circulares. O momento de Zernike com CNN obteve desempenho em meio termo, mas próximo ao momento de Zernike com ResNet, com bom resultado na precisão em algumas variações, mas em alguns casos perdeu em *recall*. Em qualidade geral das métricas, momentos de Zernike com CNN teve melhores resultados quando comparado aos momentos de Zernike com MLP.

A base de dados IITD apresentou resultados um pouco mais baixos em todos os algoritmos, tendo em vista o cenário mais desafiador para a detecção por apresentar oclusão, uma mistura de imagens. Isso pode estar relacionado com as

imagens selecionadas para os experimentos que, por se enquadrarem na categoria "desafiadora", proporcionaram um ambiente de teste com imagens mais complexas ou *spoofing* menos visíveis.

A comparação direta entre sistema de referência e sistema proposto usando classificador MLP. O desempenho da rede MLP com momentos de Zernike obteve performance pouco inferior ao sistema de referência, especialmente em cenários em que o momento de Zernike teve baixíssimas ordens do polinômio, sendo o valor mais próximo para ordem de  $n$  igual à 12 com acurácia de 79% em comparação aos 82,20% de acurácia do sistema de referência. A viabilização de uma abordagem com classificadores mais dedicados ao processamento de imagens se mostrou mais assertiva. Foi o caso em que utilizou-se CNN e ResNet na classificação, obtendo, respectivamente, valores de 90% e 93% de acurácia.

## 7.1 Proposta para Trabalhos Futuros

Nesta seção, apresentam-se algumas propostas para trabalhos futuros relacionados ao estudo de caso descrito nesta dissertação. As sugestões estão elencadas a seguir:

- Além das bases IITD e CASIA, que já apresentam desafios distintos, explorar outras bases com características específicas de imagens geométricas que seja possível contribuir para validar a robustez das técnicas em relação a classificadores de formas para cenários complexos;
- Implementar técnica de rede neural artificial GAN (do inglês, *generative adversarial network*) como classificador no sistema proposto, fazendo os devidos ajustes necessários na rede generativa. Realizar um estudo onde a rede discriminadora terá como entrada características das imagens de iris reais e de imagens geradas como ataque *spoofing* de iris sintética entregues pela própria rede generativa;
- Realizar estudo no sistema proposto variando a quantidade de termos dos números de repetições de momentos de Zernike à serem utilizados na classificação. Avaliar a influência do termo de repetições de momentos de Zernike na

acurácia final obtida após classificar em imagem de íris real ou imagem de íris de ataque de *spoofing*;

- Implementar utilizando técnicas disponíveis em API's com uma gama de redes neurais artificiais otimizadas, de modo à obter um estudo para valores ótimos de classificadores para o sistema proposto;
- Explorar o desenvolvimento de momentos de Zernike aplicados à janelamento com sobreposição, no entendimento que as informações de características circulares ou não alinhadas estarão contidas na janela de dimensão  $P \times P$  e podem ter continuidade de informação nas janelas adjacentes.
- Realizar estudo no sistema proposto do janelamento das imagens para momentos de Zernike para valores variando para valores superiores à  $16 \times 16$ . Avaliar a influência da dimensão da janela para valores extremos. A variação tem como limite quando tiver a mesma dimensão da imagem, se tornando uma extração de características globais;

# Referências Bibliográficas

- [1] ALMEIDA, H., LOBO, S., LOBÃO, P., et al., “Preparações oftálmicas - novas estratégias para a promoção da biodisponibilidade/ Ophthalmic preparations - new strategies to promote the bioavailability”, *Acta Farmacêutica Portuguesa*, v. 10, pp. 44–83, 07 2021.
- [2] CRASS, J., *The Adaptive Optics Lucky Imager: combining adaptive optics and lucky imaging*, Ph.D. Thesis, University of Notre Dame, 07 2014.
- [3] UPNEJA, R., PAWLAK, M., SAHAN, A. M., “An accurate approach for the computation of polar harmonic transforms”, *Optik*, v. 158, pp. 623–633, 2018.
- [4] KOHLI, N., YADAV, D., VATSA, M., et al., “Detecting medley of iris spoofing attacks using DESIST”. In: *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pp. 1–6, 2016.
- [5] MANZOOR, S. I., SELWAL, A., “An Analysis of Biometric Based Security Systems”. In: *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, pp. 306–311, 2018.
- [6] CHONG, C.-W., RAVEENDRAN, P., MUKUNDAN, R., “A comparative analysis of algorithms for fast computation of Zernike moments”, *Pattern Recognition*, v. 36, pp. 731–742, 03 2003.
- [7] SINGH, C., SINGH, J., “A Survey on Rotation Invariance of Orthogonal Moments and Transforms”, *Signal Processing*, v. 185, pp. 108086, 03 2021.
- [8] DAUGMAN, J., “How Iris Recognition Works”, *Circuits and Systems for Video Technology, IEEE Transactions on*, v. 14, pp. 21 – 30, 02 2004.

- [9] KIM, W.-Y., KIM, Y.-S., “Region-based shape descriptor using Zernike moments”, *Signal Processing: Image Communication*, v. 16, pp. 95–102, 09 2000.
- [10] TEAGUE, M. R., “Image analysis via the general theory of moments\*”, *J. Opt. Soc. Am.*, v. 70, n. 8, pp. 920–930, Aug 1980.
- [11] SARKAR, M., “Modular pattern classifiers: a brief survey”. In: *Smc 2000 conference proceedings. 2000 ieee international conference on systems, man and cybernetics. 'cybernetics evolving to systems, humans, organizations, and their complex interactions' (cat. no.0*, v. 4, pp. 2878–2883 vol.4, 2000.
- [12] ZHANG, J. M., HARMAN, M., MA, L., et al., “Machine Learning Testing: Survey, Landscapes and Horizons”, *IEEE Transactions on Software Engineering*, v. 48, n. 1, pp. 1–36, 2022.
- [13] SUN, S., CAO, Z., ZHU, H., et al., “A Survey of Optimization Methods From a Machine Learning Perspective”, *IEEE Transactions on Cybernetics*, v. 50, n. 8, pp. 3668–3681, 2020.
- [14] SHELKE, R., BAGAL, S. B., “Iris Recognition System: A Novel Approach for Biometric Authentication”. In: *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, pp. 1–5, 2017.
- [15] WANG, C., WANG, X., XIA, Z., et al., “Image Description With Polar Harmonic Fourier Moments”, *IEEE Transactions on Circuits and Systems for Video Technology*, v. 30, n. 12, pp. 4440–4452, 2020.
- [16] KHADE, S., GITE, S., THEPADE, S., et al., “Detection of Iris Presentation Attacks Using Hybridization of Discrete Cosine Transform and Haar Transform With Machine Learning Classifiers and Ensembles”, *IEEE Access*, v. PP, pp. 1–1, 12 2021.
- [17] RAO, K. R., YIP, P., *Discrete cosine transform: algorithms, advantages, applications*. Academic press, 2014.

- [18] ASTOLA, J., YAROSLAVSKY, L., *Advances in signal transforms: theory and applications*. v. 7. Hindawi Publishing Corporation, 2007.
- [19] AGGARWAL, C., *Neural Networks and Deep Learning: A Textbook*. Springer International Publishing, 2018.
- [20] GUPTA, M., GUPTA, S., “Classification of Gliomas Using Efficient Zernike Moments Based Shape Descriptors Extracted from Segmented MR Images”. In: *Soft Computing and Signal Processing*, pp. 445–454, Springer Singapore: Singapore, 2021.
- [21] DRONKY, M. R., KHALIFA, W., ROUSHDY, M., “Using residual images with BSIF for iris liveness detection”, *Expert Systems with Applications*, v. 182, pp. 115266, 2021.
- [22] PIETIKÄINEN, M., HADID, A., ZHAO, G., et al., “Local Binary Patterns for Still Images”, In: *Computer Vision Using Local Binary Patterns*, pp. 13–47, Springer London: London, 2011.
- [23] JUNIOR, W., DA SILVA, E., GOLDENSTEIN, S., *Reconhecimento de Padrões utilizando Filtros de Correlação com Análise de Componentes Principais*, Ph.D. Thesis, Universidade Federal do Rio de Janeiro - UFRJ, 08 2010.
- [24] S, C., G S, S., R, N., et al., “Analysis of Image Quality using Sobel Filter”. In: *Conference: 2019 Third International Conference on Inventive Systems and Control (ICISC)*, pp. 526–531, 01 2019.
- [25] KAUR, B., “Face Spoofing Detection System using Local Invariant Feature Set”. In: *IEEE Delhi Section Flagship Conference (DELCON)*, pp. 1–5, 2023.
- [26] AVERBUCH, A., COIFMAN, R., DONOHO, D., et al., “Fast and accurate Polar Fourier transform”, *Applied and Computational Harmonic Analysis*, v. 21, n. 2, pp. 145–167, 2006.
- [27] MUKUNDAN, R., ONG, S., LEE, P., “Image analysis by Tchebichef moments”, *IEEE Transactions on Image Processing*, v. 10, n. 9, pp. 1357–1364, 2001.

- [28] YAP, P.-T., PARAMESRAN, R., ONG, S.-H., “Image analysis by Krawtchouk moments”, *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, v. 12, pp. 1367–77, 02 2003.
- [29] ZHU, H., SHU, H., ZHOU, J., et al., “Image analysis by discrete orthogonal dual Hahn moments”, *Pattern Recognition Letters*, v. 28, pp. 1688–1704, 10 2007.
- [30] GAUTAM, G., RAJ, A., MUKHOPADHYAY, S., “Deep supervised class encoding for iris presentation attack detection”, *Digital Signal Processing*, v. 121, pp. 103329, 2022.
- [31] HUSSEIS, A., LIU-JIMENEZ, J., GOICOCHEA-TELLERIA, I., et al., “A Survey in Presentation Attack and Presentation Attack Detection”. In: *2019 International Carnahan Conference on Security Technology (IC-CST)*, pp. 1–13, 2019.
- [32] KHADE, S., GITE, S., THEPADE, S. D., et al., “Detection of Iris Presentation Attacks Using Feature Fusion of Thepade’s Sorted Block Truncation Coding with Gray-Level Co-Occurrence Matrix Features”, *Sensors*, v. 21, n. 21, 2021.
- [33] FANG, M., DAMER, N., BOUTROS, F., et al., “Iris Presentation Attack Detection by Attention-based and Deep Pixel-wise Binary Supervision Network”. In: *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pp. 1–8, 2021.
- [34] LI, Z., LIU, F., YANG, W., et al., “A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects”, *IEEE Transactions on Neural Networks and Learning Systems*, v. 33, n. 12, pp. 6999–7019, 2022.
- [35] ALOYSIUS, N., GEETHA, M., “A review on deep convolutional neural networks”. In: *2017 International Conference on Communication and Signal Processing (ICCSP)*, pp. 0588–0592, 2017.

- [36] AGARWAL, A., NOORE, A., VATSA, M., et al., “Generalized Contact Lens Iris Presentation Attack Detection”, *IEEE Transactions on Biometrics, Behavior, and Identity Science*, v. 4, n. 3, pp. 373–385, 2022.
- [37] KUMAR, R. M., SREEKUMAR, K., “A survey on image feature descriptors”, *Int J Comput Sci Inf Technol*, v. 5, n. 6, pp. 7668–7673, 2014.
- [38] GLASSNER, A. S., *Principles of Digital Image Synthesis*. Morgan Kaufmann Publishers Inc., 1994.
- [39] SOLOMON, C., BRECKON, T., *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*. 1st ed. Wiley Publishing, 2011.
- [40] WANG, Q., YUAN, Y., “Learning to resize image”, *Neurocomputing*, v. 131, pp. 357–367, 2014.
- [41] RUIZ-ALBACETE, V., TOME-GONZALEZ, P., ALONSO-FERNANDEZ, F., et al., “Direct Attacks Using Fake Images in Iris Verification”. In: *Biometrics and Identity Management*, pp. 181–190, Springer Berlin Heidelberg: Berlin, Heidelberg, 2008.
- [42] GOICOECHEA-TELLERIA, I., KIYOKAWA, K., JIMENEZ, J., et al., “Low-Cost and Efficient Hardware Solution for Presentation Attack Detection in Fingerprint Biometrics Using Special Lighting Microscopes”, *IEEE Access*, v. PP, pp. 1–1, 01 2019.
- [43] SOLTANE, M., BAKHTI, M., “Multi-modal biometric authentications: concept issues and applications strategies”, *International Journal of Advanced Science and Technology*, v. 48, 2012.
- [44] GORODNICHY, D. O., “Evolution and evaluation of biometric systems”. In: *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pp. 1–8, 2009.
- [45] YADAV, D., KOHLI, N., DOYLE, J. S., et al., “Unraveling the Effect of Textured Contact Lenses on Iris Recognition”, *IEEE Transactions on Information Forensics and Security*, v. 9, n. 5, pp. 851–862, 2014.

- [46] GUPTA, P., BEHERA, S., VATSA, M., et al., “On Iris Spoofing Using Print Attack”. In: *2014 22nd International Conference on Pattern Recognition*, pp. 1681–1686, 2014.
- [47] GONZALEZ, R., WOODS, R., *Digital Image Processing*. Pearson, 2018.
- [48] ROLI, F., DIDACI, L., MARCIALIS, G. L., “Adaptive Biometric Systems That Can Improve with Use”, In: *Advances in Biometrics: Sensors, Algorithms and Systems*, Springer London: London, 2008.
- [49] BARBU, T., CIOBANU, A., LUCA, M., “Multimodal biometric authentication based on voice, face and iris”. In: *2015 E-Health and Bioengineering Conference (EHB)*, pp. 1–4, 2015.
- [50] RUIZ-ALBACETE, V., TOME, P., ALONSO-FERNANDEZ, F., et al., “Direct Attacks Using Fake Images in Iris Verification”. In: *Signal Processing: Image Communication*, pp. 181–190, 05 2008.
- [51] LIU, C., WECHSLER, H., “A shape- and texture-based enhanced Fisher classifier for face recognition”, *IEEE Transactions on Image Processing*, v. 10, n. 4, pp. 598–608, April 2001.
- [52] SADKHAN, S. B., AL-SHUKUR, B. K., MATTAR, A. K., “Human voice extracted biometric features: What can be used for”. In: *2017 International Conference on Current Research in Computer Science and Information Technology (ICCIT)*, pp. 7–12, 2017.
- [53] ISO/IEC JTC1 SC37, “Iso/iec 30107-1:2023 information technology - biometric presentation attack detection - part 1: Framework. 2023.” Available at <https://www.iso.org/standard/83828.html>, 2023.
- [54] CHATTERJEE, A., *Biometric Presentation Attack Detection. 2053-2563*, IOP Publishing, 2023.
- [55] RICE, J., *Mathematical Statistics and Data Analysis. Advanced series*, Cengage Learning, 2007.

- [56] HUANG, Z., LENG, J., “Analysis of Hu’s moment invariants on image scaling and rotation”. In: *2010 2nd International Conference on Computer Engineering and Technology*, v. 7, pp. V7-476–V7-480, 2010.
- [57] GUO, L., LIU, L., SUN, H., “Focus Measure Based on the Image Moments”. In: *2018 IEEE International Conference on Mechatronics and Automation (ICMA)*, pp. 1151–1156, 2018.
- [58] ALGHAMDI, W., CALMON, F. P., “Mutual Information as a Function of Moments”. In: *2019 IEEE International Symposium on Information Theory (ISIT)*, pp. 3122–3126, 2019.
- [59] KIM, H.-K., KIM, J.-D., SIM, D.-G., et al., “A modified Zernike moment shape descriptor invariant to translation, rotation and scale for similarity-based image retrieval”. In: *2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No.00TH8532)*, v. 1, pp. 307–310 vol.1, 2000.
- [60] HU, R., XIANG, S., “Cover-Lossless Robust Image Watermarking Against Geometric Deformations”, *IEEE Transactions on Image Processing*, v. 30, pp. 318–331, 2021.
- [61] KOSAKI, R., MAEDA, N., BESSHO, K., et al., “Magnitude and Orientation of Zernike Terms in Patients with Keratoconus”, *Investigative ophthalmology & visual science*, v. 48, pp. 3062–8, 07 2007.
- [62] MOLNAR, C., *Interpretable Machine Learning*.  
<https://christophm.github.io/interpretable-ml-book/>, 2018,  
<https://christophm.github.io/interpretable-ml-book/>.
- [63] CRESWELL, A., WHITE, T., DUMOULIN, V., et al., “Generative Adversarial Networks: An Overview”, *IEEE Signal Processing Magazine*, v. 35, n. 1, pp. 53–65, 2018.
- [64] SINGH, J., BANERJEE, R., “A Study on Single and Multi-layer Perceptron Neural Network”. In: *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 35–40, 2019.

- [65] LECUN, Y., BOTTOU, L., BENGIO, Y., et al., “Gradient-based learning applied to document recognition”, *Proceedings of the IEEE*, v. 86, n. 11, pp. 2278–2324, 1998.
- [66] HE, K., ZHANG, X., REN, S., et al., “Deep Residual Learning for Image Recognition”. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.
- [67] ADRIAN, R., “Practical Python and OpenCV: An Introductory, Example Driven Guide to Image Processing and Computer Vision”, 2016.

# Apêndice A

## Artigos Publicados

### A.1 Artigos publicados em colaboração

- Detecção de Placas Veiculares em Ambientes Aéreos utilizando super Resolução por GAN e Equalização de Histograma, October 2023, DOI: 10.14209/sbrt.2023.1570908005, XLI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, São Paulo, Brasil, José Elislande Breno de Souza Linhares, Waldir S. S. Junior, Frederico Pinagé, Georgio Colares, et al.
- Reconhecimento de Placas Veiculares em Cenários Complexos utilizando o Método do Subespaço Mútuo e Redes Neurais Convolucionais, October 2022, DOI: 10.14209/sbrt.2022.1570813829, XL Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, Santa Rita do Sapucaí, MG, Brasil, Anderson S Jesus, Waldir S. S. Junior, Newton Pinheiro Bitar, Bernardo Bentes Gatto, Georgio Colares, et al.

# Apêndice B

## Fundamentos de imagem

### B.1 Representações da imagem digital

Uma imagem digital  $f(x, y)$  encontra-se no domínio espacial, e as coordenadas  $x$  e  $y$  são suas variáveis espaciais, onde  $x$  varia em valores inteiros de  $\{0, 1, \dots, M-1\}$ , e  $y$  varia também em valores inteiros de  $\{0, 1, \dots, N-1\}$ . A imagem digital é representada por uma matriz  $M \times N$  [47]. Há um valor inteiro denominado em níveis de cinza, correspondente a  $f(x, y)$  no domínio espacial de coordenadas  $(x, y)$ . A imagem é composta por um ou mais canais. Imagens com três canais são denominadas imagens coloridas e são divididas no espaço de cores RGB [47].

O menor detalhe discernível em uma imagem é denominado pixel [38]. A resolução espacial é obtida de várias maneiras na imagem digital, e a mais comum relaciona a quantidade de *pixels* da imagem, ou os pontos por unidade de distância da imagem [47]. A menor mudança discernível no nível de intensidade é definida como resolução de intensidade. O número de *pixels* e intensidades é utilizado para gerar e representar uma imagem digital.

### B.2 Componentes de imagens digitais

Dentre as informações que compõem a imagem digital estão a luminância, o brilho, a cromaticidade e o contraste [47]. A luminância é a intensidade de luminosidade em uma imagem, enquanto brilho é a luminância percebida com relação à área de sua vizinhança. A cromaticidade é informação multiespectral de frequência que é

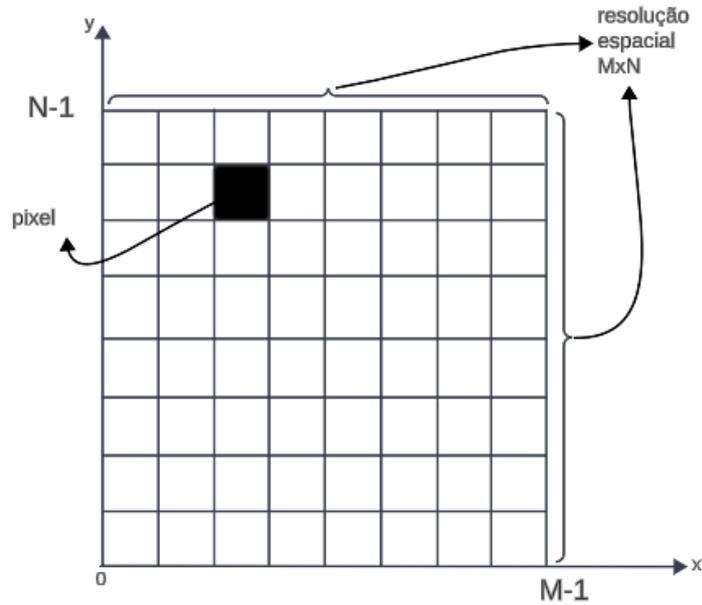


Figura B.1: Representação de uma imagem digital  $f(x, y)$  no domínio espacial.

adquirida no domínio visível. E o contraste é a diferença de níveis de cores ou tons de cinza existentes entre os recursos da imagem [47]. Existem vários modelos de representação de cores de imagens digitais, como exemplo de composição base está o RGB (vermelho, verde e azul), CMY (ciano, magenta, amarelo) e HSV (matiz, saturação e valor), que usam vários níveis de intensidade e filtragem para captura de níveis de brilho para cada uma das cores. Brilho e contraste são medidas descritivas da aparência de uma imagem; as medidas de intensidade podem ser consideradas como componentes simples de caracterização de textura [38].