

William Divino Ferreira

**Deteccção de Imagens Falsificadas Baseada em
Descritores Locais de Textura e Rede Neural
Convolutacional**

Goiânia

Junho/2020



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

[] Dissertação [X] Tese

2. Nome completo do autor

William Divino Ferreira

3. Título do trabalho

“Detecção de Imagens Falsificadas Baseada em Descritores Locais de Textura e Rede Neural Convolutacional”

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento [X] SIM [] NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

a) consulta ao(à) autor(a) e ao(à) orientador(a);

b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.

O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Gelson Da Cruz Junior, Professor do Magistério Superior**, em 25/08/2020, às 09:32, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **WILLIAM DIVINO FERREIRA, Discente**, em 25/08/2020, às 09:37, conforme horário oficial de Brasília, 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.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1507879** e o código CRC **6AB7DA34**.

Referência: Processo nº 23070.035688/2020-72

SEI nº 1507879

William Divino Ferreira

**Detecção de Imagens Falsificadas Baseada em
Descritores Locais de Textura e Rede Neural
Convolucional**

Universidade Federal de Goiás – UFG

Escola de Engenharia Elétrica, Mecânica e de Computação

Programa de Pós-Graduação em Engenharia Elétrica e de Computação

Orientador: Dr. Gelson da Cruz Júnior

Coorientador: Dr. Fabrízio A. A. de M. N. Soares

Goiânia

Junho/2020

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Ferreira, William Divino

Detecção de Imagens Falsificadas Baseada em Descritores Locais de Textura e Rede Neural Convolutiva [manuscrito] / William Divino Ferreira. - 2020.

CLI, 151 f.: il.

Orientador: Prof. Gelson da Cruz Júnior; co-orientador Fabrizzio Alphonsus Alves de Melo Nunes Soares.

Tese (Doutorado) - Universidade Federal de Goiás, Escola de Engenharia Elétrica, Mecânica e de Computação (EMC), Programa de Pós-Graduação em Engenharia Elétrica e de Computação, Cidade de Goiás, 2020.

Bibliografia.

Inclui lista de figuras, lista de tabelas.

1. Descritores Locais de Textura. 2. Detecção de Imagens Falsificadas. 3. Forense em Imagens Digitais. 4. Redes Neurais Convolutivas. I. Júnior, Gelson da Cruz, orient. II. Título.

CDU 004



UNIVERSIDADE FEDERAL DE GOIÁS

ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

ATA DE DEFESA DE TESE

Ata Nº **04/2020** da sessão de Defesa de Tese de **William Divino Ferreira** que confere o título de Doutor em **Engenharia Elétrica e de Computação** na área de concentração em **Engenharia de Computação**.

Aos **trinta dias do mês de junho de dois mil e vinte**, a partir das **14h00min**, na **Escola de Engenharia Elétrica, Mecânica e de Computação**, realizou-se a sessão pública de Defesa de Tese intitulada “**Deteção de Imagens Falsificadas Baseada em Descritores Locais de Textura e Rede Neural Convolutacional**”. Os trabalhos foram instalados pelo Orientador, Professor Doutor **Gelson da Cruz Júnior (EMC/UFG)**, com a participação dos demais membros da Banca Examinadora: **Professor Doutor Hélio Pedrini (IC/Unicamp)**, membro titular externo; Professor Doutor **Rogério Lopes Salvini (INF/UFG)**, membro titular externo, Professor Doutor **Ronaldo Martins da Costa (INF/UFG)**, membro titular externo; Professor Doutor **Fabrizio Alphonsus Alves de Melo Nunes Soares (INF/UFG)**, membro titular externo; Professor Doutor **Rodrigo Pinto Lemos (EMC/UFG)**, membro titular interno. **A banca ocorreu através de videoconferência**. Durante a arguição os membros da banca **não fizeram** sugestão de alteração do título do **trabalho**. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Tese tendo sido o candidato **aprovado** pelos seus membros. Proclamados os resultados pelo Professor Doutor **Gelson da Cruz Júnior**, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos **trinta dias do mês de junho de dois mil e vinte**.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Rogério Lopes Salvini, Professor do Magistério Superior**, em 30/06/2020, às 17:12, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Fabrizio Alphonsus Alves De Melo Nunes Soares, Professor do Magistério Superior**, em 30/06/2020, às 17:13, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Ronaldo Martins Da Costa, Professor do Magistério Superior**, em 30/06/2020, às 17:14, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Helio Pedrini, Usuário Externo**, em 30/06/2020, às 17:30, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

Documento assinado eletronicamente por **WILLIAM DIVINO FERREIRA, Discente**, em 30/06/2020, às 18:38, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do



[Decreto nº 8.539, de 8 de outubro de 2015.](#)



Documento assinado eletronicamente por **Gelson Da Cruz Junior, Professor do Magistério Superior**, em 01/07/2020, às 12:50, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015.](#)



Documento assinado eletronicamente por **Rodrigo Pinto Lemos, Professor do Magistério Superior**, em 01/07/2020, às 16:35, conforme horário oficial de Brasília, 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.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1406555** e o código CRC **11989060**.

Referência: Processo nº 23070.027574/2020-59

SEI nº 1406555

Dedico este trabalho a todos aqueles que, como o grande mestre Darcy Ribeiro, acreditam que “mais vale errar, se arrebatando, do que poupar-se para nada”.

Agradecimentos

A realização deste trabalho só foi possível, primeiramente, em virtude da permissão concedida pelo Grande Arquiteto do Universo, nosso Pai Maior, e depois graças à colaboração, direta ou indireta, de várias pessoas. Manifesto os meus agradecimentos e a minha eterna gratidão a todas elas e, de forma particular:

- à minha querida esposa Cristiane, pela ajuda, amor, companheirismo e compreensão dispensados em todos os momentos;
- aos meus amados filhos, Beatriz e Gabriel, pelo amor, carinho e muitas alegrias, que me deram forças para perseverar;
- ao Professor Fabrízio, por seu empenho, paciência, amizade, compreensão e estímulo durante todas as etapas do desenvolvimento desta tese;
- ao Professor Gélson, por todo apoio e pela amizade consolidada durante o Doutorado;
- aos meus pais, Waldir e Eunice, pelos inestimáveis valores e ensinamentos a mim transmitidos.

“A gente não faz amigos, reconhece-os.”

Vinícius de Moraes

Resumo

A manipulação de imagens digitais tem se tornado muito comum nos dias atuais. Assim, é possível copiar, clonar e redimensionar uma imagem muito facilmente, tornando-se desafiador validar a integridade e a autenticidade da mesma. Além disso, imagens digitais podem ser utilizadas por peritos criminais em suas análises forenses. Neste cenário, uma investigação criminal a partir de imagens digitais se torna algo extramente difícil, pois o uso destas imagens como prova requer garantias de sua legitimidade, sob pena de comprometer o processo legal como um todo. Diante deste contexto, este trabalho desenvolve um modelo de identificação de imagens falsificadas baseado em descritores locais de textura e em rede neural convolucional. Para isto, neste trabalho é realizada, inicialmente, uma avaliação de quatorze descritores locais de textura aplicados sobre cinco bases públicas de imagens de textura, e, então, são selecionados os descritores com melhor eficácia. Logo após, os descritores selecionados são aplicados em quatro bases públicas de identificação de imagens falsificadas para extração de características de textura. Estas características são, então, usadas no treinamento e teste de uma rede neural convolucional residual, com a classificação das imagens em autênticas ou falsificadas ocorrendo através do uso de um classificador de máquina de vetores de suporte. Os resultados obtidos pelo modelo proposto são motivadores, mesmo em bases com um número reduzido de imagens.

Palavras-chave: Descritores locais de textura. Detecção de imagens falsificadas. Forense em imagens digitais. Redes neurais convolucionais.

Abstract

Nowadays, digital image transformation has become a widespread activity. Hence, image copying, cloning, and resizing are easily performed, making it challenging to check image integrity and authenticity. Moreover, a criminal investigation from digital images becomes extremely hard, because using those images as proof demands to ensure its legitimacy, under a risk to implicate the whole legal process. In this sense, this work develops a model for forged images based on local texture descriptors with convolutional neural networks. Henceforth, in this work, firstly, we evaluated fourteen local texture descriptors in five public image texture datasets, and then we selected descriptors with the best efficacy. Second, the selected descriptors are applied to four public datasets to extract texture features from forged and legit images. Finally, those features are used to train a residual convolutional neural network, and then, classifying images as authentic or forged with a Support Vector Machine Classifier. A result of the proposed model provides enthusiasm, mainly when applied to a dataset with a small number of images.

Keywords: Local texture descriptors. Image forgery detection. Digital image forensics. Convolutional neural networks.

Lista de ilustrações

Figura 1 – Exemplo de imagem editada maliciosamente (FRIDRICH, 2009).	20
Figura 2 – Variedade de texturas.	21
Figura 3 – Visões de um gramado em função do ângulo do observador (MIRMEHD; XIE; Jasjit Suri, 2008).	22
Figura 4 – Funcionamento geral de um sistema esteganográfico (AL-ATABY; AL- NAIMA, 2011).	26
Figura 5 – Principais métodos esteganográficos.	27
Figura 6 – Principais métodos de marca d’água.	31
Figura 7 – Métodos baseados em SCI.	33
Figura 8 – Exemplo de <i>copy-move</i> (FRIDRICH, 2009).	35
Figura 9 – Métodos baseados em <i>copy-move</i>	36
Figura 10 – Exemplos de imagens autênticas da base CASIA.	40
Figura 11 – Exemplos de imagens falsificadas da base CASIA.	41
Figura 12 – Exemplos de imagens autênticas da base Columbia.	43
Figura 13 – Exemplos de imagens falsificadas da base Columbia.	44
Figura 14 – Exemplos de imagens autênticas da base CoMoFoD.	45
Figura 15 – Exemplos de imagens falsificadas da base CoMoFoD.	46
Figura 16 – Ilustração de rede neural.	49
Figura 17 – Relacionamento entre inteligência artificial, aprendizagem de máquina, aprendizagem profunda e redes neurais convolucionais.	50
Figura 18 – Ilustração de uma rede neural convolucional.	50
Figura 19 – Ilustração de convolução em uma imagem 2D.	51
Figura 20 – Funções de ativação.	52
Figura 21 – Exemplos de <i>pooling</i> com passo da janela igual a dois.	52
Figura 22 – Exemplo de <i>dropout</i> (SRIVASTAVA et al., 2014).	53
Figura 23 – Linha do tempo para as principais CNNs.	55
Figura 24 – Exemplos de imagens de textura.	59
Figura 25 – Organização das abordagens de extração de características de textura.	60
Figura 26 – Exemplo de cálculo da matriz de coocorrência baseado em (MIRMEHD; XIE; Jasjit Suri, 2008).	60
Figura 27 – Exemplo de cálculo da GLRLM considerando a direção 0° e quatro níveis de cinza (THIBAUT et al., 2013).	61
Figura 28 – Exemplo da aplicação de SVD.	62
Figura 29 – Exemplo da aplicação de transformada de Fourier.	63

Figura 30	– Exemplo da aplicação de transformada <i>wavelet</i>	64
Figura 31	– Organização dos principais descritores de textura da literatura.	68
Figura 32	– Linha do tempo dos principais descritores de textura.	68
Figura 33	– Funcionamento do descritor LBP.	69
Figura 34	– Processo de cálculo do MBP (HAFIANE et al., 2008).	70
Figura 35	– Processo de cálculo do MTP (KHAN et al., 2013).	72
Figura 36	– Ilustração da geração do descritor MTP (KHAN et al., 2013).	72
Figura 37	– Máscaras de Sobel.	74
Figura 38	– Processo de cálculo do WLD (CHEN et al., 2010).	76
Figura 39	– Máscaras de Kirsch em todas as oito direções.	77
Figura 40	– Máscaras de gradiente local.	79
Figura 41	– Processo de cálculo do descritor LGIP (ZHOU; WANG, 2012).	79
Figura 42	– Processo de cálculo do LDTP (Rivera, Castillo e Chae (2015)). (a) Imagem original. (b) Respostas após a aplicação das máscaras de Kirsh. (c) Imagem codificada. (d) Exemplo de valores de intensidade da vizinhança. (e) Respostas para a vizinhança apresentada em (d). (f) Código para a vizinhança apresentada em (d).	81
Figura 43	– Processo de cálculo do LFD (LEI et al., 2011).	84
Figura 44	– Exemplos de descritores locais de textura extraídos de uma imagem.	86
Figura 45	– Exemplos da base de imagens Brodatz.	88
Figura 46	– Exemplo de classe da base de imagens Brodatz.	89
Figura 47	– Exemplos da base de imagens CURET.	90
Figura 48	– Exemplos da base de imagens Kylberg.	91
Figura 49	– Exemplos da base de imagens ALOT.	91
Figura 50	– Exemplos da base de imagens DTD.	92
Figura 51	– Visão geral do modelo proposto.	98
Figura 52	– Visão geral da avaliação dos descritores locais de textura.	98
Figura 53	– Ilustração do processo de transferência de aprendizado.	100
Figura 54	– Acurácia dos descritores locais de textura.	104
Figura 55	– Acurácia média dos descritores locais de textura.	105
Figura 56	– Maiores valores de acurácia obtidos nas bases de imagens.	106
Figura 57	– Exemplos de imagens pertencentes à classe <i>freckled</i> da base DTD.	107
Figura 58	– Imagem original e resultado da composição dos descritores da imagem normalizada para 256×256	109
Figura 59	– Modelo proposto para detecção de imagens falsificadas.	110
Figura 60	– Desempenho das principais CNNs na base ImageNet (MATLABR2020A, 2020).	110
Figura 61	– Desempenho das principais CNNs na base ImageNet (KAWATSU et al., 2017).	111

Figura 62	– Visão geral da matriz de confusão.	114
Figura 63	– Matrizes de confusão para o descritor LTP (TAN; TRIGGS, 2010). . .	116
Figura 64	– Matrizes de confusão para o descritor MTP (KHAN et al., 2013). . . .	117
Figura 65	– Matrizes de confusão para o descritor WLD (CHEN et al., 2010). . . .	118
Figura 66	– Matrizes de confusão para o descritor Combinado.	119
Figura 67	– Eficácia média dos descritores.	121
Figura 68	– Exemplo de classificação de imagens falsas-negativas da base CASIA v1.0. (a) Imagem falsificada (Sp_D_NNN_A_ani0033_ani0034_0285.jpg) formada a partir de (b) e (c). (b) imagem autêntica (Au_ani_0033.jpg). (c) imagen autêntica (Au_ani_0034.jpg).	122
Figura 69	– Exemplo de classificação de imagens falsas-positivas da base CASIA v1.0. (a) Imagem autêntica (Au_cha_0025.jpg). (b) Imagem autêntica (Au_cha_0036.jpg).	123
Figura 70	– Exemplo de classificação de imagens falsas-negativas da base CASIA v2.0. (a) Imagem falsificada (Tp_D_CRN_M_N_nat00019_nat00024_11047.jpg) formada a partir de (b) e (c). (b) imagem autêntica (Au_nat_00019.jpg). (c) imagen autêntica (Au_nat_00024.jpg).	123
Figura 71	– Exemplo de classificação de imagens falsas-positivas da base CASIA v2.0. (a) Imagem autêntica (Au_cha_00070.jpg). (b) Imagem autêntica (Au_cha_00075.jpg).	124
Figura 72	– Exemplo de classificação de imagens falsas-negativas da base Columbia. (a) Imagem falsificada (SP_T_053.bmp) obtida a partir de (b). (b) Imagem autêntica (AU_T_053.bmp).	124
Figura 73	– Exemplo de classificação de imagens falsas-positivas da base Columbia. (a) Imagem autêntica (AU_TS_O_081.bmp). (b) Imagem autêntica (AU_SS_O_047.bmp).	125
Figura 74	– Exemplo de classificação de imagens falsas-negativas da base CoMoFoD. (a) Imagem falsificada (032_F_BC1.png) obtida a partir de (b). (b) Imagem autêntica (032_O_BC1.png).	125
Figura 75	– Exemplo de classificação de imagens falsas-positivas da base CoMo- FoD. (a) Imagem autêntica (090_O_JC7.jpg). (b) Imagem autêntica (028_O_JC9.jpg).	126

Lista de tabelas

Tabela 1	– Definições de manipulação, falsificação e adulteração de imagem (ZHENG; ZHANG; THING, 2019).	25
Tabela 2	– Análise comparativa entre esteganografia e criptografia (JOHRI; MISHRA; DAS, 2016).	26
Tabela 3	– Análise comparativa entre marca d’água e esteganografia (NAGARAJ; VIJAYALAKSHMI; ZAYARAZ, 2013).	33
Tabela 4	– Resumo das características das bases de imagens.	39
Tabela 5	– Características das principais CNNs.	55
Tabela 6	– Trabalhos de detecção de imagens falsificadas baseados <i>deep learning</i> .	94
Tabela 7	– Resumo das características das bases de imagens de textura.	99
Tabela 8	– CNNs utilizadas no experimento.	100
Tabela 9	– Configurações do computador utilizado.	101
Tabela 10	– Resultados para a rede AlexNet.	102
Tabela 11	– Resultados para a rede GoogLeNet.	103
Tabela 12	– Resultados para o descritor LTP (TAN; TRIGGS, 2010).	116
Tabela 13	– Resultados para o descritor MTP (KHAN et al., 2013).	117
Tabela 14	– Resultados para o descritor WLD (CHEN et al., 2010).	118
Tabela 15	– Resultados para o descritor Combinado.	119
Tabela 16	– Eficácia dos descritores locais de textura analisados.	120

Sumário

1	Introdução	19
1.1	Descrição do Problema	19
1.2	Motivação	20
1.3	Hipótese	21
1.4	Objetivos e Contribuições	22
1.5	Estrutura do Texto	23
2	Detecção de Imagens Falsificadas	24
2.1	Introdução	24
2.2	Métodos Ativos	25
2.2.1	Esteganografia	25
2.2.1.1	Domínio da Espacial	27
2.2.1.2	Domínio da Frequência	29
2.2.2	Marca D'Água	30
2.2.2.1	Domínio Espacial	31
2.2.2.2	Domínio da Frequência	31
2.3	Métodos Passivos	33
2.3.1	Identificação da Câmera de Origem	33
2.3.1.1	Não Uniformidade da Foto-resposta	33
2.3.1.2	Extração de Características	34
2.3.1.3	Tabelas de Quantização JPEG	34
2.3.1.4	Aberração Cromática	34
2.3.1.5	Função de Resposta da Câmera	35
2.3.2	<i>Copy-move</i>	35
2.3.2.1	Baseados em Características	36
2.3.2.2	Baseados em Blocos	36
2.3.3	Reamostragem	38
2.4	Bases de Imagens de FID	39
2.4.1	CASIA	39
2.4.2	Columbia	39
2.4.3	CoMoFoD	42
2.5	Considerações Finais do Capítulo	47
3	Redes Neurais Convolucionais	48
3.1	Introdução	48
3.2	Arquitetura	49

3.2.1	Convolução	50
3.2.2	Função de Ativação	50
3.2.3	<i>Pooling</i>	52
3.2.4	Regularização	52
3.2.5	Camada Totalmente Conectada	53
3.3	Evolução Histórica	54
3.4	Considerações Finais do Capítulo	56
4	Análise de Textura e Descritores Locais	57
4.1	Introdução	58
4.2	Análise de Textura	58
4.2.1	Abordagem Estatística	60
4.2.1.1	Matriz de Coocorrência	60
4.2.1.2	Matrizes de Comprimento de Corridas de Cinza	61
4.2.2	Abordagem Baseada em Processamento de Sinais	61
4.2.2.1	Decomposição de Valor Singular	62
4.2.2.2	Transformada de Fourier	62
4.2.2.3	Transformada Wavelet	64
4.2.3	Abordagem Baseada em Modelo Paramétrico	65
4.2.3.1	Campos Aleatórios de Markov	65
4.2.3.2	Fractais	65
4.2.4	Abordagens <i>Deep Learning</i>	66
4.3	Descritores Locais de Textura	67
4.3.1	Padrão Binário Local	67
4.3.2	Padrão Binário da Mediana	69
4.3.3	Padrão Ternário Local	70
4.3.4	Padrão Ternário da Mediana	70
4.3.5	Padrão Direcional do Gradiente	73
4.3.6	Padrão Ternário Local do Gradiente	73
4.3.7	Descritor Local de Weber	74
4.3.8	Codificação Binária Melhorada de Weber	75
4.3.9	Variância do Padrão Direcional Local	77
4.3.10	Padrão de Crescimento do Gradiente Local	78
4.3.11	Padrão de Textura Direcional Local	79
4.3.12	Quantização de Fase Local	82
4.3.13	Descritor de Frequência Local	83
4.3.14	Codificação Binária Monogênica	83
4.3.15	Exemplo Prático de Descritores Locais de Textura	86
4.4	Bases de Imagens de Textura	87
4.4.1	Brodatz	87

4.4.2	CUReT	87
4.4.3	Kylberg	88
4.4.4	ALOT	88
4.4.5	DTD	89
4.5	Considerações Finais do Capítulo	90
5	Trabalhos Relacionados	93
5.1	Introdução	93
5.2	Abordagens <i>Deep Learning</i> para Detecção de Imagens Falsificadas	93
5.3	Considerações Finais do Capítulo	96
6	Modelo Proposto para Detecção de Imagens Falsificadas	97
6.1	Introdução	97
6.2	Extração das Características de Textura	98
6.2.1	Avaliação dos Descritores Locais de Textura	98
6.2.2	Descrição do Experimento de Seleção de Descritores	99
6.2.3	Resultados da Seleção de Descritores	101
6.2.4	Discussão	107
6.2.5	Descritores Selecionados	108
6.3	Classificação	108
6.4	Considerações Finais do Capítulo	112
7	Resultados Experimentais	113
7.1	Introdução	113
7.2	Configurações do Experimento	115
7.3	Resultados	115
7.3.1	Resultados com o Descritor LTP	116
7.3.2	Resultados com o Descritor MTP	117
7.3.3	Resultados com Descritor WLD	118
7.3.4	Resultados com o Descritor Combinado	119
7.3.5	Análise Comparativa dos Descritores	120
7.4	Discussão	121
7.5	Considerações Finais do Capítulo	125
8	Conclusões, Contribuições e Trabalhos Futuros	127
8.1	Conclusões e Contribuições	127
8.2	Trabalhos Futuros	128
	Referências	129

CAPÍTULO 1

Introdução

Sumário

1.1	Descrição do Problema	19
1.2	Motivação	20
1.3	Hipótese	21
1.4	Objetivos e Contribuições	22
1.5	Estrutura do Texto	23

1.1 Descrição do Problema

Atualmente, câmeras digitais são muito populares, sendo possível encontrá-las em uma variedade de dispositivos, tais como telefones celulares, câmeras de monitoramento e vigilância, dentre outros.

A abundância de imagens provenientes das câmeras digitais acaba por favorecer o uso destas imagens por peritos forenses para auxiliar a elucidar crimes. Por outro lado, a disponibilidade de *software* de edição de imagens, que permite a realização de operações que contemplem desde a alteração de cores e nitidez de uma imagem até aquelas que realizem montagens sofisticadas, permitiu que imagens maliciosamente editadas para prejudicar a percepção real de uma cena se tornassem muito comuns. Um exemplo de imagem editada com fins maliciosos é apresentado na Figura 1, onde é possível observar que ambas as imagens existentes apresentam o mesmo cenário, mas uma pessoa que está presente na imagem original foi retirada da cena na imagem editada.



(a) Original

(b) Editada

Figura 1 – Exemplo de imagem editada maliciosamente (FRIDRICH, 2009).

Em meio a esse contexto, nós encontramos a forense em imagens digitais (FID), uma área de conhecimento focada na recuperação e análise de evidências digitais em um processo de investigação criminal. É possível verificar que a FID tem se concentrado basicamente em dois problemas: a identificação da procedência de uma imagem digital e também a sua integridade. A identificação da origem de uma imagem consiste em reconhecer, por exemplo, o modelo da câmera responsável por gerar aquela imagem. A verificação da integridade de uma imagem digital promove a realização da avaliação do conteúdo de seu arquivo para verificar se houve algum processo de adulteração com o objetivo de produzir uma imagem manipulada.

Na prática, é fácil constatar que o número de imagens manipuladas tem se espalhado pelas redes sociais e aplicativos de mensagens, contribuindo, por exemplo, para impulsionar o fenômeno das *fake news*, propondo, assim, constantes desafios para a área de FID, e, ainda, evidenciando a importância de se identificar se uma imagem digital foi ou não manipulada.

A seção seguinte traz os elementos motivadores para o desenvolvimento deste trabalho.

1.2 Motivação

Estudos realizados por Sacchi, Agnoli e Loftus (2007) sugerem que imagens podem influenciar as pessoas e até mesmo modificar suas memórias a respeito de um fato real. Já o trabalho de Frenda et al. (2013) evidencia que imagens falsificadas podem ser mais bem lembradas pelas pessoas do que as imagens consideradas autênticas.

Merece destaque o fato de que o desenvolvimento de uma análise manual de um grande volume de imagens digitais é uma tarefa árdua, cansativa e tediosa. Em uma pesquisa típica que dura cerca de 6 a 18 meses, peritos forenses podem avaliar mais de

100.000 arquivos digitais (SARANTINOS; AL-NEMRAT; NAEEM, 2013). Já no estudo forense feito por Monteiro e Pereira (2009), foi reportado que, em uma base de dados contendo mais do que 300.000 arquivos de imagens e 1.100 vídeos, somente 148 imagens com conteúdo ilícito (abuso sexual) foram encontradas. Além disso, como apresentado por Sunde e Dror (2019), decisões forenses podem ser afetadas por diversos fatores como arquitetura cognitiva do cérebro, treinamento e motivação, fatores organizacionais, expectativas iniciais, materiais de referência e evidências do caso.

A seção a seguir apresenta a hipótese considerada para a elaboração desta tese de doutoramento.

1.3 Hipótese

A textura de uma imagem é uma das características essenciais que permite ao sistema visual humano identificar informações sobre a distribuição espacial e variação de luminosidade de uma imagem. Além disso, ela descreve o arranjo estrutural de superfícies e relações entre regiões vizinhas de uma cena, contribuindo para descrever propriedades como, por exemplo, suavidade, rugosidade e regularidade. Assim sendo, é possível deduzir que a textura se apresenta como sendo uma característica desafiadora.

A textura pode, também, pode ser interpretada como uma característica composta por variações de intensidade ou de cor, tipicamente originadas a partir da rugosidade da superfície dos objetos, conforme descrito por Mirmehd, Xie e Jasjit Suri (2008). Diante disso, dependendo da superfície analisada, estas variações podem apresentar regularidade ou aleatoriedade, conforme pode ser observado na Figura 2.

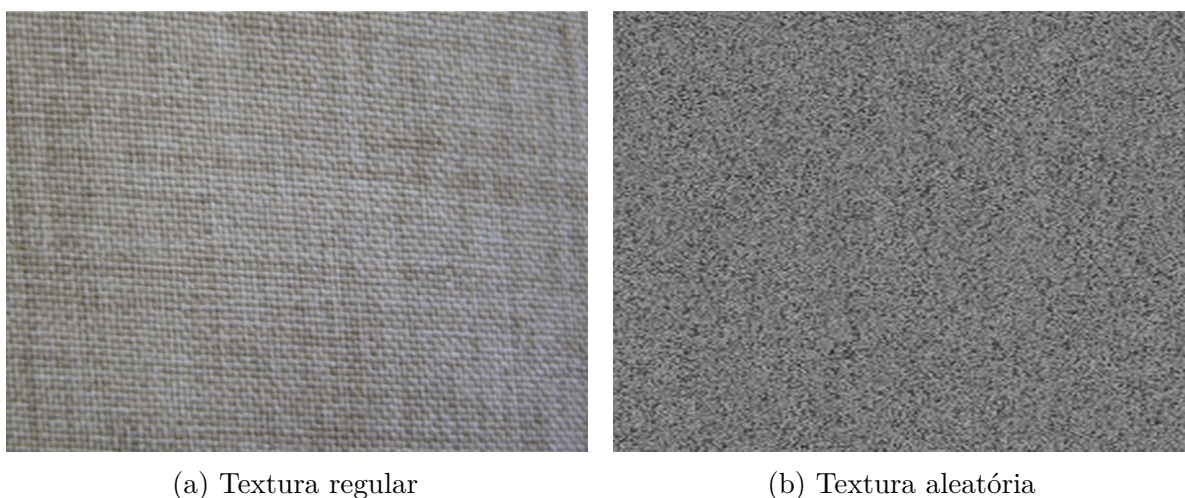


Figura 2 – Variedade de texturas.

Destaca-se, ainda, a possibilidade de ocorrer variação na textura em função do ângulo do observador. Este fato pode ser verificado na Figura 3.



(a) Gramado visto de cima

(b) Gramado visto com 40° de inclinação

Figura 3 – Visões de um gramado em função do ângulo do observador (MIRMEHD; XIE; Jasjit Suri, 2008).

Dessa forma, tendo em vista a natureza desafiadora da textura, a hipótese considerada neste trabalho é que é muito complexo promover manipulações em imagens digitais sem deixar traços nas suas informações de textura. Assim, através da análise da textura das imagens, é possível detectar se a mesma foi ou não manipulada.

Diante disso, considerando o problema da abundância de imagens maliciosamente manipuladas, o impacto que elas ocasionam ao serem vistas pelas pessoas, as dificuldades encontradas pelos peritos e, ainda, a hipótese descrita, a seção seguinte apresenta os objetivos que se pretende alcançar com desenvolvimento do presente trabalho.

1.4 Objetivos e Contribuições

O objetivo geral deste trabalho compreende o desenvolvimento de um modelo de suporte à decisão na identificação de imagens falsificadas a partir da análise de suas texturas, permitindo, dessa forma, determinar a autenticidade das mesmas. Já como objetivos específicos pretende-se:

- Realizar um estudo do estado da arte em métodos de identificação de imagens digitais falsificadas;
- Estudar os descritores locais de textura, clássicos e mais modernos;
- Avaliar o uso dos principais descritores locais de textura da literatura com redes neurais convolucionais;
- Analisar a utilização de redes neurais convolucionais na identificação de imagens falsificadas.

A seção seguinte apresenta uma visão geral da estrutura do texto do presente trabalho.

1.5 Estrutura do Texto

O restante deste documento encontra-se organizado da seguinte maneira: o Capítulo 2 versa sobre detecção de imagens falsificadas, abordando os métodos ativos e passivos, bem como apresentando as bases de imagens utilizadas em trabalhos de detecção de imagens falsificadas. O Capítulo 3, por sua vez, apresenta as redes neurais convolucionais, os componentes de sua arquitetura e sua evolução histórica. No Capítulo 4, são discutidas as abordagens para análise de textura, bem como os principais descritores de textura da literatura. Além disso, são apresentadas as principais bases de imagens de texturas. O Capítulo 5 discute os trabalhos relacionados. Já o Capítulo 6, traz o detalhamento do modelo proposto nesta tese. O Capítulo 7 apresenta e discute os resultados experimentais obtidos com a aplicação do modelo proposto. Por fim, o Capítulo 8 relaciona as conclusões contribuições deste trabalho.

Detecção de Imagens Falsificadas

Sumário

2.1	Introdução	24
2.2	Métodos Ativos	25
2.2.1	Esteganografia	25
2.2.2	Marca D'Água	30
2.3	Métodos Passivos	33
2.3.1	Identificação da Câmera de Origem	33
2.3.2	<i>Copy-move</i>	35
2.3.3	Reamostragem	38
2.4	Bases de Imagens de FID	39
2.4.1	CASIA	39
2.4.2	Columbia	39
2.4.3	CoMoFoD	42
2.5	Considerações Finais do Capítulo	47

2.1 Introdução

O trabalho de [Zheng, Zhang e Thing \(2019\)](#) detalha as diferenças entre uma imagem falsificada, ou forjada, e uma imagem adulterada, conforme apresentado na Tabela 1. Assim, é possível verificar que a produção de uma imagem falsificada envolve, necessariamente,

uma intenção maliciosa, o que potencializa o desafio de identificar a autenticidade de uma imagem digital.

Terminologia	Definição
Manipulação	Operações desenvolvidas na imagem para modificá-la.
Falsificação	Manipulações desenvolvidas na imagem inteira com intenção maliciosa.
Adulteração	Somente partes da imagem são intencionalmente modificadas.
Geração	Utiliza algoritmos para criar imagens com cenas de simulação do mundo real.

Tabela 1 – Definições de manipulação, falsificação e adulteração de imagem (ZHENG; ZHANG; THING, 2019).

Dessa maneira, tendo em vista o desafio que é identificar se uma imagem foi ou não falsificada, vários métodos têm sido desenvolvidos e aperfeiçoados a fim de se assegurar a autenticidade de imagens digitais (LIAN; KANELLOPOULOS; RUFFO, 2009). Tais métodos encontram-se organizados em dois grandes grupos, a saber: os ativos e os passivos. Estes grupos serão detalhados nas subseções seguintes.

2.2 Métodos Ativos

Os métodos ativos requerem a inserção de elementos na imagem original (SUBHEDAR; MANKAR, 2014). Consequentemente, informações prévias da imagem são vitais para o seu processo de autenticação. Assim, os métodos ativos são aqueles baseados em esteganografia e também aqueles que fazem uso de marca d'água. Ambos os métodos serão discutidos a seguir.

2.2.1 Esteganografia

Esteganografia consiste de um método para comunicação secreta que esconde informação confidencial através da exploração das limitações do sistema visual humano (RAFAT; HUSSAIN, 2016). Assim, com a inserção de uma chave através do uso da esteganografia, é possível verificar se uma imagem é ou não falsificada. Entretanto, é importante considerar que a esteganografia difere substancialmente de criptografia, uma vez que ela não visa tornar ilegível a mensagem entre as partes. Contudo, a esteganografia possibilita o sigilo de informações mesmo em mensagens que podem ser lidas por qualquer pessoa. A Tabela 2 apresenta uma análise comparativa entre esteganografia e criptografia.

	Esteganografia	Criptografia
Mecanismo de Segurança	A mídia transportadora incorpora a mensagem secreta.	O texto puro é convertido em um texto criptografado pelo uso de uma chave secreta.
Segurança de Dados	Depende de como uma mensagem é incorporada no meio e de quão imperceptível ela é.	Depende da robustez do algoritmo de criptografia e da sofisticação da chave de segurança.
Robustez	Depende do nível de imperceptibilidade dos dados confidenciais.	Depende da complexidade do algoritmo de criptografia.
Ataque	Menos propenso a ataques porque os dados não são detectáveis.	Mais propenso a ataques porque o texto cifrado pode levantar suspeitas.

Tabela 2 – Análise comparativa entre esteganografia e criptografia (JOHRI; MISHRA; DAS, 2016).

O processo de implementação do método de detecção de imagens falsificadas que faz uso da esteganografia procura, então, esconder a informação estratégica na mídia, chamada de transportadora, através do uso da chave esteganográfica (KESSLER, 2004). A Figura 4 ilustra o funcionamento geral de um sistema esteganográfico.

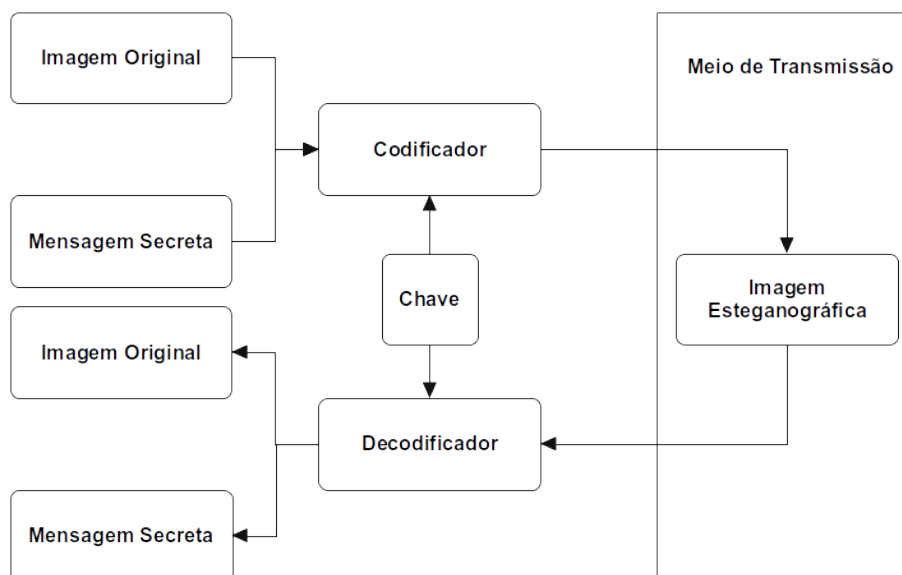


Figura 4 – Funcionamento geral de um sistema esteganográfico (AL-ATABY; AL-NAIMA, 2011).

A Figura 5 apresenta uma taxonomia das principais técnicas esteganográficas, considerando o domínio espacial e da frequência (JAIN; BOADDH, 2016).

Na próxima seção, são descritos os principais métodos esteganográficos que atuam no domínio espacial da imagem.

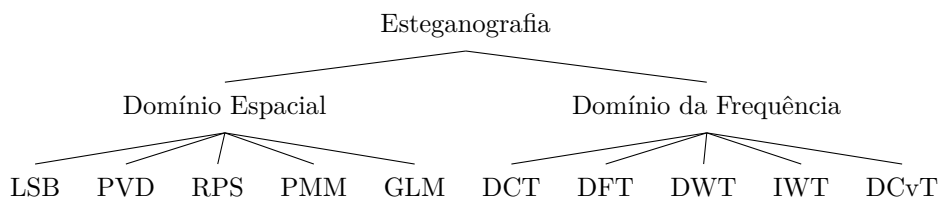


Figura 5 – Principais métodos esteganográficos.

2.2.1.1 Domínio da Espacial

A seguir, serão discutidas os seguintes métodos esteganográficos do domínio espacial: LSB, PVD, RPS, PMM e GLM.

A. Bits Menos Significantes

Também conhecido como LSB, do inglês *least significant bits*, este método preconiza a substituição dos bits dos pixels menos significativos da imagem por bits da informação secreta.

Os autores [Tavares e Madeiro \(2016\)](#) desenvolveram um projeto que habilita a redução do número esperado de modificações por pixel através da proposição de um método inspirado no jogo de caça-palavras, o LSB Word-Hunt. O método proposto reduziu a carga útil do sistema em 62,5% e obteve uma melhora de desempenho. Já o trabalho de [Danuputri, Mantoro e Hardjianto \(2016\)](#), apresenta uma abordagem híbrida que integra a cifra Vigenere com esteganografia a fim de garantir confidencialidade em ambiente baseado no sistema operacional Android.

Um método aperfeiçoado foi apresentado por [Singh e Singh \(2015\)](#). Este método embutiu a informação secreta nos três canais RGB de imagens coloridas, melhorando a qualidade da imagem e sua capacidade de carga. O trabalho de [Huang, Zhong e Huang \(2010\)](#) mostrou uma técnica LSB adaptativa que pôde escolher regiões embutidas de acordo com o tamanho da informação secreta, enquanto que [Singh et al. \(2007\)](#) propuseram um algoritmo para embutir mensagens secretas escondidas em localizações de pixels não adjacentes em bordas da imagem.

B. Diferença de Valor do Pixel

Neste método, a diferença de valores entre dois pixels adjacentes (PVD, do inglês, *pixel value differencing*) é usada para determinar o número de bits da informação secreta que pode ser inserida.

Considerando o método PVD, o trabalho de [Al-dmour, Al-ani e Nguyen \(2014\)](#) propõe um método que objetiva garantir a confidencialidade de dados médicos de pacientes usando PVD para identificar regiões de contraste na imagem que geralmente têm

informação relevante para o médico e, portanto, não podem receber dados secretos. No trabalho desenvolvido por [Thanekar e Pawar \(2013\)](#), é possível encontrar uma variação do Tri-PVD (TPVD) que permite que mais dados sejam escondidos na imagem. O artigo escrito por [Wu e Tsai \(2003\)](#) apresenta um método esteganográfico eficiente para embutir dados secretos em imagens em níveis de cinza.

C. Seleção Aleatória de Pixels

O método de seleção aleatória de pixels (RPS, do inglês, *random pixel selection*) faz uma seleção aleatória dos bits menos significativos da imagem, os quais serão substituídos por bits que contém as informações a serem ocultadas.

O trabalho de [Venus e Rachna \(2016\)](#) propõe uma técnica que pode esconder texto ou pixels aleatoriamente em uma imagem. Já [Laskar e Hemachandran \(2013\)](#) propoem um método RPS que esconde mensagens secretas em bits menos significativos no plano vermelho de uma imagem RGB. O trabalho escrito por [Bassil \(2012\)](#) sugere o uso de uma gramática livre de contexto para codificar as coordenadas dos pixels selecionados aleatoriamente em sentenças da língua inglesa, permitindo duas maneiras para entregar o dado secreto: uma imagem transportadora e um texto em inglês que codifica a localização dos pixels aleatórios transportadores.

D. Método de Mapeamento de Pixels

O método de mapeamento de pixels (PMM, do inglês, *pixel mapping method*) faz uso de uma função que relaciona a intensidade do pixel central com a sua vizinhança. A incorporação é, então, realizada através do mapeamento de cada dois ou quatro bits da mensagem secreta em pixels da vizinhança.

As posições de incorporação são selecionadas por alguma função e dependem do valor da intensidade do pixel inicial e de seus 8 vizinhos. A incorporação de dados é feita através do mapeamento de cada dois ou quatro bits da mensagem secreta em cada pixel vizinho, com base em alguns recursos desse pixel.

Os autores [Mandai e Bhattacharyya \(2015\)](#) fazem uso do PMM em conjunto com algoritmos genéticos para criptografar e descriptografar dados secretos e, em seguida, aumentar a segurança do compartilhamento de dados em um ambiente de nuvem. O trabalho desenvolvido por [Banerjee, Bhattacharyya e Sanyal \(2013\)](#) estende o método PMM baseado no valor de intensidade e posição do pixel.

E. Modificação do Nível de Cinza

Este método, também conhecido como GLM, do inglês *gray level modification*,

modifica o nível de cinza para ocultar informação, mapeando dados secretos dentro de uma imagem. Assim, o trabalho desenvolvido por [Shobana e Manikandan \(2013\)](#) usa as intensidades dos pixels dos planos RGB da imagem para armazenar a informação sigilosa.

Na sequência, são descritos os métodos esteganográficos do domínio da frequência.

2.2.1.2 Domínio da Frequência

Na sequência, serão abordados os seguintes métodos do domínio da frequência: DCT, DFT, DWT, IWT e DCvT.

A. Transformada Discreta do Cosseno

Através do uso transformada discreta do cosseno (DCT, do inglês, *discrete cosine transform*), este método utiliza a informação da representação de uma imagem através da soma de sinusoides de diferentes magnitudes e frequências para promover a seleção de pixels que armazenarão as informações secretas.

O método DCT é utilizado por [Ni, Hu e Shi \(2018\)](#) para propor um novo esquema esteganográfico que visa incorporar as estatísticas do domínio espacial e DCT. Já no artigo escrito por [Lie e Lin \(2005\)](#), propriedades estatísticas no domínio espacial e DCT são propostas para checar se uma imagem digital possui uma mensagem secreta embutida.

B. Transformada Discreta de Fourier

Este método usa a informação extraída da DFT para selecionar os pixels da imagem que receberão os dados secretos. Assim, os trabalhos de [Ghoshal e Mandal \(2011\)](#) e [Bhattacharyya et al. \(2009\)](#) fazem uso da DFT para melhorar o processo de autenticação, enquanto o trabalho desenvolvido por [Alturki e Mersereau \(2001\)](#) mostra uma técnica esteganográfica para comunicações seguras baseadas em DFT.

C. Transformada Wavelet Discreta

Este método também é conhecido como DWT, do inglês *discrete wavelet transform*, e promove o armazenamento dos dados secretos através do resultado da decomposição *wavelet* da imagem.

O método DWT usado por [Elhoseny et al. \(2018\)](#) visa oferecer segurança na transmissão de diagnóstico médico embutido em exames. O artigo escrito por [Ibaida e Khalil \(2013\)](#) propõe um método esteganográfico baseado em *wavelet* que faz uso de criptografia e embaralhamento para oferecer segurança para dados confidenciais de pacientes. Por outro lado, o trabalho desenvolvido por [Lyu e Farid \(2006\)](#) descreve uma abordagem para detectar mensagens escondidas embutidas em imagens digitais baseada em DWT.

D. Transformada Wavelet Inteira

A transformada *wavelet* inteira (IWT, do inglês *integer wavelet transform*) é utilizada neste método para determinar os pixels da imagem que irão receber a informação sensível.

O método IWT é empregado por [Jothy e Anusuyya \(2016\)](#) para criar imagens esteganográficas com alta capacidade de ocultação. O trabalho desenvolvido por [Al-dmour e Al-ani \(2016\)](#) mostra um algoritmo esteganográfico que combina detecção de borda com codificação XOR. O estudo proposto por [Thanikaiselvan e Arulmozhivarman \(2013\)](#) aplica IWT aos planos RGB, embutindo dados secretos aleatórios através da seleção de coeficientes *wavelet* baseado na teoria de grafos. Além disso, o modelo proposto por [Zou et al. \(2006\)](#) usa o IWT para ocultar dados confidenciais nas imagens que podem ser recuperados mesmo após a compactação.

E. Transformada Curvelet Discreta

Este método trabalha com a transformada *curvelet* discreta (DCvT, do inglês, *discret curvelet transform*) para determinar os pixels da imagem que irão armazenar a informação estratégica.

A técnica descrita por [Jero e Ramu \(2016\)](#) usa DCvT para embutir dados confidenciais de paciente em sinais de eletrocardiogramas. O trabalho proposto por [Al-ataby e Al-naima \(2011\)](#) também usa DCvT e propõe uma técnica de esteganografia de imagens de alta capacidade. Também relacionado a DCvT, o artigo de [Hermanth et al. \(2017\)](#) analisa o uso de diferentes transformações (*ridgelet*, *bandelet* e *curvelet*) para implementar a esteganografia.

A seção seguinte apresenta os métodos ativos implementados através do uso de marca d'água.

2.2.2 Marca D'Água

A marca d'água é usada para oferecer conteúdo multimídia com o princípio de autenticidade. Para informação de identificação de sua origem, o produtor do conteúdo insere uma marca, que deve ser discreta o suficiente para que a informação primária não seja deturpada e, também, que seja facilmente verificável, para que não haja dúvida de sua origem. Além disso, conforme descrito por [Urvoy, Goudia e Autrusseau \(2014\)](#), as técnicas de marca d'água concentram-se em quatro requisitos essenciais: robustez, segurança, capacidade e invisibilidade.

A Figura 6 mostra os métodos de identificação de imagens falsificadas baseados em

marca d'água, considerando o domínio espacial e o da frequência.

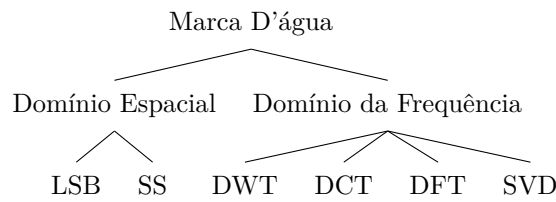


Figura 6 – Principais métodos de marca d'água.

A seguir, são descritos os principais métodos de marca d'água dos domínios espacial e da frequência:

2.2.2.1 Domínio Espacial

Na sequência, serão abordados os seguintes métodos do domínio espacial relacionados à marca d'água: LSB e SS.

A. Bits Menos Significantes

Os bits menos significantes da imagem também podem ser utilizados para a implementação de marca d'água. Esse método usa os bits menos significantes da imagem para armazenar informações de autenticidade. Assim, o trabalho de [Liu, Lin e Yuan \(2018\)](#) implementa um mecanismo cego e robusto de marca d'água para direitos autorais de imagens coloridas.

B. Espalhamento Espectral

O espalhamento espectral (SS, do inglês, *spread spectrum*) é uma técnica de modulação do sinal onde a largura de banda é maior do que o necessário para transmitir a informação, e com esta largura de banda adicional sendo obtida por um código independente da informação. O sinal transmitido é, então, espalhado sobre esta largura de banda maior.

Assim, o trabalho desenvolvido por [Bose e Maity \(2017\)](#) implementa um modelo de detecção de marca d'água em imagens compactadas, procurando atacar dois problemas: a autenticidade e a facilidade de transmissão das imagens com largura de banda reduzida.

Logo a seguir, serão descritos os métodos de marca d'água do domínio da frequência.

2.2.2.2 Domínio da Frequência

Na sequência, serão abordados os seguintes métodos de marca d'água do domínio da frequência: DWT, DCT, DFT e SVD.

A. Transformada Wavelet Discreta

Esse método faz uso da transformada *wavelet* para encontrar regiões de interesse na imagem que podem armazenar a marca d'água. Assim, os autores [Chen e Huang \(2015\)](#) propuseram um método para embutir marcas d'água em uma região de interesse obtida através do uso da DWT para oferecer direitos autorais seguros. O trabalho organizado por [Totla e Bapat \(2013\)](#) apresenta uma análise comparativa de marcas d'água usando DWT e também DCT.

B. Transformada Discreta do Cosseno

O trabalho de [Liu, Pan e Song \(2017\)](#) descreve um algoritmo que usa codificação DCT e fractais para criar uma marca d'água robusta. No artigo escrito por [Vizireanu e Preda \(2015\)](#), é proposto um modelo de marca d'água DCT robusto para compressão JPEG. O artigo escrito por [Ernawan e Kabir \(2018\)](#) apresenta uma técnica de marca d'água DCT para proteção de direitos autorais.

C. Transformada Discreta de Fourier

A DFT é usada por [HAMIDI et al. \(2015\)](#) para propor uma abordagem cega de marca d'água como uma solução para o problema de proteção de direitos autorais. No trabalho de [Perez-Meana et al. \(2012\)](#), podemos encontrar um método de marca d'água usando pontos de interesse SURF, do inglês *speeded-up robust features* e domínio DFT.

D. Decomposição de Valor Singular

A decomposição de valor singular (SVD, do inglês, *singular value decomposition*) é usada junto com a DWT por [Khoo, Makbol e Rassem \(2016\)](#) para propor uma técnica de marca d'água em blocos. O trabalho desenvolvido por [Prasad, Rao e Kannan \(2016\)](#) propõe um sistema híbrido de marca d'água baseado em SVD e em BND, do inglês *Binocular Notification Difference*. O artigo escrito por [Rosiyadi et al. \(2012\)](#) sugere uma abordagem de marca d'água para implementar a segurança de direitos autorais em documentos e imagens eletrônicas do governo baseada em DCT e SVD.

Por fim, a Tabela 3 mostra uma análise comparativa entre marca d'água e esteganografia baseada no trabalho de [Nagaraj, Vijayalakshmi e Zayaraz \(2013\)](#).

	Marca d'água	Esteganografia
Foco	Direitos autorais e autenticação.	Comunicação secreta usando ambiente público.
Carga	Marca visível que identifica a origem da imagem para todos.	Mensagem secreta oculta que pode ser lida apenas pelo destinatário.
Comunicação	Um para muitos.	Um para um.

Tabela 3 – Análise comparativa entre marca d'água e esteganografia (NAGARAJ; VIJAYALAKSHMI; ZAYARAZ, 2013).

A próxima seção abordará os principais métodos passivos para detecção de imagens falsificadas da literatura.

2.3 Métodos Passivos

Diferentemente dos métodos de autenticação ativa, os passivos não precisam de informações sobre a imagem original. Estes métodos exploram as informações intrínsecas das imagens, realizando a análise de seu conteúdo sem recorrer a nenhuma marcação inserida anteriormente. As subseções seguintes detalham os principais métodos passivos, a saber: identificação da câmera de origem, *copy-move* e reamostragem.

2.3.1 Identificação da Câmera de Origem

Os métodos de identificação da câmera de origem (SCI, do inglês, *source camera identification*) identificam traços deixados na imagem pela câmera que a gerou. A Figura 7 apresenta uma visão geral dos métodos de detecção de imagens falsificadas baseados em SCI.

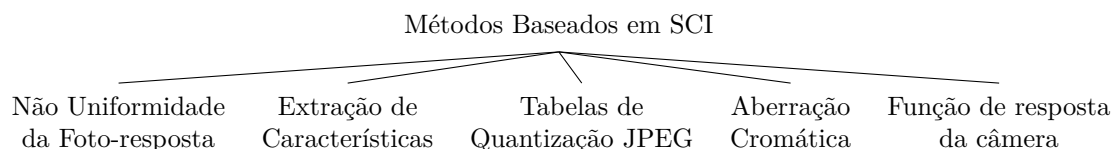


Figura 7 – Métodos baseados em SCI.

2.3.1.1 Não Uniformidade da Foto-resposta

O método de não uniformidade de foto-resposta (PRNU, do inglês, *photo-response non-uniformity*) utiliza ruídos gerados pelos sensores da câmera durante o processo de aquisição para produzir uma assinatura que identifique a câmera.

Os autores [Lawgaly e Khelifi \(2017\)](#) apresentaram um processo baseado em PRNU que usa DCT na etapa de filtragem. Já o trabalho desenvolvido por [Kee, Johnson e Farid \(2011\)](#) apresenta um modelo que faz uso do PRNU para encontrar contas de redes sociais que possivelmente tenham originado uma determinada imagem. Uma visão geral interessante do PRNU sob a ótica da forense de imagens digitais é apresentada por [Fridrich \(2009\)](#). O trabalho desenvolvido por [Chen, Fridrich e Goljan \(2008\)](#), por sua vez, traz uma proposta de estrutura unificada para identificação de câmeras digitais.

2.3.1.2 Extração de Características

Os recursos extraídos de imagens digitais são amplamente utilizados para identificar uma câmera digital que produziu a imagem. Assim, os autores [Xu et al. \(2016\)](#) usam características de textura extraídas de imagens coloridas para obter uma assinatura da câmera que as adquiriu. Os trabalhos de [Jin, Huang e Fan \(2012\)](#) e de [Tuama, Comby e Chaumont \(2016\)](#) empregam redes neurais convolucionais para identificar os modelos das câmeras digitais.

2.3.1.3 Tabelas de Quantização JPEG

Este processo faz uso de características intrínsecas do processo de montagem da imagem JPEG para construir uma assinatura de câmera digital.

O trabalho de [Qiao et al. \(2017\)](#) abrange uma análise estatística da distribuição de pixels ruidosos para identificar a assinatura da câmera digital. Os autores [Thai et al. \(2017\)](#) apresentam um processo baseado no DCT para estimar as tabelas de quantização normalmente criadas por uma câmera digital. Os autores [Kee, Johnson e Farid \(2011\)](#) propõem um ambiente para autenticação de imagens baseado em cabeçalhos JPEG enquanto que [Kornblum \(2008\)](#) usa tabelas de quantização JPEG para localizar imagens que foram falsificadas.

2.3.1.4 Aberração Cromática

Esse método trabalha com as aberrações produzidas pelo processo de construção das lentes usadas nas câmeras digitais para encontrar assinaturas para estas mesmas lentes.

No estudo feito por [Van, Emmanuel e Kankanhalli \(2007\)](#) as aberrações cromáticas são estimadas nos canais R , G e B . Em seguida, elas são usadas no classificador SVM, do inglês *support vector machines*, para identificar o *smartphone* que produziu as imagens. Em um trabalho desenvolvido por [Johnson e Farid \(2006\)](#), é descrita uma técnica para identificar aberrações cromáticas laterais e, com isso, as imagens digitais são analisadas para identificar traços de falsificação. Por fim, [Battiato, Giudice e Paratore \(2016\)](#) apresentam uma pesquisa sobre detecção de imagens falsificadas e dedicam uma subseção em seu artigo exclusivamente para tratar de técnicas de aberração de lentes.

2.3.1.5 Função de Resposta da Câmera

Este método também é conhecido como CRF, do inglês, *camera response function*, e analisa os estímulos luminosos que são capturados por sensores em uma função não linear que se comporta como uma assinatura para as câmeras digitais. Assim, [Rodrigues e Barreto \(2015\)](#) apresentam um processo para estimar a CRF a partir de uma única imagem de uma superfície bicolor geral. No trabalho desenvolvido por [Hsu e Chang \(2010\)](#), é mostrado um ambiente que estima a CRF a partir de pontos de irradiância planares invariantes localmente.

2.3.2 Copy-move

A detecção *copy-move* se concentra na parte da imagem que foi copiada e colada em uma ou mais partes da própria imagem, podendo a parte copiada ter sido ou não editada. Portanto, o objetivo do *copy-move* é ocultar partes da imagem, eliminando a autenticidade da mesma. A Figura 8 traz um exemplo ilustrativo de uma imagem que passou por uma operação de *copy-move*.

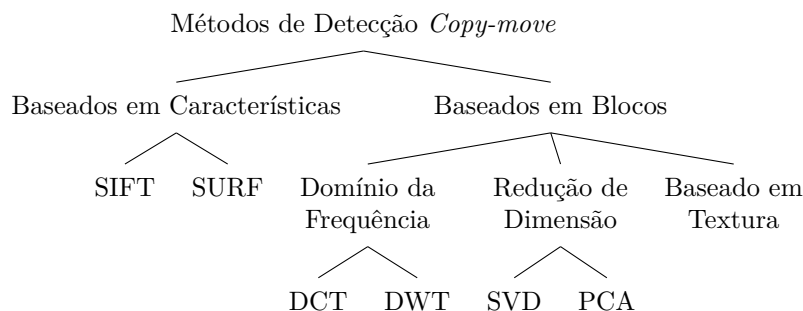


(a) Imagem original

(b) Imagem falsificada

Figura 8 – Exemplo de *copy-move* ([FRIDRICH, 2009](#)).

Como pode ser visto na Figura 9 e descrito nos trabalhos de [Warif et al. \(2016\)](#) e [Soni, Das e Thounaojam \(2018\)](#), os métodos que identificam a ocorrência de *copy-move* em uma imagem podem ser classificados em duas categorias, a saber: baseados em características e baseados em blocos.

Figura 9 – Métodos baseados em *copy-move*.

2.3.2.1 Baseados em Características

Na sequência, serão discutidos os métodos baseados em características SIFT e SURF.

A. SIFT

O trabalho de [Amerini et al. \(2011\)](#) propôs um modelo SIFT, do inglês *scale-invariant feature transform*, que identifica uma operação *copy-move* e a transformação geométrica usada para executá-la. O trabalho de [Costanzo et al. \(2014\)](#) também desenvolveu uma abordagem baseada em SIFT e assume a presença de ações contra-forenses. Já [Alberry et al. \(2018\)](#) propuseram uma abordagem baseada em SIFT e em *Fuzzy* para desenvolver um detector de *copy-move* rápido e confiável. Por fim, o trabalho desenvolvido por [Li e Zhou \(2019\)](#) propõe uma técnica de detecção e localização de falsificação *copy-move* baseada em número reduzido de pontos-chaves para melhorar o desempenho.

B. SURF

Características robustas SURF também são muito utilizadas para construir detectores de operações *copy-move*. Assim, [Shivakumar e Baboo \(2011\)](#) apresentam uma técnica baseada em SURF e KD-Tree para identificar *copy-move* em imagens de alta resolução. Uma tabela comparativa entre SIFT e SURF para os detectores de *copy-move* é apresentada por [Jaseela e Nishadha \(2016\)](#). O trabalho realizado por [Al-Hammadi e Emmanuel \(2016\)](#) apresenta um ambiente baseado em SURF aprimorado para detecção de *copy-move* usando super-resolução.

A seção seguinte, apresentará os métodos de detecção de *copy-move* baseados em blocos.

2.3.2.2 Baseados em Blocos

Os métodos baseados em blocos são organizados em três categorias, a saber: domínio da frequência, redução de dimensão e baseados em textura. Estes métodos serão

discutidos a seguir.

A. Domínio da Frequência

O domínio da frequência, através da transformação de sinal, pode oferecer as assinaturas para os blocos de imagem e, então, permitir a identificação de regiões duplicadas. O artigo escrito por [Mahmood et al. \(2018\)](#) usou DWT e DCT para detectar e localizar *copy-move* em imagens digitais. Além disso, [Yan, Pun e Yuan \(2016\)](#) trabalharam com a transformada de Fourier-Mellin para eliminar eventuais distorções geométricas e criar assinaturas de regiões com o objetivo de identificar regiões clonadas. Já [Pun e Chung \(2018\)](#) usaram a transformada analítica discreta de Fourier-Mellin para extrair características de bloco em um segundo estágio de um modelo para identificar *copy-move*. No trabalho desenvolvido por [Dixit, Naskar e Mishra \(2017\)](#), a transformada estacionária *wavelet* (SWT) e SVD são usadas para implementar um detector de imagens forjadas, enquanto o trabalho de [Wo et al. \(2017\)](#) propôs um detector de *copy-move* baseado em transformada complexa polar multi-raio.

B. Redução de Dimensão

Os métodos de redução de dimensão são amplamente utilizados para reduzir o número de blocos das imagens a serem processados, contribuindo, dessa maneira, para a melhoria do desempenho dos ambientes. Os trabalhos de [Kang e Wei \(2008\)](#) e [Zhao e Guo \(2013\)](#) trabalharam com SVD para implementar um algoritmo para detectar *copy-move* e [Bashar et al. \(2010\)](#) trabalharam com a análise de componentes principais (PCA, do inglês, *principal component analysis*).

C. Baseados em Textura

O LBP é um operador em escala de cinza bastante utilizado. Assim, [Mahmood et al. \(2017\)](#) e [Guo, Zhang e Zhang \(2010\)](#) investigaram a variação do padrão binário local para implementar os detectores de operações *copy-move*. Os autores [Farooq, Yousaf e Hussain \(2017\)](#), por sua vez, trabalharam com uma abordagem genérica do LBP para detectar *copy-move*. Além disso, [Li et al. \(2013\)](#) trabalharam com características extraídas usando o LBP invariantes à rotação para implementar um esquema para detectar *copy-move* e [Davarzani et al. \(2013\)](#) usaram LBP multirresolução.

As informações de textura obtidas dos filtros Gabor, por exemplo, podem ser usadas para identificar regiões de imagem falsificadas. Além disso, as abordagens desenvolvidas por [Lee \(2015\)](#) e [Muzaffer G. \(2016\)](#) usaram a filtragem Gabor para implementar abordagens para identificar *copy-move*. Na pesquisa realizada por [Ardizzone, Bruno e Mazzola \(2010\)](#), vários descritores de textura são investigados, considerando o problema da detecção de

operações *copy-move*.

Existem abordagens que implementam processos mistos para atacar problemas de detecção de *copy-move*. Os artigos escritos por [Ferreira et al. \(2016\)](#) e [Li et al. \(2017\)](#) propuseram modelos que combinam processos baseados em características e em blocos para viabilizar a detecção *copy-move*.

A próxima seção tratará dos métodos de detecção de imagens falsificadas com reamostragem.

2.3.3 Reamostragem

Quando transformações espaciais, como redimensionamento, rotação e espelhamento são aplicadas a uma imagem digital, temos um tipo específico de falsificação conhecida como reamostragem (*resampling*). No entanto, as transformações aplicadas em uma imagem deixam rastros que normalmente não estão presentes nas imagens originais, e os métodos que procuram identificá-los são conhecidos como detectores de reamostragem.

A reamostragem pode ocorrer, por exemplo, na redução da escala (*downscaling*) de imagens de uma loja *on-line*, nos serviços de compartilhamento de imagens e também, maliciosamente, para ocultar elementos em uma cena, conforme apresentado no trabalho realizado por [Peng, Wu e Kang \(2017\)](#).

O trabalho desenvolvido por [Popescu e Farid \(2005\)](#) discute como identificar falsificações detectando traços de *resampling*, mostrando que as operações de reamostragem introduzem correlações estatísticas entre os pixels da imagem e descreve, então, métodos para detectar essas correlações. No entanto, a abordagem proposta por eles não trabalha com imagens que tenham sido submetidas a algum processo de compactação.

Os autores [Vázquez-Padín, Comesaña e Pérez-González \(2015\)](#) descrevem uma estratégia para identificar *resampling* baseada no SVD. Os resultados são comparados com o trabalho de [Kirchner \(2010\)](#), apresentando melhores desempenho e resultados.

No artigo escrito por [Peng, Wu e Kang \(2017\)](#) são apresentados filtros passa-alta multidirecionais que são aplicados a fim de criar diferenças multidirecionais. Eles se encaixam em um modelo auto-regressivo para produzir coeficientes e histogramas normalizados, que são usados para extrair as características que revelam traços de reamostragem, mesmo quando a imagem foi submetida a técnicas anti-forenses. No entanto, é observado que a abordagem proposta teve seu desempenho afetado pela compactação JPEG com perdas.

Um modelo para identificar *resampling* através da realização da decomposição do subespaço e dos princípios da teoria da matriz aleatória foi investigado por [Vazquez-Padin, Perez-Gonzalez e Comesana-Alfaro \(2017\)](#).

A próxima seção apresentará as principais bases de imagens de forense em imagens

digitais da literatura que podem ser utilizadas para testes e validação de modelos de identificação de imagens falsificadas.

2.4 Bases de Imagens de FID

Existem bases de imagens que foram idealizadas para permitir o desenvolvimento de trabalhos relacionados à detecção de imagens falsificadas (bases de forense em imagens digitais). Nesta seção, serão detalhadas as bases mais modernas (a partir do ano 2000) e que ainda são, de alguma forma, mantidas. A Tabela 4 apresenta um resumo das características das bases de imagens escolhidas.

Bases de Imagens	Número de Imagens			Resolução (em pixels)
	Autênticas	Falsificadas	Total	
CASIA v1.0 (DONG; WANG; TAN, 2013)	800	921	1.721	384 × 256
CASIA v2.0 (DONG; WANG; TAN, 2013)	7.491	5.123	12.614	320 × 240 a 800 × 600
Columbia (NG; CHANG, 2004)	933	912	1.845	128 × 128
CoMoFoD (TRALIC et al., 2013)	5.000	5.000	10.000	512 × 512
Total	14.224	11.956	26.180	

Tabela 4 – Resumo das características das bases de imagens.

As subseções seguintes detalham as quatro bases de imagens, destacando suas características e referências.

2.4.1 CASIA

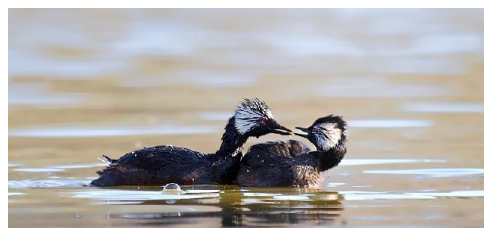
Essa base de imagens foi introduzida por Dong, Wang e Tan (2013) e se apresenta como sendo uma base pública de imagens com duas versões denominadas CASIA v1.0 e CASIA v2.0. A primeira versão possui 1.721 imagens JPEG coloridas com tamanho de 384 × 256 pixels, sendo 800 imagens autênticas e 921 imagens falsificadas. Já na segunda versão, existem 12.614 imagens coloridas, sendo seus formatos JPEG, BMP e TIFF. As imagens na base CASIA v2.0 têm resoluções diferentes, variando de 320 × 240 a 800 × 600 pixels, e existem 7.491 imagens autênticas e 5.123 imagens falsificadas. As duas versões da CASIA têm suas imagens agrupadas em oito categorias, a saber: cena, animal, arquitetura, personagem, plantas, itens, natureza e textura. A Figura 10 mostra exemplos de imagens autênticas dessa base, enquanto que a Figura 11 apresenta as imagens falsificadas equivalentes.

2.4.2 Columbia

O relatório escrito por Ng e Chang (2004) descreve a base de imagens de avaliação de detecção de operações de *copy-move* Columbia. Esta base é composta por 1.845 blocos



(a) cena



(b) animal



(c) arquitetura



(d) personagem



(e) plantas



(f) itens

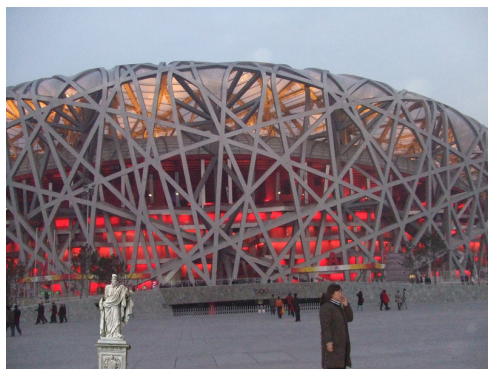


(g) natureza

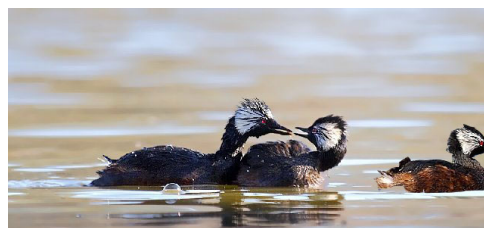


(h) textura

Figura 10 – Exemplos de imagens autênticas da base CASIA.



(a) cena



(b) animal



(c) arquitetura



(d) personagem



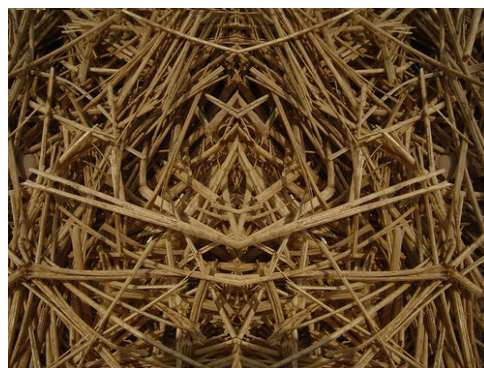
(e) plantas



(f) itens



(g) natureza



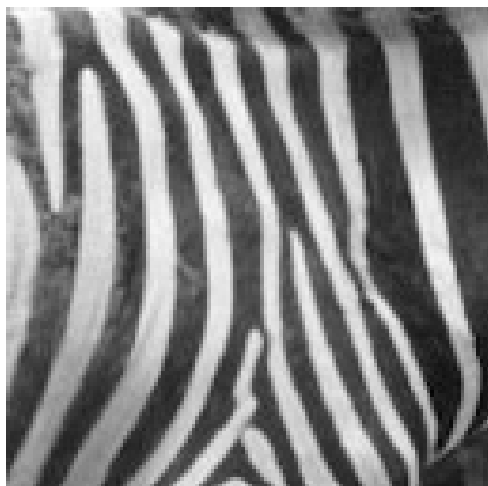
(h) textura

Figura 11 – Exemplos de imagens falsificadas da base CASIA.

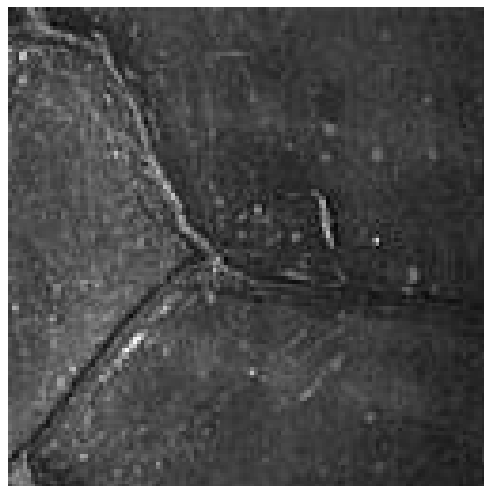
de imagens com tamanho 128×128 pixels e todos eles estão em escala de cinza BMP. Existem 933 imagens autênticas e 912 imagens falsificadas. A base Columbia encontra-se subdividida em cinco grupos: bloco de imagem com uma região inteiramente texturizada e homogênea (T), bloco de imagem com uma região suave e homogênea (S), bloco de imagem com uma fronteira de objeto entre uma região texturizada e outra suave (TS), bloco de imagem com uma fronteira de objeto entre duas regiões texturizadas (TT) e bloco de imagem com uma fronteira de objeto entre duas regiões suaves (SS). A Figura 12 mostra exemplos das imagens autênticas da base Columbia e a Figura 13 apresenta exemplos das imagens falsificadas da mesma base.

2.4.3 CoMoFoD

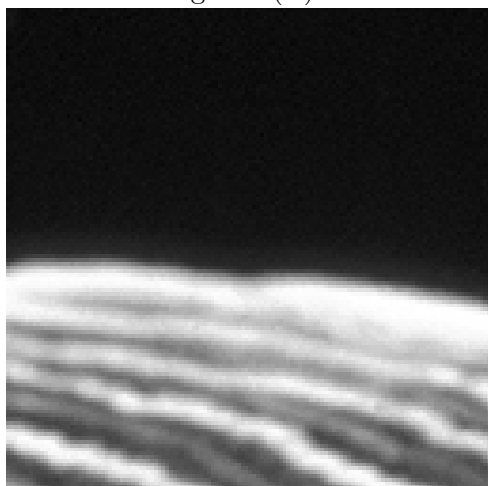
O trabalho desenvolvido por [Tralic et al. \(2013\)](#) apresenta uma base de imagens (CoMoFoD, do inglês, *copy-move forgery detection*) para detecção de falsificação baseada em operações de *copy-move* composta por 10.000 imagens de 512×512 pixels, sendo 5.000 imagens autênticas e 5.000 imagens falsificadas. As imagens estão separadas em cinco grupos de acordo com a manipulação aplicada, a saber: translação (uma região copiada é transladada para uma nova localização sem ser desenvolvida qualquer transformação), rotação (uma região copiada é rotacionada e transladada para uma nova localização), escala (uma região copiada é escalada e transladada para uma nova localização), combinação (duas ou mais transformações são aplicadas em uma região copiada antes de copiá-la para uma nova localização) e distorção (uma região copiada é distorcida e transladada para uma nova localização). Diferentes tipos de operações de pós-processamento também foram executados nas imagens da base, como compactação JPEG, borramento, adição de ruído, redução de cor, etc. A Figura 14 mostra exemplos de imagens autênticas da CoMoFoD e a Figura 15, por outro lado, mostra exemplos de imagens falsificadas da CoMoFoD.



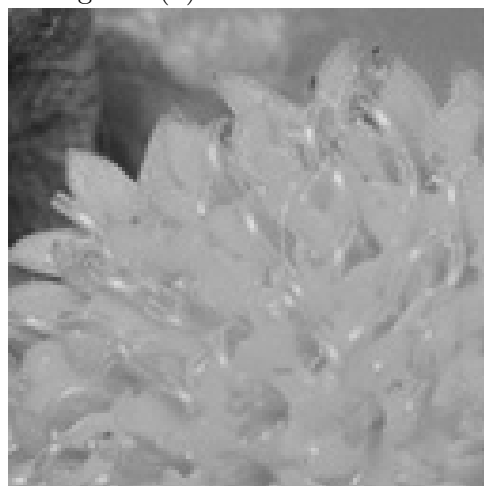
(a) bloco com região inteiramente texturizada homogênea (T)



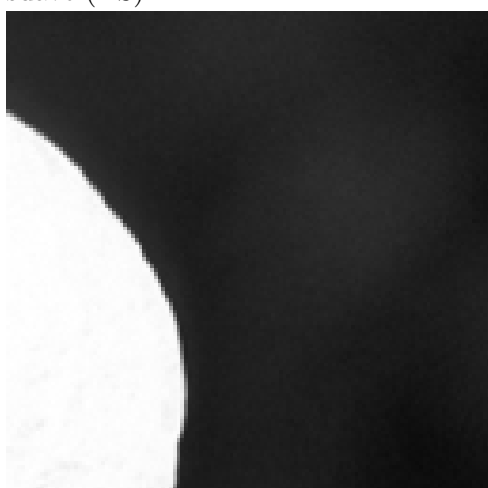
(b) bloco com região inteiramente suave homogênea (S)



(c) objeto entre região texturizada e suave (TS)

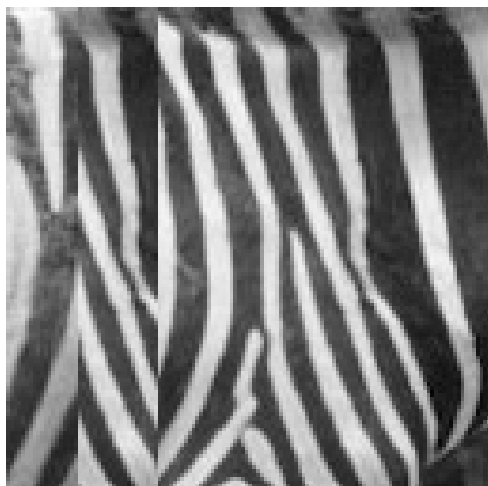


(d) fronteira de objeto entre duas regiões texturizadas (TT)

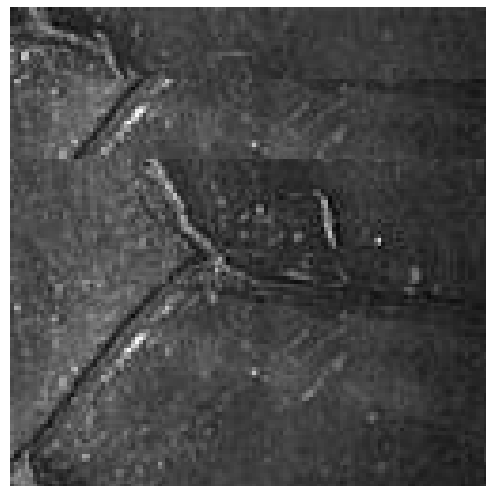


(e) fronteira de objeto entre duas regiões suaves (SS)

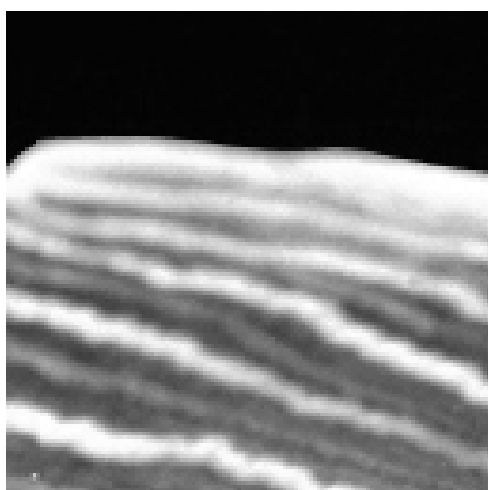
Figura 12 – Exemplos de imagens autênticas da base Columbia.



(a) bloco com região inteiramente texturizada homogênea (T)



(b) bloco com região inteiramente suave homogênea (S)



(c) objeto entre região texturizada e suave (TS)



(d) fronteira de objeto entre duas regiões texturizadas (TT)



(e) fronteira de objeto entre duas regiões suaves (SS)

Figura 13 – Exemplos de imagens falsificadas da base Columbia.



Figura 14 – Exemplos de imagens autênticas da base CoMoFoD.



Figura 15 – Exemplos de imagens falsificadas da base CoMoFoD.

A seção seguinte trará as considerações finais para este capítulo.

2.5 Considerações Finais do Capítulo

Este capítulo apresentou a definição de imagens falsificadas, bem como um detalhamento dos métodos ativos e passivos para detecção de imagens falsificadas. Foram apresentadas e descritas, também, as principais bases públicas de forense em imagens digitais que são utilizadas para testes e validação de modelos de identificação de imagens falsificadas.

O próximo capítulo tratará sobre redes neurais convolucionais, enfocando definições, elementos de sua arquitetura e seu histórico de evolução.

Redes Neurais Convolucionais

Sumário

3.1	Introdução	48
3.2	Arquitetura	49
3.2.1	Convolução	50
3.2.2	Função de Ativação	50
3.2.3	<i>Pooling</i>	52
3.2.4	Regularização	52
3.2.5	Camada Totalmente Conectada	53
3.3	Evolução Histórica	54
3.4	Considerações Finais do Capítulo	56

3.1 Introdução

As redes neurais artificiais, ou simplesmente redes neurais, de uma maneira geral, são algoritmos adequados para realizar o reconhecimento de padrões para se identificar ou classificar objetos. Elas constituem um modelo de computação cuja estrutura em camadas se inspirou na organização em rede de neurônios do cérebro, com camadas de neurônios conectados ([HAYKIN, 2001](#)).

Assim, as redes neurais constituem uma das ferramentas de aprendizagem de máquina (ML, do inglês, *machine learning*) mais utilizadas nos dias atuais. A ML constitui

um paradigma da inteligência artificial (IA) que permite o aprendizado diretamente a partir dos dados, e não através de recursos de programação explícita.

Uma rede neural combina várias camadas de processamento que fazem uso de elementos simples que trabalham em paralelo, tal como ocorre no sistema nervoso biológico. Assim, a sua arquitetura básica consiste de uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. As camadas, por sua vez, apresentam-se interconectadas através de nós ou neurônios, onde cada camada faz uso da saída da camada anterior como entrada, conforme ilustrado pela Figura 16.

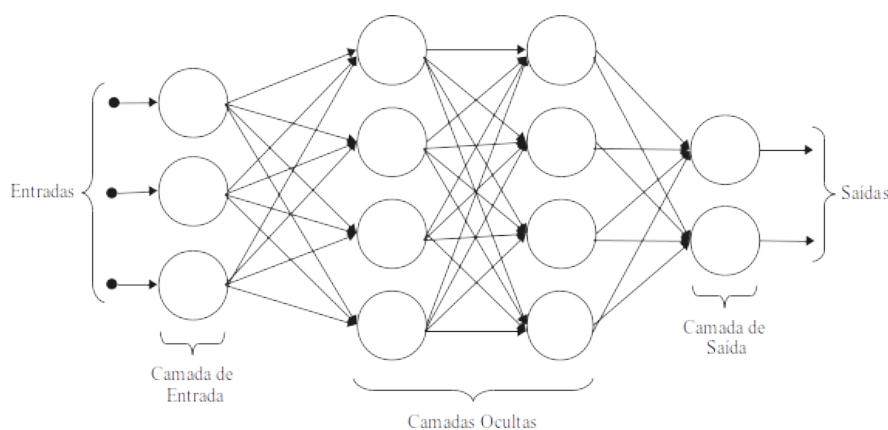


Figura 16 – Ilustração de rede neural.

As redes neurais que operam com duas ou três camadas ocultas conectadas entre si são conhecidas como redes neurais superficiais (*shallow neural networks*). Já as redes neurais de aprendizagem profunda (*deep learning*), podem ter até dezenas de camadas ocultas.

A rede neural convolucional (CNN, do inglês, *convolutional neural network*) ou ConvNet, originalmente proposta por [Lecun et al. \(1998\)](#), se apresenta como um dos algoritmos mais populares de *deep learning*, sendo especialmente desenvolvida para trabalhar com imagens ([AGGAR; WAL, 2018](#)). Dessa forma, a CNN é normalmente treinada com uma grande coleção de dados e pode aprender características dos mesmos. A Figura 17 ilustra o relacionamento entre AI, *machine learning*, *deep learning* e *convolutional neural network*.

A próxima seção detalhará os elementos presentes na arquitetura das redes neurais convolucionais.

3.2 Arquitetura

A arquitetura geral de uma CNN é formada por uma camada de entrada seguida por camadas de convolução com aplicação de função de ativação e também camadas de *pooling*. Toda esta estrutura oferece entrada para uma rede neural multicamadas (MLP,

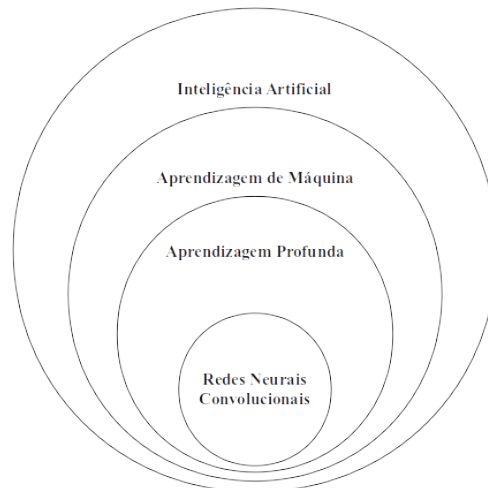


Figura 17 – Relacionamento entre inteligência artificial, aprendizagem de máquina, aprendizagem profunda e redes neurais convolucionais.

do inglês, *multilayer perceptron*) totalmente conectada e, por fim, tem-se a camada de saída. Esta arquitetura geral é ilustrada pela Figura 18 e melhor detalhada a seguir.

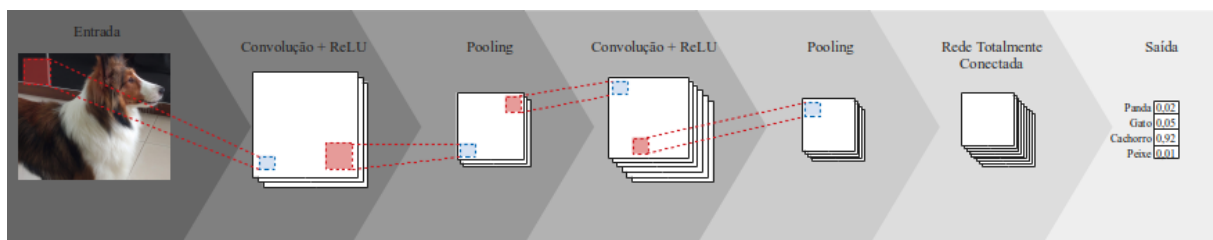


Figura 18 – Ilustração de uma rede neural convolucional.

3.2.1 Convolução

A convolução é o principal componente de uma CNN e consiste de um conjunto de filtros que são aplicados a uma imagem de entrada a fim de promover o destaque de algum tipo de característica como, por exemplo, bordas e orientação. A convolução viabiliza, então, a redução da imagem a um formato mais fácil de ser processado e procura preservar as características essenciais da imagem. A Figura 19 ilustra a convolução de uma imagem binária (composta apenas de zeros e uns) com dimensão $5 \times 5 \times 1$ com um filtro $3 \times 3 \times 1$.

3.2.2 Função de Ativação

Existem três tipos de função de ativação $\sigma(x)$ que são normalmente utilizados, a saber: a ReLU, do inglês, *rectified linear unit*, a sigmoide e a tangente hiperbólica (*tanh*). As equações seguintes detalham estas funções, respectivamente:

$$\sigma(x) = \max(0, x) \quad (3.1)$$

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Imagem binária

1	0	1
0	1	0
1	0	1

Filtro

Convolução		Resultado																																		
<table border="1" style="margin: 5px auto;"> <tr><td>1_{x1}</td><td>1_{x0}</td><td>1_{x1}</td><td>0</td><td>0</td></tr> <tr><td>0_{x0}</td><td>1_{x1}</td><td>1_{x0}</td><td>1</td><td>0</td></tr> <tr><td>0_{x1}</td><td>0_{x0}</td><td>1_{x1}</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td></tr> </table>	1 _{x1}	1 _{x0}	1 _{x1}	0	0	0 _{x0}	1 _{x1}	1 _{x0}	1	0	0 _{x1}	0 _{x0}	1 _{x1}	1	1	0	0	1	1	0	0	1	1	0	0	⇒	<table border="1" style="margin: 5px auto;"> <tr><td>4</td><td></td><td></td></tr> <tr><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td></tr> </table>	4								
1 _{x1}	1 _{x0}	1 _{x1}	0	0																																
0 _{x0}	1 _{x1}	1 _{x0}	1	0																																
0 _{x1}	0 _{x0}	1 _{x1}	1	1																																
0	0	1	1	0																																
0	1	1	0	0																																
4																																				
<table border="1" style="margin: 5px auto;"> <tr><td>1</td><td>1_{x1}</td><td>1_{x0}</td><td>0_{x1}</td><td>0</td></tr> <tr><td>0</td><td>1_{x0}</td><td>1_{x1}</td><td>1_{x0}</td><td>0</td></tr> <tr><td>0</td><td>0_{x1}</td><td>1_{x0}</td><td>1_{x1}</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td></tr> </table>	1	1 _{x1}	1 _{x0}	0 _{x1}	0	0	1 _{x0}	1 _{x1}	1 _{x0}	0	0	0 _{x1}	1 _{x0}	1 _{x1}	1	0	0	1	1	0	0	1	1	0	0	⇒	<table border="1" style="margin: 5px auto;"> <tr><td>4</td><td>3</td><td></td></tr> <tr><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td></tr> </table>	4	3							
1	1 _{x1}	1 _{x0}	0 _{x1}	0																																
0	1 _{x0}	1 _{x1}	1 _{x0}	0																																
0	0 _{x1}	1 _{x0}	1 _{x1}	1																																
0	0	1	1	0																																
0	1	1	0	0																																
4	3																																			
<table border="1" style="margin: 5px auto;"> <tr><td>1</td><td>1</td><td>1_{x1}</td><td>0_{x0}</td><td>0_{x1}</td></tr> <tr><td>0</td><td>1</td><td>1_{x0}</td><td>1_{x1}</td><td>0_{x0}</td></tr> <tr><td>0</td><td>0</td><td>1_{x1}</td><td>1_{x0}</td><td>1_{x1}</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td></tr> </table>	1	1	1 _{x1}	0 _{x0}	0 _{x1}	0	1	1 _{x0}	1 _{x1}	0 _{x0}	0	0	1 _{x1}	1 _{x0}	1 _{x1}	0	0	1	1	0	0	1	1	0	0	⇒	<table border="1" style="margin: 5px auto;"> <tr><td>4</td><td>3</td><td>4</td></tr> <tr><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td></tr> </table>	4	3	4						
1	1	1 _{x1}	0 _{x0}	0 _{x1}																																
0	1	1 _{x0}	1 _{x1}	0 _{x0}																																
0	0	1 _{x1}	1 _{x0}	1 _{x1}																																
0	0	1	1	0																																
0	1	1	0	0																																
4	3	4																																		
<table border="1" style="margin: 5px auto;"> <tr><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>0_{x1}</td><td>1_{x0}</td><td>1_{x1}</td><td>1</td><td>0</td></tr> <tr><td>0_{x0}</td><td>0_{x1}</td><td>1_{x0}</td><td>1</td><td>1</td></tr> <tr><td>0_{x1}</td><td>0_{x0}</td><td>1_{x1}</td><td>1</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td></tr> </table>	1	1	1	0	0	0 _{x1}	1 _{x0}	1 _{x1}	1	0	0 _{x0}	0 _{x1}	1 _{x0}	1	1	0 _{x1}	0 _{x0}	1 _{x1}	1	0	0	1	1	0	0	⇒	<table border="1" style="margin: 5px auto;"> <tr><td>4</td><td>3</td><td>4</td></tr> <tr><td>2</td><td></td><td></td></tr> <tr><td></td><td></td><td></td></tr> </table>	4	3	4	2					
1	1	1	0	0																																
0 _{x1}	1 _{x0}	1 _{x1}	1	0																																
0 _{x0}	0 _{x1}	1 _{x0}	1	1																																
0 _{x1}	0 _{x0}	1 _{x1}	1	0																																
0	1	1	0	0																																
4	3	4																																		
2																																				
...																																				
<table border="1" style="margin: 5px auto;"> <tr><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1_{x1}</td><td>1_{x0}</td><td>1_{x1}</td></tr> <tr><td>0</td><td>0</td><td>1_{x0}</td><td>1_{x1}</td><td>0_{x0}</td></tr> <tr><td>0</td><td>1</td><td>1_{x1}</td><td>0_{x1}</td><td>0_{x1}</td></tr> </table>	1	1	1	0	0	0	1	1	1	0	0	0	1 _{x1}	1 _{x0}	1 _{x1}	0	0	1 _{x0}	1 _{x1}	0 _{x0}	0	1	1 _{x1}	0 _{x1}	0 _{x1}	⇒	<table border="1" style="margin: 5px auto;"> <tr><td>4</td><td>3</td><td>4</td></tr> <tr><td>2</td><td>4</td><td>3</td></tr> <tr><td>2</td><td>3</td><td>4</td></tr> </table>	4	3	4	2	4	3	2	3	4
1	1	1	0	0																																
0	1	1	1	0																																
0	0	1 _{x1}	1 _{x0}	1 _{x1}																																
0	0	1 _{x0}	1 _{x1}	0 _{x0}																																
0	1	1 _{x1}	0 _{x1}	0 _{x1}																																
4	3	4																																		
2	4	3																																		
2	3	4																																		

Figura 19 – Ilustração de convolução em uma imagem 2D.

$$\sigma(x) = \frac{1}{a + e^{-x}} \tag{3.2}$$

$$\sigma(x) = \tanh(x) \tag{3.3}$$

A Figura 20 apresenta os respectivos gráficos para as funções de ativação anteriormente mencionadas.

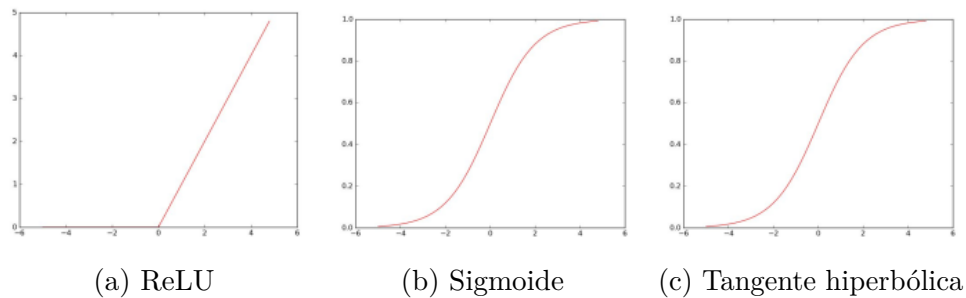


Figura 20 – Funções de ativação.

3.2.3 Pooling

Assim como as camadas de convolução, as camadas de *pooling* são responsáveis por promover a redução do domínio espacial (*downsampling*). O objetivo dessa redução é diminuir o número de parâmetros necessários para o treinamento, mantendo as características mais importantes. Além disso, as camadas de *pooling* permitem extrair características que são invariantes à translação, escala e rotação, o que favorece o processo de treinamento eficaz do modelo.

Os principais tipos de *pooling* que podem ser aplicados são o *max* e o *average*. A Figura 21 traz exemplos de *pooling*.

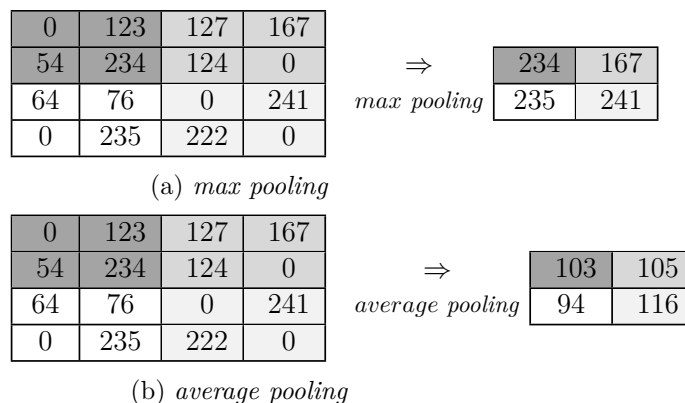


Figura 21 – Exemplos de *pooling* com passo da janela igual a dois.

3.2.4 Regularização

O sobreajuste (*overfitting*) é um problema de treinamento frequente em algoritmos de ML. Ele ocorre quando modelos excessivamente complexos acabam por ser influenciados

por ruídos durante a etapa de treinamento. Contudo, um modelo sobreajustado não irá generalizar bem para uma base de testes desconhecida e com um volume menor.

Objetivando mitigar o problema do sobreajuste, vários métodos de regularização foram propostos, a saber:

- a) **decaimento de pesos:** este método propõe a penalização dos pesos de valor elevado e foi apresentado no trabalho de [Hinton \(1987\)](#);
- b) **parada antecipada:** a ideia proposta por este método é fazer uso de um conjunto de validação independente durante o processo de treinamento para avaliar o desempenho da rede. Quando o erro no conjunto de dados de validação começa a aumentar, o treinamento é simplesmente interrompido;
- c) **expansão artificial da base:** este método propõe que os dados de treinamento sejam artificialmente expandidos através do acréscimo aleatório de dados com escalas e rotações variadas, visto que tais variações podem não estar bem representadas no conjunto de dados de treinamento. É importante ressaltar que as redes neurais normalmente devem promover o reconhecimento de padrões invariantes à escala e à rotação;
- d) **descarte (*dropout*):** este método é um dos mais utilizados e foi proposto por [Srivastava et al. \(2014\)](#). Ele preconiza o descarte aleatório de certos neurônios durante o processo de treinamento a fim de reduzir o sobreajuste. Assim, a cada iteração de treinamento, os neurônios individuais e suas conexões são descartados da rede, conforme ilustrado na Figura 22.

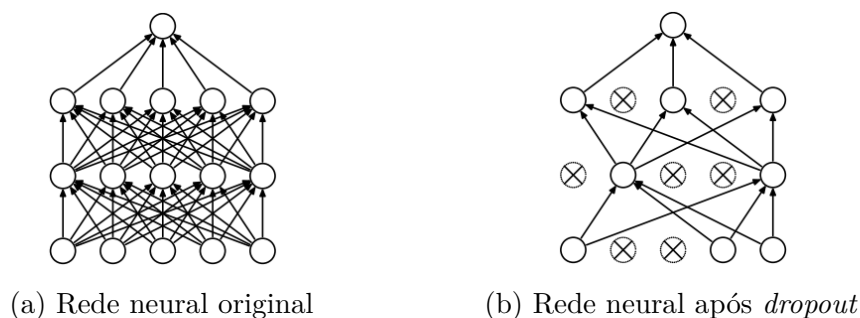


Figura 22 – Exemplo de *dropout* ([SRIVASTAVA et al., 2014](#)).

3.2.5 Camada Totalmente Conectada

A camada que constitui a rede totalmente conectada (FC, do inglês, *fully-connected layer*) é responsável por atribuir probabilidades ($p_1, p_2, p_3, \dots, p_n$) a cada uma das classes existentes no modelo, com a soma de todas as probabilidades igual a 1, ou seja $\sum_{i=1}^n p_i = 1$.

Assim sendo, em um modelo de classificação, a camada FC possui o número de neurônios igual ao número de classes.

A próxima seção traz um pouco da história das redes neurais, bem como sua evolução.

3.3 Evolução Histórica

Redes neurais clássicas como a MLP não são perfeitamente adaptadas ao trabalho com imagens digitais, pois o número de pixels de uma imagem tende a ser elevado, promovendo o correspondente aumento no número de pesos.

Através dos resultados do trabalho de [Hubel e Wiesel \(1968\)](#) foi possível observar a existência de um agrupamento de células do córtex visual humano responsáveis pela identificação de padrões de borda e, ainda, de outras células que possuem a capacidade de detectar características invariantes à escala. As observações deste trabalho apoiaram o desenvolvimento do “neocognitron”, apresentado em [Fukushima \(1980\)](#), que é tido como modelo pai das CNNs, pois já apresenta elementos como duas camadas ocultas, conectividade local e *pooling*. Contudo, o “neocognitron” não fazia uso de *backpropagation*.

A primeira CNN treinada com *backpropagation*, a LeCun, foi apresentada no trabalho de [LeCun et al. \(1989\)](#). Após praticamente uma década, o trabalho de [Lecun et al. \(1998\)](#) trouxe uma versão melhorada da LeCun, a LeNet.

O sucesso conceitual da LeNet acabou não se refletindo em aplicações mais práticas, pois as limitações de *hardware* da época se apresentavam como um obstáculo difícil de ser superado. Contudo, com a popularização da computação paralela realizada através de unidade de processamento gráfico (GPU, do inglês, *graphics processing unit*), as CNNs receberam um novo impulso e várias outras redes vêm sendo desenvolvidas desde então, conforme pode ser verificado pela linha do tempo das principais redes neurais convolucionais ilustrada na Figura 23.

A Tabela 5 apresenta o total aproximado de parâmetros e também a profundidade das redes neurais convolucionais.

Os trabalhos de [Alom et al. \(2019\)](#) e de [Khan et al. \(2020\)](#) apresentam o estado da arte de redes neurais convolucionais enfocando suas arquiteturas, principais características e aplicações.

A seguir, serão apresentadas as considerações finais deste capítulo.

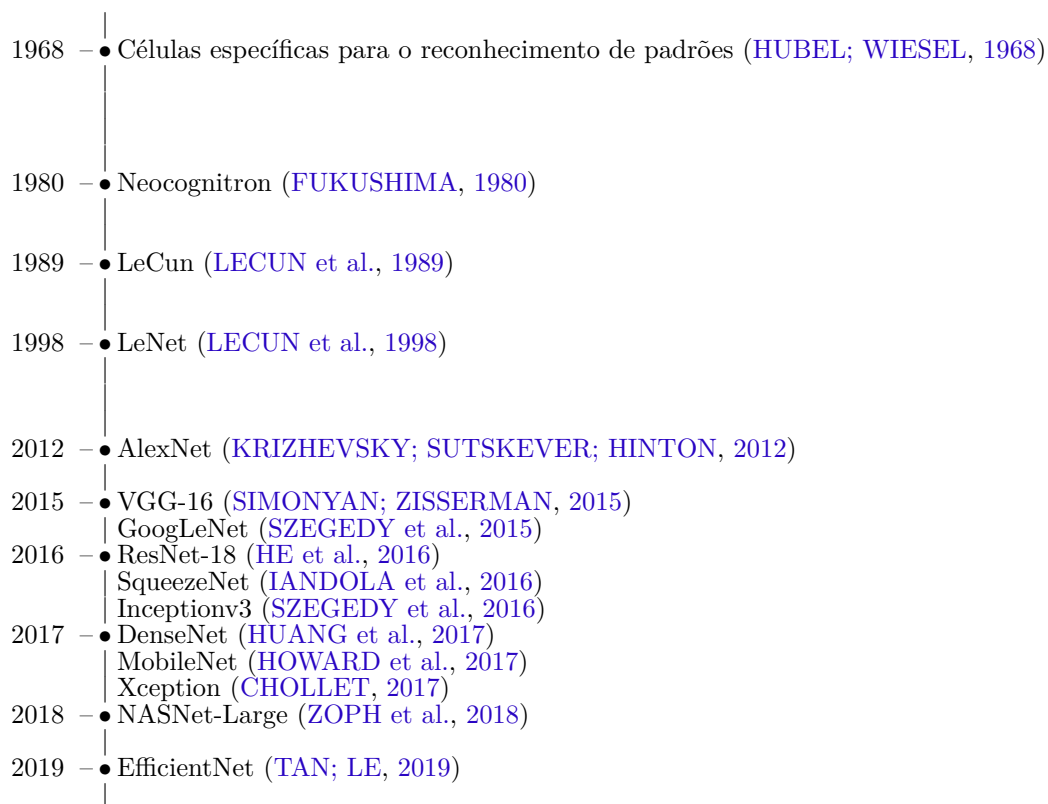


Figura 23 – Linha do tempo para as principais CNNs.

Rede Neural Convocucional	Parâmetros	Profundidade
LeNet (LECUN et al., 1998)	62 mil	5
AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012)	61 milhões	8
VGG-16 (SIMONYAN; ZISSERMAN, 2015)	138 milhões	16
GoogLeNet (SZEGEDY et al., 2015)	7 milhões	22
ResNet-18 (HE et al., 2016)	11,7 milhões	18
SqueezeNet (IANDOLA et al., 2016)	1,24 milhões	18
Inceptionv3 (SZEGEDY et al., 2016)	23,9	48
DenseNet (HUANG et al., 2017)	15 milhões	250
MobileNet (HOWARD et al., 2017)	4,2 milhões	53
Xception (CHOLLET, 2017)	22,9 milhões	71
NASNet-Large (ZOPH et al., 2018)	88,9 milhões	*
EfficientNet (TAN; LE, 2019)	7,8 milhões	**

* não possui uma sequência linear de camadas.

** apresenta arquitetura escalável.

Tabela 5 – Características das principais CNNs.

3.4 Considerações Finais do Capítulo

Neste capítulo, foram apresentados os fundamentos de redes neurais convolucionais, contemplando o funcionamento de seus componentes. Foi apresentado, também, um breve histórico das CNNs, com os respectivos trabalhos associados.

O capítulo seguinte discutirá a análise de textura e os principais descritores de textura da literatura. Serão apresentadas, ainda, as principais bases de imagens de textura.

Análise de Textura e Descritores Locais

Sumário

4.1	Introdução	58
4.2	Análise de Textura	58
4.2.1	Abordagem Estatística	60
4.2.2	Abordagem Baseada em Processamento de Sinais	61
4.2.3	Abordagem Baseada em Modelo Paramétrico	65
4.2.4	Abordagens <i>Deep Learning</i>	66
4.3	Descritores Locais de Textura	67
4.3.1	Padrão Binário Local	67
4.3.2	Padrão Binário da Mediana	69
4.3.3	Padrão Ternário Local	70
4.3.4	Padrão Ternário da Mediana	70
4.3.5	Padrão Direcional do Gradiente	73
4.3.6	Padrão Ternário Local do Gradiente	73
4.3.7	Descritor Local de Weber	74
4.3.8	Codificação Binária Melhorada de Weber	75
4.3.9	Variância do Padrão Direcional Local	77
4.3.10	Padrão de Crescimento do Gradiente Local	78
4.3.11	Padrão de Textura Direcional Local	79
4.3.12	Quantização de Fase Local	82

4.3.13	Descritor de Frequência Local	83
4.3.14	Codificação Binária Monogênica	83
4.3.15	Exemplo Prático de Descritores Locais de Textura	86
4.4	Bases de Imagens de Textura	87
4.4.1	Brodatz	87
4.4.2	CUReT	87
4.4.3	Kylberg	88
4.4.4	ALOT	88
4.4.5	DTD	89
4.5	Considerações Finais do Capítulo	90

4.1 Introdução

Este capítulo descreve a análise de textura, definindo as principais abordagens para se extrair características de textura de imagens digitais. O capítulo apresenta, ainda, detalhes do funcionamento dos principais descritores locais de textura da literatura, bem como descreve as principais bases de imagens de textura.

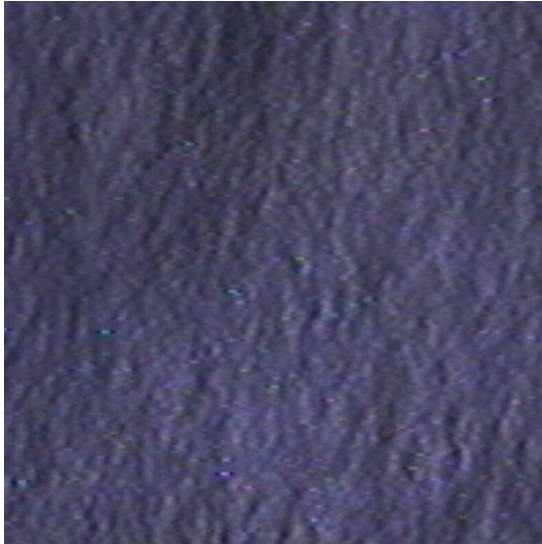
4.2 Análise de Textura

A análise da textura em imagens digitais é tradicionalmente dividida em quatro problemas: classificação, segmentação, síntese e forma da textura (TUCERYAN; JAIN, 1993). Uma etapa importante do processamento que é compartilhada pela maioria dos métodos de análise de texturas é a extração das características que possam melhor descrever as texturas (ANDREARCZYK; WHELAN, 2017).

Entretanto, o que vem a ser a textura? A textura de uma imagem é uma das características essenciais que permite ao sistema visual humano diferenciar uma imagem de outra, e vem sendo bastante estudada desde o trabalho de Julesz (1962). Além disso, a textura é um dos elementos utilizados pelo sistema visual humano que contempla informações sobre a distribuição espacial e variação de luminosidade, além de descrever o arranjo estrutural de superfícies e relações entre regiões vizinhas (PEDRINI; SCHWARTZ, 2008).

Observando-se as imagens de textura da Figura 24, extraídas de quatro bases públicas, é possível verificar que o processo de classificação da textura de uma imagem é complexo, compreendendo operações de processamento de imagens e também de visão computacional. Entretanto, a classificação automatizada da textura de uma imagem pode

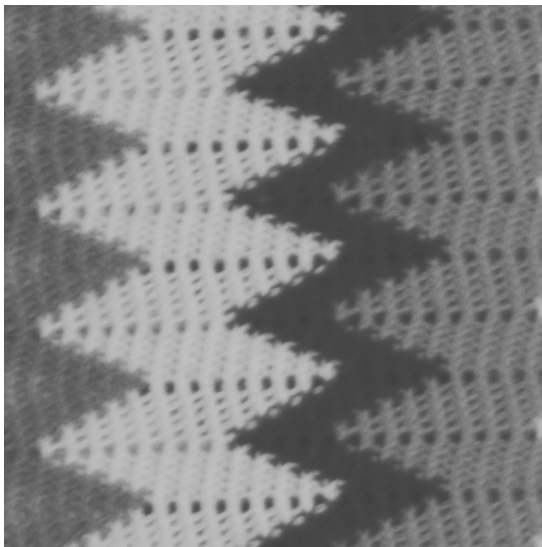
auxiliar na solução de diversos problemas como, por exemplo, o diagnóstico de doenças a partir de imagens médicas digitais (MENG et al., 2018; KAWASHIMA et al., 2019), o sensoriamento remoto (GUO; REES, 2019; GUIJARRO et al., 2015), o aumento da eficiência agrícola (MORAIS et al., 2019), dentre outros.



(a) CURET (DANA et al., 1999)



(b) ALOT (BURGHOUTS; GEUSEBROEK, 2009)



(c) Kylberg (KYLBERG, 2011)



(d) DTD (CIMPOI et al., 2014)

Figura 24 – Exemplos de imagens de textura.

Pela natureza desafiadora da textura, observa-se que a literatura da área apresenta um grande número de abordagens que objetivam analisar e extrair características de textura de imagens digitais. Embora não se verifique consenso, e apoiado nos trabalhos de Pedrini e Schwartz (2008), Andrearczyk e Whelan (2017) e Liu et al. (2019), neste trabalho as abordagens são organizadas em quatro grupos principais: estatística, baseada em processamento de sinais, baseada em modelo paramétrico e baseada em *deep learning*. A Figura 25 ilustra esta organização.

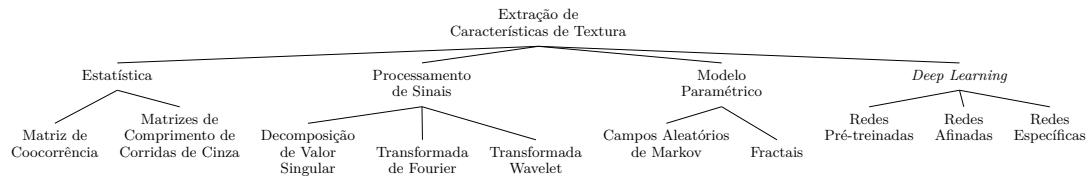


Figura 25 – Organização das abordagens de extração de características de textura.

4.2.1 Abordagem Estatística

Os métodos que fazem uso da abordagem estatística descrevem a textura através do mapeamento de relações existentes entre os níveis de cinza dos pixels pertencentes a uma imagem. Este mapeamento pode ser realizado através de medidas estatísticas de primeira ordem como média, variância, desvio padrão, assimetria e curtose.

Dentre as técnicas que utilizam abordagem estatística, destacam-se aquelas que fazem uso de matriz de coocorrência e também aquelas que utilizam matrizes de comprimento de corridas de cinza, as quais serão abordadas nas subseções seguintes.

4.2.1.1 Matriz de Coocorrência

Estatísticas de segunda ordem também podem ser utilizadas e, dessa forma, é considerada a distribuição de valores de pares de pixels através do uso de matriz de coocorrência (GLCM, do inglês, *gray level co-occurrence matrix*). Assim, considerando-se um relacionamento espacial entre pixels, a matriz de coocorrência descreve a frequência em que ocorrem as transições de nível de cinza entre os pares de pixels. A Figura 26 apresenta uma ilustração para o cálculo da matriz de coocorrência.

0	0	0	1
1	1	1	1
2	2	2	3
3	3	4	5

(a) Imagem original

	0	1	2	3	4	5
0	2	1	0	0	0	0
1	1	3	0	0	0	0
2	0	0	2	1	0	0
3	0	0	1	1	1	0
4	0	0	0	1	0	1
5	0	0	0	0	1	0

(b) GLCM

Figura 26 – Exemplo de cálculo da matriz de coocorrência baseado em (MIRMEHD; XIE; Jasjit Suri, 2008).

Um exemplo do uso de GLCM como descritor de textura pode ser encontrado no trabalho desenvolvido por Dhevana e Jayasri (2015), o qual faz uso de GLCM e de uma rede neural simplificada para detectar imagens forjadas. Já Zhang et al. (2019) utilizam GLCM e rede neural convolucional para identificar imagens forjadas e também aquelas geradas artificialmente.

4.2.1.2 Matrizes de Comprimento de Corridas de Cinza

Matrizes de comprimento de corridas de cinza (GLRLM, do inglês, *grey level run length matrices*) buscam resumir a ocorrência de um dado padrão de cinza em uma determinada direção, permitindo, desta maneira, a extração de características que permitam evidenciar o grau de aspereza de uma superfície. A Figura 27 ilustra o processo de cálculo da GLRLM.

1	2	3	4
1	3	4	4
3	2	2	2
4	1	4	1

(a) Imagem original

	Comprimento da Corrida			
	1	2	3	4
1	4	0	0	0
2	1	0	1	0
3	3	0	0	0
4	3	1	0	0

(b) GLRLM

Figura 27 – Exemplo de cálculo da GLRLM considerando a direção 0° e quatro níveis de cinza (THIBAUT et al., 2013).

É possível verificar que a abordagem estatística vem sendo muito utilizada em trabalhos recentes relacionadas à área de biomedicina (VAMVAKAS et al., 2018; CLARO et al., 2019; ABRAHAM; NAIR, 2018). Também pode-se observar que é comum a utilização de GLCM e GLRLM no mesmo *framework* de caracterização de textura, como ocorre no trabalho de Muinuddin et al. (2020), o qual faz uso da análise de textura aliada a redes neurais artificiais para identificar defeitos em couro e com as características de textura sendo extraídas das imagens através do uso de GLCM e também de GLRLM.

4.2.2 Abordagem Baseada em Processamento de Sinais

As abordagens baseadas em processamento de sinais trabalham com características da imagem extraídas a partir dos domínios da frequência e espacial. No domínio da frequência, é possível verificar abordagens que, por exemplo, fazem uso da transformada de Fourier. Já no domínio espacial, existem abordagens que usam filtragem direcional. Entretanto, também existem as abordagens que fazem uso de características do domínio espacial e da frequência ao mesmo tempo como aquelas baseadas em filtros de Gabor e *wavelets*.

Nas subseções seguintes serão descritas e exemplificadas as abordagens baseadas em processamento de sinais.

4.2.2.1 Decomposição de Valor Singular

A SVD consiste em decompor uma matriz considerando a álgebra linear conforme:

$$A = USV^T \quad (4.1)$$

onde A é uma matriz $M \times N$ com $M > N$, U é uma matriz ortogonal também $M \times N$, S é uma matriz diagonal $N \times N$ e V^T é a transposta de uma matriz ortogonal V de dimensão $N \times N$.

A Figura 28 apresenta um exemplo de aplicação de SVD em uma imagem considerando apenas os 30 primeiros valores singulares.



(a) Imagem original

(b) Imagem com os 30 primeiros valores singulares

Figura 28 – Exemplo da aplicação de SVD.

Assim, [Vazquez-Padin, Perez-Gonzalez e Comesana-Alfaro \(2017\)](#) fazem uso da SVD para detectar traços de operações de redimensionamento aplicadas sobre imagens. Já [Rosiyadi et al. \(2012\)](#) usam a transformada discreta de cosseno juntamente com SVD para implementar um esquema de proteção de *copyright* em imagens digitais.

4.2.2.2 Transformada de Fourier

A transformada de Fourier foi desenvolvida pelo francês Jean Baptiste Joseph Fourier (1768-1830) e permite a observação de uma imagem através do domínio da frequência. Assim sendo, a transformada de Fourier de uma função contínua $f(x)$ de uma variável real x é definida como:

$$F(u) = \int_{-\infty}^{\infty} f(x)e^{-j2\pi ux} dx, \text{ onde } j = \sqrt{-1} \quad (4.2)$$

a partir de $F(u)$ é possível se obter $f(x)$ através da transformada inversa de Fourier:

$$f(x) = \int_{-\infty}^{\infty} F(u)e^{j2\pi ux} du \quad (4.3)$$

A Figura 29 apresenta um exemplo de aplicação da transformada de Fourier sobre uma imagem.

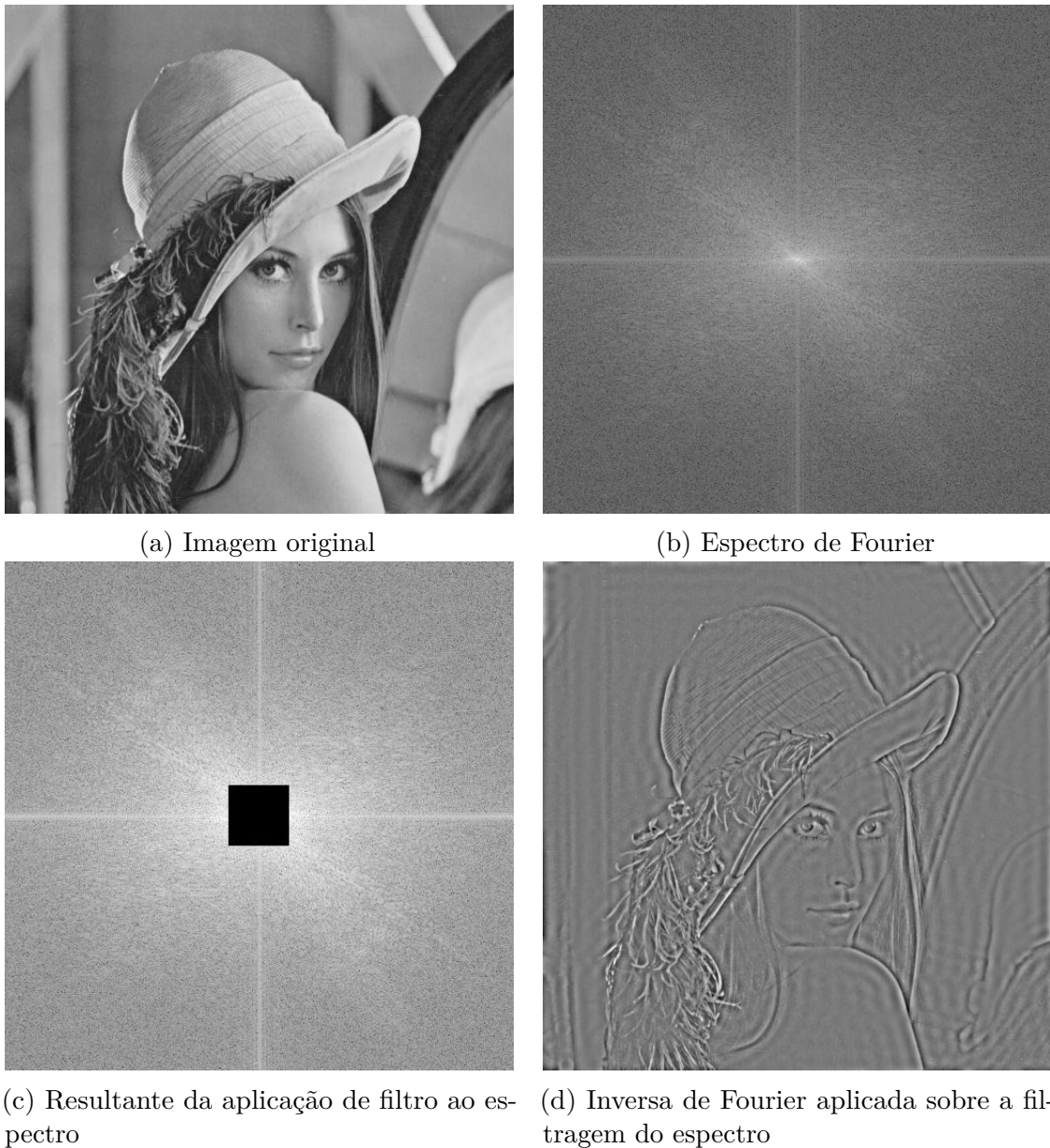


Figura 29 – Exemplo da aplicação de transformada de Fourier.

No contexto da área de forense em imagens digitais, é possível observar o exemplo de uso direto da transformada de Fourier para implementação de marcas d'água, conforme pode ser verificado no trabalho de [Urvoy, Goudia e Autrusseau \(2014\)](#) e de [Perez-Meana et al. \(2012\)](#). Além disso, [Popescu e Farid \(2005\)](#) utilizaram a transformada de Fourier no

desenvolvimento de um *framework* para identificação de imagens que foram submetidas a operações de *resampling*.

4.2.2.3 Transformada Wavelet

Diferentemente da transformada de Fourier, a transformada *wavelet* permite analisar a frequência de um sinal ao longo do tempo. Assim, pequenas mudanças de frequência na transformada de Fourier produzirão mudanças em todos os pontos no domínio do tempo. Na transformada *wavelet*, por sua vez, as mudanças são locais no domínio da frequência e do tempo. Esta característica é importante quando algumas aplicações são consideradas, como, por exemplo, processamento de vídeo (ATTO; BENOIT; LAMBERT, 2020), análise multirresolução (UNSER; SAGE; VILLE, 2009) e imagens biomédicas (IBAIDA; KHALIL, 2013).



(a) Coeficientes de aproximação



(b) Coeficientes de detalhe horizontal



(c) Coeficientes de detalhe vertical



(d) Coeficientes de detalhe diagonal

Figura 30 – Exemplo da aplicação de transformada *wavelet*.

Para exemplificar, [Dixit, Naskar e Mishra \(2017\)](#) realizam a decomposição de uma imagem através do uso de transformada *wavelet* estacionária (SWT, do inglês, *stationary wavelet transform*), e extraem características das sub-bandas SWT utilizando SVD a fim de realizar a detecção de imagens forjadas que tenham sido submetidas a operações de borramento para ocultar traços de adulteração. Já o trabalho de [Li, Duan e Zhong \(2015\)](#), utiliza o domínio Gabor-*wavelet* para propor um modelo de recuperação de textura invariante à rotação.

A Figura 30 apresenta um exemplo de decomposição *wavelet* sobre uma imagem, evidenciando os coeficientes provenientes do processo de decomposição.

4.2.3 Abordagem Baseada em Modelo Paramétrico

Na abordagem baseada em modelo paramétrico, as características de textura são obtidas a partir de um modelo de estimação de parâmetros iniciais não determinístico. As abordagens mais comuns incluem os campos aleatórios de Markov e fractais, e serão detalhadas a seguir.

4.2.3.1 Campos Aleatórios de Markov

A abordagem baseada em campos aleatórios de Markov (MRF, do inglês, *Markov random fields*) considera a propriedade de Markov que estabelece que o estado futuro de um sistema depende apenas de seu estado corrente ([MARKOV, 1954](#)). Esta propriedade aplicada a imagens digitais estabelece uma relação local de dependência entre os pixels.

Modelos baseados em MRF consideram, então, que a intensidade de um pixel depende apenas das intensidades dos pixels de sua vizinhança ([KASHYAP; CHELLAPPA; KHOTANZAD, 1982](#)). Assim, conforme descrito por [Chen e Huang \(1993\)](#), através do uso de MRF é possível obter a direcionalidade e a aspereza presentes nas texturas.

Diante disso, [Chierchia et al. \(2014\)](#) propuseram um modelo de identificação de imagens falsificadas baseado na não uniformidade da foto-resposta, normalmente oferecido pelas câmeras fotográficas, definindo o problema em termos de uma estimativa bayesiana e usando, em seguida, um campo aleatório de Markov para modelar as fortes dependências espaciais da imagem. Já [Ravi et al. \(2014\)](#) propuseram uma técnica para identificar vídeos MPEG falsificados baseada em campos aleatórios de Huber-Markov (HMRF, do inglês, *Huber Markov random fields*)

4.2.3.2 Fractais

Os fractais são baseados no conceito de auto-similaridade entre escalas, possuindo padrões repetidos em várias escalas. Eles foram utilizados como descritores de textura por [Mandelbrot \(1983\)](#), [Peleg et al. \(1984\)](#) e, em um primeiro momento, acabaram não

sendo muito usados em outros trabalhos, pois a dimensão fractal não captura informações estruturais suficientes, uma vez que texturas com aparências diferentes podem ter dimensões fractais semelhantes, caracterizando, assim, o seu baixo poder discriminativo (KELLER; CHEN; CROWNOVER, 1989; VARMA; GARG, 2007).

Recentemente, os autores Xu, Ji e Fermüller (2009) revisitaram a análise de textura através do modelo de fractais e propuseram um método de caracterização da textura chamado de espectro multi-fractal (MFS, do inglês, *multifractal spectrum*), o qual é invariante a mudanças de ponto de vista e a deformações não rígidas. O trabalho de Quan et al. (2014) explorou o conceito de lacunaridade na geometria dos fractais para propor um modelo estatístico de descrição da textura, alcançando bons resultados na classificação de textura em diversos bancos de dados de textura públicos.

4.2.4 Abordagens *Deep Learning*

Vários trabalhos de representação de textura baseados em *deep learning* têm sido propostos desde que Krizhevsky, Sutskever e Hinton (2012) conseguiram vencer o desafio de classificação de imagens ImageNet no ano de 2012 (RUSSAKOVSKY et al., 2015). Observa-se que o principal algoritmo de *deep learning* utilizado pelos trabalhos é o CNN e, segundo Liu et al. (2019), os métodos de representação de textura baseados em CNN podem ser organizados naqueles que fazem uso de redes CNN pré-treinadas, aqueles que utilizam CNNs afinadas e aqueles que usam CNNs específicas.

Os métodos que se baseiam em CNNs pré-treinadas surgem a partir da verificação de que o processo de transferência da aprendizagem de uma rede neural treinada com uma base de imagens muito grande para outros problemas é muito simplificado e pode oferecer bons resultados (CIMPOI et al., 2014; NAPOLETANO, 2017). Assim, o trabalho de Claro et al. (2019) utiliza redes pré-treinadas VGG-16 em um modelo de classificação de tipos de glaucoma a partir da análise da textura de imagens médicas.

As principais redes pré-treinadas utilizadas incluem AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), ResNet (HE et al., 2013), VGG-16 (SIMONYAN; ZISSERMAN, 2015), GoogLeNet (SZEGEDY et al., 2015) e DenseNet (HUANG et al., 2017).

Os bancos de dados de imagens normalmente utilizados pelas redes pré-treinadas incluem o ImageNet *image dataset*, evidenciado em Russakovsky et al. (2015), com mais de um milhão de imagens organizadas em mil classes diferentes, e o MIT Places - *the scene recognition database* que possui mais dez milhões de cenas distribuídas em 439 categorias, mostrado em Zhou et al. (2018).

Dessa forma, é possível observar que as redes pré-treinadas são capazes de discriminar imagens de diferentes objetos ou classes, mas podem, contudo, ser menos eficazes na

diferenciação de diferentes texturas, pois as imagens dos bancos de dados normalmente utilizados (ImageNet e MIT Places) podem conter diferentes tipos de texturas em uma única imagem.

Assim, as redes afinadas partem do princípio de promover ajustes na arquitetura das redes pré-treinadas a fim de que elas possam ser utilizadas especificamente para o reconhecimento de textura e possam, também, ser treinadas com bancos de dados de textura com um montante de imagens bem mais reduzido.

O trabalho de [Andrearczyk e Whelan \(2016\)](#) realiza um afinamento na AlexNet e propõe a rede neural convolucional de textura (T-CNN, do inglês, *texture convolutional neural network*) para o reconhecimento de textura. Os autores [Lin, Roychowdhury e Maji \(2018\)](#), por sua vez, também propõem alterações na arquitetura da AlexNet para formar a rede neural convolucional bilinear (BCNN, do inglês, *bilinear convolutional neural networks*).

O desenvolvimento de arquiteturas de redes CNN específicas (*handcrafted deep convolutional networks*) para o reconhecimento de textura também tem sido avaliado ([LIU et al., 2019](#)). Assim, [Bruna e Mallat \(2013\)](#) propõem uma rede chamada de ScatNet, do inglês *scattering network*, que faz uso de *wavelet* para realizar filtragens que preservam as informações de alta frequência das imagens. O trabalho de [Chan et al. \(2015\)](#) propõe uma CNN que faz uso da análise de componentes principais, a PCANet.

A seção seguinte discutirá os principais descritores de textura utilizados pela literatura.

4.3 Descritores Locais de Textura

Os descritores locais de textura têm como objetivo fornecer representações locais de textura invariantes ao contraste, rotação, escala etc. Estes descritores consideram normalmente as características de magnitude, orientação e fase do sinal. Os principais descritores da literatura podem ser organizados conforme apresentado pela Figura 31.

A Figura 32 apresenta a linha do tempo dos principais descritores de textura evidenciados na Figura 31.

As seções seguintes descrevem com mais detalhes os descritores de textura anteriormente mencionados.

4.3.1 Padrão Binário Local

O descritor de padrão binário local (LBP, do inglês, *local binary pattern*) foi apresentado no trabalho de [Ojala, Pietikäinen e Harwood \(1996\)](#) e caracteriza-se como um descritor de textura para imagens em nível de cinza que é invariante à rotação da

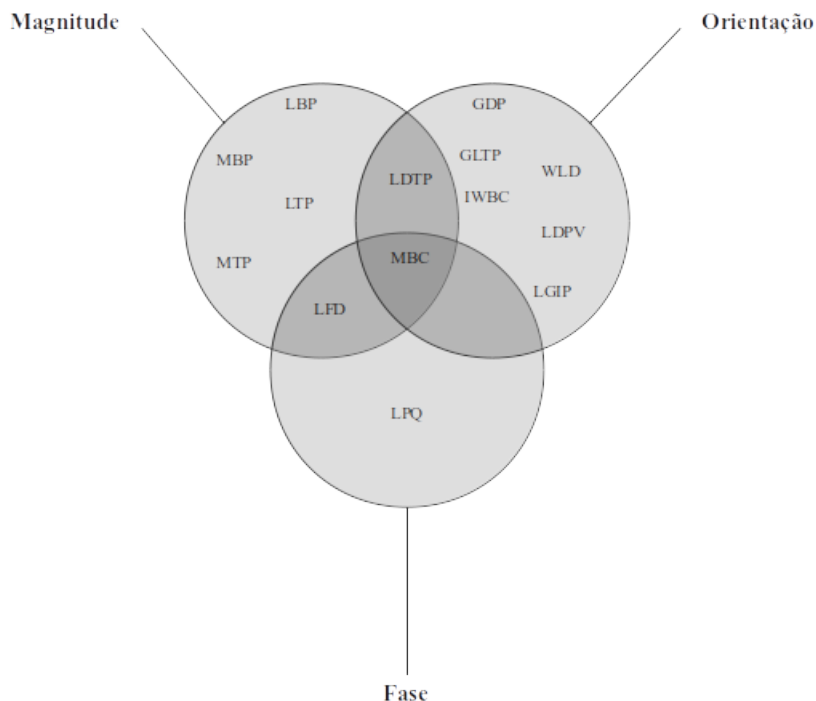


Figura 31 – Organização dos principais descritores de textura da literatura.

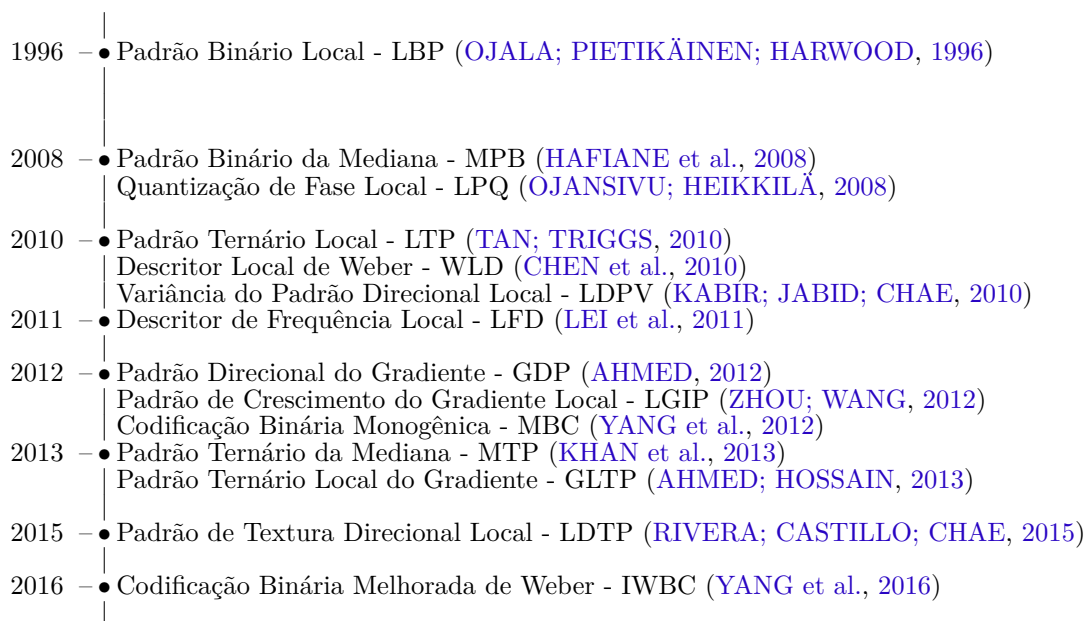


Figura 32 – Linha do tempo dos principais descritores de textura.

imagem. Ele se baseia no reconhecimento de certos padrões locais binários e possui grande importância para a caracterização de textura.

O LBP foi originalmente definido considerando-se uma janela de 3×3 , onde o valor do descritor é dado pela somatória das diferenças entre o pixel central e os pixels da

vizinhança, conforme apresentado pela equação a seguir:

$$LBP_{(x_c, y_c)} = \sum_{p=0}^7 s(g_p - g_c) 2^p \quad (4.4)$$

onde o nível de cinza g_c equivale ao nível de cinza do píxel do centro de uma vizinhança local de 8 píxels, e $g_p (p = 0, \dots, 7)$ corresponde aos valores dos níveis de cinza da vizinhança (x_c, y_c) . Além disso, temos o valor de $s(x)$ definido como:

$$s(x) = \begin{cases} 1, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases} \quad (4.5)$$

A Figura 33 ilustra o funcionamento básico do processo de cálculo do descritor LBP.

$$\begin{array}{|c|c|c|} \hline 95 & 65 & 50 \\ \hline 55 & 70 & 74 \\ \hline 85 & 40 & 17 \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & & 1 \\ \hline 1 & 0 & 0 \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|c|} \hline s(x_1) & s(x_2) & s(x_3) \\ \hline s(x_0) & & s(x_4) \\ \hline s(x_7) & s(x_6) & s(x_5) \\ \hline \end{array} \Rightarrow LBP = 10010010_2 = 146$$

Figura 33 – Funcionamento do descritor LBP.

Dessa forma, o LBP codifica primitivas de textura locais em um padrão binário, levando-se em consideração o valor do píxel central da janela (AHMED; HOSSAIN, 2013), o qual acaba atuando como *threshold*. Além disso, o LBP é um descritor de textura considerado simples de ser implementado, possuindo um bom desempenho computacional. Entretanto, ele possui uma limitação no que se refere à sensibilidade a ruídos, uma vez que uma pequena alteração no nível de intensidade do píxel central, altera inteiramente o código binário resultante (TAN; TRIGGS, 2010).

4.3.2 Padrão Binário da Mediana

O trabalho de Hafiane et al. (2008) propôs o descritor de padrão binário da mediana (MBP, do inglês, *median binary pattern*), o qual realiza o mapeamento da intensidade local para um código binário, considerando como *threshold* o valor da mediana da janela, conforme apresentado pela equação seguinte:

$$MBP(i, j) = \sum_{k=0}^{L-1} f(a_k) 2^k \quad (4.6)$$

onde L representa o tamanho da vizinhança, (i, j) definem a posição do centro da janela, a_k é a intensidade do píxel e $f(a_k)$ compreende a função de quantização que trabalha

conforme evidenciado a seguir:

$$f(a_k) = \begin{cases} 1, & \text{se } a_k \geq \text{Mediana Local} \\ 0, & \text{para qualquer outro caso} \end{cases} \quad (4.7)$$

A Figura 34 apresenta um exemplo do cálculo do MBP.

120	130	135
120	127	88
90	30	43

 \Rightarrow

1	1	1
1	1	0
0	0	0

 \Rightarrow

a_3	a_2	a_1
a_4	a_8	a_0
a_5	a_6	a_7

 $\Rightarrow MBP = 100011110_2 = 286$

Mediana = 120

Figura 34 – Processo de cálculo do MBP (HAFIANE et al., 2008).

Observa-se, então, que o descritor MBP apresenta invariabilidade às alterações eventualmente realizadas na escala da imagem.

4.3.3 Padrão Ternário Local

A fim de desenvolver um descritor de textura mais discriminante e menos sensível às variações de ruído e de iluminação, os autores Tan e Triggs (2010) introduziram o descritor de padrão ternário local (LTP, do inglês, *local ternary pattern*), que constitui uma generalização do LBP, que foi discutido na Seção 4.3.1.

No descritor LTP os valores de $s(x)$ são quantizados para três valores $(1, 0, -1)$, e não apenas $(1, 0)$ como no LBP. Assim sendo, temos:

$$s(x) = \begin{cases} 1, & \text{se } x \geq g_c + t \\ 0, & \text{se } |x - g_c| < t \\ -1, & \text{se } x \leq g_c - t \end{cases} \quad (4.8)$$

onde (t) é um *threshold* especificado pelo usuário.

Por fim, o trabalho de Shaik e Mahaboob (2017) apresenta uma interessante comparação de performance de reconhecimento de expressões faciais dos descritores LBP e LTP, tendo concluído que o LTP apresentou realmente, para o caso analisado, uma performance superior.

4.3.4 Padrão Ternário da Mediana

Tendo em vista que os filtros de mediana são efetivos na presença de ruídos (GONZALEZ; WOODS, 2015), os autores Khan et al. (2013) propuseram o descritor do padrão ternário da mediana (MTP, do inglês, *median ternary pattern*) que faz uso do valor

mediano das células vizinhas em torno de cada pixel, quantizando os valores em três níveis diferentes.

O cálculo do MTP é iniciado com a definição da vizinhança de cada um dos pixels da imagem para uma janela 3×3 e, então, é realizado o cálculo da mediana M_c da janela. Em seguida, é realizada a quantização seguinte:

$$S_{MTP}(v) = \begin{cases} 1, & v > M_c + t \\ 0, & M_c - t \leq v \leq M_c + t \\ -1, & v < M_c - t \end{cases} \quad (4.9)$$

onde v é o nível de cinza da vizinhança, M_c é a mediana local e t é um *threshold* definido.

Na sequência, são calculados os valores do MTP positivo (P_{MTP}) e negativo (N_{MTP}) conforme apresentado a seguir:

$$P_{MTP} = \sum_{p=0}^7 S_P(S_{MTP}(i_p)) \times 2^p \quad (4.10)$$

onde i_p corresponde ao nível de intensidade do pixel de uma janela e a quantização de S_p ocorre conforme:

$$S_P = \begin{cases} 1, & v = 1 \\ 0, & \text{para qualquer outro caso} \end{cases} \quad (4.11)$$

$$N_{MTP} = \sum_{p=0}^7 S_N(S_{MTP}(i_p)) \times 2^p \quad (4.12)$$

com a quantização de N_P ocorrendo conforme:

$$N_P = \begin{cases} 1, & v = -1 \\ 0, & \text{para qualquer outro caso} \end{cases} \quad (4.13)$$

A Figura 35 apresenta um exemplo de cálculo dos códigos MTP , P_{MTP} e N_{MTP} para uma janela de uma imagem.

Por fim, para a formação do descritor de textura MTP, ocorre a concatenação dos histogramas P_{MTP} e N_{MTP} , conforme ilustrado na Figura 36.

Diante disso, é possível observar que o descritor MTP procura combinar as vantagens da filtragem pela mediana e da quantização em três níveis diferentes.

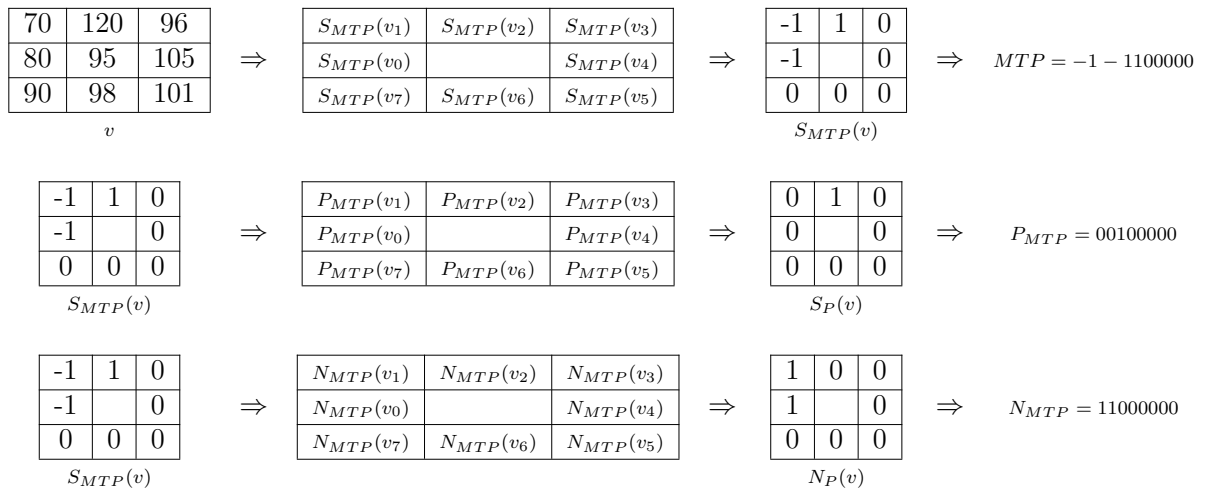


Figura 35 – Processo de cálculo do MTP (KHAN et al., 2013).

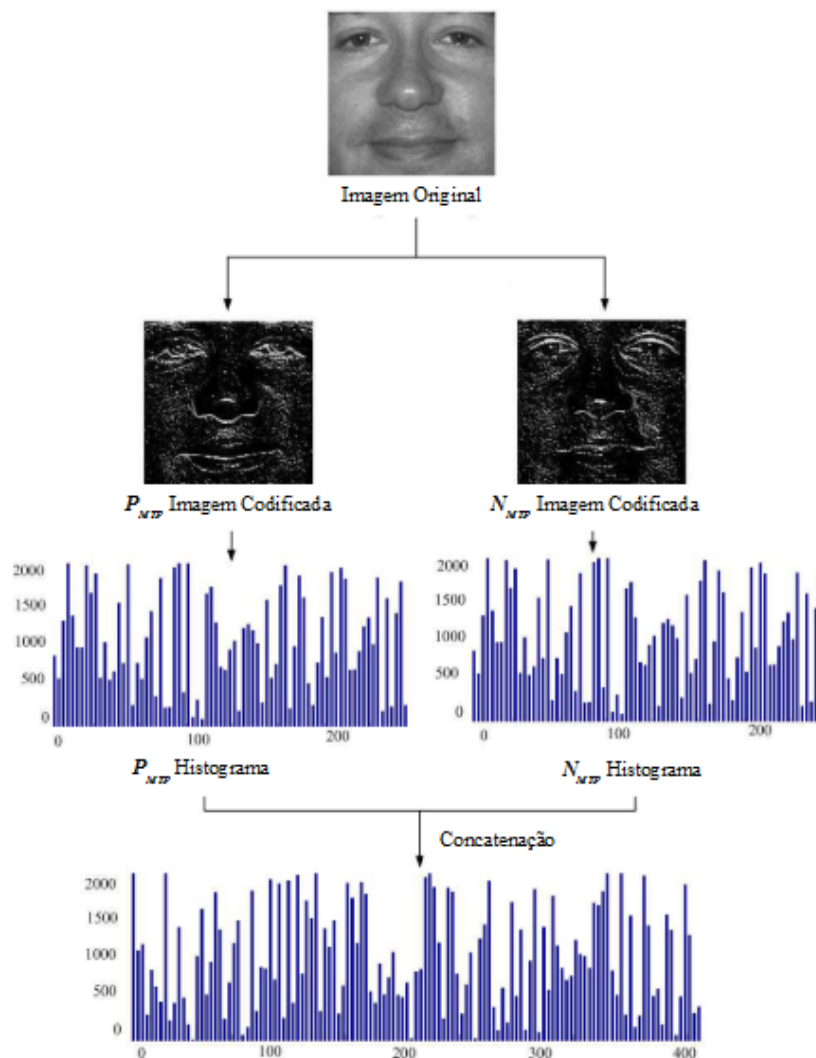


Figura 36 – Ilustração da geração do descritor MTP (KHAN et al., 2013).

4.3.5 Padrão Direcional do Gradiente

O descritor do padrão direcional do gradiente (GDP, do inglês, *gradient directional pattern*), proposto por [Ahmed \(2012\)](#), codifica a informação de textura quantizando os ângulos da direção do gradiente dos pixels (α) da imagem, a fim de reter ainda mais informações locais da imagem. Assim sendo, o GDP busca, primeiramente, calcular a direção do gradiente para cada pixel da imagem, conforme a equação seguinte:

$$\alpha(x, y) = \tan^{-1}(G_y/G_x) \quad (4.14)$$

onde G_x e G_y são obtidos através de uma operação de convolução das máscaras de Sobel com a imagem. A Figura 37 mostra as máscaras de Sobel.

Em seguida, após o cômputo dos ângulos de direção do gradiente (α), o GDP é calculado conforme equação a seguir:

$$GDP(x_c, y_c) = \sum_{p=0}^{P-1} s(GD_p, GD_c)2^p \quad (4.15)$$

onde GD_c é o ângulo de direção do gradiente do pixel central da janela (x_c, y_c) e GD_p representa os ângulos para a sua vizinhança.

No descritor em discussão os valores de $s(GD_p, GD_c)$ são quantizados conforme um *threshold* (t), especificado pelo usuário, segundo a equação seguinte:

$$s(GD_p, GD_c) = \begin{cases} 1, & \text{se } GD_c - t \leq GD_p \leq GD_c + t \\ 0, & \text{para qualquer outro caso} \end{cases} \quad (4.16)$$

Assim sendo, é possível observar que o descritor GDP busca extrair mais informações de borda da imagem para caracterizar melhor a sua textura.

4.3.6 Padrão Ternário Local do Gradiente

O trabalho desenvolvido por [Ahmed e Hossain \(2013\)](#) apresenta o descritor do padrão local ternário do gradiente (GLTP, do inglês, *gradient local ternary pattern*), o qual codifica a informação de textura local quantizando os valores de magnitude do gradiente de uma vizinhança local, usando três níveis de discriminação diferentes. O GLTP combina as vantagens do Sobel-LBP, proposto por [Zhao, Gao e Zhang \(2008\)](#), e do LTP, discutido na Seção 4.3.3.

O GLTP, primeiramente, calcula as magnitudes do gradiente $G_{x,y}$ para cada pixel da imagem, conforme a equação seguinte:

$$G_{x,y} = |G_x| + |G_y| \quad (4.17)$$

onde G_x e G_y são dois elementos do vetor gradiente e podem ser obtidos pela aplicação do operador de Sobel na imagem, o qual realiza uma operação de convolução da imagem com as máscaras apresentadas na Figura 37.

$$x = \begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \quad y = \begin{array}{|c|c|c|} \hline -1 & 2 & -1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

Figura 37 – Máscaras de Sobel.

No descritor GLTP, os valores de $s(x)$ também são quantizados para três valores $(1, 0, -1)$, conforme apresentado na Equação 4.18 evidenciada a seguir:

$$s(G_c, G_p) = \begin{cases} 1, & \text{se } G_p > G_c + t \\ 0, & \text{se } G_c - t \leq G_p \leq G_c + t \\ -1, & \text{se } G_p < G_c - t \end{cases} \quad (4.18)$$

onde G_c é a magnitude do gradiente do centro da vizinhança de uma janela 3×3 , G_p corresponde à magnitude do gradiente dos pixels em volta do centro da janela e t é um *threshold* definido pelo usuário.

O GLTP se propõe, portanto, a definir um padrão de textura que seja consistente com regiões de imagem suaves ou até mesmo altamente texturizadas.

4.3.7 Descritor Local de Weber

A Lei de Weber baseia-se no fato de que a percepção humana de um padrão depende não apenas da mudança de um estímulo, mas também da intensidade original do estímulo. Assim, ela descreve que a razão entre o limiar de incremento (ΔI) e a intensidade de fundo da cena I é uma constante k , conforme explicitado pela equação seguinte:

$$\frac{\Delta I}{I} = k \quad (4.19)$$

Dessa forma, temos que, para um determinado estímulo, a razão entre a menor mudança perceptível e o *background* da imagem é uma constante. Assim, a Lei de Weber permite verificar que os estímulos sensoriais são percebidos não em termos absolutos, mas em termos relativos. Diante dessa constatação, os autores [Chen et al. \(2010\)](#) propuseram

o descritor local de Weber (WLD, do inglês, *Weber local descriptor*) inspirados na referida Lei, introduzindo, então, em sua proposta, dois componentes para descrever a textura de uma imagem, a saber: a excitação diferencial (ε) e a orientação (Θ). A Figura 38 ilustra o processo de cálculo do WLD, sendo possível observar que a excitação diferencial (ε) e a orientação (Θ) são dadas pelas Equações 4.20 e 4.21.

$$\varepsilon(x_c) = \arctan \left[\sum_{i=0}^{p-1} \left(\frac{x_i - x_c}{x_c} \right) \right] \quad (4.20)$$

onde x_c é o pixel central da janela e os valores de x_i representam os pixels da vizinhança.

$$\Theta_t = \frac{2t}{T}\pi, \quad e \quad t = \text{mod} \left(\left\lfloor \frac{\theta'}{2\pi/T} + \frac{1}{2} \right\rfloor, T \right) \quad (4.21)$$

Assim, conforme observado por [Ding, Wang e Tao \(2018\)](#), a excitação diferencial acaba por refletir a intensidade da variância da textura local, e a orientação, por sua vez, representa a direção do gradiente do pixel central.

Segundo [Yang et al. \(2016\)](#), o descritor WLD acaba sendo tão rápido quanto o LBP e também bastante robusto. No entanto, o LBP ignora diferentes orientações dos pares de pixels vizinhos e, portanto, não pode descrever a variação da intensidade local de maneira adequada e precisa.

4.3.8 Codificação Binária Melhorada de Weber

Diante das deficiências apresentadas pelo WLD, os autores [Yang et al. \(2016\)](#) propuseram o descritor de codificação binária melhorada de Weber (IWBC, do inglês, *improved Weber binary coding*), o qual procura representar padrões locais com mais eficácia e precisão, introduzindo diferentes componentes de magnitude e orientação.

Nesse sentido, os autores definiram magnitude de Weber em (x, y) conforme pode ser verificado nas equações seguintes:

$$\varepsilon_{m-x}(x_c) = \arctan \left(\alpha \sum_{i=0}^{p-1} \frac{x_i - x_c}{x_c} \cos\theta_i \right) \quad (4.22)$$

$$\varepsilon_{m-y}(x_c) = \arctan \left(\alpha \sum_{i=0}^{p-1} \frac{x_i - x_c}{x_c} \sin\theta_i \right) \quad (4.23)$$

onde θ_i denota o ângulo entre a direção x e a excitação diferencial $(x_i - x_c)/x_c$.

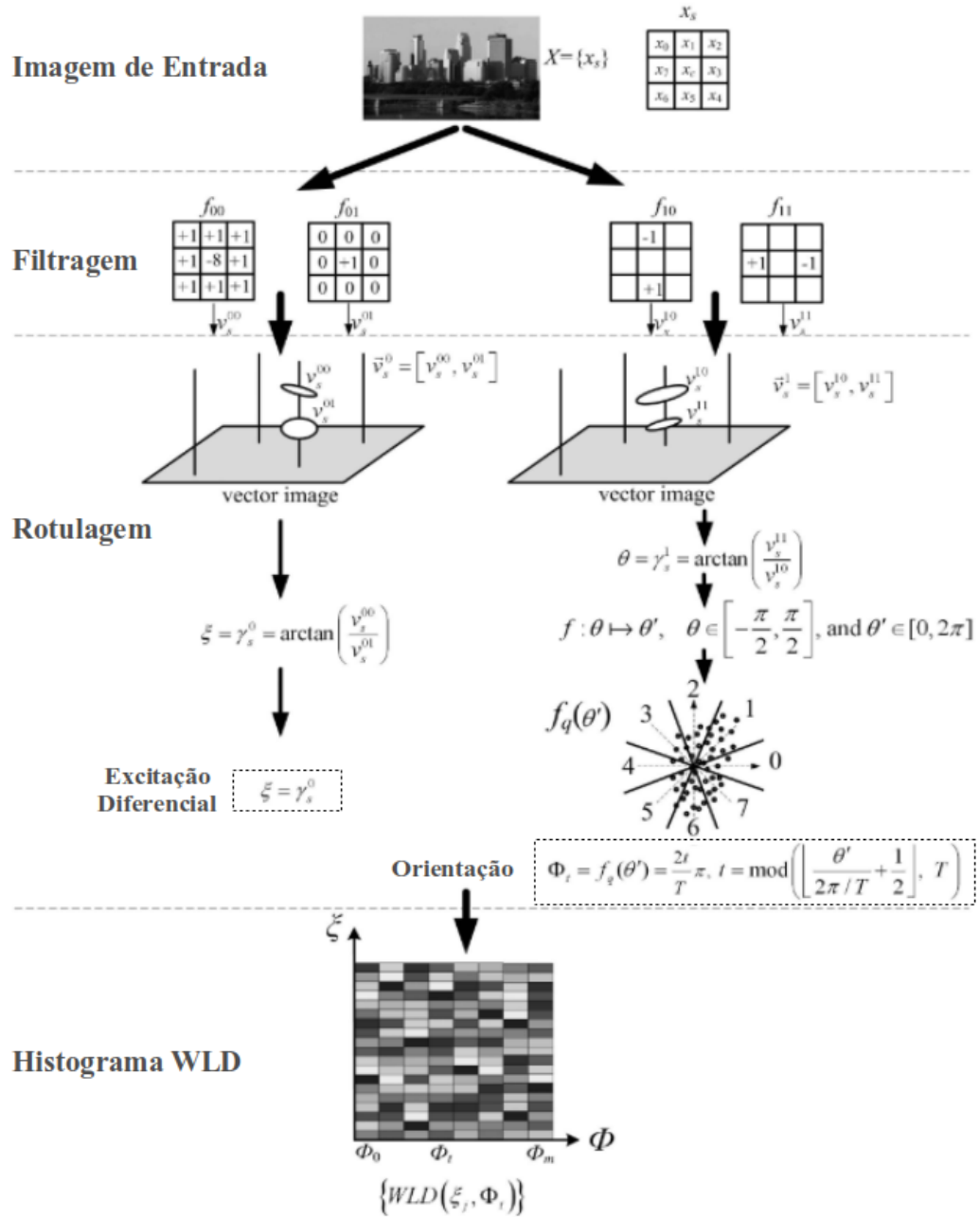


Figura 38 – Processo de cálculo do WLD (CHEN et al., 2010).

Assim, os novos componentes de Weber (magnitude ε_m e orientação ε_o) passam a ser definidos conforme as equações a seguir:

$$\varepsilon_m(x_c) = \sqrt{\varepsilon_{m-x}(x_c)^2 + \varepsilon_{m-y}(x_c)^2} \quad (4.24)$$

$$\varepsilon_o(x_c) = \arctan\left(\frac{\varepsilon_{m-y}(x_c)}{\varepsilon_{m-x}(x_c)}\right) \quad (4.25)$$

Dessa forma, os autores apresentam o seu descritor de textura como sendo mais eficaz e mais preciso que o WLD.

4.3.9 Variância do Padrão Direcional Local

O trabalho de [Kabir, Jabid e Chae \(2010\)](#) apresentou o descritor da variância do padrão direcional local (LDPV, do inglês, *local directional pattern variance*) que introduz a variação local de respostas direcionais para codificar as informações de contraste dentro do descritor.

Para o cálculo do LDPV, primeiramente se faz necessário o cálculo do padrão direcional local (LDP, do inglês, *local directional pattern*), proposto por [Jabid, Kabir e Chae \(2010\)](#). Assim, para o LDP, dado um pixel central da janela 3×3 , são computadas as oito respostas direcionais (m_0, m_1, \dots, m_7) com base nas máscaras de Kirsch (M_i) apresentadas por [Kirsch \(1971\)](#) e exibidas na Figura 39.

$$\begin{array}{cccc}
 M_0 = \begin{array}{|c|c|c|} \hline -3 & -3 & 5 \\ \hline -3 & 0 & 5 \\ \hline -3 & -3 & 5 \\ \hline \end{array} &
 M_1 = \begin{array}{|c|c|c|} \hline -3 & 5 & 5 \\ \hline -3 & 0 & 5 \\ \hline -3 & -3 & -3 \\ \hline \end{array} &
 M_2 = \begin{array}{|c|c|c|} \hline 5 & 5 & 5 \\ \hline -3 & 0 & -3 \\ \hline -3 & -3 & -3 \\ \hline \end{array} &
 M_3 = \begin{array}{|c|c|c|} \hline 5 & 5 & -3 \\ \hline 5 & 0 & -3 \\ \hline -3 & -3 & -3 \\ \hline \end{array} \\
 M_4 = \begin{array}{|c|c|c|} \hline 5 & -3 & -3 \\ \hline 5 & 0 & -3 \\ \hline 5 & -3 & -3 \\ \hline \end{array} &
 M_5 = \begin{array}{|c|c|c|} \hline -3 & -3 & -3 \\ \hline 5 & 0 & -3 \\ \hline 5 & 5 & -3 \\ \hline \end{array} &
 M_6 = \begin{array}{|c|c|c|} \hline -3 & -3 & -3 \\ \hline -3 & 0 & -3 \\ \hline 5 & 5 & 5 \\ \hline \end{array} &
 M_7 = \begin{array}{|c|c|c|} \hline -3 & -3 & -3 \\ \hline -3 & 0 & 5 \\ \hline -3 & 5 & 5 \\ \hline \end{array}
 \end{array}$$

Figura 39 – Máscaras de Kirsch em todas as oito direções.

Então, o descritor LDP é calculado conforme as equações seguintes:

$$LDP_k = \sum_{i=0}^7 b_i(m_i - m_k)2^i \quad (4.26)$$

$$b_i(a) = \begin{cases} 1, & \text{se } a \geq 0 \\ 0, & \text{se } a < 0 \end{cases} \quad (4.27)$$

onde m_k é a k – ésima resposta direcional mais significativa.

Após o cálculo do LDP para cada um dos pixels (r, c) de uma imagem (I) , de tamanho $M \times N$, temos o histograma resultante (H) computado conforme descrito pela seguinte equação:

$$H(\tau) = \sum_{r=1}^M \sum_{c=1}^N f(LDP_k(r, c), \tau) \quad (4.28)$$

onde τ é o código LDP e a função $f(a, \tau)$ é definida conforme a seguir:

$$f(a, \tau) = \begin{cases} 1, & a = \tau \\ 0, & \text{para qualquer outro caso} \end{cases} \quad (4.29)$$

Por fim, o descritor LDPV é calculado conforme apresentado pela equação:

$$LDPV(\tau) = \sum_{r=1}^M \sum_{c=1}^N \omega(LDP_k(r, c), \tau) \quad (4.30)$$

onde função ω é definida como:

$$\omega(LDP_k(r, c), \tau) = \begin{cases} \sigma(LDP_k(r, c)), & \text{para } LDP_k(r, c) = \tau \\ 0, & \text{para qualquer outro caso} \end{cases} \quad (4.31)$$

com σ definido conforme a seguir:

$$\sigma(LDP_k(r, c)) = \frac{1}{8} \sum_{i=0}^7 (m_i - \bar{m})^2 \quad (4.32)$$

onde \bar{m} é a média de todas as respostas direcionais m_i calculada para cada uma das posições (r, c) da imagem.

Os autores [Kabir, Jabid e Chae \(2010\)](#) observam que o poder discriminativo do descritor LDPV reside principalmente na integração do padrão de resposta de borda local e das informações de contraste, tornando-o, dessa maneira, robusto a ruídos e alterações de iluminação.

4.3.10 Padrão de Crescimento do Gradiente Local

Os autores [Zhou e Wang \(2012\)](#) propuseram o descritor do padrão de crescimento do gradiente local (LGIP, do inglês, *local gradient increasing pattern*), o qual expressa a direção e magnitude da tendência de crescimento da intensidade local, considerando oito direções em cada pixel e, em seguida, um código decimal é designado para descrever a tendência de crescimento global.

Para o cálculo do LGIP, primeiramente, as máscaras de Sobel são aplicadas na imagem a fim de se obter o gradiente em oito direções (M_0, M_1, \dots, M_7) para cada um dos pixels da imagem, conforme mostrado pela Figura 40.

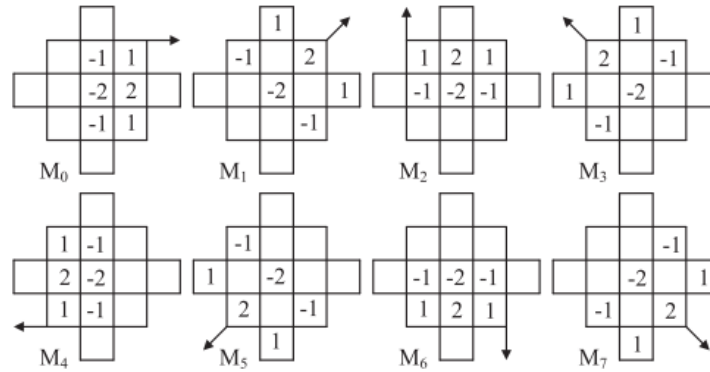


Figura 40 – Máscaras de gradiente local.

Na sequência, são calculados os vetores de tendência de crescimento da intensidade (TVs, do inglês, *trend vectors*) e o vetor de tendência de crescimento geral (OTV, do inglês, *overall trend vector*), conforme apresentado na Figura 41.

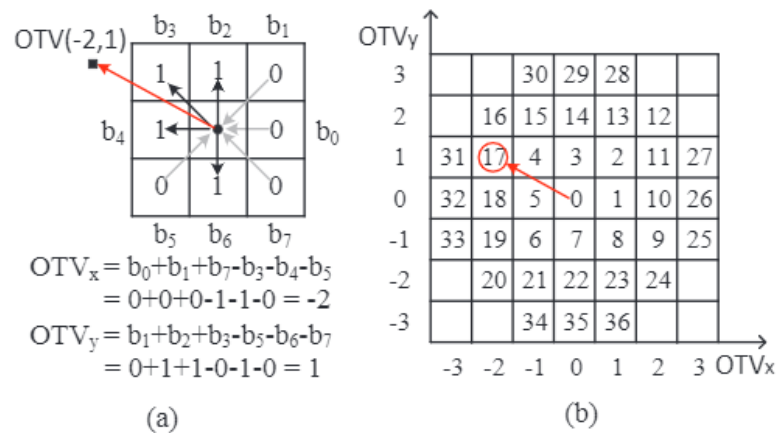


Figura 41 – Processo de cálculo do descritor LGIP (ZHOU; WANG, 2012).

O descritor LGIP acaba por combinar os pontos fortes da análise do gradiente, dos padrões uniformes e do LDP, possuindo boa estabilidade a ruídos e variações de iluminação.

4.3.11 Padrão de Textura Direcional Local

O descritor do padrão de textura direcional local (LDTP, do inglês *local directional texture pattern*), proposto por Rivera, Castillo e Chae (2015), explora as vantagens das informações direcionais e de magnitude na imagem.

O LDTP, assim como o LDPV, faz uso das máscaras de Kirsch, apresentadas na Figura 39, para extrair da imagem informações direcionais que são definidas como mostrado na equação seguinte:

$$P_{dir}^1 = arg\ max_i \{\mathbb{I}_i \mid 0 \leq i \leq 7\} \quad (4.33)$$

onde P_{dir}^1 é o principal número de informação direcional, e \mathbb{I}_i é o resultado da convolução da imagem de entrada I com a máscara de Kirsch M_i , tal como evidenciado pela Equação 4.34.

$$\mathbb{I}_i = |I * M_i| \quad (4.34)$$

Um segundo número de informação direcional P_{dir}^2 também é calculado conforme a Equação 4.33, com a ressalva de que é utilizado não o maior valor de \mathbb{I}_i , mas, sim o segundo maior valor. Dessa forma, existem dois números de informação direcional P_{dir}^1 e P_{dir}^2 . Estas duas informações de direção sinalizam o eixo principal da textura local e, em cada uma das mesmas, é calculada a diferença de intensidade dos pixels opostos na vizinhança, conforme a equação seguinte:

$$d_n^{(x,y)} = I(x_{P_{dir}^n,+}, y_{P_{dir}^n,+}) - I(x_{P_{dir}^n,-}, y_{P_{dir}^n,-}) \quad (4.35)$$

onde d_n é a n – ésima diferença para o pixel (x, y) na n – ésima principal direção. Já $I(x_{P_{dir}^n,+}, y_{P_{dir}^n,+})$ corresponde ao valor da intensidade do pixel $(x_{P_{dir}^n,+}, y_{P_{dir}^n,+})$ e $I(x_{P_{dir}^n,-}, y_{P_{dir}^n,-})$ é o valor da intensidade do pixel anterior. Assim sendo, temos que $x_{P_{dir}^n,\pm}$ e $y_{P_{dir}^n,\pm}$ são definidos como:

$$x_{P_{dir}^n,\pm} = \begin{cases} x_{\pm}, & \text{se } P_{dir}^n \in \{0, 1, 7\}, \\ x, & \text{se } P_{dir}^n \in \{2, 6\}, \\ x_{\mp}, & \text{se } P_{dir}^n \in \{3, 4, 5\} \end{cases} \quad (4.36)$$

$$y_{P_{dir}^n,\pm} = \begin{cases} y_{\pm}, & \text{se } P_{dir}^n \in \{1, 2, 3\}, \\ y, & \text{se } P_{dir}^n \in \{0, 4\}, \\ y_{\mp}, & \text{se } P_{dir}^n \in \{5, 6, 7\} \end{cases} \quad (4.37)$$

A diferença $D_f(d)$ é, então, codificada em três níveis, conforme apresentado pela equação

seguinte:

$$D_f(d) = \begin{cases} 0, & \text{se } -\varepsilon \leq d \leq \varepsilon \\ 1, & \text{se } d < -\varepsilon \\ 2, & \text{se } d > \varepsilon \end{cases} \quad (4.38)$$

onde d é a diferença atual de intensidade e ε é um valor de *threshold*. Em seguida, o valor do descritor de textura $LDTP(x, y)$ pode ser calculado conforme:

$$LDTP(x, y) = 16P_{dir}^{1(x,y)} + 4D_f(d_1^{(x,y)}) + D_f(d_2^{(x,y)}) \quad (4.39)$$

onde o valor do descritor que representa o padrão de textura direcional local para o pixel (x, y) é dado por $LDTP(x, y)$. Já os valores de $d_1^{(x,y)}$ e $d_2^{(x,y)}$ são, respectivamente, a primeira e a segunda diferença codificada para os pixels da vizinhança de (x, y) . A Figura 42, extraída do trabalho de Rivera, Castillo e Chae (2015), apresenta uma ilustração do processo de cálculo do LDTP.

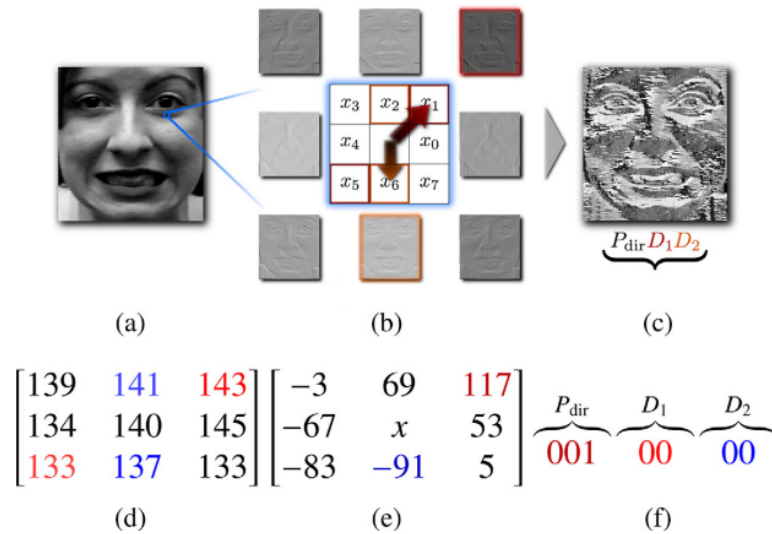


Figura 42 – Processo de cálculo do LDTP (Rivera, Castillo e Chae (2015)). (a) Imagem original. (b) Respostas após a aplicação das máscaras de Kirsh. (c) Imagem codificada. (d) Exemplo de valores de intensidade da vizinhança. (e) Respostas para a vizinhança apresentada em (d). (f) Código para a vizinhança apresentada em (d).

Por fim, observa-se que o descritor LDTP procura extrair características das principais direções da vizinhança local e, então, codifica as diferenças de intensidade dessas direções, apresentando robustez quanto à variações de iluminação e também a ruídos.

4.3.12 Quantização de Fase Local

O descritor de quantização de fase local (LPQ, do inglês, *local phase quantization*) proposto por [Ojansivu e Heikkilä \(2008\)](#) foi concebido para superar o LBP em aplicações onde as imagens estão borradas ou afetadas por iluminação não uniforme ([MUSCI et al., 2011](#)).

O LPQ trabalha com a informação local de fase da imagem que é obtida através do uso da transformada 2-D STFT, do inglês *short-term Fourier transform*, aplicada localmente em cada uma das janelas por sobre a imagem. Assim, para cada uma das janelas de cada um dos pixels de posição x da imagem $f(x)$ é aplicada a STFT conforme apresentado pela equação seguinte:

$$F(u, x) = \sum_{y \in N_x} f(x - y) e^{-j2\pi u^T y} \quad (4.40)$$

onde $x \in \{x_1, x_2, \dots, x_n\}$ simplesmente compõe a convolução 1-D das linhas e colunas. Já os coeficientes locais de Fourier $F(u, x)$ são calculados para os ângulos $[0, \pi/2, \pi, 3\pi/2]$. Nas frequências 2-D, os ângulos são $u_1 = [a, 0]^T$, $u_2 = [0, a]^T$, $u_3 = [a, a]^T$, e $u_4 = [a, -a]^T$, com $a = 1/m$, onde m é o tamanho da janela.

Assim, para cada pixel os resultados são representados como a seguir:

$$F_x^c = [F(u_1, x), F(u_2, x), F(u_3, x), F(u_4, x)] \quad (4.41)$$

$$F_x = [Re\{F(x)\}, Im\{F(x)\}]^T \quad (4.42)$$

onde $Re\{\cdot\}$ e $Im\{\cdot\}$ correspondem às partes real e complexa do número imaginário.

A transformada discreta de Fourier (DFT, do inglês, *discrete Fourier transform*) G_x da imagem borrada é, então, quantizada, considerando a seguinte regra:

$$q_j = \begin{cases} 1, & \text{se } g_j \geq 0 \\ 0, & \text{para qualquer outro caso} \end{cases} \quad (4.43)$$

Por fim, o LPQ é calculado da seguinte forma:

$$f_{LPQ}(x) = \sum_{j=1}^8 q_j 2^{j-1} \quad (4.44)$$

Diante disso, os autores [Ustubioglu et al. \(2016\)](#) exploraram as potencialidades desse descritor e propuseram um método de detecção de imagens forjadas utilizando-se do descritor LPQ.

4.3.13 Descritor de Frequência Local

O descritor LPQ, anteriormente tratado, desconsidera as informações de magnitude e, dessa forma, apresenta dificuldades ao ser empregado em imagens de baixa resolução. Diante deste fato, os autores [Lei et al. \(2011\)](#) propuseram o descritor de frequência local (LFD, do inglês, *local frequency descriptor*), o qual explora tanto as informações de magnitude invariantes a operações de borrimento, quanto as de fase.

O LFD adota uma metodologia semelhante ao LBP para descrever a magnitude da vizinhança. Assim, para uma dada janela i da imagem, tem-se que $M(u, i)$ é a magnitude de i para uma resposta u no domínio da frequência. Dessa maneira, considerando-se o centro da janela como sendo k , tem-se $M(u, k)$ e, com isso, é possível estabelecer a seguinte regra de quantização:

$$S(M(u, k), M(u, i)) = \begin{cases} 1, & \text{se } M(u, k) \geq M(u, i) \\ 0, & \text{se } M(u, k) < M(u, i) \end{cases} \quad (4.45)$$

O descritor de magnitude local (LMD, do inglês, *local magnitude descriptor*) para uma janela i e uma resposta de frequência u é obtido conforme apresentado pela equação seguinte:

$$f_{LMD}(u, i) = \sum_{k=1}^8 S(M(u, k), M(u, i)) 2^{k-1} \quad (4.46)$$

Considerando $P(i, u)$ como sendo a informação de fase de uma janela i com uma resposta de frequência u , é possível estabelecer a seguinte quantização:

$$S(P(u, k), P(u, i)) = \begin{cases} 1, & \text{se } P(u, k), P(u, i) \text{ estão no mesmo quadrante} \\ 0, & \text{para qualquer outro caso} \end{cases} \quad (4.47)$$

O descritor local de fase (LDP, do inglês, *local phase descriptor*) é, então, descrito como:

$$f_{LDP}(u, i) = \sum_{k=1}^8 S(P(u, k), P(u, i)) 2^{k-1} \quad (4.48)$$

A Figura 43 apresenta uma ilustração para o processo de cálculo do LFD.

4.3.14 Codificação Binária Monogênica

Considerando que a representação de sinal monogênico foi usada com sucesso em muitas aplicações de classificações de textura, conforme discutido no artigo escrito por

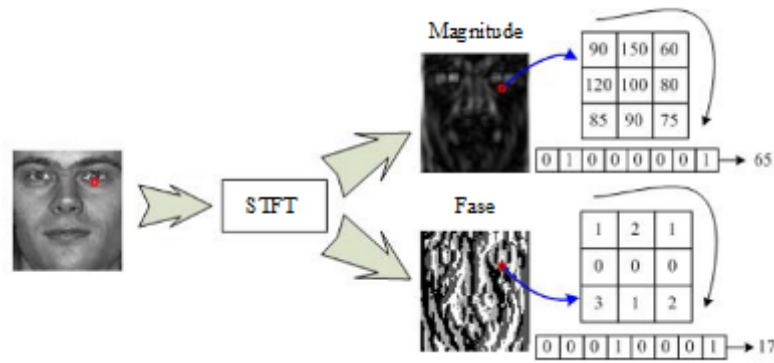


Figura 43 – Processo de cálculo do LFD (LEI et al., 2011).

Zhang et al. (2010), os autores Yang et al. (2012) propuseram um descritor de codificação binária monogênica (MBC, do inglês, *monogenic binary coding*).

O sinal monogênico foi introduzido por Felsberg e Sommer (2001) com a finalidade de generalizar o sinal analítico de 1-D para 2-D. A representação monogênica 2-D é obtida a partir da transformada de Riesz, sendo que o sinal monogênico é utilizado para decompor a imagem em três componentes complementares: orientação, fase e magnitude (ZHOU et al., 2019; FU; PENG, 2015).

A orientação o , a fase p e a magnitude a são calculadas conforme as equações seguintes:

$$a = \sqrt{g^2 + h_x^2 + h_y^2} \quad (4.49)$$

$$p = -\text{sign}(h_x) \arctan \left(\frac{\sqrt{h_x^2 + h_y^2}}{g} \right) \quad (4.50)$$

$$o = \arctan \left(\frac{h_y}{h_x} \right) \quad (4.51)$$

onde g e h são definidos por:

$$g = I * F^{-1}(G(w)) \quad (4.52)$$

$$h_i = F^{-1} \left(\frac{\sqrt{-1} w_i}{\sqrt{w_x^2 + w_y^2}} \right), i \in \{x, y\} \quad (4.53)$$

com I denotando a imagem original, $*$ o operador de convolução, F^{-1} a 2-D inversa da transformada de Fourier e $G(w)$ o filtro Log-Gabor no domínio de Fourier, conforme

apresentado a seguir:

$$G(w) = \exp \left\{ -\frac{\left[\ln \frac{w}{w_0} \right]^2}{\sqrt{2} \ln \left[\ln \left(\frac{\sigma}{w_0} \right) \right]} \right\} \quad (4.54)$$

onde w_0 é a frequência central, σ é o fator de escala da largura de banda e w_x e w_y , as frequências horizontal e vertical.

A representação de sinal monogênico tem a vantagem de que seu processo de extração de características (orientação, fase e amplitude) é invariante à rotação (UNSER; SAGE; VILLE, 2009).

4.3.15 Exemplo Prático de Descritores Locais de Textura

A Figura 44 apresenta o exemplo de uma imagem da base de textura ALOT (HOSSAIN; SERIKAWA, 2013) que teve suas características de textura extraídas através da aplicação de cada um dos descritores locais anteriormente discutidos.

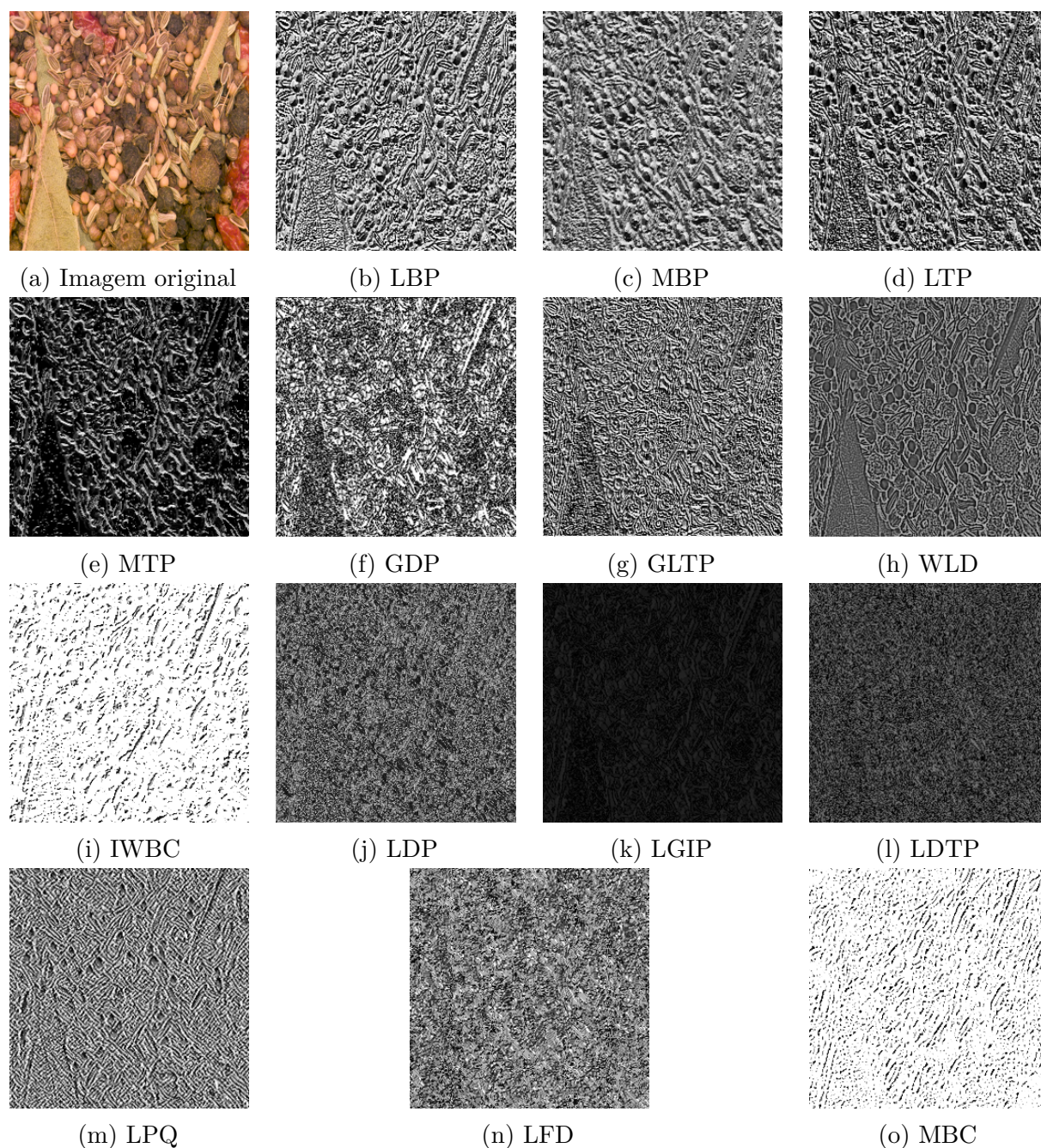


Figura 44 – Exemplos de descritores locais de textura extraídos de uma imagem.

A seção seguinte detalhará as principais bases de textura públicas.

4.4 Bases de Imagens de Textura

Várias bases de textura foram desenvolvidas ao longo dos anos e vêm contribuindo para o desenvolvimento da área de análise de textura. Os trabalhos de [Hossain e Serikawa \(2013\)](#), [Bianconi e Fernández \(2014\)](#) e [Cavalin e Oliveira \(2017\)](#) apresentam uma descrição abrangente do estado da arte das bases de imagens de textura que são públicas.

Para os propósitos deste trabalho, foram escolhidas as bases de imagens de textura mais modernas (a partir do ano 2000) e que ainda são, de alguma maneira, mantidas. Conforme será visto a seguir, a base Brodatz ([BRODATZ, 1966](#)) constitui uma exceção, pois é uma base bastante explorada pela literatura, embora tenha sido montada a partir de um livro editado no ano de 1966. Também foi aberta uma exceção para a base CURET, que também é uma base bastante comum nos trabalhos da área, tendo sido apresentada à comunidade no ano de 1999.

As seções seguintes descrevem, de maneira objetiva, as principais bases de imagens de textura escolhidas, suas características e referências.

4.4.1 Brodatz

Esta base de imagens é uma das mais tradicionais da área de análise de textura, tendo sido construída a partir da digitalização de páginas do livro de [Brodatz \(1966\)](#). Ela contempla 112 imagens em níveis de cinza de superfícies, tecidos, fibras e grãos. Neste trabalho, as imagens desta base foram obtidas a partir do sítio da tese de [Randen \(1997\)](#), tendo resolução de 640×640 pixels e extensão GIF. A Figura 45 apresenta exemplos de imagens da base Brodatz.

Tendo em vista o fato de a base original contemplar apenas 112 imagens, a mesma foi organizada de maneira análoga àquela apresentada no trabalho de [Shang et al. \(2006\)](#). Assim, cada uma das imagens da base foi subdividida em nove outras a fim de compor uma classe. Dessa forma, a base ficou organizada em 1.008 imagens agrupadas em 112 classes com 9 imagens. A Figura 46 apresenta um exemplo das classes obtidas.

4.4.2 CURET

A base de imagens de textura da Universidade de Columbia-Utrecht (CURET, do inglês, Columbia-Utrecht *reflectance and texture*) foi apresentada no trabalho de [Dana et al. \(1999\)](#) e compreende um total de 12.000 imagens coloridas de extensão PNG com resolução de 512×512 pixels e organizadas em 61 classes. A base possui imagens de diferentes materiais (tecidos, pisos, folhas de árvores etc.) e a Figura 47 traz exemplos das imagens presentes nessa base.

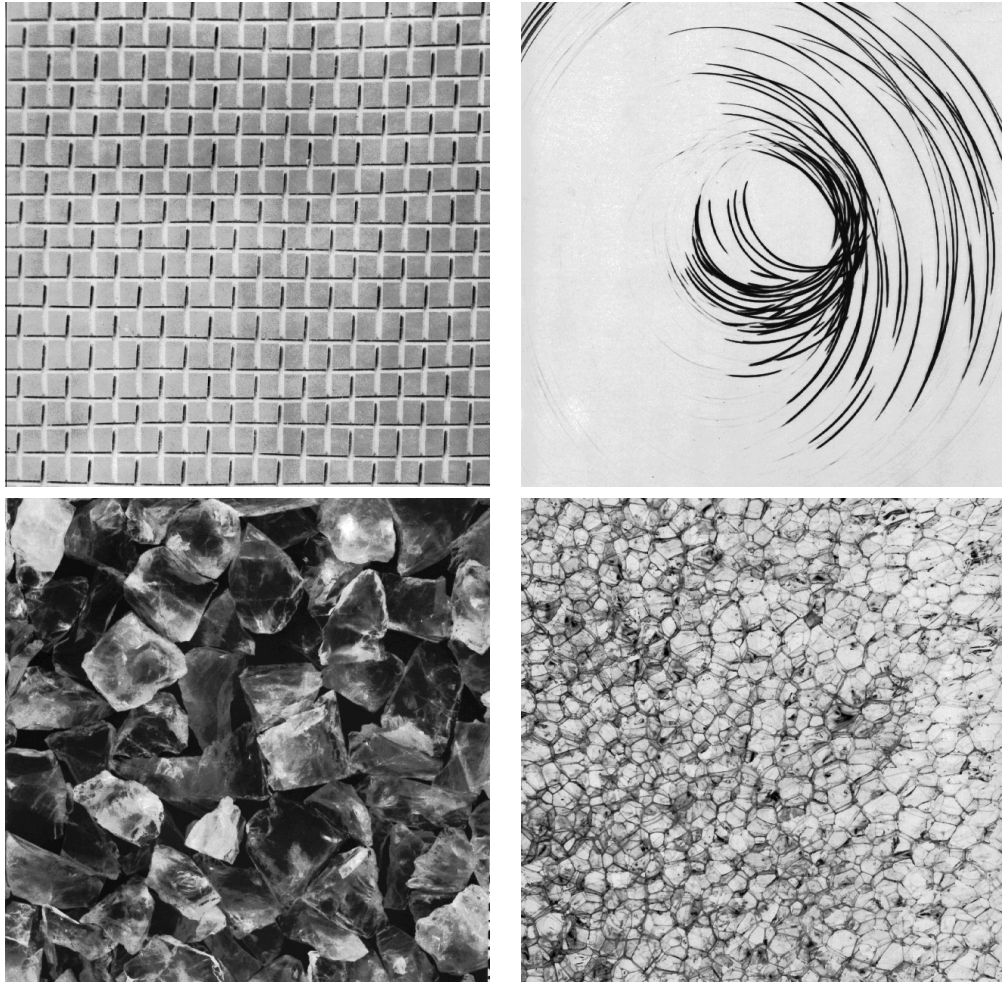


Figura 45 – Exemplos da base de imagens Brodatz.

4.4.3 Kylberg

Apresentada no trabalho de [Kylberg \(2011\)](#), esta base possui imagens de superfícies de rochas, pisos, de tecidos e grãos de alimentos. Elas encontram-se organizadas em 28 classes com 160 imagens cada, totalizando 4.480 imagens PNG em níveis de cinza e com resolução de 576×576 pixels. A Figura 48 exhibe exemplos de imagens da base.

4.4.4 ALOT

As imagens constantes da base de imagens ALOT, do inglês Amsterdam *library of textures*, estão organizadas em 250 classes, sendo um total de 27.000 imagens coloridas de extensão PNG e com resolução de 1.536×1.024 pixels. A base foi apresentada pelo trabalho de [Hossain e Serikawa \(2013\)](#) e possui imagens de objetos variados como de tecidos, pães e rochas. A Figura 49 exhibe exemplos de suas imagens.

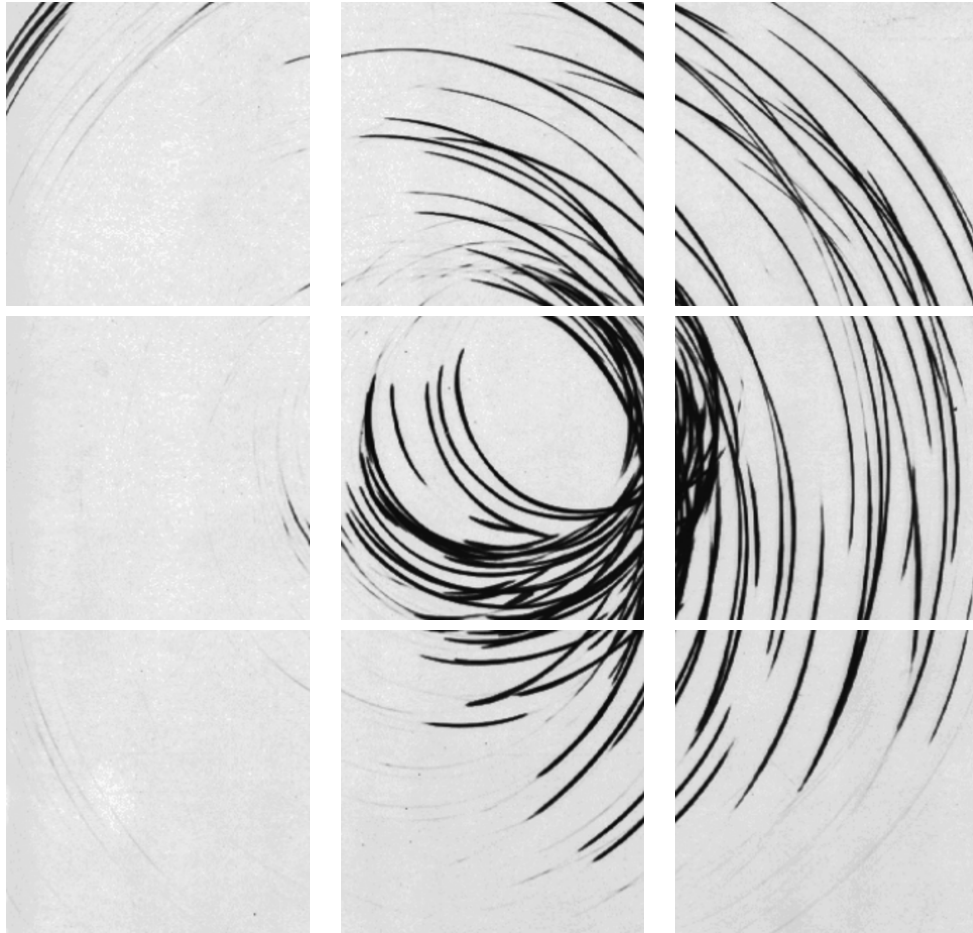


Figura 46 – Exemplo de classe da base de imagens Brodatz.

4.4.5 DTD

O conjunto de imagens da base de texturas descritíveis (DTD, do inglês, *describable textures dataset*) foi introduzido no trabalho de (CIMPOI et al., 2014) e consiste de uma coleção de imagens de texturas do mundo real rotuladas com adjetivos típicos (manchado, arranhado, ziguezagueado etc.). Assim, é possível observar que, diferentemente das bases anteriormente apresentadas (CURET, ALOT e Kylberg), a DTD foi construída com foco no reconhecimento de atributos de textura descritíveis e não em reconhecimento de materiais a partir de suas texturas.

A base DTD possui suas imagens organizadas em 47 classes com 120 imagens cada, totalizando 5.640 ao todo. As imagens são coloridas, com extensão JPEG e com diferentes resoluções (300×300 , 490×407 , 472×378 etc.). A Figura 50 traz exemplos da base DTD.

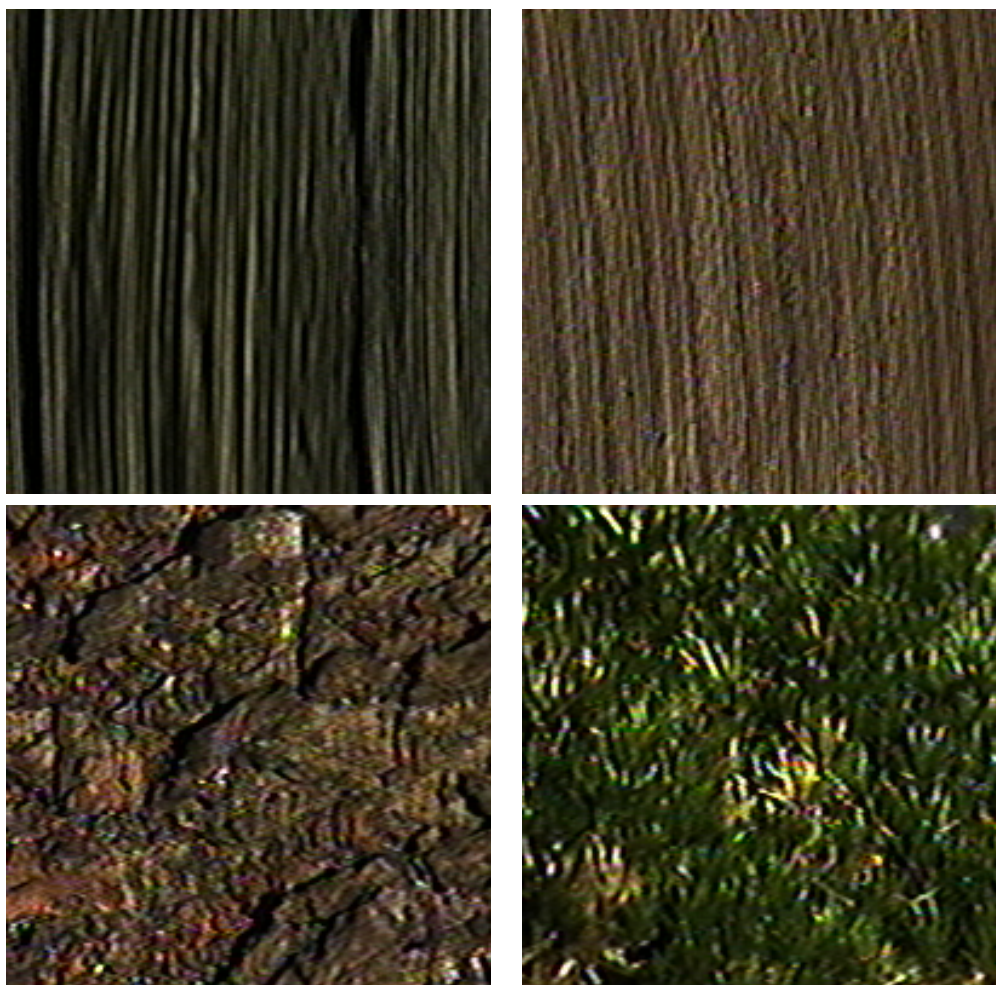


Figura 47 – Exemplos da base de imagens CURET.

4.5 Considerações Finais do Capítulo

Este capítulo apresentou a definição de textura e as principais abordagens de extração de características de textura de uma imagem digital. O capítulo também trouxe detalhamentos sobre os descritores locais de textura mais utilizados na literatura. Além disso, foi apresentada uma linha do tempo dos principais descritores de textura.

O capítulo ainda descreveu as principais bases de imagens textura que são utilizadas na literatura para fins de testes e validação de modelos.

O próximo capítulo discutirá trabalhos de forense em imagens digitais baseados em redes neurais convolucionais.

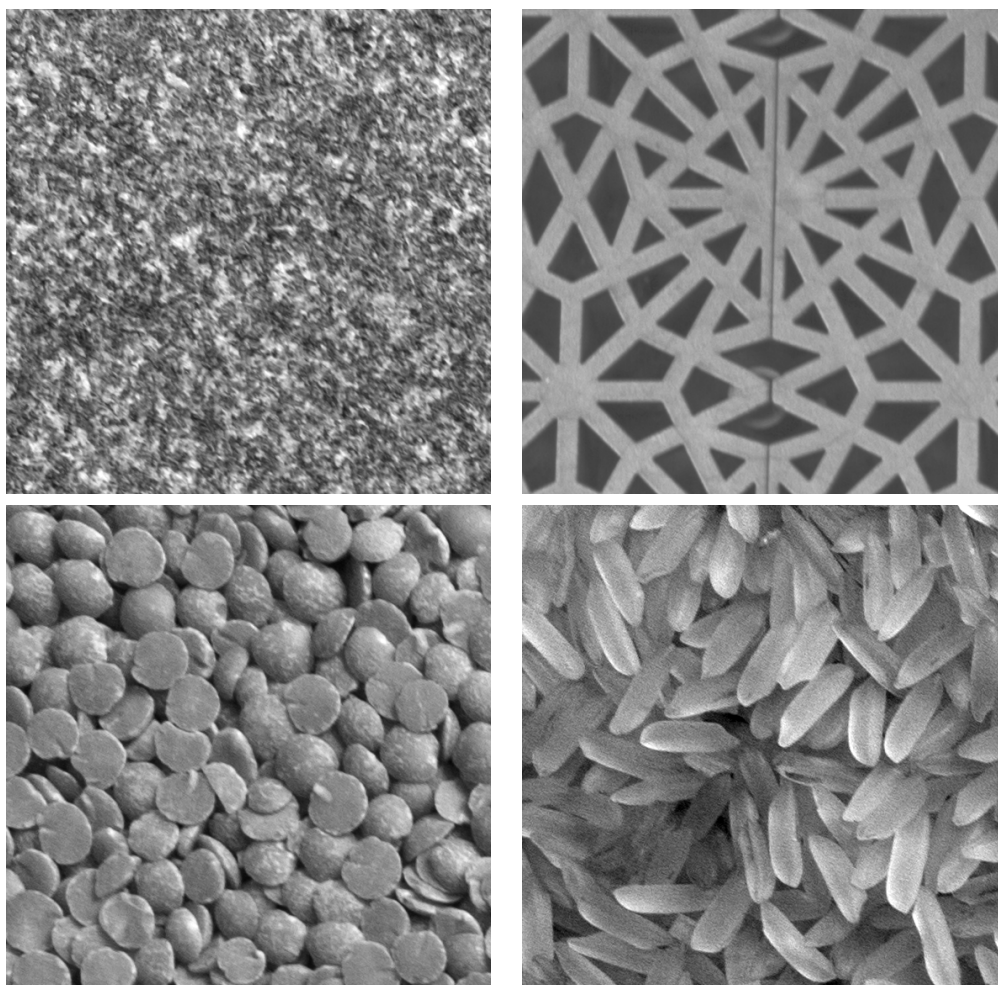


Figura 48 – Exemplos da base de imagens Kylberg.



Figura 49 – Exemplos da base de imagens ALOT.



Figura 50 – Exemplos da base de imagens DTD.

Sumário

5.1	Introdução	93
5.2	Abordagens <i>Deep Learning</i> para Detecção de Imagens Falsificadas . .	93
5.3	Considerações Finais do Capítulo	96

5.1 Introdução

Atualmente, vários trabalhos da área de forense em imagens digitais têm alcançado bons resultados com a utilização de redes neurais profundas. A seção seguinte descreve alguns destes trabalhos.

5.2 Abordagens *Deep Learning* para Detecção de Imagens Falsificadas

Atualmente, abordagens que fazem uso de *deep learning*, juntamente com métodos ativos e passivos de análise forense de imagens digitais, têm produzido resultados motivadores. A Tabela 6 apresenta exemplos desses trabalhos, através dos quais é possível verificar o uso frequente de redes neurais convolucionais

Os autores [Zeng et al. \(2018\)](#) propõem um ambiente híbrido baseado em *deep learning* que utiliza quantização e truncamento para realizar a análise esteganográfica

Método	Referências
Esteganografia	(ZENG et al., 2018; SALOMON et al., 2017; YE; NI; YI, 2017; XU; WU; SHI, 2016; ZOU; ZHANG; LIU, 2019)
Marca D'Água	(KANDI; MISHRA; GORTHI, 2017; HARIBABU; SUBRAHMANYAM; MISHRA, 2015; CHEN; CHEN, 2010)
Identificação da Câmera de Origem	(MAYER; STAMM, 2020; YAO et al., 2018; BAYA; STAMM, 2017; GUERA et al., 2017; BONDI et al., 2017)
<i>Copy-move</i>	(OUYANG; LIU; LIAO, 2017; RAO; NI, 2016; BAPPY et al., 2017)
Reamostragem	(QUAN et al., 2018; BUNK et al., 2017; BAYAR; STAMM, 2017; NASONOV; CHESNAKOV; KRYLOV, 2016)

Tabela 6 – Trabalhos de detecção de imagens falsificadas baseados *deep learning*.

de imagens JPEG, tendo apresentado evidências teóricas importantes para apoiar sua utilização.

O trabalho desenvolvido por [Salomon et al. \(2017\)](#) usa chaves esteganográficas incorporadas às imagens digitais e uma rede CNN para realizar a análise esteganográfica. O modelo proposto utiliza uma filtragem global com um único filtro na primeira camada da rede CNN.

No artigo escrito por [Ye, Ni e Yi \(2017\)](#) foram usados filtros passa-alta para extrair características do modelo no domínio espacial da imagem, que são aplicados como pesos da primeira camada de uma rede CNN. O trabalho apresenta uma comparação de desempenho entre o modelo proposto e aqueles tratados por [Holub e Fridrich \(2012\)](#), [Holub, Fridrich e Denemark \(2014\)](#) e [Li et al. \(2014\)](#).

O desempenho da análise esteganográfica usando aprendizado profundo é enfatizado por [Xu, Wu e Shi \(2016\)](#) e é apresentada uma comparação de um modelo proposto por [Holub, Fridrich e Denemark \(2014\)](#) e também por [Li et al. \(2014\)](#). No entanto, o autor enfatiza que o desempenho do modelo CNN proposto ainda não pode ser comparado com métodos esteganográficos avançados como aqueles apresentados nos trabalhos de [Denemark et al. \(2014\)](#), [Tang et al. \(2014\)](#), [Tang et al. \(2016\)](#), [Denemark, Fridrich e Comesaña \(2016\)](#).

Com relação à marca d'água, os trabalhos de [Kandi, Mishra e Gorthi \(2017\)](#) e [Haribabu, Subrahmanyam e Mishra \(2015\)](#) apresentam métodos baseados em uma rede CNN de codificador automático para marca d'água robusta e não cega, com melhor desempenho do que as técnicas que usam transformadas. Assim, o auto-codificador proposto obteve uma taxa de pico de sinal para ruído (PSNR) de 58,91 contra 56,74 do modelo

descrito por Anju (2013), que faz uso de um modelo baseado em DCT combinado com DWT .

A pesquisa desenvolvida por Chen e Chen (2010) trabalha com uma rede neural para criar um esquema de marca d'água cego que incorpora mensagens de marca d'água em diferentes blocos *wavelets* da imagem. O esquema mostrou robustez a ataques comuns de manipulação de imagem, como, por exemplo, filtro passa-baixa Gaussiano, ruído de pimenta e filtro da mediana.

As redes CNN também são usadas em ambientes que identificam a câmera que capturou uma imagem específica. Assim, o trabalho de Mayer e Stamm (2020) promove adaptações na rede neural proposta por Bayar e Stamm (2018a), a fim de viabilizar a identificação das câmeras de origem das imagens e, também, a localização de eventuais regiões que foram editadas, mas não realiza testes do modelo em bases públicas de forense em imagens digitais.

O trabalho desenvolvido por Yao et al. (2018) propõem um multi-classificador baseado em CNN que tem robustez quando as imagens são submetidas a ataques de pós-processamento, como compactação JPEG e adição de ruído. O modelo proposto apresenta melhor precisão do que o modelo CNN proposto por Bondi et al. (2017). No entanto, Guera et al. (2017) descrevem um método anti-forense que pode modificar imagens digitais para ofuscar classificadores CNN.

A abordagem proposta por Baya e Stamm (2017) trabalha com características de baixo nível da imagem para criar mapas de características que são aprendidos por uma camada convolucional restrita. O resultado é aumentado usando características residuais de filtro da mediana não linear (MFR, do inglês, *median filter residual*). A abordagem provou ser robusta para operações de *resampling* e de compactação.

Especialmente considerando a operação *copy-move*, a abordagem proposta por Ouyang, Liu e Liao (2017) usa uma rede ajustada obtida de um modelo treinado existente, demonstrando bom desempenho com imagens falsificadas geradas automaticamente, mas não sendo recomendável para imagens reais falsificadas. O trabalho desenvolvido por Rao e Ni (2016) propõe um modelo baseado em CNN que não usa estratégia aleatória. Ele trabalha com pesos na primeira camada da rede que são inicializados com 30 filtros passa-alta básicos. O trabalho desenvolvido por Bappy et al. (2017) propôs um modelo híbrido que usa as redes CNN e LSTM, do inglês *Long-Short Term Memory*, combinadas para identificar regiões de imagem que foram falsificadas.

Os trabalhos escritos por Bunk et al. (2017), Bayar e Stamm (2017) e Nasonov, Chesnakov e Krylov (2016) usam *deep learning* para identificar imagens falsificadas por meio de operações de *resampling*. Uma comparação entre CNN e LSTM é estabelecida por Bunk et al. (2017) e foi observado que ambas as redes são eficazes na exploração de

características de *resampling* para detectar regiões adulteradas na imagem. O modelo proposto por [Bayar e Stamm \(2017\)](#) usa uma rede CNN para identificar operações de *resampling* em imagens novamente compactadas. No entanto, [Nasonov, Chesnakov e Krylov \(2016\)](#) propõem uma nova abordagem baseada em uma rede CNN que pode detectar *resampling* em imagens com ou sem ruído. O trabalho de [Quan et al. \(2018\)](#) propõe um ambiente genérico CNN para classificar imagens como naturais ou geradas por computador. Eles sugerem um arranjo com duas camadas convolucionais em cascata na parte inferior de uma CNN, tendo obtido bons resultados de robustez para *resampling* e compactação JPEG.

A seção seguinte apresenta as considerações finais o presente capítulo.

5.3 Considerações Finais do Capítulo

Este capítulo discutiu trabalhos de forense em imagens digitais baseados em redes neurais de aprendizagem profunda, destacando aqueles relacionados a esteganografia, marca d'água, identificação da câmera de origem, *copy-move* e reamostragem.

O próximo capítulo abordará detalhes do modelo de identificação de imagens falsificadas proposto neste trabalho.

Modelo Proposto para Detecção de Imagens Falsificadas

Sumário

6.1	Introdução	97
6.2	Extração das Características de Textura	98
6.2.1	Avaliação dos Descritores Locais de Textura	98
6.2.2	Descrição do Experimento de Seleção de Descritores	99
6.2.3	Resultados da Seleção de Descritores	101
6.2.4	Discussão	107
6.2.5	Descritores Selecionados	108
6.3	Classificação	108
6.4	Considerações Finais do Capítulo	112

6.1 Introdução

O modelo proposto neste trabalho objetiva classificar imagens digitais como falsificadas ou não, a partir da análise de suas texturas. A Figura 51 apresenta uma visão geral do modelo.

Observa-se a existência de duas etapas distintas no modelo proposto, a saber: a etapa de extração das características de textura e a de classificação. Estas etapas serão detalhadas nas seções seguintes.

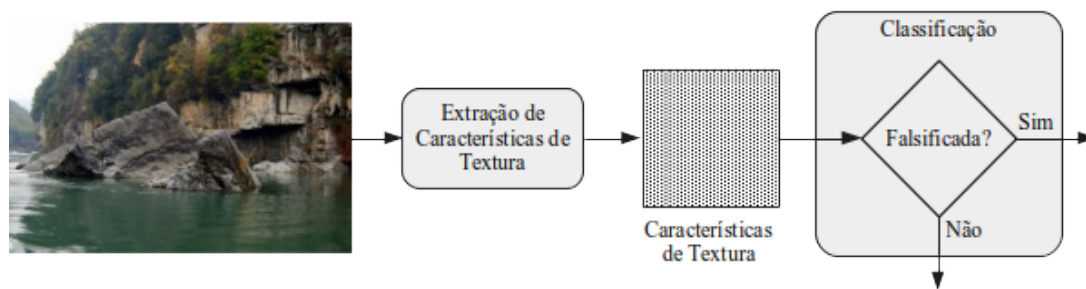


Figura 51 – Visão geral do modelo proposto.

6.2 Extração das Características de Textura

A extração das características constitui uma importante operação para se definir a textura de imagens digitais, pois a mesma é utilizada para adquirir elementos intrínsecos das imagens a fim de se levantar, por exemplo, a distribuição espacial dos pixels e a variação de luminosidade da imagem, conforme evidenciado em [Pedrini e Schwartz \(2008\)](#). Assim, a extração das características promove a realização de transformações na imagem de entrada a fim de permitir que a mesma possa ser descrita de maneira mais simplificada, com foco em sua textura.

Nesta etapa, o modelo está concentrado nas informações de textura da imagem, tendo em vista que é muito complexo realizar uma edição em uma imagem e não deixar vestígios na textura das mesmas. Entretanto, quais descritores de textura devem ser utilizados para representar as características de textura das imagens? Diante deste questionamento, a seção seguinte apresenta uma avaliação dos principais descritores da literatura, a fim de encontrar aqueles com melhor eficácia.

6.2.1 Avaliação dos Descritores Locais de Textura

Esta seção objetiva avaliar quais dos principais descritores de textura da literatura discutidos no Capítulo 4 apresentam melhor eficácia. Para tal, foi elaborado experimento conforme ilustrado pela Figura 52.

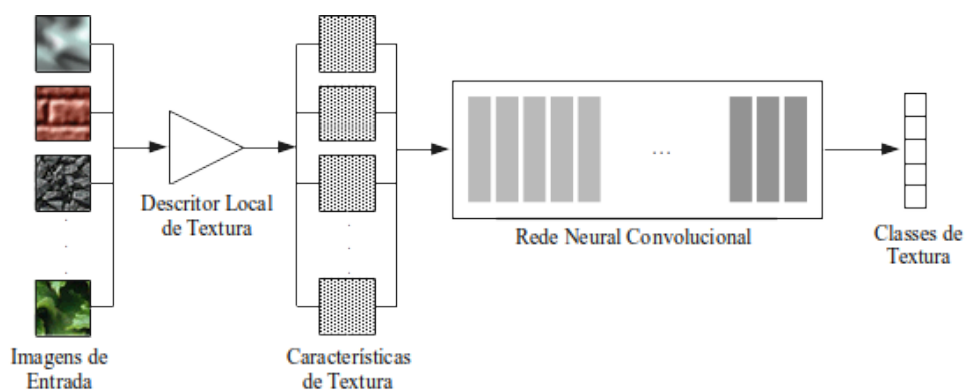


Figura 52 – Visão geral da avaliação dos descritores locais de textura.

A avaliação de performance dos descritores será realizada com base na acurácia, tendo como referência um modelo apoiado em redes neurais convolucionais motivado pelos resultados obtidos nos trabalhos desenvolvidos por [Andrearczyk e Whelan \(2016\)](#), [Kandi, Mishra e Gorthi \(2017\)](#), [Quan et al. \(2018\)](#) e [Bayar e Stamm \(2018b\)](#). Assim, o experimento realizado não se encontra focado em competir com o estado da arte em reconhecimento de textura, mas, sim, em avaliar como as redes neurais convolucionais aprendem as características de textura a partir dos descritores locais estudados.

Para fins da avaliação de eficácia dos descritores de textura, foram utilizadas bases públicas de textura, as quais foram descritas na Seção 4.4.

As seções seguintes apresentarão a descrição do experimento de avaliação dos descritores, os resultados obtidos e uma discussão a respeito dos mesmos.

6.2.2 Descrição do Experimento de Seleção de Descritores

O experimento foi realizado considerando-se as cinco principais bases que foram detalhadas no Capítulo 4 e são descritas na Tabela 7.

Tendo em vista que as imagens encontram-se em diferentes formatos (RGB e níveis de cinza) e resoluções, as imagens RGB foram, primeiramente, convertidas para níveis de cinza conforme a seguinte equação:

$$I = 0,299R + 0,587G + 0,114B \quad (6.1)$$

onde I representa a imagem em níveis de cinza e R , G e B os valores individuais da imagem original para vermelho, verde e azul, respetivamente.

Em seguida, as imagens em níveis de cinza foram todas redimensionadas para a resolução de 256×256 pixels.

Base	Imagens	Classes	Resolução (em pixels)
Brodatz (BRODATZ, 1966)	1.008	112	640×640
CUReT (DANA et al., 1999)	12.000	61	$512 \times 512 \times 3$
Kylberg (KYLBERG, 2011)	4.480	28	576×576
ALOT (HOSSAIN; SERIKAWA, 2013)	27.000	250	$1.536 \times 1.024 \times 3$
DTD (CIMPOI et al., 2014)	5.640	47	variada

Tabela 7 – Resumo das características das bases de imagens de textura.

Para fins de treinamento e teste, as bases utilizadas foram divididas aleatoriamente em **70%** das imagens para treinamento e o restante para validação. Além disso, as imagens

contendo as características de textura foram redimensionadas de acordo com a necessidade da arquitetura da rede neural utilizada, conforme apresentado na Tabela 8.

No que se refere às redes neurais convolucionais trabalhadas no experimento, foram utilizadas as CNNs apresentadas na Tabela 8. O treinamento destas redes foi definido com, no máximo, **50 épocas** para cada caso, com taxa inicial de aprendizagem igual a **0,00001** e com o gradiente descendente estocástico (SGD, do inglês, *stochastic gradient descent*) como otimizador.

CNN	Entrada	Profundidade	Parâmetros
AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012)	$227 \times 227 \times 3$	8	61 milhões
GoogLeNet (SZEGEDY et al., 2015)	$224 \times 224 \times 3$	22	7 milhões

Tabela 8 – CNNs utilizadas no experimento.

Considerando-se o tamanho reduzido das bases de textura disponíveis e as redes pré-treinadas genéricas, optou-se por utilizar transferência de aprendizado das CNNs previamente treinadas na base de imagens ImageNet (RUSSAKOVSKY et al., 2015). Este processo de transferência é comumente usado em soluções baseadas em *deep learning*, pois é mais simples do que treinar do zero uma rede com pesos inicializados aleatoriamente (LI; HOIEM, 2018; LI; GRANDVALET; DAVOINE, 2020). A Figura 53 apresenta uma ilustração do processo de transferência de aprendizado.

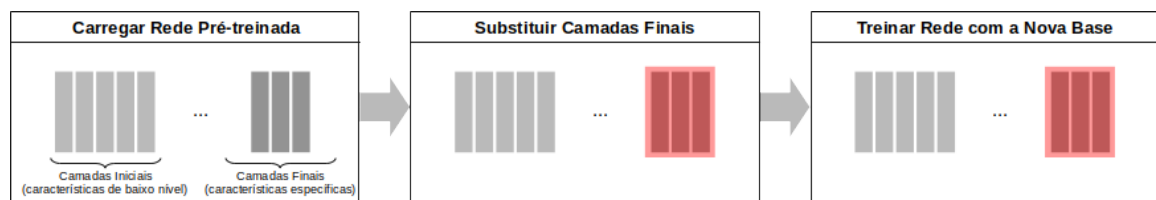


Figura 53 – Ilustração do processo de transferência de aprendizado.

Considerando que as redes pré-treinadas tiveram o seu treinamento em bases organizadas para identificação de objetos e não de textura, as redes neurais convolucionais do experimento também foram treinadas do zero (*from scratch*) com as bases de textura deste experimento, para fins de comparação.

O experimento foi executado no computador com as configurações apresentadas na Tabela 9.

Os resultados do experimento considerando-se as CNNs com transferência de aprendizado e também treinadas do zero, são apresentados na seção seguinte.

Descrição	Configuração
Equipamento	Dell Precision Tower 7910
Processador	1 × Intel® Xeon® CPU E5-2609 v3 @ 1.90GHz com 6 núcleos
Memória RAM	16 Gigabytes - DDR4-2133
Placa de Vídeo (Suporte ao ambiente Gráfico)	Placa de Vídeo VGA EVGA GeForce GT 610 com 1 GB
GPU (Exclusiva para Experimentos)	EVGA / NVIDIA GeForce GTX TITAN Z com 12 GB (Equivalente a 2 × NVIDIA GeForce GTX TITAN Black)
Ambiente de desenvolvimento	MATLAB R2018a
Sistema operacional	Ubuntu Linux 16.04.6 LTS

Tabela 9 – Configurações do computador utilizado.

6.2.3 Resultados da Seleção de Descritores

Os resultados do experimento foram obtidos considerando-se as bases de imagens de textura descritas na Tabela 7, as redes neurais convolucionais referenciadas na Tabela 8 e as configurações descritas.

As Tabelas 10 e 11 apresentam, respectivamente, os resultados obtidos com o uso das redes neurais convolucionais AlexNet e GoogLeNet. As Figuras 54, 55 e 56, por sua vez, exibem os resultados organizados em sua forma gráfica.

A seção seguinte deste capítulo discutirá os resultados que foram aqui apresentados.

	Descritor	Acurácia nas Bases de Imagens					Média	Desvio Padrão
		Brodatz	CUReT	Kylberg	ALOT	DTD		
AlexNet	LBP (OJALA; PIETIKÄINEN; HARWOOD, 1996)	0,12	0,48	0,93	0,61	0,26	0,48	0,32
	MBP (HAFIANE et al., 2008)	0,12	0,51	0,92	0,63	0,26	0,49	0,31
	LTP (TAN; TRIGGS, 2010)	0,18	0,59	0,88	0,71	0,29	0,53	0,29
	MTP (KHAN et al., 2013)	0,22	0,47	0,90	0,66	0,34	0,52	0,27
	GDP (AHMED, 2012)	0,07	0,28	0,76	0,43	0,21	0,35	0,26
	GLTP (AHMED; HOSSAIN, 2013)	0,13	0,42	0,86	0,58	0,28	0,45	0,28
	WLD (CHEN et al., 2010)	0,22	0,45	0,88	0,72	0,33	0,52	0,27
	IWBC (YANG et al., 2016)	0,08	0,31	0,78	0,42	0,19	0,36	0,27
	LDPV (KABIR; JABID; CHAE, 2010)	0,02	0,07	0,72	0,42	0,13	0,27	0,29
	LGIP (ZHOU; WANG, 2012)	0,38	0,26	0,73	0,57	0,22	0,43	0,22
	LDTP (RIVERA; CASTILLO; CHAE, 2015)	0,18	0,03	0,47	0,26	0,12	0,21	0,17
	LPQ (OJANSIVU; HEIKKILÄ, 2008)	0,49	0,38	0,81	0,61	0,22	0,50	0,22
	LFD (LEI et al., 2011)	0,10	0,06	0,54	0,21	0,10	0,20	0,20
	MBC (YANG et al., 2012)	0,24	0,21	0,67	0,49	0,20	0,36	0,21
	Média	0,18	0,32	0,78	0,52	0,23	0,41	0,26
Maior Valor	0,49	0,59	0,93	0,72	0,34	0,53	0,32	
Menor Valor	0,02	0,03	0,47	0,21	0,10	0,20	0,17	

(a) Pré-treinada

	Descritor	Acurácia nas Bases de Imagens					Média	Desvio Padrão
		Brodatz	CUReT	Kylberg	ALOT	DTD		
AlexNet	LBP (OJALA; PIETIKÄINEN; HARWOOD, 1996)	0,63	0,71	0,94	0,79	0,34	0,68	0,22
	MBP (HAFIANE et al., 2008)	0,62	0,77	0,92	0,81	0,33	0,69	0,23
	LTP (TAN; TRIGGS, 2010)	0,70	0,77	0,90	0,80	0,37	0,71	0,20
	MTP (KHAN et al., 2013)	0,76	0,68	0,92	0,80	0,40	0,71	0,19
	GDP (AHMED, 2012)	0,39	0,49	0,82	0,65	0,28	0,53	0,21
	GLTP (AHMED; HOSSAIN, 2013)	0,66	0,62	0,88	0,75	0,33	0,65	0,20
	WLD (CHEN et al., 2010)	0,79	0,60	0,89	0,83	0,39	0,70	0,20
	IWBC (YANG et al., 2016)	0,49	0,57	0,83	0,66	0,26	0,56	0,21
	LDPV (KABIR; JABID; CHAE, 2010)	0,17	0,34	0,76	0,51	0,18	0,39	0,25
	LGIP (ZHOU; WANG, 2012)	0,38	0,45	0,79	0,65	0,28	0,51	0,21
	LDTP (RIVERA; CASTILLO; CHAE, 2015)	0,14	0,25	0,65	0,45	0,17	0,33	0,21
	LPQ (OJANSIVU; HEIKKILÄ, 2008)	0,52	0,56	0,88	0,73	0,31	0,60	0,22
	LFD (LEI et al., 2011)	0,12	0,32	0,66	0,41	0,17	0,34	0,22
	MBC (YANG et al., 2012)	0,33	0,44	0,73	0,59	0,27	0,47	0,19
	Média	0,48	0,54	0,83	0,67	0,29	0,56	0,21
Maior Valor	0,79	0,77	0,94	0,83	0,40	0,71	0,25	
Menor Valor	0,12	0,25	0,65	0,41	0,17	0,33	0,19	

(b) Treinada do zero

Tabela 10 – Resultados para a rede AlexNet.

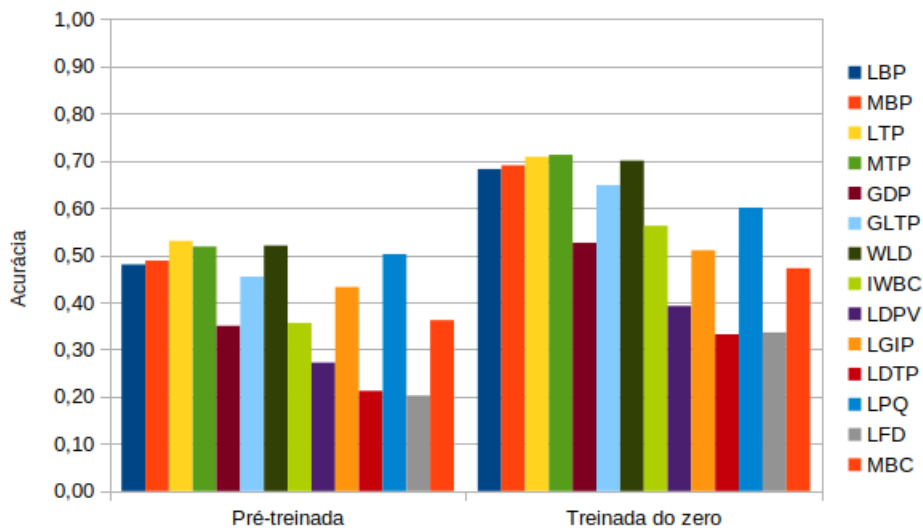
	Descritor	Acurácia nas Bases de Imagens					Média	Desvio Padrão
		Brodatz	CUReT	Kylberg	ALOT	DTD		
GoogLeNet	LBP (OJALA; PIETIKÄINEN; HARWOOD, 1996)	0,40	0,46	0,90	0,76	0,34	0,57	0,24
	MBP (HAFIANE et al., 2008)	0,45	0,56	0,89	0,77	0,34	0,60	0,23
	LTP (TAN; TRIGGS, 2010)	0,47	0,54	0,85	0,81	0,37	0,61	0,21
	MTP (KHAN et al., 2013)	0,57	0,45	0,84	0,77	0,41	0,61	0,19
	GDP (AHMED, 2012)	0,24	0,28	0,74	0,62	0,28	0,43	0,23
	GLTP (AHMED; HOSSAIN, 2013)	0,42	0,45	0,85	0,73	0,35	0,56	0,22
	WLD (CHEN et al., 2010)	0,48	0,52	0,90	0,85	0,42	0,63	0,22
	IWBC (YANG et al., 2016)	0,21	0,31	0,82	0,62	0,25	0,44	0,27
	LDPV (KABIR; JABID; CHAE, 2010)	0,07	0,11	0,58	0,43	0,16	0,27	0,22
	LGIP (ZHOU; WANG, 2012)	0,19	0,28	0,82	0,66	0,29	0,45	0,28
	LDTP (RIVERA; CASTILLO; CHAE, 2015)	0,06	0,11	0,44	0,33	0,12	0,21	0,16
	LPQ (OJANSIVU; HEIKKILÄ, 2008)	0,28	0,37	0,73	0,68	0,26	0,46	0,22
	LFD (LEI et al., 2011)	0,08	0,15	0,55	0,27	0,12	0,23	0,19
	MBC (YANG et al., 2012)	0,14	0,24	0,71	0,54	0,23	0,37	0,24
	Média	0,29	0,35	0,76	0,63	0,28	0,46	0,22
Maior Valor	0,57	0,56	0,90	0,85	0,42	0,63	0,28	
Menor Valor	0,06	0,11	0,44	0,27	0,12	0,21	0,16	

(a) Pré-treinada

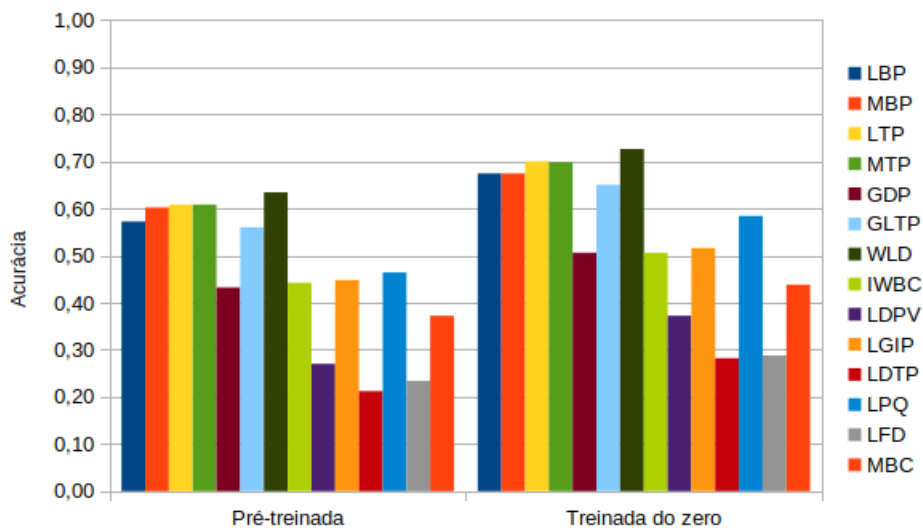
	Descritor	Acurácia nas Bases de Imagens					Média	Desvio Padrão
		Brodatz	CUReT	Kylberg	ALOT	DTD		
GoogLeNet	LBP (OJALA; PIETIKÄINEN; HARWOOD, 1996)	0,59	0,68	0,92	0,81	0,37	0,67	0,21
	MBP (HAFIANE et al., 2008)	0,61	0,68	0,90	0,81	0,37	0,67	0,20
	LTP (TAN; TRIGGS, 2010)	0,66	0,70	0,89	0,85	0,40	0,70	0,19
	MTP (KHAN et al., 2013)	0,77	0,62	0,87	0,81	0,42	0,70	0,18
	GDP (AHMED, 2012)	0,32	0,40	0,82	0,69	0,30	0,51	0,23
	GLTP (AHMED; HOSSAIN, 2013)	0,56	0,62	0,88	0,80	0,39	0,65	0,19
	WLD (CHEN et al., 2010)	0,70	0,65	0,93	0,89	0,46	0,73	0,19
	IWBC (YANG et al., 2016)	0,31	0,45	0,84	0,67	0,26	0,51	0,25
	LDPV (KABIR; JABID; CHAE, 2010)	0,13	0,31	0,69	0,54	0,19	0,37	0,24
	LGIP (ZHOU; WANG, 2012)	0,31	0,39	0,85	0,71	0,32	0,52	0,25
	LDTP (RIVERA; CASTILLO; CHAE, 2015)	0,08	0,20	0,54	0,44	0,15	0,28	0,20
	LPQ (OJANSIVU; HEIKKILÄ, 2008)	0,46	0,55	0,81	0,77	0,33	0,58	0,20
	LFD (LEI et al., 2011)	0,11	0,21	0,60	0,39	0,13	0,29	0,21
	MBC (YANG et al., 2012)	0,22	0,35	0,75	0,60	0,27	0,44	0,23
	Média	0,42	0,49	0,81	0,70	0,31	0,54	0,21
Maior Valor	0,77	0,70	0,89	0,89	0,46	0,73	0,25	
Menor Valor	0,08	0,20	0,39	0,39	0,13	0,28	0,18	

(b) Treinada do zero

Tabela 11 – Resultados para a rede GoogLeNet.

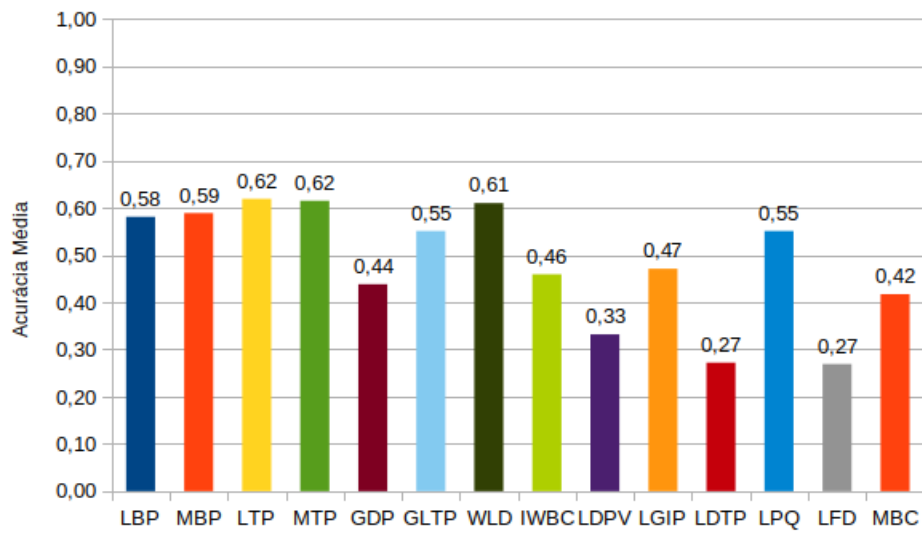


(a) AlexNet

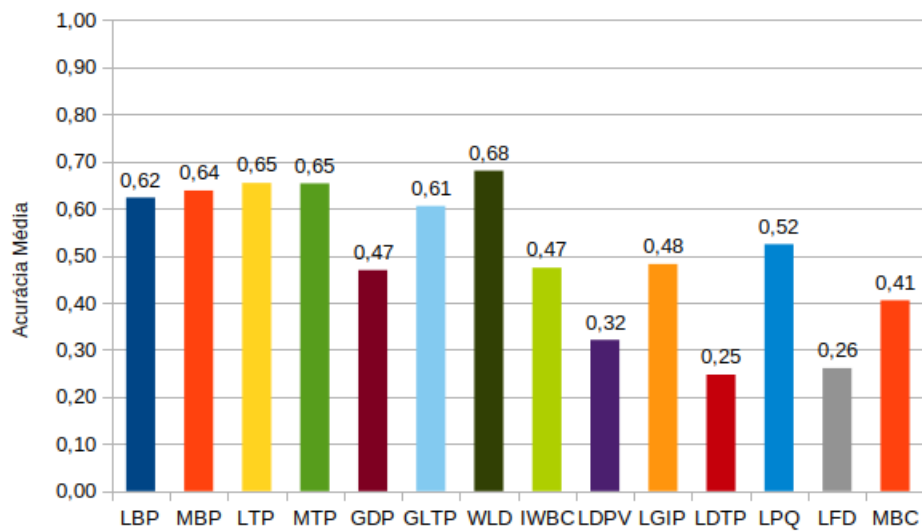


(b) GoogLeNet

Figura 54 – Acurácia dos descritores locais de textura.

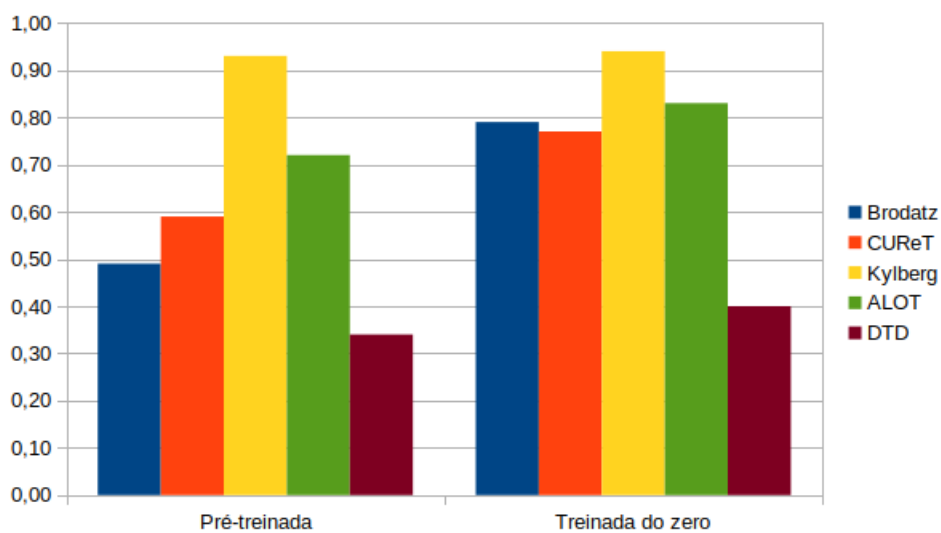


(a) AlexNet

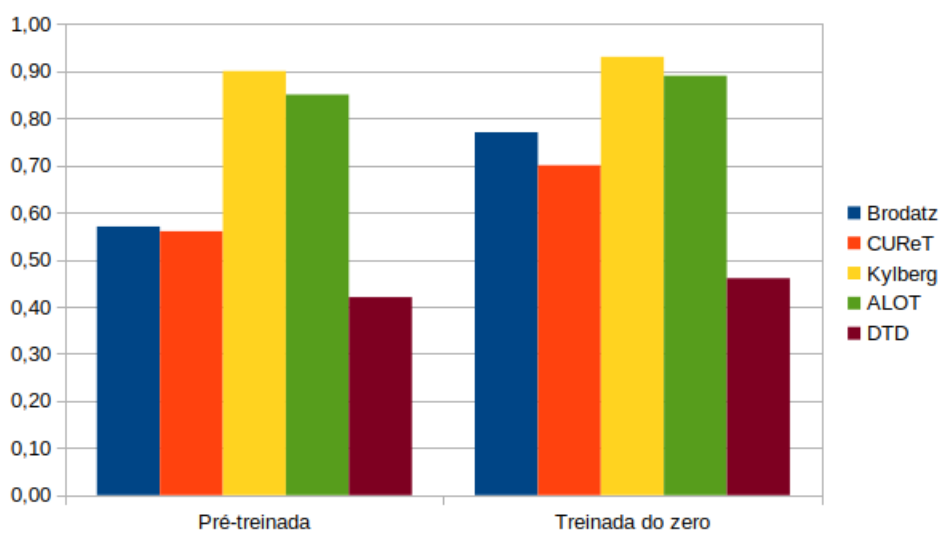


(b) GoogLeNet

Figura 55 – Acurácia média dos descritores locais de textura.



(a) AlexNet



(b) GoogLeNet

Figura 56 – Maiores valores de acurácia obtidos nas bases de imagens.

6.2.4 Discussão

Esta seção discute as conclusões do experimento que tiveram os seus resultados apresentados nas Tabelas 10 e 11 e, também, nas Figuras 54, 55 e 56. Assim, através da análise dos resultados apresentados pelas referidas tabelas, é possível verificar que os descritores LTP (TAN; TRIGGS, 2010), MTP (KHAN et al., 2013) e WLD (CHEN et al., 2010) apresentaram os melhores resultados de acurácia média. Considerando que os descritores LTP (TAN; TRIGGS, 2010) e MTP (KHAN et al., 2013) trabalham com informações de magnitude e o WLD (CHEN et al., 2010) com informações de orientação, os resultados sugerem que as informações de fase não se mostraram interessantes para a caracterização da textura das bases de imagens utilizadas.

Outro ponto importante de ser destacado é que o uso das redes pré-treinadas, com consequente transferência de aprendizado, apresentou resultados de acurácia inferiores aos das redes neurais convolucionais que foram treinadas do início, sem transferência de aprendizado, apenas fazendo uso da arquitetura predefinida da CNN. Este fato evidencia que o aprendizado obtido pelas redes na base ImageNet (RUSSAKOVSKY et al., 2015) está relacionado ao reconhecimento de objetos, e não propriamente à identificação de características de textura.

Destacam-se, ainda, os modestos resultados de acurácia obtidos com a base de imagens DTD. Esta base é realmente desafiadora para classificadores de uma maneira geral (CAVALIN; OLIVEIRA, 2017), pois as imagens de suas classes são organizadas de acordo com adjetivos típicos, que podem produzir situações como a apresentada pela Figura 57, onde é possível verificar que texturas de uma mesma classe podem pertencer a objetos totalmente distintos.



Figura 57 – Exemplos de imagens pertencentes à classe *freckled* da base DTD.

A próxima seção apresentará os descritores locais de textura que foram selecionados a partir dos resultados obtidos e das discussões aqui realizadas.

6.2.5 Descritores Selecionados

Considerando a discussão realizada na seção anterior, temos que os descritores locais de textura que apresentaram as melhores eficácias foram: o LTP (TAN; TRIGGS, 2010), o MTP (KHAN et al., 2013) e o WLD (CHEN et al., 2010). Também foi utilizado um descritor combinado resultante da composição destes três descritores, conforme mostrado pela Figura 58.

Para obtenção de uma imagem do descritor Combinado, primeiramente foi realizada uma composição das imagens resultantes da aplicação dos descritores LTP (I_{LTP}) e MTP (I_{MTP}) sobre a imagem original (I). Assim, foi obtida uma terceira imagem resultante desta composição, a ($I_{LTP-MTP}$). Em seguida, realizou-se a composição desta terceira imagem com aquela resultante da aplicação do descritor WLD (I_{WLD}) sobre I , resultado na imagem do descritor Combinado ($I_{LTP-MTP-WLD}$). Nas duas operações de composição realizadas, foi utilizada a técnica de *alpha blending*, a qual trabalha conforme a seguinte equação:

$$I_C = \alpha \times I_A + (1 - \alpha) \times I_B \quad (6.2)$$

onde I_A e I_B são as imagens de entrada, I_C é a imagem resultante e α é o fator de transparência.

A seção seguinte realiza o detalhamento do processo de classificação segundo o modelo proposto.

6.3 Classificação

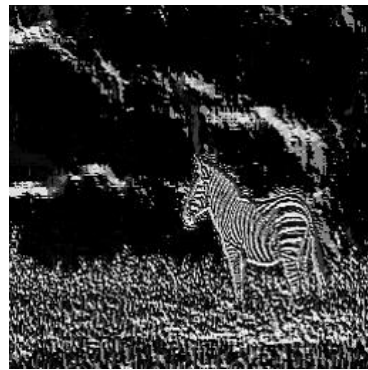
Após o processo de extração de características de textura utilizando-se os descritores anteriormente selecionados, as mesmas são inseridas na camada de entrada de uma rede neural convolucional, conforme ilustrado pela Figura 59.

O modelo utiliza a rede neural convolucional residual ResNet (HE et al., 2016) com 50 camadas (ResNet-50), 25,6 milhões de parâmetros e resolução de $224 \times 224 \times 3$ na camada de entrada da rede. Em nosso experimento, a ResNet-50 foi treinada do zero (*from scratch*) com as bases de imagens descritas no Capítulo 2.

A escolha da rede ResNet-50 se deu em virtude da boa acurácia média das redes residuais e também por possuir um tempo estimado de GPU mediano, conforme pode ser observado na Figura 60, a qual apresenta uma visão geral das principais redes neurais



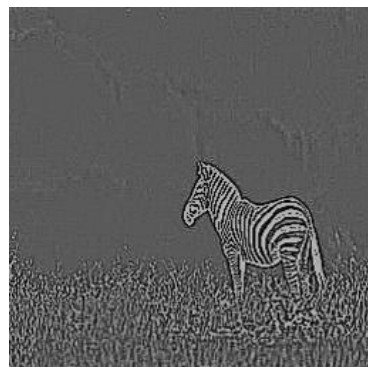
(a) Imagem original



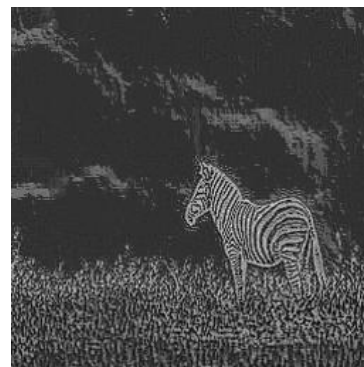
(b) LTP



(c) MTP



(d) WLD



(e) Combinado

Figura 58 – Imagem original e resultado da composição dos descritores da imagem normalizada para 256×256 .

convolucionais no que se refere à acurácia, com tempo de execução estimado de uma GPU NVIDIA Tesla P100, trabalhando com *minibatch* de 128, e, ainda, com volume das redes estimado em função do número de seus parâmetros. Estes dados foram obtidos considerando-se a base de imagens ImageNet (RUSSAKOVSKY et al., 2015).

O trabalho de Kawatsu et al. (2017), por sua vez, apresenta a Figura 61 que traz uma comparação das CNNs mais populares, considerando no eixo vertical a acurácia obtida na base ImageNet e no eixo horizontal o número de operações necessárias para classificar uma imagem. A figura apresenta, ainda, informações a respeito do tamanho das

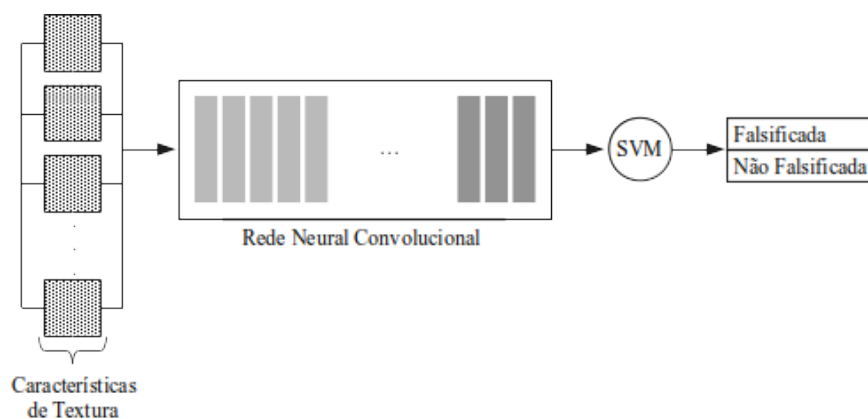


Figura 59 – Modelo proposto para detecção de imagens falsificadas.

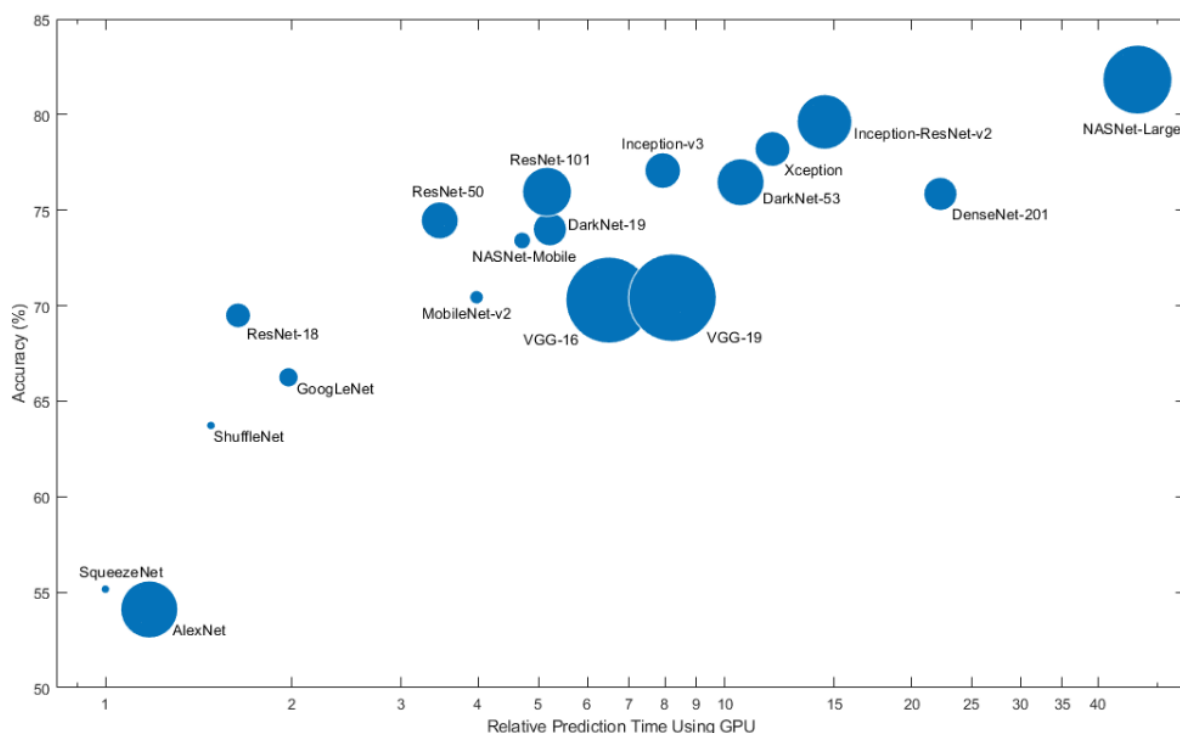


Figura 60 – Desempenho das principais CNNs na base ImageNet (MATLABR2020A, 2020).

redes em função de seus parâmetros.

Após o treinamento da rede ResNet-50, ocorre a classificação das imagens em duas classes apenas: falsificadas ou não falsificadas. Para tal, é utilizado um classificador SVM.

Conforme apresentado por Cristianini e Shawe-Taylor (2000), a SVM constitui uma ferramenta de aprendizagem estatística treinada através do uso de um algoritmo de otimização matemática. Assim, a SVM é baseada na aprendizagem estatística, a qual foi introduzida por Vapnik (1998), e vem sendo muito utilizada em diversos estudos de reconhecimento de padrões, especialmente aqueles que possuem exatamente duas classes, sendo, portanto, considerado um classificador binário (BYUN; LEE, 2002). Dessa forma,

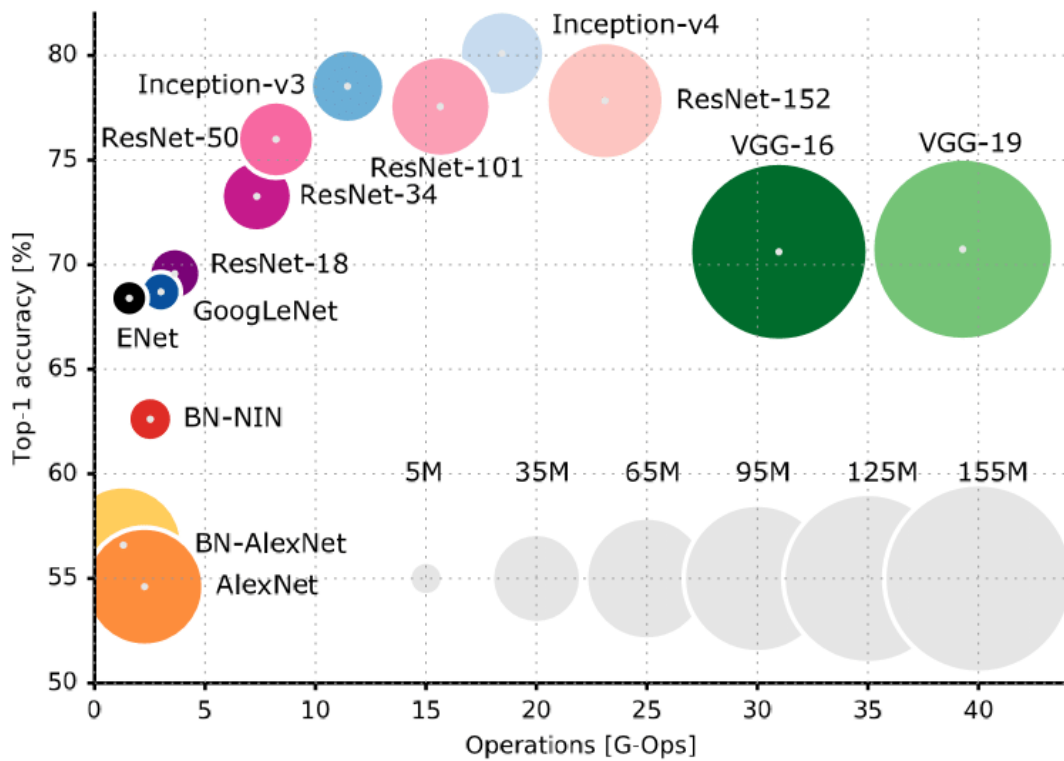


Figura 61 – Desempenho das principais CNNs na base ImageNet (KAWATSU et al., 2017).

tendo em vista que o objetivo deste trabalho é classificar imagens digitais em duas classes (falsificadas ou não falsificadas), o uso da SVM se apresenta justificável.

O classificador SVM mapeia os dados das características em um espaço de alta dimensão e, em seguida, encontra o melhor hiperplano que separa todos os pontos de dados de uma classe dos pontos de outra classe. Assim, considerando uma amostra x , a mesma é classificada pela função:

$$f(x) = \text{sign} \left(\sum_{n=1}^l \alpha_n y_n K(x_i, x) + b \right) \quad (6.3)$$

onde (x_i, y_i) são rotulados como exemplos de treinamento, α_i são multiplicadores de Lagrange do problema de otimização dupla, b é o limiar do hiperplano, e K é uma função de *kernel* como a função linear de base radial (RBF, do inglês, *radial basis function*).

O classificador SVM tem sido utilizado com sucesso na área forense de imagens digitais, como pode ser observado nos trabalhos de Korus e Huang (2016) e Boulkenafet, Komulainen e Hadid (2016).

Na seção seguinte, serão apresentadas as considerações finais deste capítulo.

6.4 Considerações Finais do Capítulo

Este capítulo apresentou o modelo proposto, detalhando o processo de seleção dos descritores de textura com melhor eficácia que são utilizados pelo modelo (LTP (TAN; TRIGGS, 2010), MTP (KHAN et al., 2013) e WLD (CHEN et al., 2010)) e, ainda, o processo de classificação das imagens em duas classes (falsificadas e não falsificadas).

No capítulo seguinte, serão apresentados os resultados experimentais obtidos com a aplicação do modelo proposto.

Resultados Experimentais

Sumário

7.1	Introdução	113
7.2	Configurações do Experimento	115
7.3	Resultados	115
7.3.1	Resultados com o Descritor LTP	116
7.3.2	Resultados com o Descritor MTP	117
7.3.3	Resultados com Descritor WLD	118
7.3.4	Resultados com o Descritor Combinado	119
7.3.5	Análise Comparativa dos Descritores	120
7.4	Discussão	121
7.5	Considerações Finais do Capítulo	125

7.1 Introdução

O experimento foi realizado considerando o modelo proposto para detecção de imagens falsificadas introduzido pela Figura 59, sendo os resultados de cada caso avaliados considerando-se as matrizes de confusão geradas. A Figura 62 ilustra uma matriz de confusão e a definição de seus componentes.

Assim, considerando as matrizes de confusão dos casos executados, as métricas analisadas foram: a acurácia (A), a *recall* (R), a precisão (P) e a *F1-Score*. A acurácia é

		Esperado	
		Falsificada	Não Falsificada
Predito	Falsificada	Verdadeiros-positivos (VP)	Falsos-positivos (FP)
	Não Falsificada	Falsos-negativos (FN)	Verdadeiros-negativos (VN)

Figura 62 – Visão geral da matriz de confusão.

razão entre a soma de VP com VN e o total de imagens. Já a *recall*, é a razão entre VP e a soma de VP com FN . A precisão é dada pela divisão de VP e a soma de VP com FP . Por último, a *F1-Score* é a média harmônica entre a precisão e a *recall* (ADNANE, 2020; AMIDI; AMIDI, 2018; NIGHANIA, 2018; HOSSIN; M.N, 2015; PROVOST; FAWCETT, 2013). As Equações 7.1, 7.2, 7.3 e 7.4 apresentam, respectivamente, as fórmulas para a acurácia, a *recall*, a precisão e a *F1-Score*:

$$A = \frac{VP + VN}{VP + VN + FP + FN} \quad (7.1)$$

$$R = \frac{VP}{VP + FN} \quad (7.2)$$

$$P = \frac{VP}{VP + FP} \quad (7.3)$$

$$F1-Score = \frac{2 \times P \times R}{P + R} \quad (7.4)$$

O experimento considerou o uso de quatro bases para forense em imagens digitais organizadas em conjuntos de treinamento e teste, conforme método de validação cruzada (*cross-validation*) **k-folds**, onde o número de partes (n) foi igual a cinco ($n = 5$). Assim, as equações 7.5, 7.6, 7.7 e 7.8 descrevem, respectivamente, a média para a acurácia (\bar{A}), a *recall* (\bar{R}), a precisão (\bar{P}) e a *F1-Score* ($\overline{F1-Score}$).

$$\bar{A} = \frac{\sum_{k=1}^n \frac{VP_k + VN_k}{VP_k + VN_k + FP_k + FN_k}}{n} \quad (7.5)$$

$$\bar{R} = \frac{\sum_{k=1}^n \frac{VP_k}{VP_k + FN_k}}{n} \quad (7.6)$$

$$\bar{P} = \frac{\sum_{k=1}^n \frac{VP_k}{VP_k + FP_k}}{n} \quad (7.7)$$

$$\overline{F1-Score} = \frac{\sum_{k=1}^n \frac{2 \times P_k \times R_k}{P_k + R_k}}{n} \quad (7.8)$$

A seção seguinte detalhará as configurações que foram utilizadas no experimento realizado.

7.2 Configurações do Experimento

O experimento considerou o uso da rede neural convolucional **ResNet-50** e das quatro bases de forense em imagens digitais que foram detalhadas na Seção 2.4, a saber: CASIA v1.0 (DONG; WANG; TAN, 2013), CASIA v2.0 (DONG; WANG; TAN, 2013), Columbia (NG; CHANG, 2004) e CoMoFoD (TRALIC et al., 2013).

O treinamento da CNN foi realizado com no **máximo 50 épocas** para cada caso, com taxa inicial de aprendizagem igual a **0,00001** e, ainda, com o gradiente descendente estocástico utilizado como otimizador.

Para execução dos algoritmos do experimento, foi utilizado o mesmo computador e ambiente descritos na Tabela 9.

7.3 Resultados

Este estudo foi dividido em quatro casos, em que os três primeiros avaliam os descritores LTP (TAN; TRIGGS, 2010), MTP (KHAN et al., 2013) e WLD (CHEN et al., 2010), respectivamente, enquanto no quarto, e último, é avaliado o desempenho a partir da composição dos três descritores anteriores.

As seções seguintes apresentam os resultados para cada um dos casos realizados.

7.3.1 Resultados com o Descritor LTP

A Figura 63 exibe as matrizes de confusão obtidas através do uso do descritor LTP (TAN; TRIGGS, 2010).

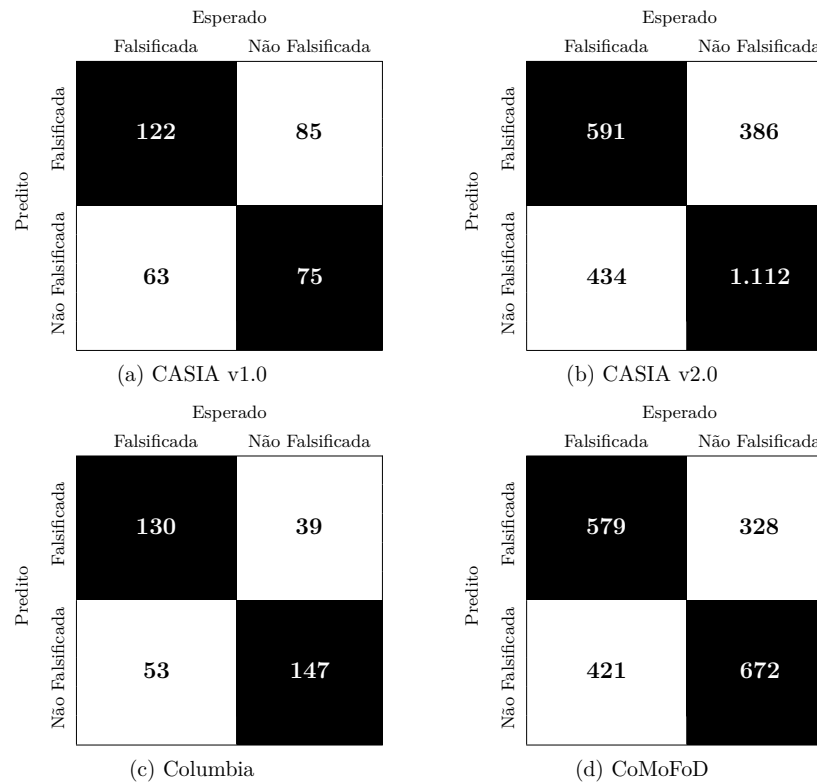


Figura 63 – Matrizes de confusão para o descritor LTP (TAN; TRIGGS, 2010).

A Tabela 12 traz os dados de acurácia, *recall*, precisão e *F1-Score* do descritor em discussão, obtidos a partir das matrizes de confusão do caso.

Base de Imagens	\bar{A}	\bar{R}	\bar{P}	$\overline{F1-Score}$
CASIA v1.0 (DONG; WANG; TAN, 2013)	0,57	0,66	0,59	0,62
CASIA v2.0 (DONG; WANG; TAN, 2013)	0,67	0,58	0,60	0,59
Columbia (NG; CHANG, 2004)	0,75	0,71	0,77	0,74
CoMoFoD (TRALIC et al., 2013)	0,63	0,58	0,64	0,61
Média	0,66	0,63	0,65	0,64

Tabela 12 – Resultados para o descritor LTP (TAN; TRIGGS, 2010).

A seção seguinte, apresentará os resultados obtidos com o uso do descritor MTP (KHAN et al., 2013).

7.3.2 Resultados com o Descritor MTP

A Figura 64 a seguir apresenta as matrizes de confusão obtidas através do uso do descritor MTP (KHAN et al., 2013).

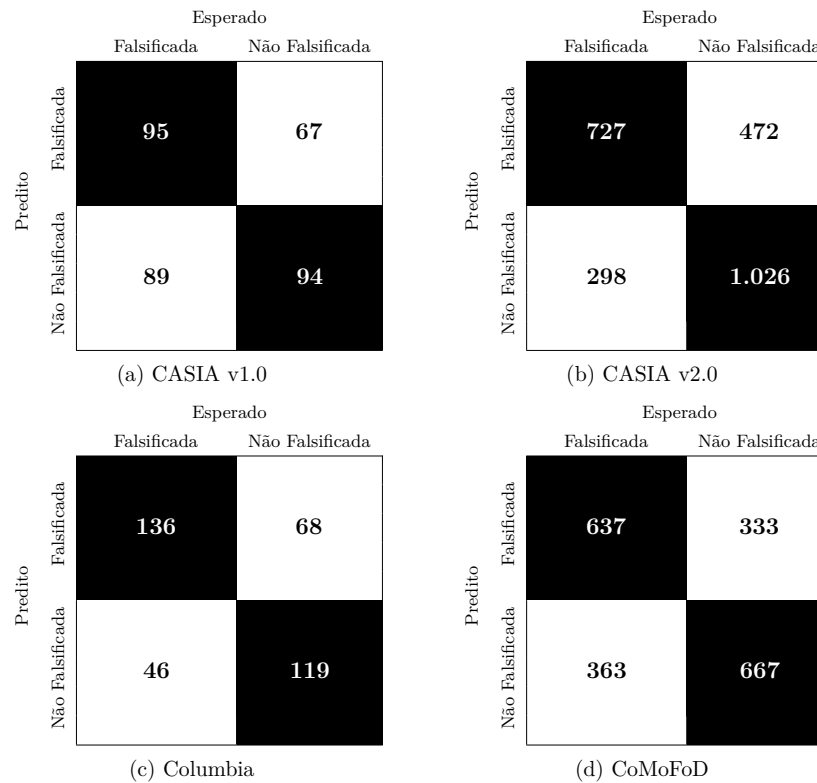


Figura 64 – Matrizes de confusão para o descritor MTP (KHAN et al., 2013).

A Tabela 13 exhibe os resultados de eficácia do descritor MTP (KHAN et al., 2013).

Base de Imagens	\bar{A}	\bar{R}	\bar{P}	$\overline{F1-Score}$
CASIA v1.0 (DONG; WANG; TAN, 2013)	0,55	0,52	0,59	0,55
CASIA v2.0 (DONG; WANG; TAN, 2013)	0,69	0,71	0,61	0,65
Columbia (NG; CHANG, 2004)	0,69	0,75	0,67	0,70
CoMoFoD (TRALIC et al., 2013)	0,65	0,64	0,66	0,65
Média	0,65	0,65	0,63	0,64

Tabela 13 – Resultados para o descritor MTP (KHAN et al., 2013).

A próxima seção, trará os resultados obtidos com a utilização do descritor WLD (CHEN et al., 2010).

7.3.3 Resultados com Descritor WLD

A Figura 65 apresenta as matrizes de confusão obtidas através do uso do modelo proposto com o descritor WLD (CHEN et al., 2010).

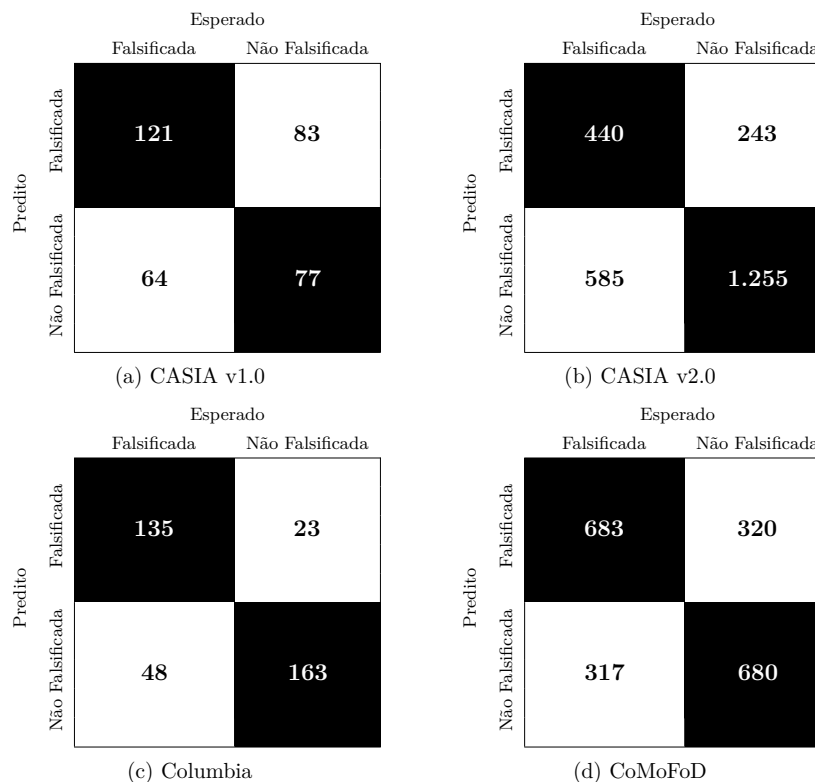


Figura 65 – Matrizes de confusão para o descritor WLD (CHEN et al., 2010).

Tendo em vista as matrizes de confusão para o descritor WLD (CHEN et al., 2010), a Tabela 14 apresenta a eficácia deste descritor em função das bases de imagens utilizadas.

Base de Imagens	\bar{A}	\bar{R}	\bar{P}	$\overline{F1-Score}$
CASIA v1.0 (DONG; WANG; TAN, 2013)	0,58	0,65	0,59	0,62
CASIA v2.0 (DONG; WANG; TAN, 2013)	0,67	0,43	0,64	0,52
Columbia (NG; CHANG, 2004)	0,81	0,74	0,85	0,79
CoMoFoD (TRALIC et al., 2013)	0,68	0,68	0,68	0,68
Média	0,68	0,63	0,69	0,65

Tabela 14 – Resultados para o descritor WLD (CHEN et al., 2010).

A seção seguinte apresentará os resultados obtidos com a utilização do descritor Combinado.

7.3.4 Resultados com o Descritor Combinado

A Figura 66 apresenta as matrizes de confusão que foram geradas a partir do uso do modelo proposto com o descritor formado pela composição dos descritores LTP (TAN; TRIGGS, 2010), MTP (KHAN et al., 2013) e WLD (CHEN et al., 2010), conforme detalhado na Seção 6.2.5.

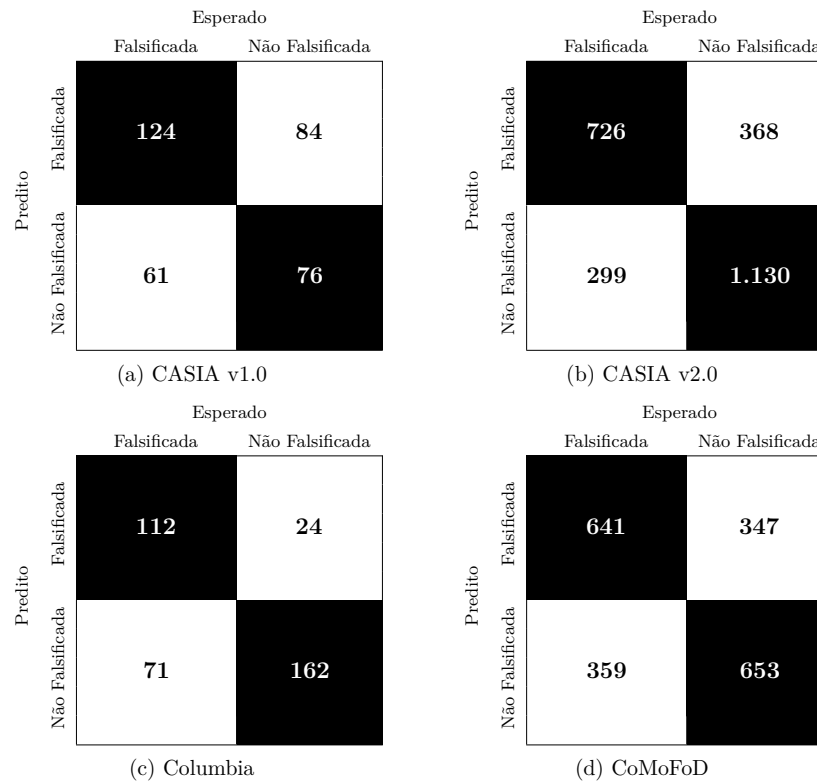


Figura 66 – Matrizes de confusão para o descritor Combinado.

A Tabela 15 apresenta a eficácia do descritor Combinado em função das bases de imagens.

Base de Imagens	\bar{A}	\bar{R}	\bar{P}	$\overline{F1-Score}$
CASIA v1.0 (DONG; WANG; TAN, 2013)	0,58	0,67	0,60	0,63
CASIA v2.0 (DONG; WANG; TAN, 2013)	0,74	0,71	0,66	0,68
Columbia (NG; CHANG, 2004)	0,74	0,61	0,82	0,70
CoMoFoD (TRALIC et al., 2013)	0,65	0,64	0,65	0,64
Média	0,68	0,66	0,68	0,67

Tabela 15 – Resultados para o descritor Combinado.

A próxima seção descreve uma análise comparativa dos descritores utilizados neste experimento.

7.3.5 Análise Comparativa dos Descritores

Para fins de análise comparativa, a Tabela 16 apresenta a eficácia dos descritores organizada por bases de imagens utilizadas no experimento.

	Descritor	Base de Imagem				Média
		CASIA v1.0	CASIA v2.0	Columbia	CoMoFoD	
Acurácia (\bar{A})	LTP (TAN; TRIGGS, 2010)	0,57	0,67	0,75	0,63	0,66
	MTP (KHAN et al., 2013)	0,55	0,69	0,69	0,65	0,65
	WLD (CHEN et al., 2010)	0,58	0,67	0,81	0,68	0,68
	Combinado	0,58	0,74	0,74	0,65	0,68
Recall (\bar{R})	LTP (TAN; TRIGGS, 2010)	0,66	0,58	0,71	0,58	0,63
	MTP (KHAN et al., 2013)	0,52	0,71	0,75	0,64	0,65
	WLD (CHEN et al., 2010)	0,65	0,43	0,74	0,68	0,63
	Combinado	0,67	0,71	0,61	0,64	0,66
Precisão (\bar{P})	LTP (TAN; TRIGGS, 2010)	0,59	0,60	0,77	0,64	0,65
	MTP (KHAN et al., 2013)	0,59	0,61	0,67	0,66	0,63
	WLD (CHEN et al., 2010)	0,59	0,64	0,85	0,68	0,69
	Combinado	0,60	0,66	0,82	0,65	0,68
$\bar{F1}$ -Score	LTP (TAN; TRIGGS, 2010)	0,62	0,59	0,74	0,61	0,64
	MTP (KHAN et al., 2013)	0,55	0,65	0,70	0,65	0,64
	WLD (CHEN et al., 2010)	0,62	0,52	0,79	0,68	0,65
	Combinado	0,63	0,68	0,70	0,64	0,67

Tabela 16 – Eficácia dos descritores locais de textura analisados.

A Figura 67 traz a representação gráfica da eficácia de cada um dos descritores locais de textura utilizados.

A seção a seguir, discutirá os resultados obtidos com o experimento realizado.

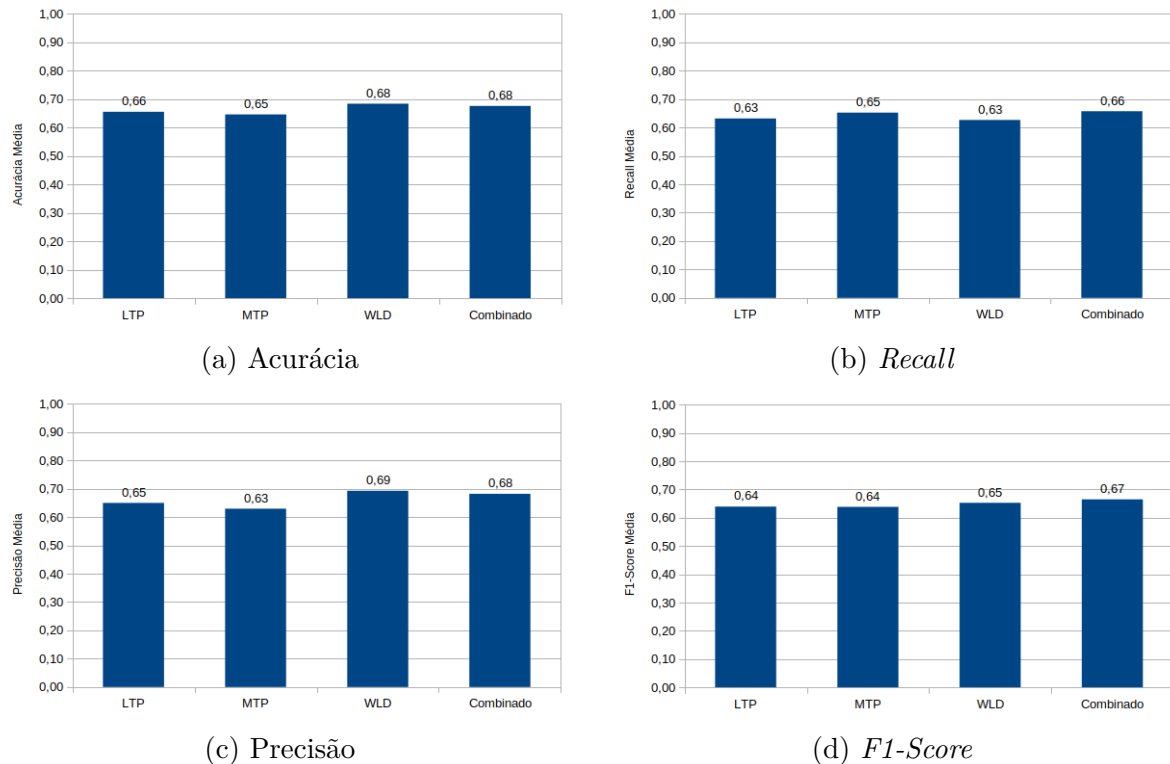


Figura 67 – Eficácia média dos descritores.

7.4 Discussão

Neste capítulo foram apresentados os resultados do experimento da aplicação do modelo proposto em quatro bases públicas, a CASIA v1.0 e a v2.0 (DONG; WANG; TAN, 2013), a Columbia (NG; CHANG, 2004) e a CoMoFoD (TRALIC et al., 2013), as quais tiveram suas principais características aqui evidenciadas.

Ao todo, o experimento fez uso de 26.180 imagens digitais, sendo 14.224 autênticas e 11.956 falsificadas. Contudo, é possível observar que, apesar do baixo volume de imagens, se comparado, por exemplo, com a base ImageNet (RUSSAKOVSKY et al., 2015), com mais de 1 milhão de imagens, o modelo conseguiu 0,81 de acurácia na base Columbia (NG; CHANG, 2004) com o descritor WLD (CHEN et al., 2010).

Outro ponto interessante de ser observado, reside no fato de que o descritor formado pela composição dos descritores LTP (TAN; TRIGGS, 2010), MTP (KHAN et al., 2013) e WLD (CHEN et al., 2010), o descritor Combinado, conseguiu alcançar, juntamente com o WLD (CHEN et al., 2010), a maior acurácia média, 0,68, enquanto que os descritores LTP (TAN; TRIGGS, 2010) e MTP (KHAN et al., 2013) alcançaram, respectivamente, 0,66 e 0,65.

No que se refere à taxa de verdadeiros positivos, a métrica *recall*, observa-se que o seu maior valor médio, de 0,66, é obtido com o descritor Combinado.

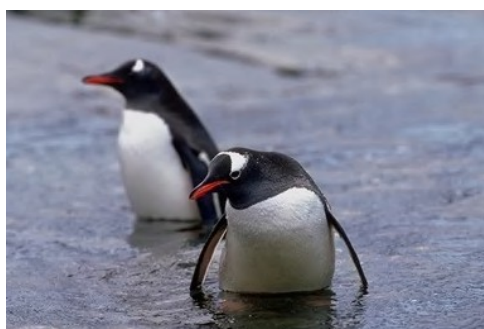
Já em relação à precisão média, é possível verificar que o descritor WLD (CHEN et al., 2010) alcança o maior valor dentre os outros descritores analisados. Dessa forma, a precisão média do descritor WLD (CHEN et al., 2010) é de 0,69, enquanto que os descritores LTP (TAN; TRIGGS, 2010), MTP (KHAN et al., 2013) e Combinado apresentaram, respectivamente, os valores 0,65, 0,63 e 0,68.

No que tange à média harmônica entre a precisão e a *recall*, a métrica *F1-Score*, é possível observar que o descritor Combinado apresentou o maior valor médio, 0,67. Os descritores LTP (TAN; TRIGGS, 2010) e MTP (KHAN et al., 2013) apresentaram, igualmente, 0,64, e o descritor WLD (CHEN et al., 2010) obteve um valor de 0,65.

Contudo, foi possível observar com os resultados do experimento que 65 imagens falsificadas da base CASIA v1.0 foram classificadas como não falsificadas (falsas-negativas) utilizando-se qualquer um dos quatro descritores de textura. A Figura 68 apresenta um exemplo destas imagens.



(a)



(b)



(c)

Figura 68 – Exemplo de classificação de imagens falsas-negativas da base CASIA v1.0. (a) Imagem falsificada (Sp_D_NNN_A_ani0033_ani0034_0285.jpg) formada a partir de (b) e (c). (b) imagem autêntica (Au_ani_0033.jpg). (c) imagem autêntica (Au_ani_0034.jpg).

Em se tratando de imagens falsas-positivas, foi possível verificar que 168 imagens autênticas da base CASIA v1.0 foram classificadas como falsificadas por todos os descritores

do experimento. A Figura 69 apresenta dois exemplos destas imagens.



Figura 69 – Exemplo de classificação de imagens falsas-positivas da base CASIA v1.0. (a) Imagem autêntica (Au_cha_0025.jpg). (b) Imagem autêntica (Au_cha_0036.jpg).

No que se refere à base CASIA v2.0, temos a existência de 534 diferentes imagens que foram classificadas como falsas-negativas por todos os descritores utilizados no experimento. A Figura 70 apresenta um exemplo destas figuras.

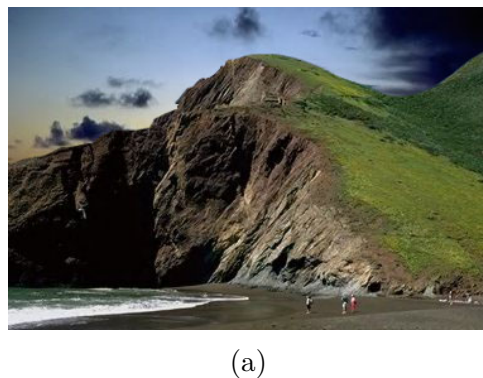


Figura 70 – Exemplo de classificação de imagens falsas-negativas da base CASIA v2.0. (a) Imagem falsificada (Tp_D_CRN_M_N_nat00019_nat00024_11047.jpg) formada a partir de (b) e (c). (b) imagem autêntica (Au_nat_00019.jpg). (c) imagen autêntica (Au_nat_00024.jpg).

Considerando-se, ainda, a base CASIA v2.0, é possível observar que 367 imagens foram classificadas como falsas-positivas pelos quatro descritores utilizados. A Figura 71 apresenta um exemplo destas figuras.



Figura 71 – Exemplo de classificação de imagens falsas-positivas da base CASIA v2.0. (a) Imagem autêntica (Au_cha_00070.jpg). (b) Imagem autêntica (Au_cha_00075.jpg).

A base Columbia, por sua vez, apresentou 72 imagens classificadas por todos os descritores do experimento como falsas-negativas, e 18 como falsas-positivas. A Figura 72 apresenta um exemplo de imagem classificada como falsa-negativa. Já a Figura 73 traz um exemplo de imagem classificada como falsa-positiva.

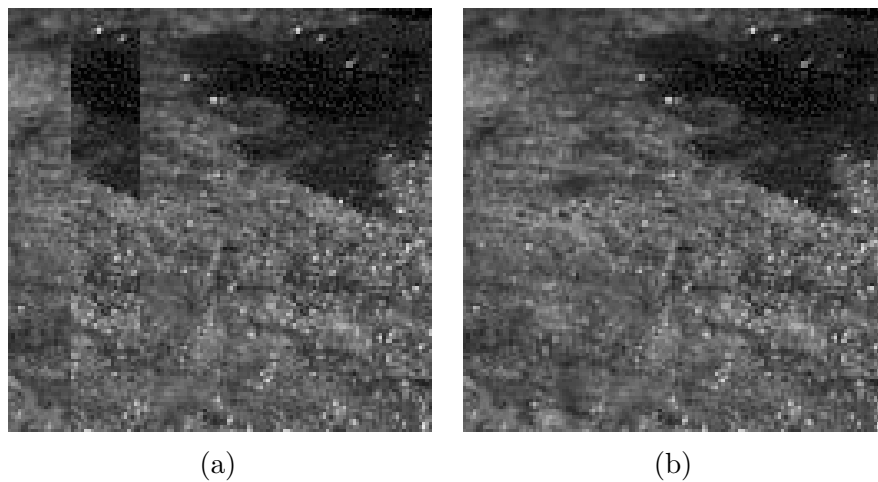


Figura 72 – Exemplo de classificação de imagens falsas-negativas da base Columbia. (a) Imagem falsificada (SP_T_053.bmp) obtida a partir de (b). (b) Imagem autêntica (AU_T_053.bmp).

Em se tratando da base CoMoFoD, é possível observar a existência de 378 imagens falsas-negativas e 214 falsas-positivas, todas elas levando-se em conta todos os descritores do experimento. A Figura 74 apresenta um exemplo de imagem falsa-negativa, enquanto que a Figura 75 traz exemplos de imagens falsas-positivas.

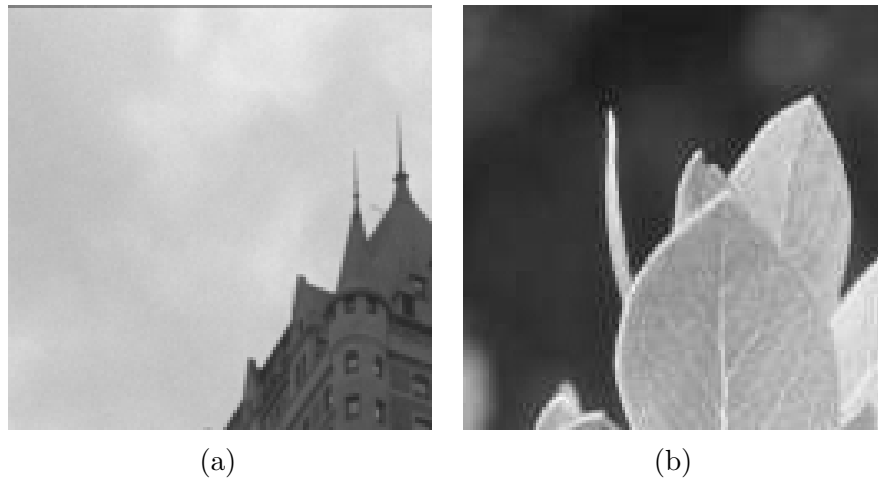


Figura 73 – Exemplo de classificação de imagens falsas-positivas da base Columbia. (a) Imagem autêntica (AU_TS_O_081.bmp). (b) Imagem autêntica (AU_SS_O_047.bmp).

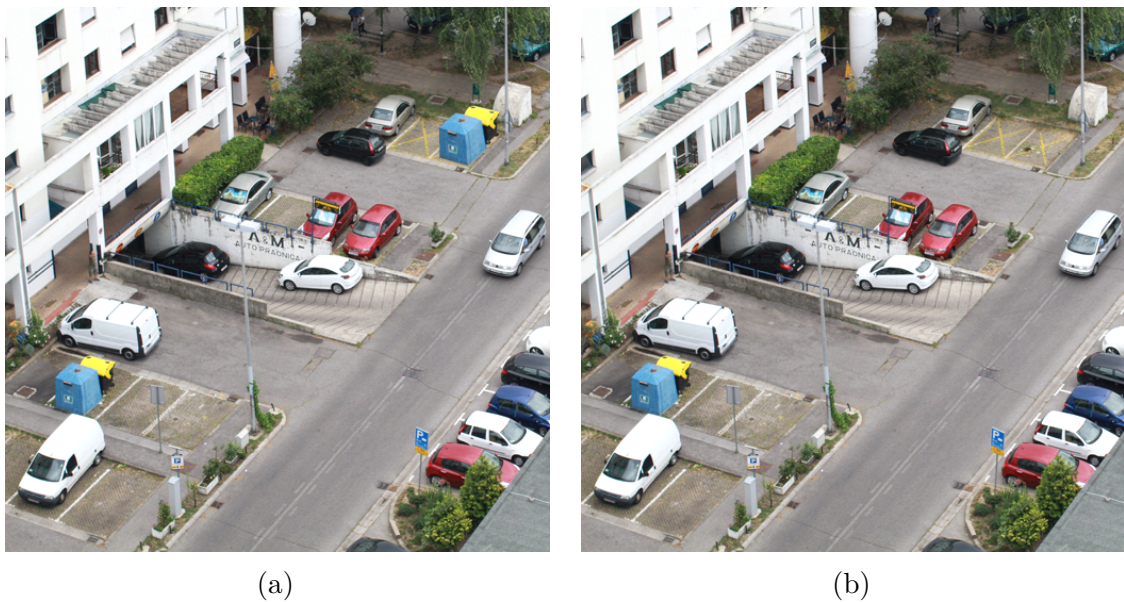


Figura 74 – Exemplo de classificação de imagens falsas-negativas da base CoMoFoD. (a) Imagem falsificada (032_F_BC1.png) obtida a partir de (b). (b) Imagem autêntica (032_O_BC1.png).

A seção seguinte apresentará as considerações finais deste capítulo de resultados experimentais.

7.5 Considerações Finais do Capítulo

Neste capítulo, foram apresentados os resultados do experimento realizado com a implementação do modelo proposto. O experimento considerou a utilização de quatro bases de imagens de digitais (CASIA v1.0 e v2.0 [Dong, Wang e Tan \(2013\)](#), Columbia ([NG; CHANG, 2004](#)) e CoMoFoD ([TRALIC et al., 2013](#))). Para a extração de características

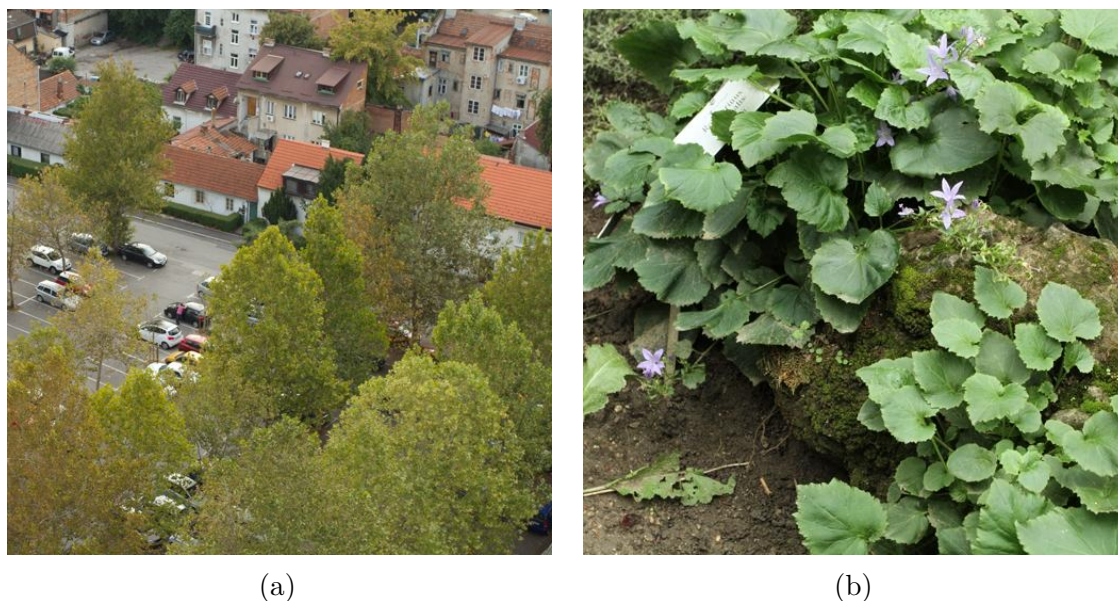


Figura 75 – Exemplo de classificação de imagens falsas-positivas da base CoMoFoD. (a) Imagem autêntica (090_O_JC7.jpg). (b) Imagem autêntica (028_O_JC9.jpg).

das imagens destas bases, foram utilizados três descritores locais de textura (LTP (TAN; TRIGGS, 2010), MTP (KHAN et al., 2013) e WLD (CHEN et al., 2010)) e também um quarto descritor formado pela composição dos três descritores mencionados anteriormente. Os resultados foram analisados a partir das matrizes de confusão geradas para cada um dos casos (descritor local de textura *versus* base de imagens) estudados.

O capítulo seguinte apresentará as conclusões auferidas com o desenvolvimento do presente trabalho, bem como as contribuições e os trabalhos futuros.

Conclusões, Contribuições e Trabalhos Futuros

8.1 Conclusões e Contribuições

Neste trabalho, foi proposto um modelo de suporte à decisão na identificação de imagens falsificadas com redes neurais convolucionais a partir de informações oferecidas por descritores locais de textura.

Para esta finalidade, foi conduzida uma revisão bibliográfica de métodos de detecção de imagens falsificadas, abrangendo métodos ativos e passivos, bem como aqueles baseados em *deep learning*. Além disso, o trabalho apresentou um vasto e atualizado conjunto de referências sintetizadas em forma textual e gráfica, permitindo uma importante contribuição para a área de forense em imagens digitais na forma de publicação do trabalho de [Ferreira et al. \(2020\)](#) em periódico internacional indexado.

O trabalho também realizou uma revisão sobre bases públicas de imagens, destacando 9 delas e, então, descrevendo as 5 principais bases de textura (Brodatz ([BRODATZ, 1966](#)), CURET ([DANA et al., 1999](#)), Kylberg ([KYLBERG, 2011](#)), ALOT ([HOSSAIN; SERIKAWA, 2013](#)) e DTD ([CIMPOI et al., 2014](#))) e as 4 bases de identificação de imagens falsificadas (CASIA v1.0 e v2.0 ([DONG; WANG; TAN, 2013](#)), Columbia ([NG; CHANG, 2004](#)) e CoMoFoD ([TRALIC et al., 2013](#))). Estas bases foram, então, descritas e organizadas em um contexto didático que pode apoiar o desenvolvimento de outras pesquisas relacionadas à identificação de imagens falsificadas que levem em consideração as informações de textura.

Outra contribuição importante deste trabalho reside na avaliação de eficácia de 14

descritores locais de textura. Esta avaliação permitiu verificar que os descritores LTP (TAN; TRIGGS, 2010), MTP (KHAN et al., 2013) e WLD (CHEN et al., 2010) apresentaram melhor capacidade de representar a textura das imagens das bases de textura utilizadas. Também foi avaliado um descritor formado pela composição dos três outros aqui citados.

Além disso, o trabalho também apresenta as abordagens para detecção de imagens falsificadas baseadas em aprendizagem profunda, discutindo trabalhos atuais e seus respectivos resultados. Foi possível observar que estes trabalhos apresentaram o uso das redes neurais convolucionais como principal algoritmo de *deep learning*.

Por fim, observa-se que, com o avanço das arquiteturas da redes neurais convolucionais, as CNNs vêm demonstrando grande capacidade de aprendizado de características intrínsecas de textura, favorecendo a sua utilização em problemas onde exista, por exemplo, uma variação muito grande na aparência das texturas. Contudo, destaca-se que o modelo proposto possui espaço para aperfeiçoamentos que podem ser implantados conforme o desenvolvimento dos trabalhos futuros relacionados na seção seguinte.

8.2 Trabalhos Futuros

Na sequência, é apresentada uma relação de trabalhos a serem desenvolvidos:

- Aplicar técnicas de aumento nas bases de imagens para viabilizar o balanceamento das mesmas conforme discutido por Johnson e Khoshgoftaar (2019);
- Testar o modelo proposto com o uso de redes neurais convolucionais de altíssimo nível de acurácia como a NASNet-Large, descrita no trabalho de Zoph et al. (2018), a qual possui mais de 88 milhões de parâmetros;
- Avaliar descritores de textura não considerados neste trabalho como o FV, do inglês, Fisher *vector*, proposto por Sánchez et al. (2013) e também o MRELBP, do inglês, *median robust extended local binary pattern*, apresentado no trabalho de Liu et al. (2016);
- Analisar a eficácia do modelo sobre a base de imagens KTH-TIPS2 proposta por Caputo, Hayman e Mallikarjuna (2005). Esta base possui imagens de materiais fotografados sob uma variedade de posições, iluminação e escala.

A seguir, são relacionadas todas as referências que foram utilizadas na confecção deste texto.

Referências

- ABRAHAM, B.; NAIR, M. S. Computer-aided classification of prostate cancer grade groups from MRI images using texture features and stacked sparse autoencoder. *Computerized Medical Imaging and Graphics*, Elsevier Ltd, v. 69, p. 60–68, 2018. ISSN 18790771. Disponível em: <<https://doi.org/10.1016/j.compmedimag.2018.08.006>>. Citado na página 61.
- ADNANE, O. *Machine Learning Model Evaluation Metrics*. 2020. Disponível em: <<https://www.kaggle.com/ishivinal/machine-learning-model-evaluation-metrics>>. Citado na página 114.
- AGGAR, C. C.; WAL. *Neural Networks and Deep Learning*. [S.l.]: Springer, 2018. ISBN 978-3-319-94463-0. Citado na página 49.
- AHMED, F. Gradient directional pattern: a robust feature descriptor for facial expression recognition. *Electronics Letters*, v. 48, n. 19, p. 1203, 2012. ISSN 00135194. Citado 4 vezes nas páginas 68, 73, 102 e 103.
- AHMED, F.; HOSSAIN, E. Automated Facial Expression Recognition Using Gradient-Based Ternary Texture Patterns. *Chinese Journal of Engineering*, v. 2013, 2013. Citado 5 vezes nas páginas 68, 69, 73, 102 e 103.
- AL-ATABY, A. A.; AL-NAIMA, F. M. High Capacity Image Steganography Based on Curvelet Transform. In: *2011 Developments in E-systems Engineering*. [S.l.: s.n.], 2011. p. 191–196. ISBN 9780769545936. Citado 3 vezes nas páginas 12, 26 e 30.
- AL-DMOUR, H.; AL-ANI, A. A steganography embedding method based on edge identification and XOR coding. *Expert Systems With Applications*, Elsevier Ltd, v. 46, p. 293–306, 2016. ISSN 0957-4174. Citado na página 30.
- AL-DMOUR, H.; AL-ANI, A.; NGUYEN, H. T. An Efficient Steganography Method for Hiding Patient Confidential Information. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. [S.l.: s.n.], 2014. p. 222–225. ISBN 9781424479290. Citado na página 27.

AL-HAMMADI, M. M.; EMMANUEL, S. Improving SURF based copy-move forgery detection using super resolution. In: *2016 IEEE International Symposium on Multimedia (ISM)*. [S.l.: s.n.], 2016. p. 341–344. Citado na página 36.

ALBERRY, H. A.; HEGAZY, A.; COMPUTING, F.; JOURNAL, I.; GI, S.; DETECTION, F. A Fast SIFT Based Method for Copy Move Forgery Detection. *Future Computing and Informatics Journal*, 2018. Citado na página 36.

ALOM, M. Z.; TAHA, T. M.; YAKOPCIC, C.; WESTBERG, S.; SIDIKE, P.; NASRIN, M. S.; HASAN, M.; Van Essen, B. C.; AWWAL, A. A.; ASARI, V. K. A state-of-the-art survey on deep learning theory and architectures. *Electronics (Switzerland)*, v. 8, n. 3, p. 1–67, 2019. ISSN 20799292. Citado na página 54.

ALTURKI, F.; MERSEREAU, R. Secure Blind Image Steganographic Technique Using Discrete Fourier. In: *Proceedings 2001 International Conference on Image Processing*. [S.l.: s.n.], 2001. v. 2, p. 542–545. Citado na página 29.

AMERINI, I.; BALLAN, L.; CALDELLI, R.; Del Bimbo, A.; SERRA, G. A SIFT-based forensic method for copy-move attack detection and transformation recovery. *IEEE Transactions on Information Forensics and Security*, v. 6, n. 3 PART 2, p. 1099–1110, 2011. ISSN 15566013. Citado na página 36.

AMIDI, A.; AMIDI, S. *Dicas e truques de aprendizado de máquina*. 2018. Disponível em: <<https://stanford.edu/~shervine/1/pt/teaching/cs-229/dicas-truques-aprendizado-maquinaclassification-metrics>>. Citado na página 114.

ANDREARCZYK, V.; WHELAN, P. F. Using filter banks in Convolutional Neural Networks for texture classification. *Pattern Recognition Letters*, v. 84, p. 63–69, 2016. ISSN 01678655. Citado 2 vezes nas páginas 67 e 99.

ANDREARCZYK, V.; WHELAN, P. F. *Texture Segmentation with Fully Convolutional Networks*. Tese (Doutorado), 2017. Disponível em: <<http://arxiv.org/abs/1703.05230>>. Citado 2 vezes nas páginas 58 e 59.

ANJU, R. Modified Algorithm for Digital Image Watermarking Using Combined DCT and DWT. *International Journal of Information and Computation Technology*, v. 3, n. 7, p. 691–700, 2013. Citado na página 95.

ARDIZZONE, E.; BRUNO, A.; MAZZOLA, G. Copy-move forgery detection via texture description. *Proceedings of the 2nd ACM workshop on Multimedia in forensics, security and intelligence - MiFor '10*, p. 59, 2010. Citado na página 37.

ATTO, A. M.; BENOIT, A.; LAMBERT, P. Timed-image based deep learning for action recognition in video sequences. *Pattern Recognition*, Elsevier Ltd, n. xxxx, 2020. ISSN 00313203. Citado na página 64.

BANERJEE, I.; BHATTACHARYYA, S.; SANYAL, G. Hiding & Analyzing Data in Image Using Extended PMM. *Procedia Technology*, Elsevier B.V., v. 10, p. 157–166, 2013. ISSN 22120173. Citado na página 28.

BAPPY, J. H.; ROY-CHOWDHURY, A. K.; BUNK, J.; NATARAJ, L.; MANJUNATH, B. S. Exploiting Spatial Structure for Localizing Manipulated Image Regions. In: *Proceedings of the IEEE International Conference on Computer Vision*. [S.l.: s.n.], 2017. v.

- 2017-Octob, p. 4980–4989. ISBN 9781538610329. ISSN 15505499. Citado 2 vezes nas páginas 94 e 95.
- BASHAR, M.; NODA, K.; OHNISHI, N.; MORI, K. Exploring duplicated regions in natural images. *IEEE Transactions on Image Processing*, n. 99, p. 1–40, 2010. ISSN 19410042. Citado na página 37.
- BASSIL, Y. An Image Steganography Scheme using Randomized Algorithm and Context-Free Grammar. *Journal of Advanced Computer Science & Technology*, v. 1, n. 4, p. 291–305, 2012. ISSN 2227-4332. Citado na página 28.
- BATTIATO, S.; GIUDICE, O.; PARATORE, A. Multimedia Forensics: discovering the history of multimedia contents. In: *Multimedia Forensics: Discovering the History of Multimedia Contents*. [S.l.: s.n.], 2016. p. 23–24. ISBN 9781450341820. Citado na página 34.
- BAYA, B.; STAMM, M. C. Augmented convolutional feature maps for robust CNN-based camera model identification. In: *2017 IEEE International Conference on Image Processing (ICIP)*. [S.l.: s.n.], 2017. p. 4098–4102. ISBN 9781509021758. Citado 2 vezes nas páginas 94 e 95.
- BAYAR, B.; STAMM, M. C. On the robustness of constrained convolutional neural networks to JPEG post-compression for image resampling detection. In: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.: s.n.], 2017. p. 2152–2156. ISBN 9781509041176. ISSN 15206149. Citado 3 vezes nas páginas 94, 95 e 96.
- BAYAR, B.; STAMM, M. C. Constrained Convolutional Neural Networks : Image Manipulation Detection. *IEEE Transactions on Information Forensics and Security*, v. 13, n. 11, p. 2691–2706, 2018. Citado na página 95.
- BAYAR, B.; STAMM, M. C. Constrained Convolutional Neural Networks: A New Approach Towards General Purpose Image Manipulation Detection. *IEEE Transactions on Information Forensics and Security*, IEEE, v. 13, n. 11, p. 2691–2706, 2018. ISSN 15566013. Citado na página 99.
- BHATTACHARYYA, D.; DUTTA, J.; DAS, P.; BANDYOPADHYAY, R.; BANDYOPADHYAY, S.; KIM, T.-h. Discrete Fourier Transformation based Image Authentication technique. *2009 8th IEEE International Conference on Cognitive Informatics*, n. 3, p. 196–200, 2009. Citado na página 29.
- BIANCONI, F.; FERNÁNDEZ, A. An appendix to “texture databases - A comprehensive survey”. *Pattern Recognition Letters*, v. 45, n. 1, p. 33–38, 2014. ISSN 01678655. Citado na página 87.
- BONDI, L.; BAROFFIO, L.; GUERA, D.; BESTAGINI, P.; DELP, E. J.; TUBARO, S. First Steps Toward Camera Model Identification with Convolutional Neural Networks. *IEEE Signal Processing Letters*, v. 24, n. 3, p. 259–263, 2017. ISSN 10709908. Citado 2 vezes nas páginas 94 e 95.
- BOSE, A.; MAITY, S. P. Spread Spectrum Watermark Detection on Degraded Compressed Sensing. *IEEE Sensors Letters*, v. 1, n. 5, p. 1–4, 2017. Citado na página 31.

- BOULKENAFET, Z.; KOMULAINEN, J.; HADID, A. Face Spoofing Detection Using Colour Texture Analysis. *IEEE Transactions on Information Forensics and Security*, v. 11, n. 8, p. 1818–1830, 2016. ISSN 1556-6013. Citado na página 111.
- BRODATZ, P. *A Photographic Album for Artists and Designers*. Dover, NY, USA: [s.n.], 1966. Citado 3 vezes nas páginas 87, 99 e 127.
- BRUNA, J.; MALLAT, S. Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 35, n. 8, p. 1872–1886, 2013. ISSN 01628828. Citado na página 67.
- BUNK, J.; BAPPY, J. H.; MOHAMMED, T. M.; NATARAJ, L.; FLENNER, A.; MANJUNATH, B. S.; CHANDRASEKARAN, S.; ROY-CHOWDHURY, A. K.; PETERSON, L. Detection and Localization of Image Forgeries Using Resampling Features and Deep Learning. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. [S.l.: s.n.], 2017. v. 2017-July, p. 1881–1889. ISBN 9781538607336. ISSN 21607516. Citado 2 vezes nas páginas 94 e 95.
- BURGHOUTS, G. J.; GEUSEBROEK, J. M. Material-specific adaptation of color invariant features. *Pattern Recognition Letters*, Elsevier B.V., v. 30, n. 3, p. 306–313, 2009. Disponível em: <<http://dx.doi.org/10.1016/j.patrec.2008.10.005>>. Citado na página 59.
- BYUN, H.; LEE, S. *Applications of Support Vector Machines for Pattern Recognition: A Survey*. [S.l.]: Springer, Berlin, Heidelberg, 2002. 1–7 p. ISSN 16113349. ISBN 978-3-540-45665-0. Citado na página 110.
- CAPUTO, B.; HAYMAN, E.; MALLIKARJUNA, P. Class-specific material categorisation. In: *Proceedings of the IEEE International Conference on Computer Vision*. [S.l.: s.n.], 2005. II, p. 1597–1604. ISBN 076952334X. Citado na página 128.
- CAVALIN, P.; OLIVEIRA, L. S. A Review of Texture Classification Methods and Databases. In: *Proceedings - 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials SIBGRAPI-T 2017*. [S.l.: s.n.], 2017. v. 2018-Janua, p. 1–8. ISBN 9781538606193. Citado 2 vezes nas páginas 87 e 107.
- CHAN, T. H.; JIA, K.; GAO, S.; LU, J.; ZENG, Z.; MA, Y. PCANet: A Simple Deep Learning Baseline for Image Classification? *IEEE Transactions on Image Processing*, v. 24, n. 12, p. 5017–5032, 2015. ISSN 10577149. Citado na página 67.
- CHEN, C. C.; HUANG, C. L. Markov random fields for texture classification. *Pattern Recognition Letters*, v. 14, n. 11, p. 907–914, 1993. ISSN 01678655. Citado na página 65.
- CHEN, J.; SHAN, S.; HE, C.; ZHAO, G.; PIETIKÄINEN, M.; CHEN, X.; GAO, W. WLD: A Robust Local Image Descriptor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 32, n. 9, p. 1705–1720, sep 2010. ISSN 0162-8828. Disponível em: <<http://ieeexplore.ieee.org/document/5204092/>>. Citado 20 vezes nas páginas 13, 14, 15, 68, 74, 76, 102, 103, 107, 108, 112, 115, 117, 118, 119, 120, 121, 122, 126 e 128.
- CHEN, M.; FRIDRICH, J.; GOLJAN, M. Determining Image Origin and Integrity Using Sensor Noise. *IEEE Transactions on Information Forensics and Security*, v. 3, n. 1, p. 74–90, 2008. Citado na página 34.

- CHEN, Y.; CHEN, J. A novel blind watermarking scheme based on neural networks for image. In: *2010 IEEE International Conference on Information Theory and Information Security*. [S.l.: s.n.], 2010. p. 548–552. ISBN 978-1-4244-6942-0. Citado 2 vezes nas páginas 94 e 95.
- CHEN, Y. H.; HUANG, H. C. Progressive Watermarking for Images with Regions of Interest. *Proceedings - 2015 International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIH-MSP 2015*, p. 25–28, 2015. Citado na página 32.
- CHIERCHIA, G.; POGGI, G.; SANSONE, C.; VERDOLIVA, L. A bayesian-MRF approach for PRNU-based image forgery detection. *IEEE Transactions on Information Forensics and Security*, IEEE, v. 9, n. 4, p. 554–567, 2014. ISSN 15566013. Citado na página 65.
- CHOLLET, F. Xception: Deep learning with depthwise separable convolutions. In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. [S.l.: s.n.], 2017. v. 2017-Janua, p. 1800–1807. ISBN 9781538604571. Citado na página 55.
- CIMPOI, M.; MAJI, S.; KOKKINOS, I.; MOHAMED, S.; and A. Vedaldi. Describing Textures in the Wild. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2014. Citado 5 vezes nas páginas 59, 66, 89, 99 e 127.
- CLARO, M.; VERAS, R.; SANTANA, A.; ARAÚJO, F.; SILVA, R.; ALMEIDA, J.; LEITE, D. An hybrid feature space from texture information and transfer learning for glaucoma classification. *Journal of Visual Communication and Image Representation*, v. 64, 2019. ISSN 10959076. Citado 2 vezes nas páginas 61 e 66.
- COSTANZO, A.; AMERINI, I.; CALDELLI, R.; BARNI, M. Forensic analysis of SIFT keypoint removal and injection. *IEEE Transactions on Information Forensics and Security*, v. 9, n. 9, p. 1450–1464, 2014. ISSN 15566013. Citado na página 36.
- CRISTIANINI, N.; SHAWE-TAYLOR, J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge: [s.n.], 2000. Citado na página 110.
- DANA, K. J.; KOENDERINK, J. J.; DANA, K. J.; NAYAR, S. K.; KOENDERINK, J. J. Reflectance and Texture of Real-World Surfaces. *ACM Transactions on Graphics*, v. 18, n. 1, p. 1–34, 1999. Citado 4 vezes nas páginas 59, 87, 99 e 127.
- DANUPUTRI, C.; MANTORO, T.; HARDJIANTO, M. Data Security Using LSB Steganography and Vigenere Chiper in an Android Environment. *Proceedings - 4th International Conference on Cyber Security, Cyber Warfare, and Digital Forensics, CyberSec 2015*, p. 22–27, 2016. Citado na página 27.
- DAVARZANI, R.; YAGHMAIE, K.; MOZAFFARI, S.; TAPAK, M. Copy-move forgery detection using multiresolution local binary patterns. *Forensic Science International*, Elsevier Ireland Ltd, v. 231, n. 1-3, p. 61–72, 2013. ISSN 03790738. Citado na página 37.
- DENEMARK, T.; FRIDRICH, J.; COMESAÑA, P. Improving Selection-Channel-Aware Steganalysis Features. In: *Electronic Imaging*. [S.l.: s.n.], 2016. p. 1–8. Citado na página 94.

- DENEMARK, T.; SEDIGHI, V.; HOLUB, V.; COGRANNE, R.; FRIDRICH, J. Selection-channel-aware rich model for Steganalysis of digital images. In: *IEEE International Workshop on Information Forensics and Security, WIFS 2014*. [S.l.: s.n.], 2014. p. 48–53. ISBN 9781479988822. Citado na página 94.
- DHEVANA, S.; JAYASRI, C. A Bayesian Classifier Approach for GLCM Based Image Forgery Detection. v. 4, n. 02, p. 255–259, 2015. Citado na página 60.
- DING, Y. J.; WANG, H. B.; TAO, L. Weber Local Descriptor with Double Orientation Line Feature for Finger Vein Recognition. In: *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*. [S.l.: s.n.], 2018. ISBN 9781538668115. Citado na página 75.
- DIXIT, R.; NASKAR, R.; MISHRA, S. Blur-invariant copy-move forgery detection technique with improved detection accuracy utilising SWT-SVD. *IET Image Processing*, v. 11, n. 5, p. 301–309, 2017. ISSN 1751-9659. Citado 2 vezes nas páginas 37 e 65.
- DONG, J.; WANG, W.; TAN, T. CASIA image tampering detection evaluation database. In: *2013 IEEE China Summit and International Conference on Signal and Information Processing, ChinaSIP 2013 - Proceedings*. [S.l.]: IEEE, 2013. p. 422–426. ISBN 9781479910434. Citado 9 vezes nas páginas 39, 115, 116, 117, 118, 119, 121, 125 e 127.
- ELHOSENY, M.; RAMIREZ-GONZALEZ, G.; ABU-ELNASR, O.; SHAWKAT, S.; N, A.; FAROUK, A. Secure Medical Data Transmission Model for IoT-based Healthcare Systems. *IEEE Access*, v. 6, 2018. ISSN 21693536. Citado na página 29.
- ERNAWAN, F.; KABIR, M. N. A Robust Image Watermarking Technique With an Optimal DCT-Psychovisual Threshold. *IEEE Access*, v. 6, p. 20464–20480, 2018. ISSN 21693536. Citado na página 32.
- FAROOQ, S.; YOUSAF, M. H.; HUSSAIN, F. A generic passive image forgery detection scheme using local binary pattern with rich models. *Computers and Electrical Engineering*, Elsevier Ltd, v. 62, p. 459–472, 2017. ISSN 00457906. Citado na página 37.
- FELSBERG, M.; SOMMER, G. The monogenic signal. *IEEE Transactions on Signal Processing*, IEEE, v. 49, n. 12, p. 3136–3144, 2001. ISSN 1053587X. Citado na página 84.
- FERREIRA, A.; FELIPUSSI, S. C.; ALFARO, C.; FONSECA, P.; VARGAS-MUNOZ, J. E.; Dos Santos, J. A.; ROCHA, A. Behavior Knowledge Space-Based Fusion for Copy-Move Forgery Detection. *IEEE Transactions on Image Processing*, v. 25, n. 10, p. 4729–4742, 2016. ISSN 10577149. Citado na página 38.
- FERREIRA, W. D.; FERREIRA, C. B. R.; JÚNIOR, G. d. C.; SOARES, F. A review on digital image forensics. *Computers Electrical Engineering*, Elsevier Ltd, v. 85, p. 106685, 2020. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0045790620305401>>. Citado na página 127.
- FREND, S. J.; KNOWLES, E. D.; SALETAN, W.; LOFTUS, E. False Memories of Fabricated Political Events. *Journal of Experimental Social Psychology*, v. 49, 2013. Citado na página 20.
- FRIDRICH, J. Digital Image Forensics. *IEEE Signal Processing Magazine*, v. 26, n. March, p. 26–37, 2009. Citado 4 vezes nas páginas 12, 20, 34 e 35.

- FU, Y.; PENG, L. Face Recognition Algorithm Collaborative Representation. *Computer Aided Drafting, Design and Manufacturing*, v. 25, n. 2, p. 14–18, 2015. Citado na página 84.
- FUKUSHIMA, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, v. 36, n. 4, p. 193–202, 1980. ISSN 03401200. Citado 2 vezes nas páginas 54 e 55.
- GHOSHAL, N.; MANDAL, J. K. A steganographic scheme for colour image authentication (SSCIA). *International Conference on Recent Trends in Information Technology, ICRTIT 2011*, p. 826–831, 2011. Citado na página 29.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento Digital de Imagens*. 3a ed.. ed. [S.l.]: Pearson, 2015. Citado na página 70.
- GUERA, D.; WANG, Y.; BONDI, L.; BESTAGINI, P.; TUBARO, S.; DELP, E. J. A Counter-Forensic Method for CNN-Based Camera Model Identification. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, v. 2017-July, p. 1840–1847, 2017. ISSN 21607516. Citado 2 vezes nas páginas 94 e 95.
- GUIJARRO, M.; RIOMOROS, I.; PAJARES, G.; ZITINSKI, P. Discrete wavelets transform for improving greenness image segmentation in agricultural images. *Computers and Electronics in Agriculture*, Elsevier B.V., v. 118, p. 396–407, 2015. ISSN 01681699. Disponível em: <<http://dx.doi.org/10.1016/j.compag.2015.09.011>>. Citado na página 59.
- GUO, W.; REES, W. G. Altitudinal forest-tundra ecotone categorization using texture-based classification. *Remote Sensing of Environment*, Elsevier, v. 232, n. December 2018, p. 111312, 2019. ISSN 00344257. Disponível em: <<https://doi.org/10.1016/j.rse.2019.111312>>. Citado na página 59.
- GUO, Z.; ZHANG, L.; ZHANG, D. Rotation invariant texture classification using LBP variance (LBPV) with global matching. *Pattern Recognition*, Elsevier, v. 43, n. 3, p. 706–719, 2010. ISSN 00313203. Citado na página 37.
- HAFIANE, A.; ; SEETHARAMAN, G.; ; PALANIAPPAN, K.; ; ZAVIDOVIQUE, B. Rotationally Invariant Hashing of Median Binary Patterns for Texture Classification. In: CAMPILHO, A.; ; KAMEL, M. (Ed.). *Lecture Notes in Computer Science*. [S.l.]: Springer Berlin Heidelberg, 2008. v. 9, n. 3, p. 242–242. ISBN 3540249796. Citado 6 vezes nas páginas 13, 68, 69, 70, 102 e 103.
- HAMIDI, M.; HAZITI, M. E.; CHERIFI, H.; ABOUTAJDINE, D. A blind Robust Image Watermarking Approach exploiting the DFT Magnitude. In: *2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA)*. [S.l.: s.n.], 2015. p. 1–6. ISBN 9781509004782. Citado na página 32.
- HARIBABU, K.; SUBRAHMANYAM, G. R. K. S.; MISHRA, D. A robust digital image watermarking technique using auto encoder based convolutional neural networks. In: *2015 IEEE Workshop on Computational Intelligence: Theories, Applications and Future Directions (WCI)*. [S.l.: s.n.], 2015. p. 1–6. ISBN 978-1-4673-8215-1. Citado na página 94.
- HAYKIN, S. *Redes neurais: princípios e prática*. 2.ed.. ed. Porto Alegre: Bookman, 2001. Citado na página 48.

HE, C.; LI, S.; LIAO, Z.; LIAO, M. Texture classification of PolSAR data based on sparse coding of wavelet polarization textons. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, v. 51, n. 8, p. 4576–4590, 2013. ISSN 01962892. Citado na página 66.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE, 2016. v. 2016-Decem, p. 770–778. ISBN 9781467388504. ISSN 10636919. Citado 2 vezes nas páginas 55 e 108.

HERMANTH, D. J.; POPESCU, D. E.; MITTAL, M.; MAHESWARI, U. S. Analysis of wavelet, ridgelet, curvelet and bandelet transforms for QR code based image steganography. In: *2017 14th International Conference on Engineering of Modern Electric Systems (EMES)*. [S.l.: s.n.], 2017. p. 121–126. ISBN 9781509060733. Citado na página 30.

HINTON, G. E. Learning translation invariant recognition in a massively parallel networks. In: BAKKER, J. W. de; NIJMAN, A. J.; TRELEAVEN, P. C. (Ed.). *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Berlin: Springer Berlin Heidelberg, 1987. v. 258 LNCS, p. 1–13. ISBN 9783540179436. ISSN 16113349. Citado na página 53.

HOLUB, V.; FRIDRICH, J. Designing steganographic distortion using directional filters. In: *2012 IEEE International Workshop on Information Forensics and Security*. [S.l.: s.n.], 2012. p. 234–239. ISBN 9781467322874. Citado na página 94.

HOLUB, V.; FRIDRICH, J.; DENEMARK, T. Universal distortion function for steganography in an arbitrary domain. In: *Eurasip Journal on Information Security*. [S.l.: s.n.], 2014. v. 2014, p. 1–13. ISSN 2510523X. Citado na página 94.

HOSSAIN, S.; SERIKAWA, S. Texture databases - A comprehensive survey. *Pattern Recognition Letters*, Elsevier B.V., v. 34, n. 15, p. 2007–2022, 2013. ISSN 01678655. Disponível em: <<http://dx.doi.org/10.1016/j.patrec.2013.02.009>>. Citado 5 vezes nas páginas 86, 87, 88, 99 e 127.

HOSSIN, M.; M.N, S. A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining Knowledge Management Process*, v. 5, n. 2, p. 01–11, 2015. ISSN 2231007X. Citado na página 114.

HOWARD, A. G.; ZHU, M.; CHEN, B.; KALENICHENKO, D.; WANG, W.; WEYAND, T.; ANDREETTO, M.; ADAM, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. 2017. Disponível em: <<http://arxiv.org/abs/1704.04861>>. Citado na página 55.

HSU, Y.-f.; CHANG, S.-f. Camera Response Functions for Image Forensics: An Automatic Algorithm for Splicing Detection. *IEEE Transactions on Information Forensics and Security*, v. 5, n. 4, p. 816–825, 2010. Citado na página 35.

HUANG, F.; ZHONG, Y.; HUANG, J. Edge Adaptive Image Steganography Based on LSB Matching Revisited. *IEEE Transactions on Information Forensics and Security*, v. 5, n. 2, p. 201–214, 2010. ISSN 16113349. Citado na página 27.

- HUANG, G.; LIU, Z.; Van Der Maaten, L.; WEINBERGER, K. Q. Densely connected convolutional networks. In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. [S.l.]: IEEE, 2017. v. 2017-Janua, p. 2261–2269. ISBN 9781538604571. Citado 2 vezes nas páginas 55 e 66.
- HUBEL, D. H.; WIESEL, T. N. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, v. 195, n. 1, p. 215–243, 1968. ISSN 14697793. Citado 2 vezes nas páginas 54 e 55.
- LANDOLA, F. N.; HAN, S.; MOSKEWICZ, M. W.; ASHRAF, K.; DALLY, W. J.; KEUTZER, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. p. 1–13, 2016. Disponível em: <<http://arxiv.org/abs/1602.07360>>. Citado na página 55.
- IBAIDA, A.; KHALIL, I. Wavelet-Based ECG Steganography for Protecting Patient Confidential Information in Point-of-Care Systems. *IEEE Transactions on Biomedical Engineering*, v. 60, n. 12, p. 3322–3330, 2013. Citado 2 vezes nas páginas 29 e 64.
- JABID, T.; KABIR, M. H.; CHAE, O. Local directional pattern (LDP) for face recognition. In: *Digest of Technical Papers International Conference on Consumer Electronics (ICCE)*. [S.l.]: IEEE, 2010. p. 329–330. ISBN 9781424443161. ISSN 13494198. Citado na página 77.
- JAIN, R.; BOADDH, J. Advances in Digital Image Steganography. In: *2016 International Conference on Innovation and Challenges in Cyber Security (ICICCS-INBUSH)*. [S.l.: s.n.], 2016. p. 163–171. ISBN 9781509020843. Citado na página 26.
- JASEELA, S.; NISHADHA, S. G. Copy Move Image Forgery Detection Using SURF Feature Point Extraction. *International Journal Of Scientific & Engineering Research*, v. 7, n. 7, p. 653–657, 2016. Citado na página 36.
- JERO, S. E.; RAMU, P. Curvelets-based ECG Steganography for Data Security. *Electronics Letters*, v. 52, n. 4, p. 283–285, 2016. ISSN 1540-3459. Citado na página 30.
- JIN, Q.; HUANG, Y.; FAN, N. Learning images using compositional pattern-producing neural networks for source camera identification and digital demographic diagnosis. *Pattern Recognition Letters*, Elsevier B.V., v. 33, n. 4, p. 381–396, 2012. ISSN 0167-8655. Citado na página 34.
- JOHNSON, J. M.; KHOSHGOFTAAR, T. M. Survey on deep learning with class imbalance. *Journal of Big Data*, Springer International Publishing, v. 6, n. 1, 2019. ISSN 21961115. Disponível em: <<https://doi.org/10.1186/s40537-019-0192-5>>. Citado na página 128.
- JOHNSON, M. K.; FARID, H. Exposing Digital Forgeries Through Chromatic Aberration. In: *MM&Sec '06 Proceedings of the 8th workshop on Multimedia and security*. [S.l.: s.n.], 2006. p. 48–55. ISBN 1595934936. Citado na página 34.
- JOHRI, P.; MISHRA, A.; DAS, S. Survey on steganography methods (text, image, audio, video, protocol and network steganography). In: *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*. [S.l.: s.n.], 2016. p. 2906–2909. ISBN 9789380544212. Citado 2 vezes nas páginas 15 e 26.

- JOTHY, N.; ANUSUYYA, S. A secure color image steganography using integer wavelet transform. *2016 10th International Conference on Intelligent Systems and Control (ISCO)*, p. 1–4, 2016. Citado na página 30.
- JULESZ, B. Visual Pattern Discrimination. *IRE Transactions on Information Theory*, v. 8, n. 2, p. 84–92, 1962. ISSN 21682712. Citado na página 58.
- KABIR, M. H.; JABID, T.; CHAE, O. A Local Directional Pattern Variance (LDPv) based face descriptor for human facial expression recognition. In: *IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS*. [S.l.]: IEEE, 2010. p. 526–532. ISBN 9780769542645. Citado 5 vezes nas páginas 68, 77, 78, 102 e 103.
- KANDI, H.; MISHRA, D.; GORTHI, S. R. Exploring the learning capabilities of convolutional neural networks for robust image watermarking. *Computers and Security*, Elsevier Ltd, v. 65, p. 247–268, 2017. ISSN 01674048. Citado 2 vezes nas páginas 94 e 99.
- KANG, X.; WEI, S. Identifying Tampered Regions Using Singular Value Decomposition in Digital Image Forensics. In: *2008 International Conference on Computer Science and Software Engineering*. [S.l.: s.n.], 2008. p. 926–930. ISBN 978-0-7695-3336-0. Citado na página 37.
- KASHYAP, R. L.; CHELLAPPA, R.; KHOTANZAD, A. Texture classification using features derived from random field models. *Pattern Recognition Letters*, v. 1, n. 1, p. 43–50, 1982. ISSN 01678655. Citado na página 65.
- KAWASHIMA, Y.; FUJITA, A.; BUCH, K.; LI, B.; QURESHI, M. M.; CHAPMAN, M. N.; SAKAI, O. Using texture analysis of head CT images to differentiate osteoporosis from normal bone density. *European Journal of Radiology*, Elsevier, v. 116, n. January, p. 212–218, 2019. ISSN 18727727. Disponível em: <<https://doi.org/10.1016/j.ejrad.2019.05.009>>. Citado na página 59.
- KAWATSU, C.; KOSS, F.; GILLIES, A.; ZHAO, A.; CROSSMAN, J. Gesture Recognition for Robotic Control Using Deep Learning. In: *2017 Ground Vehicle Systems Engineering and Technology Symposium - GVSETS*. [s.n.], 2017. Disponível em: <<https://events.esd.org/wp-content/uploads/2017/08/Gesture-Recognition-for-Robotic-Control-Using-Deep-Learning.pdf>>. Citado 3 vezes nas páginas 13, 109 e 111.
- KEE, E.; JOHNSON, M. K.; FARID, H. Digital Image Authentication From JPEG Headers. *IEEE Transactions on Information Forensics and Security*, v. 6, n. 3, p. 1066–1075, 2011. Citado na página 34.
- KELLER, J. M.; CHEN, S.; CROWNOVER, R. M. Texture description and segmentation through fractal geometry. *Computer Vision, Graphics and Image Processing*, v. 45, n. 2, p. 150–166, 1989. ISSN 0734189X. Citado na página 66.
- KESSLER, G. C. An Overview of Steganography for the Computer Forensics Examiner. *Forensic Science Communications*, v. 6, 2004. Citado na página 26.
- KHAN, A.; BASHAR, F.; AHMED, F.; KABIR, M. H. Median ternary pattern (MTP) for face recognition. In: *2013 International Conference on Informatics, Electronics and Vision, ICIEV 2013*. [S.l.]: IEEE, 2013. p. 1–5. ISBN 9781479903979. Citado 20 vezes nas

páginas 13, 14, 15, 68, 70, 72, 102, 103, 107, 108, 112, 115, 116, 117, 119, 120, 121, 122, 126 e 128.

KHAN, A.; SOHAIL, A.; ZAHOORA, U.; QURESHI, A. S. A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, p. 1–70, 2020. ISSN 0269-2821. Citado na página 54.

KHOO, B. E.; MAK BOL, N. M.; RASSEM, T. H. Block-based discrete wavelet transform-singular value decomposition image watermarking scheme using human visual system characteristics. *IET Image Processing*, v. 10, n. 1, p. 34–52, 2016. ISSN 1751-9659. Citado na página 32.

KIRCHNER, M. Linear Row and Column Predictors for the Analysis of Resized Images. In: *12th ACM Workshop on Multimedia and Security*. [S.l.: s.n.], 2010. p. 13–18. ISBN 9781450302869. Citado na página 38.

KIRSCH, R. A. Computer Determination Biological of the Constituent Images. *Computers and Biomedical Research*, v. 4, p. 315–328, 1971. Citado na página 77.

KORNBLUM, J. D. Using JPEG quantization tables to identify imagery processed by software. *Digital Investigation*, v. 5, p. 21–25, 2008. Citado na página 34.

KORUS, P.; HUANG, J. Improved tampering localization in digital image forensics based on maximal entropy random walk. *IEEE Signal Processing Letters*, v. 23, n. 1, p. 169–173, 2016. ISSN 10709908. Citado na página 111.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems*, v. 25, p. 1–1432, 2012. Citado 3 vezes nas páginas 55, 66 e 100.

KYLBERG, G. *The Kylberg Texture Dataset v. 1.0*. [S.l.], 2011. Disponível em: <<http://www.cb.uu.se/~gustaf/texture/>>. Citado 4 vezes nas páginas 59, 88, 99 e 127.

LASKAR, S. A.; HEMACHANDRAN, K. Steganography Based on Random Pixel Selection for Efficient Data Hiding. *International Journal of Computer Engineering and Technology (IJCET)*, v. 4, p. 31–44, 2013. Citado na página 28.

LAWGALY, A.; KHELIFI, F. Sensor Pattern Noise Estimation Based on Improved Locally Adaptive DCT Filtering and Weighted Averaging for Source Camera Identification and Verification. *IEEE Transactions on Information Forensics and Security*, v. 12, n. 2, p. 392–404, 2017. Citado na página 34.

LECUN, Y.; BOSER, B.; DENKER, J. S.; HENDERSON, D.; HOWARD, R. E.; HUBBARD, W.; JACKEL, L. D. Backpropagation applied to digit recognition. *Neural computation*, v. 1, n. 4, p. 541–551, 1989. Citado 2 vezes nas páginas 54 e 55.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, v. 86, p. 2278 – 2324, 1998. Disponível em: <<http://ieeexplore.ieee.org/document/726791/\full-text-sect>>. Citado 3 vezes nas páginas 49, 54 e 55.

LEE, J. C. Copy-move image forgery detection based on Gabor magnitude. *Journal of Visual Communication and Image Representation*, Elsevier Inc., v. 31, p. 320–334, 2015. ISSN 10959076. Citado na página 37.

- LEI, Z.; AHONEN, T.; PIETIKAINEN, M.; LI, S. Z. Local frequency descriptor for low-resolution face recognition. In: *2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, FG 2011*. [S.l.: s.n.], 2011. p. 161–166. ISBN 9781424491407. Citado 6 vezes nas páginas 13, 68, 83, 84, 102 e 103.
- LI, B.; WANG, M.; HUANG, J.; LI, X. A new cost function for spatial image steganography. In: *IEEE International Conference on Image Processing, ICIP 2014*. [S.l.: s.n.], 2014. p. 4206–4210. ISBN 9781479957514. Citado na página 94.
- LI, C.; DUAN, G.; ZHONG, F. Rotation Invariant Texture Retrieval Considering the Scale Dependence of Gabor Wavelet. *IEEE Transactions on Image Processing*, v. 24, n. 8, p. 2344–2354, 2015. Citado na página 65.
- LI, H.; LUO, W.; QIU, X.; HUANG, J. Image Forgery Localization via Integrating Tampering Possibility Maps. *IEEE Transactions on Information Forensics and Security*, v. 12, n. 5, p. 1240–1252, 2017. ISSN 15566013. Citado na página 38.
- LI, L.; LI, S.; ZHU, H.; CHU, S.-C.; RODDICK, J. F.; PAN, J.-S. An efficient scheme for detecting copy-move forged images by local binary patterns. *Journal of Information Hiding and Multimedia Signal Processing*, v. 4, n. 1, p. 46–56, 2013. ISSN 20734212 (ISSN). Citado na página 37.
- LI, X.; GRANDVALET, Y.; DAVOINE, F. A baseline regularization scheme for transfer learning with convolutional neural networks. *Pattern Recognition*, Elsevier Ltd, v. 98, p. 107049, 2020. ISSN 00313203. Disponível em: <<https://doi.org/10.1016/j.patcog.2019.107049>>. Citado na página 100.
- LI, Y.; ZHOU, J. Fast and Effective Image Copy-Move Forgery Detection via Hierarchical Feature Point Matching. *IEEE Transactions on Information Forensics and Security*, IEEE, v. 14, n. 5, p. 1307–1322, 2019. ISSN 15566013. Citado na página 36.
- LI, Z.; HOIEM, D. Learning without Forgetting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 40, n. 12, p. 2935–2947, 2018. ISSN 19393539. Citado na página 100.
- LIAN, S.; KANELLOPOULOS, D.; RUFFO, G. Recent advances in multimedia information system security. *Informatica*, v. 33, p. 3–24, 2009. Citado na página 25.
- LIE, W. N.; LIN, G. S. A feature-based classification technique for blind image steganalysis. *IEEE Transactions on Multimedia*, v. 7, n. 6, p. 1007–1020, 2005. ISSN 15209210. Citado na página 29.
- LIN, T. Y.; ROYCHOWDHURY, A.; MAJI, S. Bilinear Convolutional Neural Networks for Fine-Grained Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 40, n. 6, p. 1309–1322, 2018. ISSN 01628828. Citado na página 67.
- LIU, L.; CHEN, J.; FIEGUTH, P.; ZHAO, G.; CHELLAPPA, R.; PIETIKÄINEN, M. From BoW to CNN: Two Decades of Texture Representation for Texture Classification. *International Journal of Computer Vision*, Springer US, v. 127, n. 1, p. 74–109, 2019. ISSN 15731405. Disponível em: <<https://doi.org/10.1007/s11263-018-1125-z>>. Citado 3 vezes nas páginas 59, 66 e 67.

- LIU, L.; LAO, S.; FIEGUTH, P. W.; GUO, Y.; WANG, X. Median Robust Extended Local Binary Pattern for Texture Classification. *IEEE Transactions on Image Processing*, v. 25, n. 3, p. 1368–1381, 2016. Citado na página 128.
- LIU, S.; PAN, Z.; SONG, H. Digital image watermarking method based on DCT and fractal encoding. *IET Image Processing*, v. 11, n. 10, p. 815–821, 2017. ISSN 1751-9659. Citado na página 32.
- LIU, X.-L.; LIN, C.-C.; YUAN, S.-M. Blind dual watermarking for color images' authentication and copyright protection. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 28, n. 5, p. 1047–1055, 2018. ISSN 1051-8215. Citado na página 31.
- LYU, S.; FARID, H. Steganalysis Using Higher-Order Image Statistics. *IEEE Transactions on Information Forensics and Security*, v. 1, n. 1, p. 111–119, 2006. Citado na página 29.
- MAHMOOD, T.; IRTAZA, A.; MEHMOOD, Z.; Tariq Mahmood, M. Copy–move forgery detection through stationary wavelets and local binary pattern variance for forensic analysis in digital images. *Forensic Science International*, Elsevier Ireland Ltd, v. 279, p. 8–21, 2017. ISSN 18726283. Citado na página 37.
- MAHMOOD, T.; MEHMOOD, Z.; SHAH, M.; SABA, T. A robust technique for copy-move forgery detection and localization in digital images via stationary wavelet and discrete cosine transform. *Journal of Visual Communication and Image Representation*, Elsevier, v. 53, n. September 2017, p. 202–214, 2018. ISSN 10959076. Citado na página 37.
- MANDAI, S.; BHATTACHARYYA, S. *Secret Data Sharing in Cloud Environment Using Steganography and Encryption Using GA*. [S.l.: s.n.], 2015. 1469–1474 p. ISBN 9781467379106. Citado na página 28.
- MANDELBROT, B. B. *The fractal geometry of nature*. San Francisco, CA: Freeman, 1983. Disponível em: <<https://cds.cern.ch/record/98509>>. Citado na página 65.
- MARKOV, A. The theory of algorithms. *Acad. Sci. USSR*, v. 42, p. 3—375, 1954. Disponível em: <<http://mi.mathnet.ru/tm1178>>. Citado na página 65.
- MATLABR2020A. *Matlab Help Center - Pretrained Deep Neural Networks*. 2020. Disponível em: <<https://www.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html>>. Citado 2 vezes nas páginas 13 e 110.
- MAYER, O.; STAMM, M. C. Forensic Similarity for Digital Images. *IEEE Transactions on Information Forensics and Security*, IEEE, v. 15, p. 1331–1346, 2020. ISSN 15566021. Citado 2 vezes nas páginas 94 e 95.
- MENG, J.; LIU, S.; ZHU, L.; ZHU, L.; WANG, H.; XIE, L.; GUAN, Y.; HE, J.; YANG, X.; ZHOU, Z. Texture Analysis as Imaging Biomarker for recurrence in advanced cervical cancer treated with CCRT. *Scientific Reports*, v. 8, n. 1, p. 1–9, 2018. ISSN 20452322. Citado na página 59.
- MIRMEHD, M.; XIE, X.; Jasjit Suri. *Handbook of Texture Analysis*. London: Imperial College Press, 2008. ISBN 9781848161153. Citado 4 vezes nas páginas 12, 21, 22 e 60.

- MONTEIRO, P.; PEREIRA, M. Crimes De Abuso Sexual De Criança / Adolescente a Partir Da Análise De Arquivos Multimídia . In: *The Fourth International Conference on Forensic Computer Science - ICoFCS 2009*. Natal: [s.n.], 2009. p. 114–121. Citado na página 21.
- MORAIS, P. A. d. O.; SOUZA, D. M. de; CARVALHO, M. T. d. M.; MADARI, B. E.; OLIVEIRA, A. E. de. Predicting soil texture using image analysis. *Microchemical Journal*, Elsevier, v. 146, n. October 2018, p. 455–463, 2019. ISSN 0026265X. Disponível em: <<https://doi.org/10.1016/j.microc.2019.01.009>>. Citado na página 59.
- MUINUDDIN, K.; MOHAMMED, C.; SRINIVASKUMAR, S.; PRASAD, G. Defective Texture Classification using Optimized Neural Network Structure. *Pattern Recognition Letters*, Elsevier B.V., 2020. ISSN 0167-8655. Disponível em: <<https://doi.org/10.1016/j.patrec.2020.04.017>>. Citado na página 61.
- MUSCI, M.; FEITOSA, R. Q.; VELLOSO, M. L. F.; NOVACK, T.; COSTA, A. O. P.; COSTA, G. A. O. P. da. Uma Avaliação de Descritores de Textura Baseados em Códigos Binários Locais para Classificação de Imagens de Sensoriamento Remoto. In: *Boletim de Ciências Geodésicas*. [S.l.: s.n.], 2011. p. 549–570. Citado na página 82.
- MUZAFFER G., M. O. U. B. U. G. Copy Move Forgery Detection Using Gabor Filter and ORB. In: *Proceedings of 2016 International Conference on Image Processing, Production and Computer Science*. [S.l.: s.n.], 2016. p. 23–29. ISBN 9789384422714. Citado na página 37.
- NAGARAJ, V.; VIJAYALAKSHMI, V.; ZAYARAZ, G. Overview of Digital Steganography Methods and Its Applications. *International Journal of Advanced Science and Technology*, v. 60, p. 45–58, 2013. Citado 3 vezes nas páginas 15, 32 e 33.
- NAPOLETANO, P. Hand-Crafted vs Learned Descriptors for Color Texture Classification. In: BIANCO, S.; SCHETTINI, R.; TRÉMEAU, A.; TOMINAGA, S. (Ed.). *Computational Color Imaging*. [S.l.]: Springer International Publishing, 2017. v. 10213, p. 259—271. ISBN 978-3-319-56010-6. Citado na página 66.
- NASONOV, A.; CHESNAKOV, K.; KRYLOV, A. Convolutional neural networks based image resampling with noisy training set. In: *2016 IEEE 13th International Conference on Signal Processing (ICSP)*. [S.l.: s.n.], 2016. p. 62–66. ISBN 9781509013456. Citado 3 vezes nas páginas 94, 95 e 96.
- NG, T.; CHANG, S. *A Data Set of Authentic and Spliced Image Blocks*. [S.l.], 2004. 1–9 p. Citado 9 vezes nas páginas 39, 115, 116, 117, 118, 119, 121, 125 e 127.
- NI, J.; HU, X.; SHI, Y. Q. Efficient JPEG Steganography Using Domain Transformation of Embedding Entropy. *IEEE Signal Processing Letters*, v. 25, n. 6, p. 773–777, 2018. ISSN 10709908. Citado na página 29.
- NIGHANIA, K. *Various ways to evaluate a machine learning model's performance*. 2018. Disponível em: <<https://towardsdatascience.com/various-ways-to-evaluate-a-machine-learning-models-performance-230449055f15>>. Citado na página 114.
- OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, v. 29, n. 1, p. 51–59, 1996. ISSN 00313203. Citado 4 vezes nas páginas 67, 68, 102 e 103.

- OJANSIVU, V.; HEIKKILÄ, J. Blur Insensitive Texture Classification Using Local Phase Quantization. *Image and Signal Processing*, v. 5099, p. 236–243, 2008. ISSN 03029743. Citado 4 vezes nas páginas 68, 82, 102 e 103.
- OUYANG, J.; LIU, Y.; LIAO, M. Copy-Move Forgery Detection Based on Deep Learning. In: *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. [S.l.: s.n.], 2017. p. 1–5. ISBN 9781538619377. Citado 2 vezes nas páginas 94 e 95.
- PEDRINI, H.; SCHWARTZ, W. R. *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*. São Paulo: Thomson Learning, 2008. 528 p. ISBN 978-85-221-0595-3. Citado 3 vezes nas páginas 58, 59 e 98.
- PELEG, S.; NAOR, J.; HARTLEY, R.; AVNIR, D. Multiple Resolution Texture Analysis and Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6, n. 4, p. 518–523, 1984. ISSN 01628828. Citado na página 65.
- PENG, A.; WU, Y.; KANG, X. Revealing traces of image Resampling and Resampling Antiforensics. *Advances in Multimedia*, v. 2017, 2017. ISSN 16875699. Citado na página 38.
- PEREZ-MEANA, M. C.-H.; GARCIA-UGALDE, F.; NAKANO-MIYATAKE, M.; H. Robust digital image watermarking using interest points and DFT domain. In: *Telecommunications and Signal Processing (TSP), 2012 35th International Conference on*. [S.l.: s.n.], 2012. p. 715–719. ISBN 0115255562231. Citado 2 vezes nas páginas 32 e 63.
- POPESCU, A. C.; FARID, H. Exposing Digital Forgeries by Detecting Traces of Image Splicing. *IEEE Transactions on Signal Processing*, v. 53, n. 2, p. 758–767, 2005. ISSN 1053-587X. Disponível em: <10.1109/ICOSP.2006.345714>. Citado 2 vezes nas páginas 38 e 63.
- PRASAD, K. L.; RAO, T. M.; KANNAN, V. A Hybrid Semi-fragile Image Watermarking Technique Using SVD-BND Scheme for Tampering Detection with Dual Authentication. *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, p. 517–523, 2016. Citado na página 32.
- PROVOST, F.; FAWCETT, T. Decision Analytic Thinking I: What Is a Good Model? In: *Data Science for Business*. [S.l.]: O'Reilly Media, Inc., 2013. cap. 7. ISBN 9781449361327. Citado na página 114.
- PUN, C. M.; CHUNG, J. L. A two-stage localization for copy-move forgery detection. *Information Sciences*, Elsevier Inc., v. 463-464, p. 33–55, 2018. ISSN 00200255. Citado na página 37.
- QIAO, T.; RETRAINT, F.; COGRANNE, R.; THAI, T. H. Individual camera device identification from JPEG images. *Signal Processing : Image Communication*, Elsevier, v. 52, n. January 2016, p. 74–86, 2017. ISSN 0923-5965. Citado na página 34.
- QUAN, W.; WANG, K.; YAN, D.; ZHANG, X. Distinguishing between Natural and Computer-Generated Images Using Convolutional Neural Networks. *IEEE Transactions on Information Forensics and Security*, v. 13, n. 11, p. 2772–2787, 2018. ISSN 15566013. Citado 3 vezes nas páginas 94, 96 e 99.

- QUAN, Y.; XU, Y.; SUN, Y.; LUO, Y. Lacunarity analysis on image patterns for texture classification. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE, 2014. p. 160–167. ISBN 9781479951178. ISSN 10636919. Citado na página 66.
- RAFAT, K. F.; HUSSAIN, M. J. Secure Steganography for Digital Images. In: *International Journal of Advanced Computer Science and Applications*. [S.l.: s.n.], 2016. v. 7, n. 6, p. 45–59. Citado na página 25.
- RANDEN, T. *Filter and Filter Bank Design for Image Texture Recognition*. 138 p. Tese (Doutorado) — Norwegian University of Science and Technology, 1997. Disponível em: <<http://www.ux.uis.no/~simtranden/thesis/thesis.pdf>>. Citado na página 87.
- RAO, Y.; NI, J. A deep learning approach to detection of splicing and copy-move forgeries in images. In: *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*. [S.l.: s.n.], 2016. p. 1–6. ISBN 9781509011384. Citado 2 vezes nas páginas 94 e 95.
- RAVI, H.; SUBRAMANYAM, A. V.; GUPTA, G.; KUMAR, B. A. Compression noise based video forgery detection. In: *2014 IEEE International Conference on Image Processing, ICIP 2014*. [S.l.]: IEEE, 2014. p. 5352–5356. ISBN 9781479957514. Citado na página 65.
- RIVERA, A. R.; CASTILLO, J. R.; CHAE, O. Local Directional Texture Pattern image descriptor. *Pattern Recognition Letters*, v. 51, p. 94–100, 2015. ISSN 01678655. Citado 6 vezes nas páginas 13, 68, 79, 81, 102 e 103.
- RODRIGUES, P.; BARRETO, J. P. Single-Image Estimation of the Camera Response Function in Near-Lighting . In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2015. p. 1028–1036. ISBN 9781467369640. Citado na página 35.
- ROSIYADI, D.; HORNG, S. J.; FAN, P.; WANG, X.; KHAN, M. K.; PAN, Y. Copyright protection for E-government document images. *IEEE Multimedia*, v. 19, n. 3, p. 62–73, 2012. ISSN 1070986X. Citado 2 vezes nas páginas 32 e 62.
- RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA, S.; HUANG, Z.; KARPATY, A.; KHOSLA, A.; BERNSTEIN, M.; BERG, A. C.; FEI-FEI, L. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, Springer US, v. 115, n. 3, p. 211–252, 2015. ISSN 15731405. Disponível em: <<http://dx.doi.org/10.1007/s11263-015-0816-y>>. Citado 5 vezes nas páginas 66, 100, 107, 109 e 121.
- SACCHI, D. L. M.; AGNOLI, F.; LOFTUS, E. F. Changing History: Doctored Photographs Affect Memory for Past Public Events. *Applied Cognitive Psychology*, v. 21, n. 8, p. 1005–1022, 2007. Citado na página 20.
- SALOMON, M.; COUTURIER, R.; GUYEUX, C.; COUCHOT, J. F.; BAHY, J. M. Steganalysis via a convolutional neural network using large convolution filters for embedding process with same stego key: A deep learning approach for telemedicine. *European Research in Telemedicine*, Elsevier Masson SAS, v. 6, n. 2, p. 79–92, 2017. ISSN 2212764X. Citado na página 94.

- SÁNCHEZ, J.; PERRONNIN, F.; MENSINK, T.; VERBEEK, J. Image classification with the fisher vector: Theory and practice. *International Journal of Computer Vision*, v. 105, n. 3, p. 222–245, 2013. ISSN 09205691. Citado na página 128.
- SARANTINOS, N.; AL-NEMRAT, A.; NAEEM, U. Statistical Sampling Approach to Investigate Child Pornography Cases. *2013 Fourth Cybercrime and Trustworthy Computing Workshop*, p. 22–29, 2013. Citado na página 21.
- SHAIK, S.; MAHABOOB, T. Comparative performance analysis of LBP and LTP based Facial Expression Recognition. *International Journal of Applied Engineering Research*, v. 12, n. 17, p. 6897–6900, 2017. Citado na página 70.
- SHANG, J.; MARSZALEK, M.; LAZEBNIK, S.; SCHMID, C. Local Features and Kernels for Classification of Texture and Object Categories: An In-Depth Study. *International Journal of Computer Vision*, v. 73, p. 213–238, 2006. Citado na página 87.
- SHIVAKUMAR, B. L.; BABOO, S. S. Detection of Region Duplication Forgery in Digital Images Using SURF. *International Journal of Computer Science Issues*, v. 8, n. 4, p. 199–205, 2011. Citado na página 36.
- SHOBANA, M.; MANIKANDAN, R. Efficient Method for Hiding Data by Pixel Intensity. *International Journal of Engineering and Technology (IJET)*, v. 5, n. 1, p. 74–80, 2013. Citado na página 29.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. In: *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. [S.l.: s.n.], 2015. p. 1–14. Citado 2 vezes nas páginas 55 e 66.
- SINGH, A.; SINGH, H. An improved LSB based image steganography technique for RGB images. *Proceedings of 2015 IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2015*, p. 1–4, 2015. ISSN 20103743. Citado na página 27.
- SINGH, K. M.; SINGH, L. S.; SINGH, A. B.; DEVI, K. S. Hiding secret message in edges of the image. *ICICT 2007: Proceedings of International Conference on Information and Communication Technology*, n. March, p. 238–241, 2007. Citado na página 27.
- SONI, B.; DAS, P. K.; THOUNAOJAM, D. M. CMFD: a detailed review of block based and key feature based techniques in image copy-move forgery detection. *IET Image Processing*, v. 12, n. 2, p. 167–178, 2018. ISSN 1751-9659. Citado na página 35.
- SRIVASTAVA, N.; HINTON, G.; KRIZHEVSKY, A.; SUTSKEVER, I.; SALAKHUTDINOV, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, v. 15, p. 1929–1958, 2014. ISSN 03702693. Disponível em: <<http://jmlr.org/papers/v15/srivastava14a.html>>. Citado 2 vezes nas páginas 12 e 53.
- SUBHEDAR, M. S.; MANKAR, V. H. Current status and key issues in image steganography: A survey. *Computer Science Review*, v. 13-14, p. 95 – 113, 2014. ISSN 1574-0137. Citado na página 25.
- SUNDE, N.; DROR, I. E. Cognitive and human factors in digital forensics: Problems, challenges, and the way forward. *Digital Investigation*, Elsevier Ltd, v. 29, p. 101–108, 2019. ISSN 17422876. Citado na página 21.

- SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCKE, V.; RABINOVICH, A. Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, v. 07-12-June, p. 1–9, 2015. ISSN 10636919. Citado 3 vezes nas páginas 55, 66 e 100.
- SZEGEDY, C.; VANHOUCKE, V.; IOFFE, S.; SHLENS, J.; WOJNA, Z. Rethinking the Inception Architecture for Computer Vision. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE, 2016. v. 2016-Decem, p. 2818–2826. ISBN 9781467388504. ISSN 10636919. Citado na página 55.
- TAN, M.; LE, Q. V. EfficientNet: Rethinking model scaling for convolutional neural networks. In: *36th International Conference on Machine Learning, ICML 2019*. [S.l.: s.n.], 2019. v. 2019-June, p. 10691–10700. ISBN 9781510886988. Citado na página 55.
- TAN, X.; TRIGGS, B. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions. *IEEE Transactions on Image Processing*, v. 19, n. 6, p. 1635–1650, 2010. ISSN 19410042. Citado 18 vezes nas páginas 14, 15, 68, 69, 70, 102, 103, 107, 108, 112, 115, 116, 119, 120, 121, 122, 126 e 128.
- TANG, W.; LI, H.; LUO, W.; HUANG, J. Adaptive steganalysis against WOW embedding algorithm. In: *2nd ACM workshop on Information hiding and multimedia security - IH&MMSec '14*. [S.l.: s.n.], 2014. p. 91–96. ISBN 9781450326476. Citado na página 94.
- TANG, W.; LI, H.; LUO, W.; HUANG, J. Adaptive steganalysis based on embedding probabilities of pixels. *IEEE Transactions on Information Forensics and Security*, v. 11, n. 4, p. 734–744, 2016. ISSN 15566013. Citado na página 94.
- TAVARES, R.; MADEIRO, F. Word-Hunt: A LSB Steganography Method with Low Expected Number of Modifications per Pixel. *IEEE Latin America Transactions*, v. 14, n. 2, p. 1058–1064, 2016. ISSN 15480992. Citado na página 27.
- THAI, T. H.; COGRANNE, R.; RETRAINT, F.; DOAN, T. N. C. JPEG Quantization Step Estimation and Its Applications to Digital Image Forensics. *IEEE Transactions on Information Forensics and Security*, v. 12, n. 1, p. 123–133, 2017. Citado na página 34.
- THANEKAR, S. A.; PAWAR, S. S. OCTA (STAR) PVD: A different approach of image steganography. *2013 IEEE International Conference on Computational Intelligence and Computing Research, IEEE ICCIC 2013*, p. 2–6, 2013. Citado na página 28.
- THANIKAISELVAN, V.; ARULMOZHIVARMAN, P. High Security Image Steganography Using IWT and Graph Theory. In: *2013 IEEE International Conference on Signal and Image Processing Applications*. [S.l.: s.n.], 2013. p. 337–342. ISBN 9781479902699. Citado na página 30.
- THIBAULT, G.; FERTIL, B.; NAVARRO, C.; PEREIRA, S.; CAU, P.; LEVY, N.; SEQUEIRA, J.; MARI, J. L. Shape and texture indexes application to cell nuclei classification. *International Journal of Pattern Recognition and Artificial Intelligence*, v. 27, n. 1, p. 1–23, 2013. ISSN 02180014. Citado 2 vezes nas páginas 12 e 61.
- TOTLA, R. V.; BAPAT, K. S. Comparative Analysis of Watermarking in Digital Images Using DCT DWT. *International Journal of Scientific and Research Publications (IJSRP)*, v. 3, n. 2, 2013. Citado na página 32.

- TRALIC, D.; ZUPANCIC, I.; GRGIC, S.; GRGIC, M. CoMoFoD - New Database for Copy-Move Forgery Detection. In: *Proceedings ELMAR-2013*. [S.l.: s.n.], 2013. p. 49–54. ISBN 9789537044145. ISSN 13342630. Citado 10 vezes nas páginas 39, 42, 115, 116, 117, 118, 119, 121, 125 e 127.
- TUAMA, A.; COMBY, F.; CHAUMONT, M. Camera Model Identification With The Use of Deep Convolutional Neural Networks. In: *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*. [S.l.: s.n.], 2016. p. 1–6. ISBN 9781509011384. Citado na página 34.
- TUCERYAN, M.; JAIN, A. K. Texture Analysis. In: CHEN, C. H.; PAU, L. F.; P S P Wang (Ed.). *Handbook of Pattern Recognition and Computer Vision*. Dartmouth, USA: [s.n.], 1993. cap. Part 2 - B, p. 235–276. Citado na página 58.
- UNSER, M.; SAGE, D.; VILLE, D. V. D. Multiresolution monogenic signal analysis using the Riesz-Laplace wavelet transform. *IEEE Transactions on Image Processing*, v. 18, n. 11, p. 2402–2418, 2009. ISSN 10577149. Citado 2 vezes nas páginas 64 e 85.
- URVOY, M.; GOUDIA, D.; AUTRUSSEAU, F. Perceptual DFT watermarking with improved detection and robustness to geometrical distortions. *IEEE Transactions on Information Forensics and Security*, v. 9, n. 7, p. 1108–1119, 2014. ISSN 15566013. Citado 2 vezes nas páginas 30 e 63.
- USTUBIOGLU, B.; BAYKAL, E.; Gul Muzaffer; ULUTAS, G. Blur Invariant Image Forgery Detection Method Using Local Phase Quantization. *Journal of Energy and Power Engineering*, David Publishing Company, v. 10, n. 6, jun 2016. ISSN 19348975. Citado na página 82.
- VAMVAKAS, A.; TSOUGOS, I.; ARIKIDIS, N.; KAPSALAKI, E.; FOUNTAS, K.; FEZOULIDIS, I.; COSTARIDOU, L. Exploiting morphology and texture of 3D tumor models in DTI for differentiating glioblastoma multiforme from solitary metastasis. *Biomedical Signal Processing and Control*, Elsevier Ltd, v. 43, p. 159–173, 2018. ISSN 17468108. Disponível em: <<https://doi.org/10.1016/j.bspc.2018.02.014>>. Citado na página 61.
- VAN, L. T.; EMMANUEL, S.; KANKANHALLI, M. S. Identifying source cell phone using chromatic aberration. In: *2007 IEEE International Conference on Multimedia and Expo*. [S.l.: s.n.], 2007. p. 883–886. ISBN 1424410177. Citado na página 34.
- VAPNIK, V. N. *Statistical Learning Theory*. [S.l.]: Wiley, 1998. Citado na página 110.
- VARMA, M.; GARG, R. Locally invariant fractal features for statistical texture classification. In: *Proceedings of the IEEE International Conference on Computer Vision*. [S.l.: s.n.], 2007. ISBN 9781424416318. Citado na página 66.
- VÁZQUEZ-PADÍN, D.; COMESAÑA, P.; PÉREZ-GONZÁLEZ, F. An SVD approach to forensic image resampling detection. In: *2015 23rd European Signal Processing Conference (EUSIPCO)*. [S.l.: s.n.], 2015. p. 2067–2071. ISBN 9780992862633. Citado na página 38.
- VAZQUEZ-PADIN, D.; PEREZ-GONZALEZ, F.; COMESANA-ALFARO, P. A Random Matrix Approach to the Forensic Analysis of Upscaled Images. *IEEE Transactions on Information Forensics and Security*, v. 12, n. 9, p. 2115–2130, 2017. ISSN 15566013. Citado 2 vezes nas páginas 38 e 62.

- VENUS; RACHNA. Implementing Random Encoding for Image Steganography. *International Journal of Science and Research (IJSR)*, v. 5, n. 7, p. 954–957, 2016. Citado na página 28.
- VIZIREANU, D.; PREDA, R. Watermarking-based image authentication robust to JPEG compression. *Electronics Letters*, v. 51, n. 23, p. 1873–1875, 2015. ISSN 0013-5194. Citado na página 32.
- WARIF, N. B. A.; WAHAB, A. W. A.; IDRIS, M. Y. I.; RAMLI, R.; SALLEH, R.; SHAMSHIRBAND, S.; CHOO, K. K. R. Copy-move forgery detection: Survey, challenges and future directions. *Journal of Network and Computer Applications*, Elsevier, v. 75, p. 259–278, 2016. ISSN 10958592. Citado na página 35.
- WO, Y.; YANG, K.; HAN, G.; CHEN, H.; WU, W. Copy-move forgery detection based on multi-radius PCET. *IET Image Processing*, v. 11, n. 2, p. 99–108, 2017. ISSN 1751-9659. Citado na página 37.
- WU, D. C.; TSAI, W. H. A steganographic method for images by pixel-value differencing. *Pattern Recognition Letters*, v. 24, n. 9-10, p. 1613–1626, 2003. ISSN 01678655. Citado na página 28.
- XU, B.; WANG, X.; ZHOU, X.; XI, J.; WANG, S. Source camera identification from image texture features. *Neurocomputing*, Elsevier, v. 207, p. 131–140, 2016. ISSN 0925-2312. Citado na página 34.
- XU, G. S.; WU, H. Z.; SHI, Y. Q. Structural Design of Convolutional Neural Networks for Steganalysis. *IEEE Signal Processing Letters*, v. 23, n. 5, p. 708–712, 2016. ISSN 1070-9908. Citado na página 94.
- XU, Y.; JI, H.; FERMÜLLER, C. Viewpoint invariant texture description using fractal analysis. *International Journal of Computer Vision*, v. 83, n. 1, p. 85–100, 2009. ISSN 09205691. Citado na página 66.
- YAN, C. P.; PUN, C. M.; YUAN, X. C. Quaternion-Based Image Hashing for Adaptive Tampering Localization. *IEEE Transactions on Information Forensics and Security*, v. 11, n. 12, p. 2664–2677, 2016. ISSN 15566013. Citado na página 37.
- YANG, B. Q.; ZHANG, T.; GU, C. C.; WU, K. J.; GUAN, X. P. A novel face recognition method based on IWLD and IWBC. *Multimedia Tools and Applications*, Multimedia Tools and Applications, v. 75, n. 12, p. 6979–7002, 2016. ISSN 15737721. Citado 4 vezes nas páginas 68, 75, 102 e 103.
- YANG, M.; ZHANG, L.; SHIU, S. C. K.; ZHANG, D. Monogenic binary coding: An efficient local feature extraction approach to face recognition. *IEEE Transactions on Information Forensics and Security*, IEEE, v. 7, n. 6, p. 1738–1751, 2012. ISSN 15566013. Citado 4 vezes nas páginas 68, 84, 102 e 103.
- YAO, H.; QIAO, T.; XU, M.; ZHENG, N. Robust Multi-classifier for Camera Model Identification based on Convolution Neural Network. *IEEE Access*, v. 6, p. 24973–24982, 2018. ISSN 21693536. Citado 2 vezes nas páginas 94 e 95.
- YE, J.; NI, J.; YI, Y. Deep Learning Hierarchical Representations for Image Steganalysis. *IEEE Transactions on Information Forensics and Security*, v. 12, n. 11, p. 2545–2557, 2017. ISSN 15566013. Citado na página 94.

- ZENG, J.; TAN, S.; LI, B.; HUANG, J. Large-Scale JPEG Image Steganalysis Using Hybrid Deep-Learning Framework. *IEEE Transactions on Information Forensics and Security*, v. 13, n. 5, p. 1200–1214, 2018. ISSN 15566013. Citado 2 vezes nas páginas 93 e 94.
- ZHANG, K.; LIANG, Y.; ZHANG, J.; WANG, Z.; LI, X. No One Can Escape: A General Approach to Detect Tampered and Generated Image. *IEEE Access*, IEEE, v. 7, p. 129494–129503, 2019. ISSN 21693536. Citado na página 60.
- ZHANG, L.; ZHANG, L.; GUO, Z.; ZHANG, D. MONOGENIC-LBP: A new approach for rotation invariant texture classification. *Proceedings - International Conference on Image Processing, ICIP*, IEEE, p. 2677–2680, 2010. ISSN 15224880. Citado na página 84.
- ZHAO, J.; GUO, J. Passive forensics for copy-move image forgery using a method based on DCT and SVD. *Forensic Science International*, v. 233, n. 1-3, p. 158–166, 2013. ISSN 03790738. Citado na página 37.
- ZHAO, S.; GAO, Y.; ZHANG, B. Sobel-LBP. In: *2008 15th IEEE International Conference on Image Processing*. [S.l.: s.n.], 2008. p. 2144–2147. ISBN 9781424417643. Citado na página 73.
- ZHENG, L.; ZHANG, Y.; THING, V. L. A survey on image tampering and its detection in real-world photos. *Journal of Visual Communication and Image Representation*, Elsevier Inc., v. 58, p. 380–399, 2019. ISSN 10959076. Citado 3 vezes nas páginas 15, 24 e 25.
- ZHOU, B.; LAPEDRIZA, A.; KHOSLA, A.; OLIVA, A.; TORRALBA, A. Places: A 10 Million Image Database for Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 40, n. 6, p. 1452–1464, 2018. ISSN 01628828. Citado na página 66.
- ZHOU, L.; WANG, H. Local Gradient Increasing Pattern for Facial Expression Recognition. In: *International Conference on Image Processing, ICIP*. [S.l.]: IEEE, 2012. p. 2601–2604. ISBN 9781467325332. ISSN 15224880. Citado 6 vezes nas páginas 13, 68, 78, 79, 102 e 103.
- ZHOU, W.; YU, L.; QIAN, Y.; QIU, W.; ZHOU, Y.; LUO, T. Deep blind quality evaluator for multiply distorted images based on monogenic binary coding. *Journal of Visual Communication and Image Representation*, v. 60, p. 305–311, 2019. ISSN 10959076. Citado na página 84.
- ZOPH, B.; VASUDEVAN, V.; SHLENS, J.; LE, Q. V. Learning Transferable Architectures for Scalable Image Recognition. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City, Utah, EUA: [s.n.], 2018. Citado 2 vezes nas páginas 55 e 128.
- ZOU, D.; SHI, Y. Q.; NI, Z.; SU, W. A Semi-Fragile Lossless Digital Watermarking Scheme Based on Integer Wavelet Transform. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 16, n. 10, p. 1294–1300, 2006. Citado na página 30.
- ZOU, Y.; ZHANG, G.; LIU, L. Research on image steganography analysis based on deep learning. *Journal of Visual Communication and Image Representation*, Elsevier Inc., v. 60, p. 266–275, 2019. ISSN 10959076. Citado na página 94.

