



*Processamento de Impressões Digitais com
Texturas Utilizando Redes Neurais Pix2pix e
Xception*

Jader dos Santos Teles Cordeiro

Orientador: Prof. Dr. José Hiroki Saito

Janeiro / 2022

Dissertação de Mestrado em Ciência da
Computação

Processamento de Impressões Digitais com Texturas Utilizando Redes Neurais Pix2pix e Xception

Esse documento corresponde à dissertação apresentada à Banca Examinadora no Programa de Mestrado em Ciência da Computação do UNIFACCAMP – Centro Universitário Campo Limpo Paulista.

Campo Limpo Paulista, 13 de janeiro de 2022.

Jader dos Santos Teles Cordeiro

Prof. Dr. José Hiroki Saito (Orientador)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Ficha catalográfica elaborada pela
Biblioteca Central da Unifaccamp

C819p

Cordeiro, Jader dos Santos Teles

Processamento de impressões digitais com texturas
utilizando redes neurais *Pix2pix* e *Xception* / Jader dos Santos Teles
Cordeiro. Campo Limpo Paulista, SP: Unifaccamp, 2022.

Orientador: Profº. Dr. José Hiroki Saito

Dissertação (Programa de Mestrado Profissional em Ciência
da Computação) – Centro Universitário Campo Limpo Paulista –
Unifaccamp.

1. Detecção de imagem falsa. 2. Rede neural adversária. 3.
Gerador e discriminador. 4. Impressões digitais. 5. Texturas.
6. Réplicas. I. Saito, José Hiroki. II. Centro Universitário
Campo Limpo Paulista. III. Título.

CDD- 006.32

Agradecimentos

Agradeço a Deus por permitir superar as dificuldades em toda trajetória, pela compreensão da família, minha mãe pelo amor e carinho, ao meu pai (*in memoriam*) pelos ensinamentos.

Ao Prof. Dr. José Hiroki Saito, pelo qual, tenho uma eterna gratidão, admiração e grande respeito. Agradeço pelo apoio, orientação, ajuda em todo o trabalho, pela paciência e atenção. Obrigado por compartilhar seu conhecimento e por promover de forma significativa o crescimento acadêmico e pessoal, permitindo vivenciar e perceber a pesquisa com outro olhar.

Agradeço a todos os professores do programa de Mestrado em Ciência da Computação, que participaram direta ou indiretamente através dos conteúdos abordados nas disciplinas.

Aos colegas de curso, pelo apoio.

Resumo: A pesquisa sobre detecção de imagens de impressões digitais falsas é importante, uma vez que existem muitas tentativas de falsificações. Este trabalho teve como objetivo realizar estudos sobre a construção de réplicas de impressões digitais, para contribuir no aperfeiçoamento dos sistemas de detecção dessas falsificações. Existem técnicas de redes neurais, especificamente, a estrutura GAN (*Generative Adversarial Network*), constituída de duas redes neurais, geradora G e discriminadora D, que é usada para sintetizar imagens com realismo usando G, e aferindo o resultado usando D, que podem ser usadas para aumentar os bancos de dados de amostras para a finalidade de treinamento de classificadores. Neste trabalho de dissertação é utilizada a rede GAN Pix2pix para gerar imagens sintéticas de impressões digitais, para aumentar o número de amostras em bancos de imagens, uma vez que os algoritmos geralmente requerem uma quantidade grande de amostras para treinamento. A rede GAN também é aplicada para a obtenção de réplicas de impressões digitais, com a incorporação de texturas obtidas de amostras específicas, uma abordagem que foi denominada SIMTEX (*Simulação de textura*). Para a verificação da similaridade entre imagens de impressões digitais originais e sintetizadas, é utilizada uma rede convolucional Xception (*Extreme Inception*), além da métrica SSIM (*Structure Similarity Index Measure*) e RMSE (*Root Mean Square Error*). Para 4800 réplicas de impressões digitais de 3 pessoas, usando SIMTEX, comparadas com as respectivas originais, resultaram em similaridade média de 0,9649 na RMSE, 0,9604 na SSIM e 0,9847 com Xception, e desvio padrão de 0,0043, 0,0091 e 0,0158, respectivamente. Foi aplicado o teste de reconhecimento para as réplicas, utilizando a rede Xception, treinada com as impressões digitais originais, e variando o número de épocas de treinamento do SIMTEX.

Palavras-chave: detecção de imagem falsa, rede neural adversária, gerador e discriminador, impressões digitais, texturas, réplicas.

Abstract: Research on the detection of fake fingerprint images is important since there are many attempts at forgery. This work aimed to carry out studies on the construction of fingerprint replicas, to contribute to the improvement of detection systems for these counterfeits. There are neural network techniques, specifically, the GAN (Generative Adversarial Network) framework, consisting of two networks, generator G and discriminator D, which are used to synthesize images realistically using G, and checking the result using D, which can be used to augment the number of samples in databases for the purpose of classifiers training. In this dissertation proposal, the GAN Pix2pix network is used to generate synthetic fingerprint images to increase the image banks, since the algorithms usually require a large number of samples for training. The GAN network is also applied to obtain fingerprints replicas by incorporating textures obtained from specific samples, an approach that has been termed SIMTEX (*Texture Simulation*). For checking the similarity between originals and synthesized fingerprint images, a convolutional network Xception (*Extreme Inception*) is used, in addition to the SSIM and RMSE metric. For 4800 fingerprints replicas of 3 people, using SIMTEX, compared to the respective originals, resulted in mean similarity of 0.9649 for RMSE, 0.9604 for SSIM and 0.9847 for Xception, and standard deviation of 0.0043, 0.0091 and 0.0158, respectively. The recognition test was applied to the replicas, using the Xception network, trained with the original fingerprints, and varying the number of training times of SIMTEX.

Keywords: false image detection, adversarial neural network, generator and discriminator, fingerprints, textures, replicas.

Sumário

CAPÍTULO 1 - Introdução	1
1.1. Problemas abordados	3
1.2. Objetivos gerais e específicos	4
1.3. Principais contribuições	5
1.4. Organização do trabalho	5
CAPÍTULO 2 - Contextualização e Revisão Bibliográfica	6
2.1. Contextualização	6
2.2. Pesquisa bibliográfica sobre redes neurais profundas e redes GAN	10
2.3. Pesquisa bibliográfica sobre IDs falsas e tecnologias propostas para sua detecção	16
2.4. Pesquisa bibliográfica sobre imagens biométricas de forma geral e detecção de falsificações	21
2.5. Considerações finais	25
CAPÍTULO 3 - Estrutura das Redes Neurais Utilizadas	26
3.1. Rede adversária	26
3.2. Rede adversária condicional	29
3.3. Pix2pix	31
3.3.1. Arquitetura dos módulos gerador e discriminador da rede Pix2pix	33
3.4. Xception	39
3.4.1. Arquitetura da rede Xception	40
3.5. Considerações finais	42
CAPÍTULO 4 - Materiais e Métodos	44
4.1. Banco de dados utilizados	44
4.1.1. Banco de dados biométrico Sokoto	44
4.1.2. Banco de dados de textura	45
4.1.3. Banco de dados próprio – BDP	45
4.1.3.1. Dispositivo de posicionamento para captura biométrica	46
4.1.3.2. Obtenção das imagens de IDs originais	46
4.1.3.3. Obtenção das imagens réplicas de impressões digitais	47
4.2. Plano de Pesquisa e Metodologia	50
4.2.1. Pré-processamentos	52
4.2.2. Metodologia para ampliação utilizando Pix2pix	55
4.2.3. Metodologia SIMTEX para a obtenção de réplicas de IDs	57

4.2.4. Similaridade entre IDs	58
4.2.4.1. SSIM	59
4.2.4.2. Distância Euclidiana	60
4.2.5. Reconhecimento de IDs	61
4.3. Avaliação do desempenho	61
4.4. Considerações finais	62
CAPÍTULO 5 – Experimentos e Resultados	63
5.1. Experimento 1 - Geração de IDs sintéticas com Pix2pix a partir do banco BDP1	63
5.2. Experimento 2 - Geração de IDs sintéticas com Pix2pix a partir do banco Sokoto	65
5.3. Experimento 3 – Síntese de Réplicas utilizando SIMTEX para o banco BDP1	66
5.4. Experimento 4 – Síntese de Réplicas utilizando SIMTEX para o banco Sokoto	70
5.5. Experimento 5 – Síntese de Réplicas utilizando SIMTEX para o banco BDP2	74
5.5.1. Experimento 5.1 – Construção de réplicas de IDs do banco BDP2, associando textura com SIMTEX	74
5.5.2. Experimento 5.2 - Testes de reconhecimento das réplicas de IDs obtidas no Experimento 5.1	76
5.5.3. Experimento 5.3 - Medidas de Similaridade das réplicas de IDs, obtidas no Experimento 5.1, com as originais do BDP2	81
5.6. Recursos computacionais	84
5.7. Considerações finais	84
CAPÍTULO 6 – Conclusões e Trabalhos Futuros	85
REFERÊNCIAS	88
APÊNDICE 1 – Tabelas de Reconhecimento das réplicas do banco BDP2	103

Glossário

AdaIN	- <i>Adaptive Instance Norm</i>
Adam	- <i>Adaptive moment estimation</i>
AlexNet	- <i>Alex Network</i>
BDP	- Banco de Dados Próprio
BGP	- <i>Binary Gabor Patterns</i>
BN	- <i>Batch Normalization</i>
BRISQUE	- <i>Blind/Reference-less Image Spatial Quality Evaluator</i>
BSIF	- <i>Binary Statistical Image Features</i>
CapsNet	- <i>Capsules Network</i>
CDR	- <i>Cup-to-Disc Ratio</i>
cGAN	- <i>Conditional Generative Adversarial Network</i>
CLAHE	- <i>Contrast-Limited Adaptive Histogram Equalization</i>
CNN	- <i>Convolutional Neural Network</i>
CycleGAN	- <i>Cycle-Consistent Adversarial Network</i>
DC-GAN	- <i>Deep Convolutional Generative Adversarial Network</i>
DenseNet	- <i>Dense Convolutional Network</i>
DTD	- <i>Describable Textures Dataset</i>
ERR	- <i>Equal Error Rate</i>
FAR	- <i>False Accept Rate</i>

FCN	- <i>Fully Convolutional Network</i>
FDSNet	- <i>Finger Dorsal Spoof Detection Network</i>
FPAD	- <i>Fingerprint Presentation Attack Detection</i>
FRR	- <i>False Reject Rate</i>
GAN	- <i>Generative Adversarial Network</i>
GIMP	- <i>GNU Image Manipulation Program</i>
GLCM	- <i>Gray Level Co-occurrence Matrix</i>
GTAW	- <i>Gas Tungsten Arc Welding</i>
HOG	- <i>Histogram of Oriented Gradient</i>
HRNet	- <i>High-Resolution Net</i>
ID	- <i>Impressão Digital</i>
ILBP	- <i>Improved Local Binary Patterns</i>
IQA	- <i>Image Quality Assessment</i>
LBP	- <i>Local Binary Patterns</i>
LeakyReLU	- <i>Leaky Rectified Linear Unit</i>
LFC	- <i>Light Field Camera</i>
LiveDet2011	- <i>Liveness Detection</i>
LR	- <i>Learning rate</i>
LSCI	- <i>Laser Speckle Contrast Imaging</i>

MWI	- <i>Modified Williams Index</i>
NFIQ	- <i>NIST Fingerprint Image Quality Measure</i>
NR	- <i>No-reference</i>
PAD	- <i>Presentation Attack Detection</i>
PDE	- <i>Partial Differential Equations</i>
PIL	- <i>Python Imaging Library</i>
PIN	- <i>Personal Identification Number</i>
QDA	- <i>Quadratic Discriminant Analyzer</i>
ReLU	- <i>Rectified Linear Unit</i>
ResNet	- <i>Residual Network</i>
RMSE	- <i>Root Mean Square Error</i>
ROC	- <i>Receiver Operating Characteristic</i>
ROI	- <i>Region Of Interest</i>
ShallowNet	- <i>Shallow Convolutional Network</i>
SIMTEX	- <i>Simulação de textura</i>
SLDL	- <i>Sri Lanka Driving License</i>
Slim-ResCNN	- <i>Slim residual convolutional neural network</i>
SRC	- <i>Sparse Representation Classifier</i>
SS-OCT	- <i>Swept Source Optical Coherence Tomography</i>

- SSIM** - *Structure Similarity Index Measure*
- StyleGAN** - *Style Generative Adversarial Network*
- SVM** - *Support Vector Machine*
- VGGNet** - *Visual Geometry Group Network*
- WGAN** - *Wasserstein GAN*
- Xception** - *Extreme Inception*
- YOLOv2** - *You Only Look Once*

Lista de Tabelas

Tabela 1.	Resumo da pesquisa sobre redes neurais e redes GAN	15
Tabela 2.	Pesquisas envolvendo IDs falsas	21
Tabela 3.	Pesquisas sobre imagens biométricas de forma geral e detecção de falsificações	25
Tabela 4.	Medidas de similaridade das IDs originais com as respectivas IDs sintetizadas	64
Tabela 5.	Proporção de similaridade dos conjuntos Sokoto por dedo gerado	65
Tabela 6.	Medidas de similaridade das réplicas com as IDs originais para a textura <i>blotchy</i>	68
Tabela 7.	Medidas de similaridade das réplicas com as IDs originais para a textura <i>bumpy</i>	68
Tabela 8.	Medidas de similaridade das réplicas com as IDs originais para a textura <i>cracked</i>	69
Tabela 9.	Medidas de similaridade das réplicas com as IDs originais para a textura <i>crosshatched</i>	69
Tabela 10.	Medidas de similaridade das réplicas com as IDs originais para a textura <i>fibrous</i>	70
Tabela 11.	Medidas de similaridade das réplicas com as IDs originais para a textura <i>blotchy</i>	72
Tabela 12.	Medidas de similaridade das réplicas com as IDs originais para a textura <i>bumpy</i>	72
Tabela 13.	Medidas de similaridade das réplicas com as IDs originais para a textura <i>cracked</i>	73
Tabela 14.	Medidas de similaridade das réplicas com as IDs originais para a textura <i>crosshatched</i>	73
Tabela 15.	Medidas de similaridade das réplicas com as IDs originais para a textura <i>fibrous</i>	73
Tabela 16.	Medidas de desempenho das réplicas obtidas com a textura <i>blotchy</i>	79
Tabela 17.	Medidas de desempenho das réplicas obtidas com a textura <i>bumpy</i>	79
Tabela 18.	Medidas de desempenho das réplicas obtidas com a textura <i>crosshatched</i>	79
Tabela 19.	Medidas de desempenho das réplicas obtidas com a textura <i>fibrous</i>	79

Tabela 20.	Similaridade média e desvio padrão para as réplicas usando textura <i>blotchy</i>	81
Tabela 21.	Similaridade média e desvio padrão para as réplicas usando textura <i>bumpy</i>	81
Tabela 22.	Similaridade média e desvio padrão para as réplicas usando textura <i>crosshatched</i>	81
Tabela 23.	Similaridade média e desvio padrão para as réplicas usando textura <i>fibrous</i>	82

Lista de Figuras

Figura 1	Algumas modalidades biométricas como método de identificação.	7
Figura 2	Módulos de um sistema biométrico.	9
Figura 3	Estrutura da rede adversária GAN.	27
Figura 4	Estrutura da rede adversária condicional cGAN.	30
Figura 5	Diagrama de blocos da rede Pix2pix.	32
Figura 6	Arquitetura U-Net.	34
Figura 7	(a) representação de um kernel 3x3 e (b) kernel com os respectivos valores.	35
Figura 8	Matriz C de convolução 4 x 16.	35
Figura 9	(a) matriz de entrada original, (b) matriz de entrada em coluna 16x1, (c) resultado da multiplicação em matriz coluna 4x1 e (d) resultado da multiplicação em matriz 2x2.	36
Figura 10	(a) matriz de convolução transposta, (b) matriz coluna multiplicadora, (c) matriz 16x1, resultado da multiplicação e (d) resultado em matriz 4x4.	37
Figura 11	Arquitetura Xception.	40
Figura 12	Ilustração da aplicação de convoluções separáveis em profundidade.	41
Figura 13	Dispositivo para posicionamento de captura biométrica: (a) projeto, (b) construído.	46
Figura 14	(a) moldeira de acrílico, (b) molde de <i>biscuit</i> construído sobre a moldeira.	49
Figura 15	(a) molde de <i>biscuit</i> , (b) dedo artificial.	50
Figura 16	Diagrama de blocos – Plano de Pesquisa.	51
Figura 17	Definição da região de interesse, ROI	52
Figura 18	(a) componente-x e (b) componente-y.	55
Figura 19	Diagrama ilustrativo do treinamento da rede Pix2pix e geração das IDs sintetizadas.	56
Figura 20	Estrutura da metodologia SIMTEX para geração de IDs com textura pela rede Pix2pix.	58
Figura 21	Diagrama ilustrativo do SSIM.	59
Figura 22	Imagens de IDs original e após 3,14 e 168 épocas de aplicação da rede Pix2pix, da esquerda para a direita, respectivamente.	63
Figura 23	Réplica obtida pelo SIMTEX, em 200 épocas, a partir de uma ID original, em 5 texturas <i>blotchy</i> , <i>bumpy</i> , <i>cracked</i> , <i>crosshatched</i> e	67

fibrous, respectivamente, da esquerda para direita.

Figura 24	Textura incorporada nos conjuntos por pessoa – Sokoto.	71
Figura 25	Impressões digitais que foram sintetizadas em 6, 200, 300 e 400 épocas com textura.	75
Figura 26	(a) <i>blotchy</i> , (b) <i>bumpy</i> , (c) <i>crosshatched</i> e (d) <i>fibrous</i> – Reconhecimento de IDs com texturas em 6, 200, 300 e 400 épocas.	77
Figura 27	Quantidade de erros de reconhecimento para diferentes números de épocas de treinamento.	78
Figura 28	Curva ROC para as réplicas obtidas com as texturas: (a) <i>blotchy</i> , (b) <i>bumpy</i> , (c) <i>crosshatched</i> e (d) <i>fibrous</i> .	80
Figura 29	Gráficos de Similaridade média obtidas pela métrica RMSE das réplicas com texturas <i>blotchy</i> , <i>bumpy</i> , <i>crosshatched</i> e <i>fibrous</i> , da esquerda para a direita e de cima para baixo.	83
Figura 30	Gráficos de Similaridade média obtidas pela métrica SSIM das réplicas com texturas <i>blotchy</i> , <i>bumpy</i> , <i>crosshatched</i> e <i>fibrous</i> , da esquerda para a direita e de cima para baixo.	83
Figura 31	Gráficos de Similaridade média obtidas pela métrica Xception das réplicas com texturas <i>blotchy</i> , <i>bumpy</i> , <i>crosshatched</i> e <i>fibrous</i> , da esquerda para a direita e de cima para baixo.	84

CAPÍTULO 1 – Introdução

A verificação de Impressão Digital (ID) é amplamente adotada na identificação de indivíduos em muitos segmentos que necessitam de um certo nível de segurança. Porém, para que os ataques de apresentação nos dispositivos de captura, conhecidos em inglês como *spoof attacks*, sejam praticados, os delinquentes não necessitam de conhecimentos específicos para tal. Chugh & Jain (2020) relatam que os objetos de falsificação podem ser fabricados com materiais comuns, como gelatina, cola, silicone.

Uma das formas de verificar a autenticidade de uma ID é pela confirmação do respectivo dedo ao vivo. A técnica denominada *live finger*, em inglês, verifica o dedo real apresentado no leitor do *scanner* que faz a captura da imagem, com a implantação de hardware adicional capaz de obter informações do batimento cardíaco, fluxo de sangue, temperatura, transpiração, sensores capacitivos, entre outros. Uma outra solução seria com base em software, na qual um algoritmo pode ser acoplado ao subsistema biométrico como recurso de extração de uma ou mais características, podendo ser as da textura da pele, frequência da crista, quantidade de poros, etc. Claro que o software não funciona sem o hardware, mas é uma forma alternativa de abordagem (Maltoni et al. 2009).

Mesmo que esses métodos atinjam desempenhos razoáveis, ainda são passíveis de aperfeiçoamentos (Nikam & Agarwal 2008). Por exemplo, as especificidades de cenários e de dispositivos de captura podem prejudicar a qualidade da imagem durante a aquisição, como a falta de iluminação, o excesso de brilho, a umidade na parte analisada ou no sensor.

Algoritmos com ênfase na detecção de impressões digitais baseados em texturas, como contramedida aos ataques com materiais simulados foram pesquisados por (Nikam & Agarwal 2008) e (Abhyankar & Schuckers 2006). Na verificação do impacto dos

materiais fabricados na detecção ao vivo, Rattani & Ross (2014) projetaram um esquema de pré-processamento para melhorar a distinção dos artefatos. Além disso, He et al. (2019) empregaram as representações profundas de diferentes espaços de cores para serem classificados por CNNs (*Convolutional Neural Networks*).

Por outro lado, as técnicas de reconstrução de imagens realísticas utilizando redes neurais, recentemente desenvolvidas, permitem reconstruir imagens de IDs realísticas a partir de imagens de IDs originais (Isola et al. 2017). Essas técnicas, diferentemente da geração de imagens falsas de *spoof attacks*, podem auxiliar no incremento do banco de dados, aumentando a quantidade de amostras de IDs para o treinamento do software de reconhecimento.

Nos ensaios experimentais realizados e também em pesquisas anteriores relatadas na literatura, pudemos constatar a versatilidade das GANs (*Generative Adversarial Networks*) condicionais para gerar imagens sintéticas. GAN é uma classe de estruturas de redes neurais projetadas inicialmente por (Goodfellow et al. 2014), que possui estrutura integrada por duas redes, a rede geradora (G) e a rede discriminadora (D), que competem entre si. A rede G compete tentando gerar uma imagem sintetizada mais próxima da original, e a rede discriminadora, na função de classificar as imagens em originais ou sintetizadas. Elegemos a rede Pix2pix (Isola et al. 2017) com características semelhantes, mas que possui um fator de condição, que serve de critério para obtermos novas amostras, conhecida por GAN condicional, a qual consideramos um recurso essencial para nossa pesquisa, principalmente quando temos a necessidade de incluir características de textura na imagem alvo, além da possibilidade de poder gerar variações das peculiaridades individuais das IDs, permitindo obter outras imagens que possam ser utilizadas, resultando na ampliação do banco de dados de treinamentos e testes. O fator de condição pode ser outro tipo de informação como rótulo da classe ou outros dados como imagens. Temos no treinamento do gerador, a inclusão da imagem de condição x , além da entrada de arranjo aleatório z , para a geração da imagem sintetizada y . A imagem de saída y é avaliada pelo discriminador, um crítico que estima de forma

eficiente comparando a imagem gerada, com a imagem apresentada como condição, aferindo se as amostras produzidas estão próximas das de referência.

A rede Pix2pix é constituída de estruturas convolucionais, que pelo fato das conexões dos neurônios restritas a uma pequena área, tem custo computacional menor em relação às redes totalmente conectadas. Uma outra rede que utiliza essa estrutura convolucional é a rede Xception (*Extreme Inception*), proposto por (Chollet 2017), que devido à incorporação de aprimoramentos recentemente propostos como as conexões residuais (Szegedy et al. 2016), apresenta bom desempenho.

1.1. Problemas abordados

Neste trabalho de dissertação, foram abordados dentre outros, dois problemas na área de identificação de indivíduos por ID: 1) utilizar a técnica de reconstruções realísticas de imagens de IDs originais, para a ampliação do banco de dados de treinamento do sistema de reconhecimento de impressões digitais, e 2) caracterizar as diferenças de texturas em imagens de IDs reais e falsificadas (réplicas), utilizando redes neurais profundas.

As redes GANs são aplicadas para gerar imagens sintéticas realísticas, diferentes de imagens de IDs originais, porém reconhecíveis, para o incremento do banco de imagens para treinamentos e testes.

No pré-processamento das imagens reais e falsificadas, utiliza-se do método de equalização de histograma, CLAHE (*Contrast-Limited Adaptive Histogram Equalization*), aplicada por (Zuiderveld 1994) e em trabalhos de (Pizer et al. 1987).

Para amostras de IDs reais, foi desenvolvido um Banco de Dados Próprio (BDP) com 10 imagens originais de cada dedo da mão esquerda, de três indivíduos. Foi utilizada uma abordagem denominada SIMTEX (*Simulação de textura*) para a construção de réplicas, incorporando texturas às IDs originais, no estudo da

vulnerabilidade dos sistemas biométricos. Para isso, aferimos a similaridade e a reconhecibilidade das réplicas de IDs em relação às IDs originais.

1.2. Objetivos gerais e específicos

O objetivo geral do trabalho é pesquisar as diversas atividades em relação às impressões digitais, incluindo obtenção de amostras para banco de dados de imagens originais de IDs, amostras sintetizadas por redes Pix2pix, e réplicas obtidas por incorporação de textura, no contexto do o trabalho é pesquisar as diversas atividades em relação às impressões digitais, incluindo obtenção de amostras para banco de dados de imagens originais de IDs, amostrestudo da vulnerabilidade dos sistemas biométricos utilizando impressões digitais.

Como objetivos específicos podemos destacar os seguintes itens:

- Realizar o pré-processamento de imagens de IDs, redimensionando as imagens, aplicando o método de realce de Sobel, de Sobel & Feldman (1968) e de equalização de histograma, CLAHE aplicado por (Zuiderveld 1994) e em trabalhos de (Pizer et al. 1987);
- Utilizar modelos de construção de réplicas de IDs, com dedos artificiais, que podem resultar em imagens com texturas diferenciáveis, devido às propriedades do material utilizado;
- Utilizar um procedimento de simulação de textura em substituição aos modelos de construção de réplicas de IDs com dedos artificiais, associando imagens de um banco de dados especial de textura;
- Abordar a ampliação de banco de dados reduzidos, com o uso de redes Pix2pix para a geração de imagens realísticas a partir de amostras de imagens originais;

- Desenvolver um sistema de reconhecimento de imagens de IDs, originais, diferenciando-as das réplicas de IDs.

1.3. Principais contribuições

As principais contribuições deste trabalho podem ser descritas como:

- Aplicação do pré-processamento com o método de equalização de histograma, CLAHE, e realce de imagem com operadores de Sobel, para contribuir na precisão da classificação de IDs;
- Metodologia para aumentar a base de dados com a rede Pix2pix, gerando imagens sintéticas realísticas diferentes das originais, porém reconhecíveis;
- Metodologia para a construção de um banco de dados próprio, constituído de imagens de IDs originais e réplicas, com o uso de dedos artificiais, ou utilizando a incorporação de texturas em IDs originais;
- Utilização da rede Xception, distância Euclidiana RMSE e a métrica SSIM para o cálculo de similaridade entre IDs.

1.4. Organização do trabalho

No restante do texto, no Capítulo 2, são descritos: a contextualização, os trabalhos relacionados ao reconhecimento, detecção e classificação de imagens com variedade de redes GANs. Na sequência, no Capítulo 3, temos a estrutura das redes neurais, com detalhes das redes adversária condicionais cGAN (*Conditional Generative Adversarial Network*) e redes convolucionais; e no Capítulo 4, os materiais e métodos. No Capítulo 5 abordamos os resultados e discussões e no Capítulo 6, as conclusões e trabalhos futuros. No Apêndice 1 constam as tabelas de reconhecimento das réplicas de IDs obtidas, variando o número de épocas de treinamento.

CAPÍTULO 2 – Contextualização e Revisão Bibliográfica

Neste capítulo, apresentamos inicialmente as informações contextuais, seguidas da revisão bibliográfica sobre os trabalhos encontrados na literatura, em relação ao tema proposto. Temos na Seção 2.1 a contextualização; Seção 2.2 a pesquisa bibliográfica na área de reconhecimento de padrões explorando redes neurais profundas e redes GAN. Na Seção 2.3, temos uma pesquisa bibliográfica com enfoque nas impressões digitais falsas e na Seção 2.4, uma pesquisa bibliográfica em imagens falsas em sistemas biométricos de forma mais ampla.

2.1. Contextualização

Podemos definir grandes grupos que utilizam o sistema biométrico, como as organizações educacionais, comerciais, governamentais, forenses que englobam aplicações nas investigações e, também, medicina e odontologia legal. Alguns métodos no processo de identificação envolvem análises moleculares; odontológicas; reconhecimento facial, ocular e intraocular; identificação das impressões digitais e palmares; análise do fluxo sanguíneo nas veias; padrões de marcha ou caminhada; tom de voz, entre outros (Maltoni et al. 2009).

Gupta (2017) inclui a celeridade de medidas que minimizam a possibilidade de violações das autenticidades individuais. A biometria vem conquistando espaço sobre os métodos comuns que envolvem cartões, senhas, autenticação em duas etapas, PINs (*Personal Identification Numbers*), e outros, constituindo um mecanismo que possibilita estabelecer traços específicos de unicidade na confirmação da identificação.

O reconhecimento biométrico de acordo com Jain et al. (2004), deve abranger alguns requisitos fundamentais de referência na definição das particularidades que serão analisadas:

- Universalidade: o atributo deve ser comum a todos;

- Distinção, Reconhecibilidade, Discriminabilidade: deve ser possível distinguir as peculiaridades das pessoas;
- Permanência: o parâmetro definido deve ter seu aspecto preservado, mantendo-se estável por um longo período; e
- Coleta e quantificação: possibilidade de coleta simplificada e medidas com parâmetros quantitativos.

Devemos considerar também, de acordo com Jain et al. (2004), referindo-se ao desempenho como um fator que envolve a velocidade e precisão na verificação dos recursos pretendidos, além da necessidade de ampla aceitabilidade de todos na utilização dos parâmetros indicados, de preferência que não sejam invasivos e tecnicamente consistentes contra métodos enganosos.

Os seres humanos possuem diversas partes do corpo que podem ser utilizadas para extrair características únicas para serem implementadas na biometria. Na Figura 1 apresentamos alguns métodos no processo de identificação.

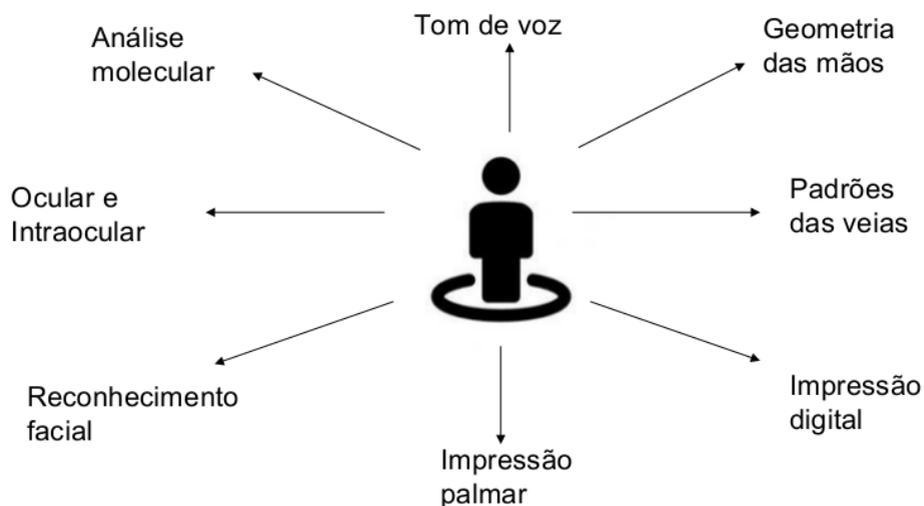


Figura 1. Algumas modalidades biométricas como método de identificação.

Destacamos a biometria na região das mãos onde temos algumas modalidades, a impressão digital, o padrão das veias, impressão palmar. Já na região da face, temos a modalidade ocular, além de outras. Quando esses critérios são especificados em apenas uma modalidade, consideramos os sistemas biométricos unimodais. Os sistemas biométricos multimodais envolvem informações complementares associadas a outros recursos, simultaneamente verificados, de forma a ampliar o potencial discriminativo obtendo melhor precisão no reconhecimento, abrangendo também um número maior de pessoas que nos sistemas unimodais. Como exemplo, citamos uma combinação da análise do tom de voz em conjunto com a biometria das mãos.

Uma estrutura de um sistema biométrico adaptada de Sivasankari & Muthukumar (2016) sobre os principais componentes e o fluxo das informações até o processo de armazenamento está ilustrada na Figura 2 (Jain et al. 2004), constando os módulos de inscrição, verificação e identificação. Essa estrutura é composta pelas camadas de 1) aquisição da imagem, com sensores e outros instrumentos; 2) pré-processamento da imagem, para adequação de necessidades peculiares; 3) extração de recursos, que fará a obtenção de informações específicas que definirão sua unicidade; 4) comparação, na verificação da igualdade do recurso apresentado com o modelo para pontuar sua avaliação e 5) armazenamento no banco de dados, no qual constam os recursos aprendidos e os modelos inscritos.

Um sistema biométrico pode ser utilizado de acordo com as fases de inscrição, validação e identificação:

- A fase de inscrição tem a finalidade de registrar as informações biométricas de um indivíduo, que serão vinculadas a uma chave exclusiva;
- Na fase de verificação, as imagens coletadas serão transformadas em parâmetros e comparadas aos dados biométricos armazenados, a fim de validar o usuário registrado, onde, normalmente está vinculado a uma codificação e

- Na fase de identificação é feita a comparação da imagem biométrica apresentada, com modelos registrados na base de dados do sistema. Esta é a fase de utilização, propriamente dita, do sistema biométrico. A fase de inscrição refere-se à formação do banco de dados do sistema.

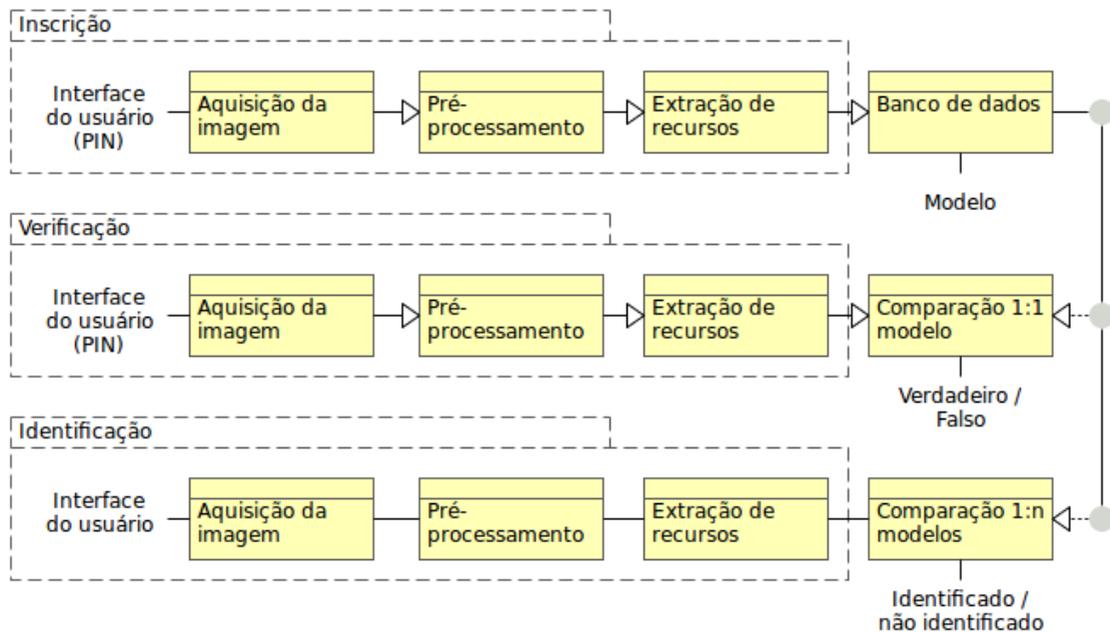


Figura 2. Módulos de um sistema biométrico - adaptado de (Jain et al. 2004).

Conforme já mencionado, os sistemas biométricos podem ser unimodais ou multimodais. Com a possibilidade de usar diferentes tecnologias, Geethanjali & Thamaraiselvi (2013) consideram múltiplos fatores de forma independente e seus atributos específicos, onde uma fusão pode ser realizada a nível de atributos, na qual várias características extraídas serão associadas, a nível de classificação, ocorrendo uma combinação de dados em diversos pontos para considerar o resultado da classificação, que será o nível de decisão. Amirthalingam & Radhamani (2013) mostra a importância da técnica de fusão para aprimorar os pontos fortes e reduzir a desvantagem dos métodos individuais. São informações obtidas com diversas tecnologias consideradas dados multibiométricos, sendo recursos decorrentes de diferentes modalidades ou obtidos de

uma variedade de atributos distintos correlacionados na mesma modalidade. Assim, um conjunto de pontuações será gerado, a fim de aprimorar a precisão do sistema biométrico.

2.2. Pesquisa bibliográfica sobre redes neurais profundas e redes GAN

Uma investigação sobre os trabalhos já realizados, para fornecer informações atuais, relevantes em relação ao tema e com importantes subsídios para a pesquisa, foi realizada no formato de estudo exploratório inicial de forma sistemática, em base de dados de repositórios com produções científicas nas palavras-chave: detecção de imagem falsa, rede neural adversária, gerador e discriminador, impressões digitais, texturas e réplicas, ajustadas para busca avançada em todo texto e metadados nos últimos cinco anos entre 2015 a 2020, em ACM Digital Library, ArXiv.org, IEEE Xplore, Science Direct e Springer Link, retornando 898 publicações. Com as adequações necessárias, seguimos os critérios de Gil (2002) e Marconi & Lakatos (2017) para um exame minucioso de verificação crítica. Assim, na seleção de trabalhos que abordavam o tema proposto, foram consideradas informações sobre a base dados, métricas e metodologia utilizada, para evitar informações confusas, distorcidas, incompletas e falhas que poderiam influenciar negativamente a pesquisa.

Zhang L. et al. (2019) abordam o reconhecimento das placas de licença de veículos em condições complexas, fundamental no monitoramento rodoviário. Relatam problemas encontrados como de sombras pela baixa iluminação, excesso de brilho, sujeira, distorções com inclinação vertical e horizontal, algumas obstruções pela neve ou chuva, além da dificuldade na coleta de placas de veículos em regiões diferentes no mesmo país. Para isso, empregam em sua estrutura o CycleGAN (*Cycle-Consistent Adversarial Networks*) de (Zhu et al. 2017), para obter placas de veículos com diferentes letras, números, sobre a influência de condições adversas, como de iluminação, clima, entre outras, além do Xception reformulado, derivado do original (Chollet 2017), na extração de atributos. Foi incorporado o YOLOv2 (*You Only Look Once*) de Redmon &

Farhadi (2017) com a finalidade de obter as caixas delimitadoras das placas de identificação veicular.

Ma et al. (2020) utilizam uma nova estrutura de classificação de imagens de células sanguíneas para o diagnóstico de doenças relacionadas ao sistema imunológico humano. Apresentam uma classificação de leucócitos utilizando uma rede adversária geradora de convoluções profundas DC-GAN (*Deep Convolutional Generative Adversarial Network*) de Radford et al. (2016) combinada com a rede residual de He et al. (2015), a ResNet (*Residual Network*). Ressaltam que amostras com dados insuficientes, não balanceados, causam influência no resultado. Para isso, introduzem o DC-GAN para sintetizar novas imagens de amostras. Também, utilizam a rede neural ResNet de 34 camadas, que supera a degradação com o aumento da profundidade, pois um canal de conexão direta, permite que uma fração da saída da camada anterior seja preservada e que informações sejam passadas diretamente para camadas posteriores.

Na abordagem de Thuy & Hoang (2020) é feita a extração de dados em imagem histopatológica e combinam aprendizagem profunda, aprendizagem de transferência e rede adversária, pois um método de diagnóstico, o qual permita a detecção precoce, passa a ser fundamental no tratamento do câncer de mama. Para este fim, incluem a rede neural convolucional, como técnica de extração de atributos. A abordagem de transferência de estilo, o StyleGAN (*Style Generative Adversarial Network*) de Karras et al. (2019) foi aplicada para gerar imagens benignas e malignas sintetizadas, que seriam variações das originais. Com a finalidade de incluir a textura cancerígena na imagem benigna sintetizada a rede Pix2pix (Isola et al. 2017) foi utilizada com a entrada condicional dessa textura. Desta forma, as novas imagens incorporaram o câncer durante o processo de aprendizagem. O trabalho demonstrou o uso de variados artifícios, com a rede GAN como gerador de imagens médicas e a rede VGGNet (*Visual Geometry Group Network*), especificamente a VGG16 e VGG19 como classificadores.

O aprimoramento dos parâmetros utilizados por Shankaranarayana et al. (2017), resultou em uma melhor triagem do glaucoma como prevenção na perda da visão. Neste sentido, exploram as redes totalmente convolucionais FCN (*Fully Convolutional Network*) em conjunto com a rede adversária, na perspectiva do aprimoramento nos resultados da segmentação. Alguns fatores importantes no diagnóstico de glaucoma devem ser levados em consideração, como a observação do CDR, em inglês *cup-to-disc ratio*, a relação do disco óptico e copo óptico, e para isso, existe a necessidade da segmentação de forma precisa. Assim, um codificador-decodificador totalmente convolucional baseado na U-Net de Ronneberger et al. (2015), incorporando conexões residuais do ResNet (He et al. 2015), foi utilizado no modelo de rede adversária GAN de Isola et al. (2017) com a tarefa de gerar representações plausíveis.

Os trabalhos de Feng et al. (2020) apresentam o potencial do aprendizado profundo na indústria de manufaturas, no aprimoramento do controle do processo de soldagem a arco de gás tungstênio, o GTAW (*Gas Tungsten Arc Welding*). O controle de qualidade com a inspeção da soldagem é fundamental para uma estrutura sólida e os defeitos seriam uma ameaça à sua integridade. Para isso, incluíram as redes neurais com arquiteturas diversas para classificar a qualidade da solda, com uma estrutura denominada *DeepWelding*, projetada com um repositório contendo o pré-processamento para recortar e remover o ruído da imagem de soldagem, o controle de qualidade que seleciona apenas as imagens melhores e por fim a classificação que acompanhará a qualidade final da solda. Na etapa de pré-processamento, o Pix2pix (Isola et al. 2017) foi empregado com o intuito de remover a reflexão da luz. As imagens originais contendo o ruído seriam as entradas e as sem ruído, seriam o alvo. Na seleção das imagens capturadas retirando as de baixa qualidade para não influenciar com resultados de pouca precisão, foi incluído um seletor com rede neural convolucional AlexNet de Krizhevsky et al. (2012). Na etapa de classificação, múltiplas redes neurais convolucionais com diversas arquiteturas foram integradas, compondo uma avaliação final. Entre elas estão a AlexNet com 5 camadas convolucionais e três camadas totalmente conectadas, a

VGGNet apresentada por Simonyan & Zisserman (2015) pela simplicidade e profundidade, a ResNet com conexões residuais nas camadas incluídas (He et al. 2015), também a DenseNet (*Dense Convolutional Network*) de Huang G. et al. (2017) com as saídas de todas camadas anteriores sendo usadas como entradas e suas próprias saídas alimentando as camadas posteriores e a CapsNet (Sabour et al. 2017; Hinton et al. 2018) pela robustez contra distorções geométricas. Com isso, temos uma estrutura *DeepWelding* que usa múltiplas fontes e tecnologias de aprendizagem profunda.

Zhang H. et al. (2019) apresentam novas técnicas para melhorar o monitoramento do processo de soldagem, na inspeção e controle de qualidade levando solidez e segurança estrutural em produtos metálicos. Exploram uma possível solução para identificar defeitos com uma abordagem em conjunto de imagens de raios-x. Para lidar com quantidades insuficientes visando um aumento da base de dados nos conjuntos desequilibrados é aplicada a WGAN (*Wasserstein GAN*) com trabalhos iniciados por Arjovsky & Bottou (2017a) e Arjovsky et al. (2017b) e aprimorados por Gulrajani et al. (2017), como um método confiável na falta de dados reais. Além disso, incluem a transferência de aprendizagem baseada na extração de recursos, que serão empregados no problema de detecção de defeito de solda. Deste modo, usam o Xception abordado por Chollet (2017) e MobileNet de Howard et al. (2017), duas redes sendo treinadas nos mesmos conjuntos de dados fazendo a classificação pelas pontuações de similaridade em seis classes.

Chugh & Jain (2020) atuam na transferência de características de textura entre imagens de ID, com a intenção de sintetizar com material desconhecido e adicionadas ao conjunto de dados para a CNN aprender os atributos e discriminar as imagens falsas com 12 materiais diferentes, utilizando o *Fingerprint Spoof Buster* de Chugh et al. (2018) e Slim-ResCNN (*Slim residual convolutional neural network*) de Zhang Y. et al. (2019). Os ataques de apresentação com a falsificação das IDs estão crescentes e podem ser realizados com uma variedade de processos e materiais. Uma das limitações na detecção dessas falsificações são os novos artefatos que não participam do treinamento. Como

solução empregam a proposta de (Huang X. & Belongie 2017) que implementam o AdaIN (*Adaptive Instance Norm*), o qual pode computar diretamente os parâmetros, a fim de gerar imagens com outras texturas similares à pesquisa de Johnson et al. (2016), que abordam a função de perda perceptiva, calculando a diferença da imagem de entrada com a imagem de saída, combinando diferentes características de estilo que são transferidas na reconstrução dos artefatos.

Wang et al. (2018) evidenciam melhorias com o aprendizado profundo no reconhecimento da impressão palmar, e para lidar com a falta de amostras em quantidades suficientes, aplicam uma rede baseada na DC-GAN de Radford et al. (2016) com modificações que incluem o SSIM (*Structure Similarity Index Measure*), o qual Wang et al. (2004) abordam como uma estrutura de similaridade e a função de ativação LeakyReLU (*Leaky Rectified Linear Unit*) utilizada por Maas et al. (2013) em todas as camadas convolucionais para gerar imagens satisfatórias e na função de reconhecimento, o Xception (Chollet 2017) devido a sua eficiência computacional aliada a precisão na classificação.

Avanços importantes foram apresentados por Tariq et al. (2018) para classificadores baseados em redes neurais e na detecção de rosto humano falso. Denotam que funcionalidades avançadas estão disponíveis nos aplicativos de edição de imagens, não necessitando de conhecimento específico. Para a finalidade, os autores desenvolveram uma arquitetura de rede neural convolucional, denominada ShallowNet (*Shallow Convolutional Network*), uma rede neural convolucional pouco profunda.

Tan et al. (2019) exploram a identificação de faces falsas que foram geradas por StyleGAN e DC-GAN, e para os experimentos são utilizadas três redes neurais convolucionais, a ResNet, VGG e GoogLeNet. O modelo ResNet-18 de He et al. (2015) é uma rede residual de 18 camadas que pode aprender a fração residual da camada anterior. Simonyan & Zisserman (2015) apresentam VGG-16, com camadas convolucionais que não necessitam reter muitos parâmetros em comparação com outros

modelos. Uma opção pelo Inception-v3 com Szegedy et al. (2015) é a terceira versão do GoogLeNet, que divide grandes convoluções bidimensionais em duas convoluções unidimensionais menores, para redução de parâmetros com eficiência. Na tarefa de criar rosto falso, o DC-GAN referido por Radford et al. (2016) contém no gerador e no discriminador camadas convolucionais que melhoram a qualidade das imagens geradas. Da mesma forma, o StyleGAN (Karras et al. 2019) adiciona informações de estilo ao gerador progressivamente.

A Tabela 1 apresenta uma visão geral da revisão de literatura sobre aplicação das redes neurais profundas no reconhecimento de padrões, contendo as referências, os assuntos abordados, a rede Geradora aplicada e o método de classificação utilizado.

Assim, salientamos o importante papel que as redes GANs exercem nos mais variados segmentos como: monitoramento, saúde, produção, identificação e outros.

Tabela 1. Resumo da pesquisa sobre redes neurais e redes GAN

Referências	Abordagens	Rede Geradora	Métodos de classificação
Chugh & Jain (2020)	Transferência de textura em impressão digital	GAN	Fingerprint Spoof Buster e Slim-ResCNN
Feng et al. (2020)	Inspeção de soldagem a arco de gás tungstênio.	Pix2pix	AlexNet, VGGNet, Resnet, DenseNet e CapsNet
Ma et al. (2020)	Classificação de leucócitos.	DC-GAN	ResNet
Shankaranarayana et al. (2017)	Triagem do glaucoma com aprimoramento da segmentação.	U-Net e Pix2pix	ResNet
Tan et al. (2019)	Identificação de faces geradas artificialmente.	DC-GAN e StyleGAN	ResNet, VGG e GoogLeNet
Tariq et al. (2018)	Deteção de rosto humano falso.	GAN	ShallowNet
Thuy & Hoang (2020)	Extração de dados em imagem histopatológica.	StyleGAN e Pix2pix	VGG16 e VGG19
Wang et al. (2018)	Reconhecimento de impressão palmar	DC-GAN	Xception
Zhang H. et al. (2019)	Controle de qualidade na soldagem de estruturas metálicas.	WGAN	MobileNet e Xception
Zhang L. et al. (2019)	Reconhecimento de placas de veículos em condições complexas.	CycleGAN	Xception

2.3. Pesquisa bibliográfica sobre IDs falsas e tecnologias propostas para sua detecção

São descritas nesta seção tecnologias empregadas na identificação de impressões digitais falsas e relatamos a complexidade em detectar simulações construídas para fins de ataques de apresentação aos sistemas biométricos.

Na identificação de materiais fabricados para falsificação de imagens de IDs, Rattani et al. (2015) alertam que os sistemas de verificação possuem certa vulnerabilidade quando sofrem ataques de identificação falsa (*spoofing attacks*), uma imitação da característica biométrica de outro indivíduo, e investigam descritores de textura para detecção do material no qual foi fabricado. Na extração de características dispõem do BGP (*Binary Gabor Patterns*), uma abordagem de (Zhang L. et al. 2012) para multi-resolução que demonstra eficiência para classificação de textura invariante de rotação, em escala de cinza.

Kannala & Rahtu (2012) utilizam BSIF (*Binary Statistical Image Features*) para obter as características estatísticas binárias de imagem, que codificam informações de textura de forma eficiente, muito pertinente nas representações baseadas em histogramas. Utilizam também o GLCM (*Gray Level Co-occurrence Matrix*) de Haralick et al. (1973), uma matriz de co-ocorrência em nível de cinza como método estatístico que examina a textura de uma imagem, a frequência de pares de pixels com valores específicos que ocorrem em uma imagem e o LBP (*Local Binary Pattern*), padrão binário local, proposto por Ojala et al. (2002), que compara cada pixel com a sua vizinhança. Tendo uma imagem de textura, cada localização pode ser atribuída com números no padrão binário.

Na possibilidade de apontar eficiência na detecção de biometria falsa, Kulkarni & Sankpal (2017) abordam a detecção ao vivo usando avaliação da qualidade da imagem. Relatam que as características das imagens reais, quando comparadas com as falsas, são diferentes e podem ser verificadas na cor, nos níveis de luminância, no grau de nitidez,

nas distorções estruturais, ocorrendo a falta de alguma propriedade que a original contém. No desenvolvimento adotaram um algoritmo em C++, que foi construído para os cálculos da qualidade de imagem, que faz a medida de índice de similaridade SSIM por influência de Liu et al. (2012), com a ênfase na similaridade do gradiente, na qual, aferem as alterações no contraste e informações estruturais da imagem.

Com a necessidade de avaliar a detecção com redes convolucionais e padrões binários nas impressões digitais apresentadas ao vivo, Nogueira et al. (2014) implementaram a extração de atributos com a técnica proposta por Chiachia (2014), que utiliza uma rede convolucional comum constituída por camadas alternadas de convolução e agrupamento local, com o intuito de extrair padrões que são frequentes nas imagens, e LBP, além de incluírem uma máquina de vetores de suporte nas duas redes, o SVM (*Support Vector Machine*) influenciados pelas abordagens de Cao et al. (2003), Lei & Govindaraju (2005) e de Krizhevsky et al. (2012) no aumento do conjunto de dados.

Com foco na imagem dorsal do dedo, Singh et al. (2019) usam câmera de campo de luz no cenário de ataques de apresentação com papel impresso, papel impresso embrulhado, digitalização e *smartphone*. Detalham uma arquitetura básica definida como FDSNet (*Finger Dorsal Spoof Detection Network*), a qual utiliza uma câmera Lytro na captura da direção dos raios de luz incidentes, permitindo a reconstrução da profundidade. Incorporam autocodificador com base na U-Net (Ronneberger et al. 2015) com a finalidade de extração de características, contendo um mapa de profundidade na imagem dorsal do dedo e para sua diferenciação empregaram a rede VGG-19 (Simonyan & Zisserman 2015), por preservar as informações de alto nível e os detalhes de baixo nível.

Patil & Dhole (2016) acreditam na qualidade da imagem para detectar a apresentação ao vivo da região analisada, recomendando o IQA (*Image Quality Assessment*) como um recurso na redução da vulnerabilidade nos sistemas biométricos,

aumentando o nível de segurança no enfrentamento de ataques fraudulentos por meio de mecanismos fabricados, amostras sintéticas ou reconstruídos. Aplicam o método proposto por Galbally et al. (2014), atuando sobre as diferenças de qualidade na comparação entre a imagem de referência e a distorcida, com 14 medidas indicadas como parâmetros na determinação da amostra real ou falsa de forma não invasiva. A implementação em Matlab R2013a foi para análise dos componentes principais na classificação, determinando os conjuntos de características discriminantes de qualidade do material de referência.

Os descritores de textura e redes neurais artificiais são utilizados por Zaghetto et al. (2017), na identificação de ataques de apresentação. Destacam que os descritores de textura extraem recursos que podem ser aplicados na diferenciação das imagens, como o LBP. No método foi aplicado o LBP aprimorado, sendo o ILBP (*Improved Local Binary Patterns*) de Hongliang et al. (2004), um padrão binário local melhorado e também a GLCM de Haralick et al. (1973). Na classificação, foi utilizada uma rede neural artificial semelhante à proposta por Demuth & Beale (1993), na qual elaboraram uma comparação de 200 imagens reais de 20 pessoas diferentes e as falsas com cera de abelha, massa, látex, silicone e cola.

No projeto de minimizar o impacto de elementos falsos no sistema de reconhecimento de IDs, Rattani & Ross (2014) incluem pré-processamento na melhoria da distinção entre materiais fabricados. Na análise de textura o LBP foi definido como avaliador de detecção ao vivo, com base em Nikam & Agarwal (2008), para a eliminação do ruído em cinco tipos de materiais diferentes na obtenção de particularidades complementares. Acompanharam os encaminhamentos do LiveDet (*Liveness Detection*) 2011 com Yambay et al. (2012) na extração de características, o SVM no treinamento com duas classes.

Khutlang & Nelwamondo (2014) realizaram a segmentação de papila aplicada na tomografia óptica SS-OCT (*Swept Source Optical Coherence Tomography*). O sistema

OCT foi utilizado para capturar a estrutura do dedo, obtendo o contorno papilar 3D que representa a biometria da ID interna. Em um trabalho específico de segmentação da ID, com a tomografia de coerência óptica de fonte varrida em imagens 3D, na etapa do pré-processamento as imagens são filtradas para minimizarem os ruídos com PDE (*Partial Differential Equations*), equações diferenciais parciais da teoria de Perona e Malik, modificadas por Catté et al. (1992) reduzindo os efeitos das manchas. Para detectar a camada mais externa da pele, o estrato córneo, cada fatia é filtrada usando limiares de imagem multinível, que Otsu (1979) explorou como um método não paramétrico e não supervisionado de seleção automática de limiar para segmentação de imagem. Na avaliação do contorno papilar utilizaram a distância de Hausdorff, empregando a técnica de Huttenlocher et al. (1993) na detecção de bordas e restauração de imagens. Também foi utilizado o MWI (*Modified Williams Index*), um índice proposto por Williams (1976) para comparar a concordância de um observador, com a concordância conjunta de outros observadores, originalmente definido apenas para dados nominais e que foi estendido por Chalana & Kim (1997) para dados multivariados numéricos.

Pala & Bhanu (2017) reformulam a rede clássica binária que envolve a verificação, utilizando aprendizado métrico profundo, na distinção da concepção entre as IDs reais e as falsas utilizando um mecanismo fabricado de falsificação. Entendem que a composição de três redes neurais convolucionais com pesos compartilhados pode beneficiar a obtenção de atributos, a fim de extrair uma distância intra-classe e uma distância inter-classe. Utilizam a rede derivada da concepção de Springenberg et al. (2015), sugerindo estruturas de redes simples contendo apenas camadas convolucionais, com a finalidade de descobrir os aspectos que as tornam originais e os atributos das falsas que são constituídas de vários materiais utilizados para fabricação.

Mirzaalian et al. (2019) conduzem uma verificação sobre a vantagem da imagem de contraste de laser, especificamente a LSCI (*Laser Speckle Contrast Imaging*) e as redes neurais convolucionais diante dos ataques de apresentação de IDs, denominado FPAD (*Fingerprint Presentation Attack Detection*). A LSCI possibilita obter a

circulação de sangue por baixo da pele, com a luz refletida pela iluminação do laser nas superfícies afetadas. Apoiam-se no estudo de Hussein et al. (2018) com diferentes arquiteturas baseadas em subimagens (*patches*), considerando pequenas partes das imagens nas convoluções, em uma estrutura modificada equivalente da AlexNet (Krizhevsky et al. 2012).

Uma combinação de informações que Silva et al. (2015) utilizam, poderia influenciar o algoritmo de reconhecimento ao vivo. São fatores que incluem a qualidade da imagem, a pressão do dedo durante aquisição de ID, algumas informações disponíveis baseadas em poros, aspectos que poderiam melhorar a classificação de recursos. Verificam as diferenças de intensidade de nível de cinza entre as IDs com o estudo de Abhyankar & Schuckers (2006). No encontro de medir a qualidade de imagem, foi adotado o NFIQ (*Nist Fingerprint Image Quality Measure*) de Tabassi et al. (2004), que obtém mapas de peculiaridades da ID, número e singularidade das minúcias. Recorreram, para mensurar o número de poros, à proposta de Zhao et al. (2010) que abordaram a modelagem e extração adaptativa de poros na ID. A frequência de poros é medida com abordagens de Derakhshani et al. (2003), que investigaram o padrão de transpiração sobre a pele da ponta dos dedos, uma concepção de que as IDs reais possuem uma maior frequência de poros, pois alguns materiais não conseguem reproduzir plenamente esta característica. Na classificação, Silva et al. (2015) apontam que na detecção das classes ao vivo e nas falsas foram usados outros métodos como, *Multilayer Perceptron*, *Optimum Path Forest*, *K-Nearest Neighbors*, mas relataram apenas o SVM utilizado por Hall et al. (2009), que obteve melhor desempenho com os ajustes indicados por Amancio et al. (2014).

Na Tabela 2 ilustra-se um resumo sobre as pesquisas que abrangem IDs falsas, suas referências, abordagens, os algoritmos ou suas técnicas e os métodos de classificação.

Tabela 2. Pesquisas envolvendo IDs falsas

Referências	Abordagens	Algoritmos / técnicas	Métodos de classificação
Khutlang & Nelwamondo (2014)	Atuam na segmentação de papila	PDE, limite de Otsu, distância de Hausdorff, MWI.	-
Kulkarni & Sankpal (2017)	Abordam a detecção ao vivo na qualidade da imagem	Gaussian Filter	SSIM
Mirzaalian et al. (2019)	Verificam o contraste a laser para obter a circulação do sangue	LSCI	AlexNet
Nogueira et al. (2014)	Detecção de ao vivo das impressões digitais	LBP	SVM
Pala & Bhanu (2017)	Os aspectos que tornam as digitais originais e as que foram fabricadas	Segmentação	CNN
Patil & Dhole (2016)	Atua sobre a qualidade da imagem baseado em 14 medidas	IQA	-
Rattani & Ross (2014)	Avaliam a detecção ao vivo no reconhecimento de impressão digital	LBP	SVM
Rattani et al. (2015)	Investigam descritores de textura	BGP, BSIF, GLCM, LBP	-
Silva et al. (2015)	Mensuram a qualidade da imagem baseados na intensidade de nível de cinza	Recursos estatísticos e NFIQ	SVM
Singh et al. (2019)	Identificação de ataque de apresentação com digital falsa	U-Net	VGG
Zaghetto et al. (2017)	Utilizam descritores de textura e redes neurais artificiais na identificação do dedo ao vivo	LBP, ILBP e GLCM	Rede neural

2.4. Pesquisa bibliográfica sobre imagens biométricas de forma geral e detecção de falsificações

Uma série de métodos abordados nesta seção, relatam pesquisas com a finalidade de verificar inconsistências nas imagens, localizar as regiões forjadas, quantificar essas adulterações, para melhorar o desempenho na distinção da imagem falsa com a original. O PAD (*Presentation Attack Detection*) de acordo com Raghavendra et al. (2015), é a

detecção de ataque de apresentação que visa enganar a identificação das características pessoais. Acontece quando alguém não autorizado apresenta para autenticação artefato construído, simulando usuário legítimo inscrito, para obter acesso a instalações, recursos ou informações. Esses artefatos são fabricados com materiais que falsificam os traços de um indivíduo que serão capturados pelo sistema de biometria.

Chen & Ross (2018) atuaram no reconhecimento da íris, cujo método pode obter atributos ópticos ou fisiológicos do olho acoplando hardware ou por algoritmos que verificam detalhes morfológicos e de textura das imagens. Utilizam a CNN de Redmon et al. (2016) e Redmon & Farhadi (2017), uma rede neural convolucional de referência que foi adaptada e aplicada para definir caixa delimitadora da imagem indicada e gerar pontuação sobre possível ataque de apresentação. Ressaltam que o método não faz a segmentação da íris, no qual este termo é amplamente usado para indicar a extração dos contornos internos e externos da íris. Implementam uma versão reduzida de CNN, baseada no YOLOv2 de Redmon & Farhadi (2017) para obter rapidez e precisão no reconhecimento, em que é gerada uma pontuação sobre a avaliação das classes verdadeiras e falsas, com as coordenadas da caixa delimitadora na região da íris.

Um crescente interesse da comunidade acadêmica nas pesquisas sobre as vulnerabilidades do reconhecimento facial é relatado por Raghavendra et al. (2015). Apresentam a introdução da câmera de campo de luz, o LFC (*Light Field Camera*) para registrar a direção da entrada de cada raio de luz e intensidade com múltipla profundidade. O ataque de apresentação é uma ameaça constante, realizada com artefatos de fácil confecção e que não necessita de conhecimentos específicos. Alguns ataques comuns aos sistemas de reconhecimento facial, usam foto impressa ou exibição na tela de dispositivos eletrônicos. Neste contexto, exploram na pesquisa características em relação ao efeito da luz, com uma câmera Lytro, obtendo informações de profundidade de toda a cena em várias imagens com diferentes regiões em foco. Consideram o algoritmo de Viola & Jones (2004) para a detecção da face por sua robustez, aplicam pré-processamento com filtro Gaussiano e 26 medidas para estimar

um valor quantitativo para imagem de campo de luz. Uma combinação da transformada de Gabor empregada por Wright et al. (2009), o classificador de representação esparsa com o SRC (*Sparse Representation Classifier*) de Berg et al. (2007) são aplicados para obter melhor pontuação. Também, com a justificativa de robustez, precisão e simplicidade no treinamento, é utilizado o SVM como opção de classificador.

Uma nova abordagem foi estudada por Li et al. (2020), para detectar falsificações, utilizando imagens de raios-x facial em tons de cinza. Relatam o sucesso em sintetizar imagens de rostos realistas e a dificuldade na distinção quando misturadas com as imagens reais, podendo essa técnica ser usada para fins maliciosos. Os raios-x fornecem uma maneira eficaz no reconhecimento de falsificações, para determinar se foi apresentado um rosto forjado ou real, pois algumas técnicas de manipulação de imagens faciais, quase não deixam vestígios das alterações e podem dificultar a identificação das modificações visuais. Para isso, adotaram a rede neural convolucional HRNet (*High-Resolution Net*) que mantém representação de alta resolução em todo o processo, anteriormente pesquisado por Sun et al. (2019) e o trabalho de Rössler et al. (2019) com o Xception (Chollet 2017), baseado em convoluções separáveis com conexões residuais, na perspectiva de um classificador de última geração.

Jayan & Aneesh (2018) abordam o reconhecimento de falsificações, o *spoofing*, com foco em fotografias nas redes sociais. Abrangem um algoritmo na extração de segmentos do rosto em diferentes espaços de cores, medindo os parâmetros de qualidade da imagem que resultam em vetores de atributos. Na etapa de identificação do rosto, utilizaram a proposta de Viola & Jones (2001) para conversão entre modelos de cores de RGB em HSV, YcbCr e LAB na obtenção de atributos de diferentes modelos. Na classificação foi empregado o QDA (*Quadratic Discriminant Analyzer*) aplicado por Bishop (2006), com grupos que representam a média e a variância de duas classes, usadas para estimar os valores pontuados na avaliação da imagem real ou falsa.

Com o aumento de extravios e adulterações, a detecção de documentos individuais chamou a atenção das autoridades do Sri Lanka, especificamente na carteira de habilitação, licença para dirigir, denominada SLDL (*Sri Lanka Driving License*). Nesse sentido, Samarasinghe et al. (2017) implementaram uma solução de baixo custo, voltada para necessidade específica do país. No pré-processamento, inicialmente a área do cartão que contém a licença para dirigir é determinada com o método ROI (*Region Of Interest*), que define a região de interesse na foto adquirida. Na avaliação da qualidade da imagem, não está disponível uma referência para comparação. Assim, foi utilizada a técnica NR (*No-reference*) de Mittal et al. (2012), sendo um modelo de avaliação de qualidade espacial sem referência. Para isso, aplicam filtros na obtenção de diferentes perspectivas que contribuem na melhor precisão. O primeiro é o algoritmo BRISQUE (*Blind/Reference-less Image Spatial Quality Evaluator*) referido por Mittal et al. (2011), que gera uma pontuação com valores próximos a zero ao indicar uma boa qualidade e foi utilizado na identificação de distorções na imagem de entrada, sendo muito adequado em aplicações de tempo real, uma vez que necessita de poucos recursos computacionais. Como segundo filtro, o algoritmo de Kurtosis aplicado por Caviedes & Oberti (2004), onde avaliam a qualidade de uma foto, se está legível e fornece valores mais altos ao indicar uma melhor nitidez e mais baixos quando são inferiores, pode ser usado como métrica de controle na captura de imagens que necessitam de alta qualidade.

Já para análise do holograma, considerado um recurso de segurança, os estudos de Lowe (2004) e Rublee et al. (2011) especificam que atributos são bastante distintos e a probabilidade de poder ter correspondência em um banco de dados com outras imagens é alta, confrontando cada amostra com a sua equivalente de referência. Na extração dos principais recursos foi recorrido ao histograma HOG (*Histogram Of Oriented Gradient*), sendo mensurado com SVM indicado por Hsu et al. (2003), para uma comparação aplicando um classificador de dados.

A Tabela 3 mostra um resumo sobre as pesquisas em imagens biométricas de forma geral e detecção de falsificações, suas referências, abordagens, os algoritmos ou suas técnicas e o métodos de classificação.

Tabela 3. Pesquisas sobre imagens biométricas de forma geral e detecção de falsificações

Referências	Abordagens	Algoritmos / técnicas	Métodos de classificação
Chen & Ross (2018)	Reconhecimento da íris	-	YOLOv2
Jayan & Aneesh (2018)	Mensuram os parâmetros de qualidade da imagem.	Algoritmo de Viola-Jones,	QDA
Li et al. (2020)	Incluíram a imagem de raios-x facial.	Escala de cinza	HRNet e Xception
Raghavendra et al. (2015)	Atuam no reconhecimento facial com a câmera de campo de luz, para registrar a intensidade com múltipla profundidade.	Algoritmo de Viola-Jones, filtro gaussiano, transformada de Gabor e SRC	SVM
Samarasinghe et al. (2017)	Detecção de falsificação na carteira de habilitação do Sri Lanka	BRISQUE, algoritmo de Kurtosis e histograma HOG	SVM

2.5. Considerações finais

Neste capítulo foram abordados os detalhes relativos à contextualização sobre o tema de pesquisa proposto e a pesquisa bibliográfica sobre esses assuntos. No próximo capítulo serão descritas as principais características das redes neurais utilizadas.

CAPÍTULO 3 – Estrutura das Redes Neurais utilizadas

Neste capítulo, são apresentados os detalhes das estruturas de redes neurais artificiais abordadas na proposta de pesquisa de dissertação. Na Seção 3.1 é apresentada a rede geradora adversária GAN; na Seção 3.2, a rede GAN condicional; na Seção 3.3, a rede Pix2pix e na Seção 3.4, rede Xception.

3.1. Rede adversária

Com o avanço na área de redes neurais artificiais, tornou-se possível a criação de imagens sintéticas de fontes reais, com qualidade realística, muitas vezes indistinguíveis pela percepção humana e podem ser utilizadas para fins diversos. GANs constituem uma classe de estruturas projetadas com duas redes neurais que se confrontam. A rede geradora (G) tenta gerar novas imagens a partir de um arranjo aleatório, e uma rede discriminatória (D) verifica se essas imagens são reais ou falsas. Após um certo número de iterações, a rede G é capaz de gerar imagens bem próximas das verdadeiras, a ponto da rede D não conseguir distinguir se as imagens geradas são falsas ou verdadeiras. A rede GAN pode aprender a gerar um conjunto de imagens sintéticas, as quais consideramos artificiais. Mesmo sendo amplamente utilizada no contexto de imagens, também é aplicada em áudio ou vídeo.

Exemplificamos a dinâmica da rede GAN fazendo uma analogia com um jogo de várias etapas, onde um oponente analisa o que outro fez de melhor, o momento que errou e quais estratégias serão necessárias para conseguir ganhar. É um conceito que pode ser incorporado como se fosse um jogo soma-zero, jogo não cooperativo da teoria dos jogos de (Nash, 1951), pois as pontuações de um determinado jogador anulam as do outro. Desta forma, leva a pontuação agregada ao jogo para o valor zero em ambos os jogadores.

O treinamento da rede adversária ocorre simultaneamente nos dois modelos, ambas as funções aprendem enquanto competem e o ruído aleatório serve como fonte que gera randomicamente diferentes informações sobre os dados, enquanto o gerador G atua para produzir imagens de aparência real na tentativa de enganar o discriminador D , que busca distinguir entre as imagens reais e as geradas.

Temos na Figura 3 os detalhes do modelo do treinamento da rede GAN e para interpretarmos mais adequadamente definimos $p_{\text{dados}}(x)$ sendo a distribuição dos dados considerados originais ou reais, x as amostras de $p_{\text{dados}}(x)$, $G(z)$ como a rede Geradora, $D(x)$ a Discriminadora, além de $p_z(z)$ a distribuição do gerador e por fim z as amostras de $p_z(z)$.

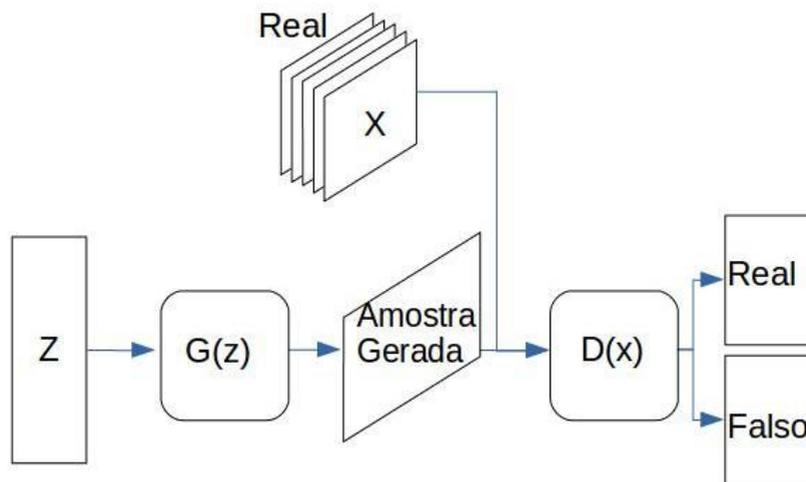


Figura 3. Estrutura da rede adversária, GAN - adaptado de Goodfellow et al. (2014).

Os modelos de redes GAN, capazes de aprender características para sintetizar amostras a partir de conjunto de dados, se iniciaram com a proposta de Goodfellow et al. (2014). Neste modelo o gerador $G(z)$, recebe z que é uma amostra do arranjo aleatório de entrada $p_z(z)$ e representa um mapeamento no espaço de dados, onde G é uma função

diferenciável representada por um perceptron multicamadas com parâmetros θ_g . Depois a rede discriminadora $D(x)$ é alimentada com os dados gerados ou pela entrada x de $p_{\text{dados}}(x)$ que são os reais, para prever se a entrada é real ou falsa.

Também, define-se um segundo perceptron multicamadas D , com parâmetros θ_d cuja saída é um valor escalar. $D(x)$ representa a probabilidade de que x tenha vindo dos dados reais, ao invés de p_g que são os gerados. Treina-se D , o discriminador para maximizar de forma que $\log D(x)$ seja próximo de 1, sendo reais e $D(G(z))$ esteja em 0, considerando como falsas, com a probabilidade de atribuir um rótulo correto para as amostras de G , e também para os exemplos de treinamento. Simultaneamente, treinamos G para minimizar $\log(1-D(G(z)))$.

Na equação 1, os parâmetros são ajustados para G minimizar $\log(1 - D(G(z)))$ e D maximizar $\log D(x)$, disputando entre si com treinamentos concomitantes. Em outras palavras, (Goodfellow et al. 2014) D e G se aplicam ao jogo minimax com a função de valor $V(G,D)$:

$$\min_G \max_D V(D, G) = E_{x \sim p_{\text{dados}}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log (1 - D(G(z)))] \quad (1)$$

O Algoritmo 1, detalha que $G(z)$ tem a intenção de convergir para uma estimacão adequada em $p_{\text{dados}}(x)$. Para otimizar D são estabelecidas k etapas a fim de não resultar em sobreajuste, enquanto G segue com apenas uma otimização.

Algoritmo 1. Treinamento da rede neural geradora G com minilote.

```

for número de iterações de treinamento
  for  $k$  etapas do
    - minilote de  $m$  amostras de ruído  $\{z^{(1)}, \dots, z^{(m)}\}$  do ruído anterior  $p_z(z)$ .
    - minilote de  $m$  exemplos  $\{x^{(1)}, \dots, x^{(m)}\}$  da distribuição geradora
      de dados  $p_{\text{dados}}(x)$ .
    - atualizar o discriminador aumentando o gradiente estocástico.
  
```

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m \left[\log D(x^{(i)}) + \log \left(1 - D(G(z^{(i)})) \right) \right]$$

end for

- minilote de m amostras de ruído $\{z^{(1)}, \dots, z^{(m)}\}$ do ruído anterior $p_g(z)$.
- **Atualizar** o gerador descendo o seu gradiente estocástico

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log \left(1 - D(G(z^{(i)})) \right)$$

end for

Espera-se que a rede aprenda as características específicas, a forma e os detalhes que são apresentados como entrada, para efetivamente produzir sínteses realísticas. É uma rede neural equivalente a um classificador binário que procura discriminar as imagens geradas enquanto a rede geradora melhora seu desempenho, competindo para produzir imagens cada vez melhores. Para isso, possui uma função de perda que atualiza os pesos da rede, alterna entre o gradiente ascendente no discriminador e o gradiente descendente no gerador, direcionando o aprimoramento da rede geradora. Goodfellow et al. (2014) sugerem para o treinamento uma quantidade significativa de imagens com suas variações intra-classe, devendo ter diversidade suficiente para modelar os dados. Além disso, indicam em trabalhos futuros a possibilidade de incluir um fator condicional que poderia ser adicionado ao gerador e ao discriminador.

3.2. Rede adversária condicional

Consideramos a GAN tradicional uma rede não supervisionada, na qual as imagens são geradas aleatoriamente acompanhando variações dentro da classe e necessitam de um grande volume de informações para o seu treinamento. Isso quer dizer que teremos resultados com referência ao conjunto de dados fornecido. Já a GAN condicional, é uma extensão da GAN tradicional, sendo incluída uma entrada que modifica a rede para um modelo supervisionado, definindo o alvo a ser alcançado. Ela pode gerar imagens com diversas características apenas indicando qual a referência

desejada, não sendo uma saída randômica e sim uma que temos expectativa, um interesse.

Mirza & Osindero (2014) e Gauthier (2014) apresentam experimentos com a cGAN, um modelo de rede adversária geradora condicional, na qual o gerador e o discriminador recebem informação adicional de entrada como condição. Nos modelos GANs que não possuem condição, os dados são gerados sem nenhum controle sobre eles. Já os GANs com informações adicionais podem direcionar o processo de geração de dados.

Na equação 2 (Mirza & Osindero 2014) , o gerador e o discriminador recebem a condição (y) como informação adicional de entrada:

$$\min_G \max_D V(D, G) = E_{x \sim p_{\text{dados}}(x)} [\log D(x|y)] + E_{z \sim p_z(z)} [\log (1 - D(G(z|y)))] \quad (2)$$

Com a finalidade de padronizar a definição, faremos uma associação das entradas e saídas. Utilizamos z como sendo um ruído de entrada ao Gerador, y como a informação incorporada para condicionar o Gerador e o Discriminador, além do x referindo-se aos dados de saída do Gerador e também aos de entrada do Discriminador.

Uma ilustração do Gerador e do Discriminador, em diagrama de blocos na Figura 4, detalha a estrutura da rede adversária condicional.

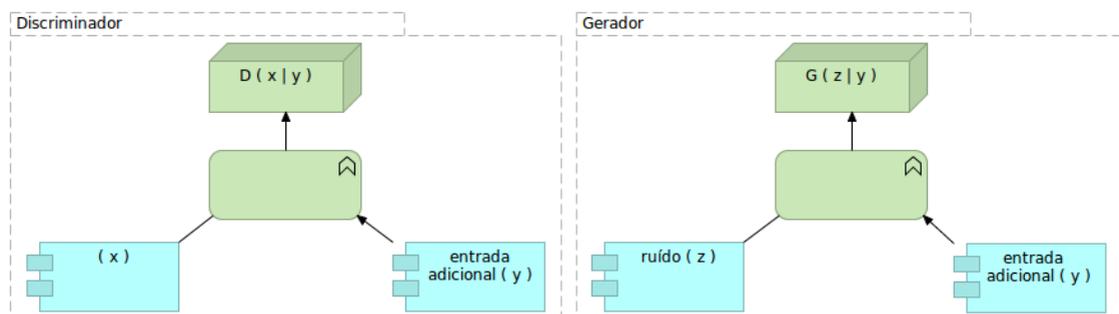


Figura 4. Estrutura da rede adversária condicional, cGAN - adaptado de Mirza & Osindero (2014).

Na entrada do discriminador é apresentada x como uma imagem sintetizada pelo gerador e y sendo a variável condicionante. No gerador, o ruído $p_z(z)$ e y são combinados na entrada da função geradora.

3.3. Pix2pix

A rede Pix2pix proposta por Isola et al. (2017), na tradução de imagem para imagem, utiliza o recurso de redes adversárias condicionais, as quais não apenas mapeiam a imagem de entrada para a imagem de saída, mas utilizam uma função de perda para treinar esse mapeamento e classificar se a imagem de saída é real ou falsa.

Pix2pix é uma rede que utiliza a classificação binária e calcula a perda em relação à imagem alvo. A função de perda procura mensurar a diferença entre a previsão da saída e a saída real desejada pelo treinamento, quando temos esses valores aumentados, estamos com uma menor precisão. Desta maneira, as duas redes, a Geradora e a Discriminadora são novamente realimentadas para ajustar os parâmetros e aprender com as perdas, que indicam as modificações necessárias para alcançar o objetivo. A cada iteração a rede G aprimora os resultados, se aproximando cada vez mais das reais e a rede D verifica se a amostra gerada é real ou falsa, indicando a necessidade ou não de seu aperfeiçoamento.

Sobre o mapeamento, consideramos que as camadas da rede neural convolucional estão em planos, denominados mapas de características, em inglês conhecido como *feature maps*. Assim, podemos entender que cada mapa de características, é equivalente à operação de convolução sobre o resultado da camada anterior.

O Pix2pix surgiu para evidenciar uma aplicação de uso geral das GANs e possui um gerador baseado na arquitetura U-Net, uma rede convolucional proposta por Ronneberger et al. (2015) e um discriminador (classificador convolucional) que utiliza

uma arquitetura PatchGAN, em que a discriminação ocorre por frações de imagens (*patch*).

A Figura 5 é uma ilustração da rede Pix2pix em diagrama de blocos. O Gerador G aprende um mapeamento do arranjo de ruído z para gerar imagem $G: \{x, z\} \rightarrow y$, e o Discriminador D classifica a imagem sintetizada y pelo Gerador G, utilizando um conjunto condicional de amostras x e determina a probabilidade de serem reais ou falsas com $D: \{x, y\} \rightarrow [0,1]$. A notação adotada por Isola et al. (2017), define y como sendo a imagem sintetizada pelo Gerador e determina x como a variável condicional.

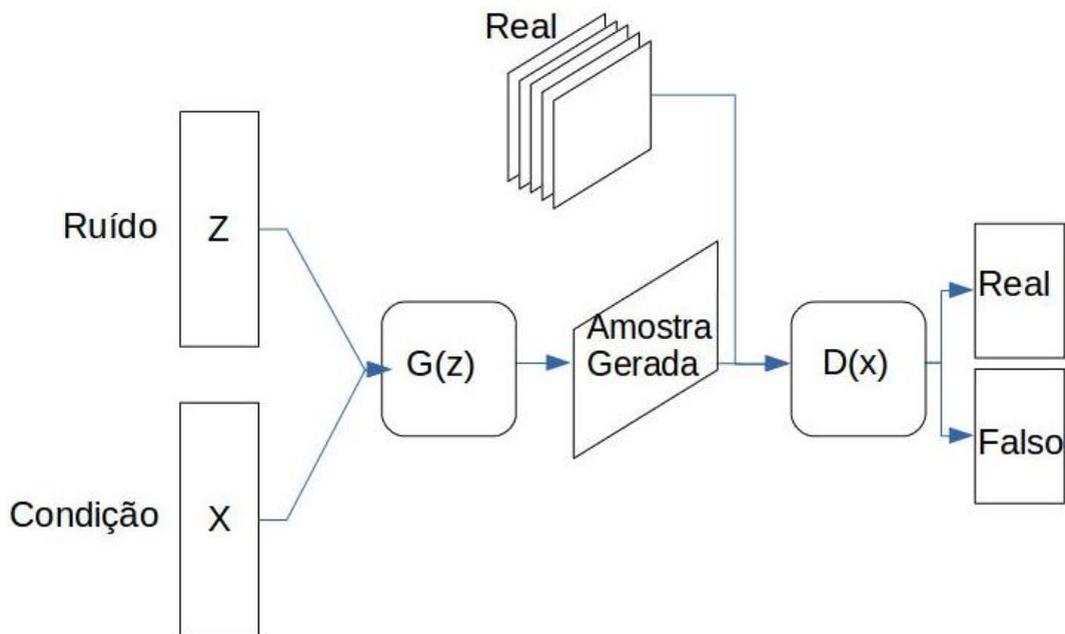


Figura 5. Diagrama de blocos da rede Pix2pix - adaptado de Isola et al. (2017).

A equação 3 (Isola et al. 2017) da função objetivo mostra que o gerador G tenta minimizar a diferença entre a imagem gerada y a partir do vetor de ruído z e a imagem x ; e o discriminador D tenta maximizar essa diferença,

$$L_{cGAN}(G,D) = E_{x,y}[\log D(x,y)] + E_{x,z}[\log(1 - D(x,G(x,z)))] \quad (3)$$

Na equação 4 (Isola et al. 2017), temos a minimização desse objetivo feita por G , contra um adversário D que tenta maximizá-lo:

$$G^* = \arg \min_G \max_D L_{cGAN}(G, D) = \lambda L_{L1}(G) \quad (4)$$

Desse modo, podemos entender o Pix2pix como um jogo de minimizador de G e maximizador de D , que acontece em etapas alternadas, com ajustes nos parâmetros na busca de otimização, até um resultado satisfatório.

3.3.1. Arquitetura dos módulos gerador e discriminador da rede Pix2pix

Arquitetura da rede geradora

A rede geradora tem sua estrutura baseada na U-Net (Ronneberger et al. 2015), sendo inteiramente convolucional e não tem camadas totalmente conectadas. Sua aparência tem um formato de "U" e apresenta uma certa simetria entre as partes. A U-Net é apropriada para associar as informações contextuais que foram obtidas nas camadas de contração, na primeira metade da rede, com os mapas de características e seus equivalentes adquiridos nas camadas de expansão, na segunda metade de rede.

As camadas de contração são uma sequência de filtros de convolução 3×3 , seguidas da função ReLU (*Rectified Linear Unit*) (Krizhevsky et al. 2012) e uma operação de *max-pooling* 2×2 com passo 2. Na camada de expansão, cada etapa consiste na amostragem do mapa de características seguido por uma convolução transposta 2×2 , *up-convolution*, uma concatenação com as camadas de contração correspondentes e duas convoluções 3×3 , cada uma seguida de uma ReLU.

Na última camada, há uma convolução 1×1 que mapeará os vetores de atributos de 64 elementos para o número desejado de classes. A Figura 6 mostra a rede geradora, com a correspondência entre as camadas de contração e expansão, de mesmas dimensões, por setas na parte inferior, da arquitetura U-Net. As matrizes passam por reduções progressivas, até um limite onde o processo é revertido. Desta forma o fluxo de informações transita por todas as camadas.

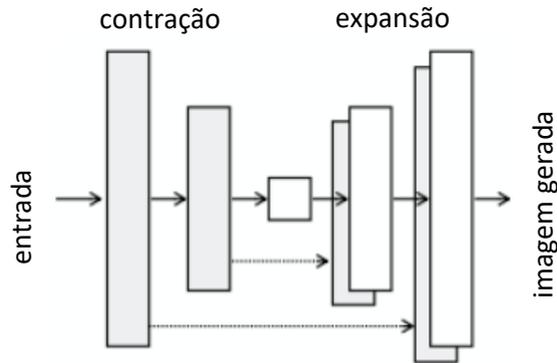


Figura 6. Arquitetura U-Net – adaptado Isola et al. (2017).

Deste modo, na tradução de imagens para imagens, existem informações que transitam entre as camadas de entrada e saída, com a necessidade de compartilhamento onde a transmissão direta, sem desvios, torna-se essencial. Assim, foi incorporado um gerador de saltos para reduzir os caminhos entre as camadas distantes.

A arquitetura da rede U-Net, detalhada anteriormente, possui semelhanças na proposta de Radford et al. (2016), tendo uma forma convolucional convencional na primeira parte, onde as dimensões das matrizes de atributos das camadas vão diminuindo. Porém, como descrito acima, existe uma segunda parte, que definimos como expansão, em que a estrutura evolui aumentando as dimensões dessas matrizes, até atingir a matriz final com as mesmas dimensões da imagem de entrada, que corresponde à imagem gerada.

A convolução transposta, que é a etapa de expansão, mantém a conectividade entre os valores de entrada com os valores de saída, para que os parâmetros possam ser aprendidos de maneira que aumente sua amostragem, seguindo um caminho oposto ao da convolução normal. Para ilustração descrevemos a convolução normal e a transposta em seguida.

Convolução convencional

Na exemplificação de Shibuya (2017), temos a operação de convolução matricial definida por uma entrada de 4x4, convolvida por kernel de 3x3, com passo 1 (*stride*) e sem preenchimento (*padding*), ou seja, teremos $i = 4$, $k = 3$, $s = 1$ e $p = 0$. A Figura 7 (a) adaptada de Dumoulin & Visin (2018), ilustra as variáveis de um kernel 3 x 3 e a Figura 7 (b), o mesmo kernel com os respectivos valores.

$w_{0,0}$	$w_{0,1}$	$w_{0,2}$
$w_{1,0}$	$w_{1,1}$	$w_{1,2}$
$w_{2,0}$	$w_{2,1}$	$w_{2,2}$

(a)

1	4	1
1	4	3
3	3	1

(b)

Figura 7. (a) representação de um kernel 3x3 e (b) kernel com os respectivos valores.

Para obtermos a matriz de convolução, devemos reestruturá-la para efetuar uma operação de convolução, conforme Figura 8. Como temos uma entrada 4x4 e kernel 3x3, faremos sua transmutação para uma matriz 4x16 preenchendo com zeros, os deslocamentos que o kernel deve realizar para compatibilização, com a posição onde a operação é aplicada. Em seguida, compomos a matriz C de convolução 4x16, considerando $w_{(ij)}$ os elementos do kernel, tendo (i) a representação da linha e (j) a coluna.

$$\begin{pmatrix} 1 & 4 & 1 & 0 & 1 & 4 & 3 & 0 & 3 & 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 & 1 & 4 & 3 & 0 & 3 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 4 & 1 & 0 & 1 & 4 & 3 & 0 & 3 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 4 & 1 & 0 & 1 & 4 & 3 & 0 & 3 & 1 \end{pmatrix}$$

Figura 8. Matriz C de convolução 4 x 16.

Para a operação de convolução, a matriz de entrada, Figura 9(a), deve ser transformada numa matriz coluna contendo todos os valores, Figura 9(b), em sequência de linhas. Neste contexto, para cada linha teremos uma operação de convolução, realizando a multiplicação da matriz de convolução pela matriz coluna de entrada. Como resultado da multiplicação temos uma matriz coluna 4x1, mostrada na Figura 9(c), que pode ser convertida para a matriz 2x2, Figura 9(d).

$$\begin{array}{cccc}
 \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & 1 & 2 & 3 \\ 4 & 0 & 1 & 2 \\ 2 & 2 & 2 & 3 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \\ 2 \\ 1 \\ 2 \\ 3 \\ 4 \\ 0 \\ 1 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 3 \end{pmatrix} & \begin{pmatrix} 29 \\ 32 \\ 29 \\ 37 \end{pmatrix} & \begin{pmatrix} 29 & 32 \\ 29 & 37 \end{pmatrix} \\
 \text{(a)} & \text{(b)} & \text{(c)} & \text{(d)}
 \end{array}$$

Figura 9. (a) matriz de entrada original, (b) matriz de entrada em coluna 16x1, (c) resultado da multiplicação em matriz coluna 4x1 e (d) resultado da multiplicação em matriz 2x2.

Convolução transposta

A função da convolução transposta é realizar a deconvolução, a partir do resultado de uma operação de convolução, por exemplo, de uma matriz de entrada 4x4, com um kernel 3x3, resultando em uma matriz de saída 2x2, para a obtenção da matriz de entrada 4x4. Para tanto utiliza-se da matriz de convolução transposta, C^T de 16x4,

para podermos multiplicar pelo vetor de tamanho 4x1, obtida pela transformação da matriz 2x2 do resultado da convolução, em uma matriz coluna 4x1. A Figura 10(a) ilustra a matriz de convolução transposta, a Figura 10(b), a matriz coluna 4x1, a Figura 10(c), a matriz coluna 16x1, resultante da multiplicação e a Figura 10(d), a matriz 4x4 resultante da deconvolução.

$$\begin{array}{ccc}
 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 4 & 1 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 4 & 1 & 4 & 1 \\ 3 & 4 & 1 & 4 \\ 0 & 3 & 0 & 1 \\ 3 & 0 & 1 & 0 \\ 3 & 3 & 4 & 1 \\ 1 & 3 & 3 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 2 \\ 1 \\ 4 \\ 4 \end{pmatrix} & \begin{pmatrix} 2 \\ 9 \\ 6 \\ 1 \\ 6 \\ 29 \\ 30 \\ 30 \\ 7 \\ 10 \\ 29 \\ 33 \\ 13 \\ 12 \\ 24 \\ 16 \\ 4 \end{pmatrix} & \begin{pmatrix} 2 & 9 & 6 & 1 \\ 6 & 29 & 30 & 7 \\ 10 & 29 & 33 & 13 \\ 12 & 24 & 16 & 4 \end{pmatrix} \\
 \text{(a)} & \text{(b)} & \text{(c)} & \text{(d)}
 \end{array}$$

Figura 10. (a) matriz de convolução transposta, (b) matriz coluna multiplicadora, (c) matriz 16x1, resultado da multiplicação e (d) resultado em matriz 4x4.

Assim, descrevemos a convolução transposta que é utilizada na arquitetura U-Net, que compõe a rede Pix2pix, com a finalidade de aumentar a amostragem da entrada, para a geração das imagens de saída. A seguir descrevemos a arquitetura do módulo discriminador da rede Pix2pix.

Arquitetura da rede discriminadora

Para a rede discriminadora, é utilizada a arquitetura PatchGAN que faz a discriminação na escala de frações da imagem (*patches*) determinando se essas frações são reais ou falsas. Para Isola et al. (2017), ao utilizar tamanhos grandes de frações de

imagens, o processamento é mais rápido e para otimizar o tempo de processamento é implementado o modelo proposto por Ioffe & Szegedy (2015), um minilote de treinamento (*minibatch*) na normalização de lotes, para conseguir altas taxas de aprendizado.

Como função de ativação é adotada a ReLU, uma unidade linear retificada de Nair & Hinton (2010) muito utilizada em redes neurais profundas, em que considera $f(x) = \max(0, x)$ um limite inferior para a saída, fazendo permanecer apenas os valores positivos da camada anterior.

Uma abordagem em rede profunda de Li & Wand (2016) com o discriminador PatchGAN sobre parte da imagem (*patch*), permite que a aprendizagem seja direcionada para o mesmo contexto com diferentes descrições. O Discriminador é treinado para distinguir as frações de imagens sintetizadas amostradas com as da imagem alvo, pontuando cada parte (*patch*) da imagem considerada como real com o valor $s = 1$, e $s = 0$, caso contrário. Podemos considerar a imagem de textura por $x_t \in R^{(w_t) \times (h_t)}$, a imagem sintetizada por $x \in R^{w \times h}$ e $\gamma(x)$ definido como um regularizador numérico. A iteração deconvolutiva para que x seja minimizada é descrita pelas equações 5 e 6 (Li & Wand 2016):

$$x = \underset{x}{\operatorname{argmin}} E_t(\Phi(x), \Phi(x_t)) + \alpha_1 E_c(\Phi(x), \Phi(x_c)) + \alpha_2 Y(x) \quad (5)$$

em que

$$E_t(\Phi(x), \Phi(x_t)) = \frac{1}{N} \sum_{i=1}^N \max(0, 1 - 1 \times s_i). \quad (6)$$

Na equação 6 temos E_t com a perda de textura e $\Phi(x)$ com a saída do mapa de características de x da camada ReLU na rede neural VGG19. Consideramos N o total de subimagens que foram amostradas em $\Phi(x)$. Já s_i possui a pontuação de classificação do i -ésima *patch* neural. A rede discriminadora inicializa seus parâmetros aleatoriamente, melhorando seu aprendizado constantemente com a atualização dos parâmetros após cada deconvolução.

3.4. Xception

Nesta seção será descrita a rede Xception (Chollet 2017), uma rede neural convolucional, selecionada para o trabalho de Dissertação de Mestrado, para ser aplicada na estimativa de similaridade entre as imagens de impressões digitais falsas e verdadeiras, pois demonstrou eficiência nas pesquisas estudadas e foi derivada de versões anteriores já conceituadas. Possui implementação de conexões residuais que resultou em robustez sem perder a precisão, com a possibilidade de mensurar a similaridade da imagem em uma escala de 0 até 1 diante da inferência sobre as diversas classes do treinamento. Essa rede teve a sua origem herdada de outras arquiteturas, que contribuíram em diversos aspectos como a VGG-16 (Simonyan & Zisserman 2015), que utiliza camadas sobrepostas com campo receptivo reduzido 3x3 e em algumas composições com filtros 1x1, conforme Szegedy et al. (2014) que exploraram o Inception-v1 na ampliação da rede com o aprofundamento das convoluções, e o Inception-v2 (Ioffe & Szegedy 2015) em que houve a normalização na camada de entrada no treinamento da rede. Já no Inception-v3 (Szegedy et al. 2015) implementaram a fatoração de convoluções na busca de eficiência computacional. Para o Inception-ResNet (Szegedy et al. 2016) verificaram as vantagens de incluir conexões residuais.

Para evitar a degradação durante o treinamento, com a perda de precisão, He et al. (2015) introduziram uma estrutura de aprendizagem residual profunda para cada camada empilhada. Uma versão modificada, mas mantendo os benefícios da abordagem original com conexão residual (*residual connections*), foi integrada na arquitetura do Xception.

3.4.1. Arquitetura da rede Xception

Nas redes neurais convolucionais, as camadas estão organizadas em três dimensões, as quais possuem largura, altura (ambas com referência nas dimensões espaciais) e a profundidade, fazendo o mapeamento ocorrer concomitantemente entre os canais e as correlações espaciais.

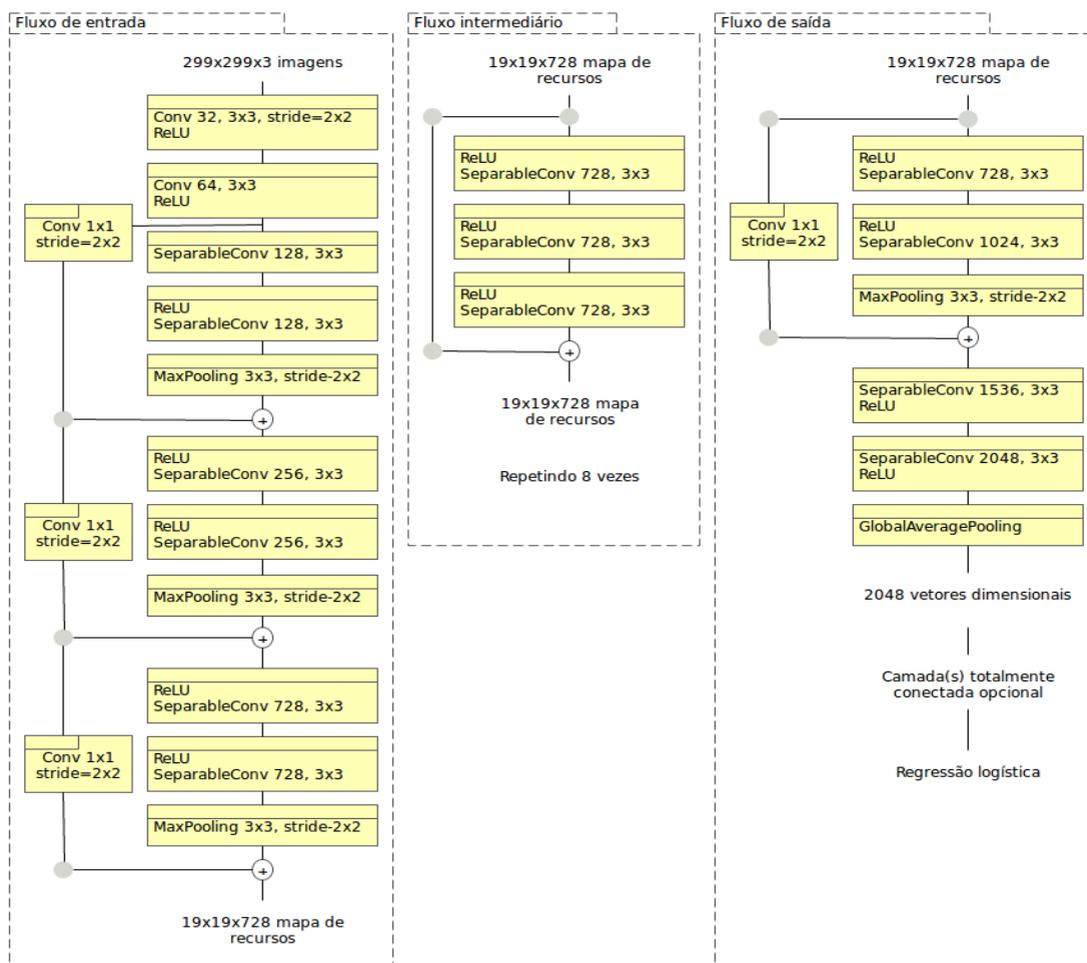


Figura 11. Arquitetura Xception – adaptado de Chollet (2017).

A vantagem da arquitetura da rede neural Xception com camadas de convoluções profundas separáveis, é realizar de forma dissociada a convolução 1x1 e a convolução espacial sobre os canais de entrada. A rede está estruturada com 36 camadas convolucionais para extração de características em 14 módulos. Com exceção do primeiro e último módulos, todos possuem conexões residuais lineares. A Figura 11 detalha a arquitetura, onde os dados transitam inicialmente pelo fluxo de entrada, depois são transferidos para o fluxo intermediário repetindo o ciclo por oito vezes e direcionados para o fluxo de saída.

No módulo Xception foram incluídas convoluções separáveis em profundidade (*depthwise separable convolutions*), uma convolução 1x1 para mapeamento das correlações entre os canais, atuando de forma dissociada para mapear as convoluções espaciais nos canais de saída. Foi inspirado nos trabalhos iniciados por Sifre & Mallat (2013) e Sifre & Mallat (2014). Posteriormente, abordagens semelhantes (Jin et al. 2015), (Wang et al. 2017) diminuíram cálculos redundantes desmembrando as convoluções e Howard et al. (2017) buscaram a eficiência de modelos móveis na computação com recursos reduzidos.

Desta forma, as modificações nos módulos com implementação de pilhas convolucionais separáveis em profundidade são mostradas na Figura 12, onde a imagem de entrada passa inicialmente por um número de convoluções 1x1, convoluções pontuais (*pointwise*), gerando o mesmo número de canais de saída (*output channels*) para posteriormente serem aplicados os respectivos kernels, no caso, 3x3. Os canais resultantes são então concatenados (*concat*) para a obtenção da saída da rede.

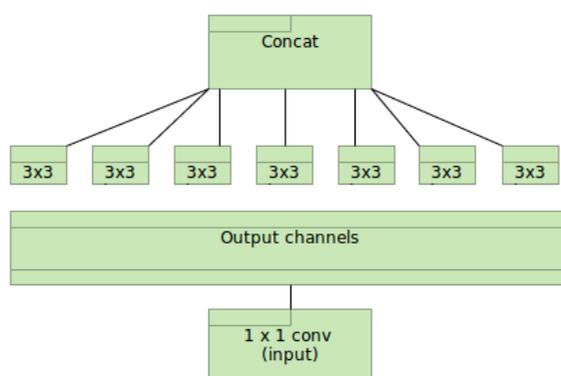


Figura 12. Ilustração da aplicação de convoluções separáveis em profundidade - adaptado de Chollet (2017).

A técnica de BN (*Batch Normalization*) de Ioffe & Szegedy (2015) é aplicada nas camadas convolucionais e convoluções separadas, seguindo uma normalização por lote que otimiza o treinamento.

A mudança de distribuição com o avanço do treinamento de cada camada, incide na camada posterior, necessitando adequar-se continuamente. Essa alternância da distribuição em rede profunda é o deslocamento covariável interno (*internal covariate shift*). Pela importância de equilibrar os parâmetros, foi proposta a normalização de lotes para estabilizar a rede, que resultará na estimativa média e variância de cada ativação das camadas.

Partindo do princípio de que a normalização é empregada individualmente a cada ativação, tem-se $x^{(k)} = x$ e um minilote β de tamanho m . Assim teríamos m valores de ativação no minilote, sendo $\beta = \{x_{1...m}\}$, levando em consideração os valores normalizados de $\hat{x}_{1...m}$ e suas transformações lineares $y_{1...m}$. Nos referindo a transformação da normalização de lote teríamos $\beta N_{\gamma,\beta}: \{x_{1...m}\} \rightarrow \{y_{1...m}\}$.

O Algoritmo 2 (Ioffe & Szegedy 2015) demonstra a dinâmica da transformação da normalização de lote.

Algoritmo 2 – Normalização de lote considerando ϵ uma constante.

Entrada: Valores de x em um minilote: $\beta = \{x_1 \dots m\}$;
 Parâmetros a serem aprendidos: γ, β ;

Resultado: $\{y_i = BN_{\gamma,\beta}(x_i)\}$

$\mu_\beta \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i)$	[média do minilote]
$\sigma_\beta^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_\beta)^2$	[variância do minilote]
$\hat{x}_i \leftarrow \frac{x_i - \mu_\beta}{\sqrt{\sigma_\beta^2 + \epsilon}}$	[normalização]
$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma,\beta}(x_i)$	[escala e deslocamento]

3.5. Considerações finais

Neste capítulo foram descritas as principais características das redes Pix2pix e Xception a serem utilizadas no trabalho de dissertação. A rede GAN condicional Pix2pix

deve ser utilizada para reproduzir imagens realísticas de impressões digitais para aumentar a quantidade de amostras no banco de dados de treinamentos e testes, e a rede Xception, que é uma rede que utiliza convoluções separáveis em profundidade (*depthwise separable convolution*) para melhoria do desempenho, deve ser utilizada no estudo da classificação das amostras reais e falsificadas. A seguir, serão descritos os bancos de dados utilizados e a metodologia.

CAPÍTULO 4 – Materiais e Métodos

Este capítulo refere-se à metodologia para o desenvolvimento do trabalho. Serão descritos inicialmente na Seção 4.1, os bancos de imagens utilizados. Além disso, abordaremos sobre a composição de uma base própria com outras amostras de IDs, que serão obtidas para o estudo de discriminação das imagens originais e réplicas. Nessa composição houve tentativa de utilização de dedos artificialmente construídos e a simulação com a alteração de texturas, prevalecendo essa última. Ao final, serão descritos os procedimentos para a detecção de similaridades entre IDs.

4.1. Bancos de dados utilizados

Para os experimentos foram utilizados três bancos de dados: a) banco de dados biométrico *Sokoto* (Shehu et al. 2018); b) banco de dados de textura (Cimpoi et al. 2014); e c) banco de dados de IDs próprio (*Banco de Dados Próprio* - BDP), obtidos com a colaboração de três indivíduos que permitiram a captura de imagens reais de impressões digitais.

4.1.1. Banco de dados biométrico Sokoto

Nos experimentos utilizamos o banco de dados biométrico *Sokoto Coventry Fingerprint Dataset* (Shehu et al. 2018), desenvolvido para fins acadêmicos, o qual é composto por 6000 imagens de impressões digitais de 600 pessoas. Possui rótulos por gênero, mão esquerda, direita e dedos. Para os ensaios foram utilizadas imagens da mão esquerda, na dimensão original de 96x103 pixels. Essas imagens foram ampliadas para 256x256 com o software GIMP (*GNU Image Manipulation Program*), para compatibilização com a rede e criados pares de imagens [z, x]. Conforme descrito no capítulo anterior, uma das configurações para a geração de novas imagens com a rede Pix2pix, é utilizar o formato de pares de imagens [z, x] onde z é uma imagem ruidosa aleatória e x é a imagem de referência.

4.1.2. Banco de dados de textura

Com a finalidade de obtermos informações representadas com texturas para fins de pesquisas, utilizamos neste trabalho de Dissertação um conjunto de dados com texturas descritíveis (Cimpoi et al. 2014), denominado DTD (*Describable Textures Dataset*), composto por 5640 imagens de objetos diversos, organizado com 47 categorias (classes) de acordo com a percepção humana, com 120 imagens por categoria, nos tamanhos de 300x300 e 640x640 pixels. Nos experimentos preliminares realizados, as classes *blotchy*, *bumpy*, *cracked*, *crosshatched* e *fibrous*, foram selecionadas devido a existência de nervuras semelhantes às existentes em IDs e redimensionamos as imagens para 256 x 256 pixels. A hipótese da pesquisa é a formação de texturas diversas nas IDs originais enquanto diferentes amostras são capturadas nas diversas instâncias dos procedimentos, diferenciando cada uma das imagens originais. Essas texturas se assemelham às classes de texturas preliminarmente selecionadas do banco DTD. Assim, seria possível obter réplicas de IDs acrescentando imagens dessas texturas às imagens de IDs originais.

4.1.3. Banco de dados próprio - BDP

Além do uso dos bancos de dados *Sokoto*, disponível na internet, foi desenvolvido, inicialmente um banco de dados próprio, BDP, com a colaboração de 3 indivíduos. Após a composição desse banco BDP, foram pesquisadas duas abordagens de serem obtidas réplicas dessas IDs. Uma abordagem com a construção de dedos artificiais utilizando diversos materiais, e uma outra abordagem utilizando reconstrução de IDs a partir das originais, com o incremento de texturas, denominado SIMTEX (*Simulação de textura*).

Para a obtenção desse banco de dados BDP foi desenvolvido um dispositivo de posicionamento dos dedos para a captura da imagem.

4.1.3.1. Dispositivo de posicionamento para captura biométrica

Um dispositivo de posicionamento do dedo, para a captura de imagem de ID utilizando câmera de um *smartphone*, permite uma padronização nos procedimentos de captura de imagem minimizando possíveis distorções. Assim, um artefato em acrílico de 1,0 mm de espessura, com dimensões internas de 130 mm de altura, 48 mm de largura e 48 mm de profundidade, com três abas de sustentação em papelão, reforçadas com fita adesiva flexível, foi construído, para ser acoplado na câmera traseira do celular utilizada como dispositivo eletrônico de captura, Figura 13.

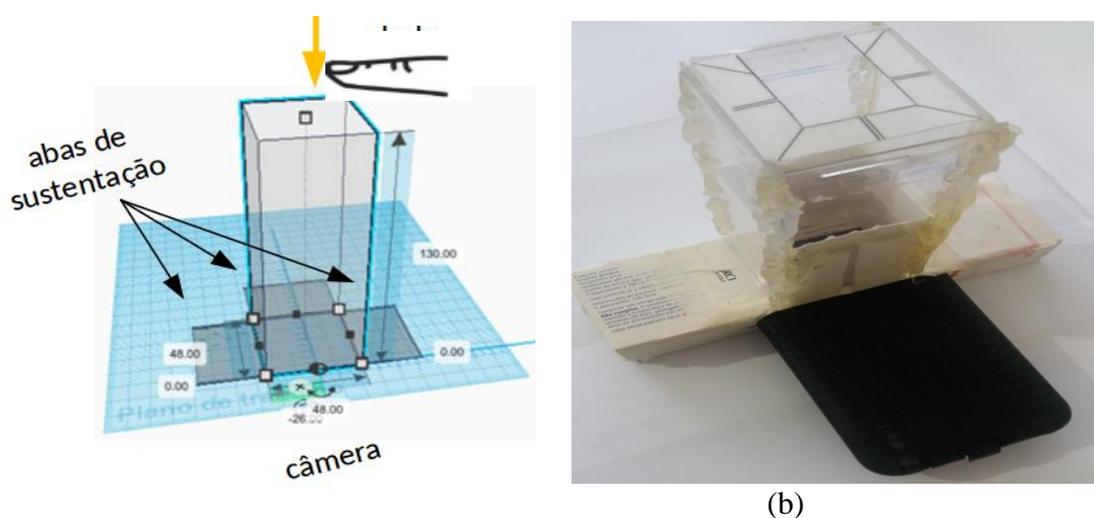


Figura 13. Dispositivo para posicionamento de captura biométrica: (a) projeto, (b) construído.

O banco de dados próprio foi construído nas duas etapas seguintes:

4.1.3.2. Obtenção das imagens de IDs originais

Na sequência temos a obtenção das imagens de impressões digitais originais do banco de dados próprio 1 (BDP1) e do banco de dados próprio 2 (BDP2).

A. Banco de dados de IDs (BDP1)

Primeiramente, é desenvolvida uma base de dados originais com dez imagens de cada dedo da mão esquerda de 3 indivíduos colaboradores, resultando em 150 imagens de 3096x3096 pixels. Sobre essas imagens, um pré-processamento denominado de seleção de região de interesse, ROI (*region of interest*), foi aplicado para a obtenção de imagens com dimensão de 1500x1500 pixels, posteriormente redimensionadas para 256x256 pixels utilizando o software GIMP (The GIMP Development Team, 2019), para se adequar ao algoritmo de processamento posterior. Essas imagens de 256x256 pixels constituem o banco de dados denominado de BDP1.

B. Banco de dados de IDs (BDP2)

Em um segundo momento, foi possível obter uma nova base de dados com 20 imagens de cada dedo da mão esquerda, dos mesmos três colaboradores, resultando em 300 imagens originais. Essas imagens foram coletadas aplicando o mesmo protocolo para cada imagem a partir do posicionamento do dedo no dispositivo de captura. Da mesma forma que no BDP1, nas imagens capturadas de 3096x3096 pixels, foram aplicadas o método ROI, para definir a região de interesse em 1500x1500 pixels e em seguida, redimensionadas para 256x256 pixels com o software GIMP (The GIMP Development Team, 2019).

4.1.3.3 Obtenção das imagens réplicas de impressões digitais

Conforme descrito no início desta seção, houve a tentativa de utilização de dedos artificialmente construídos e a simulação com a alteração de texturas, SIMTEX, para a obtenção de réplicas das imagens originais de IDs. A intenção de se construir um banco de dados de réplicas de IDs originais é em função do estudo realizado para a detecção de falsificações de impressões digitais, como objetivo do trabalho de dissertação.

O desenvolvimento das réplicas de IDs, a partir da construção de dedos artificiais, se inicia pela coleta das características da mão esquerda dos mesmos indivíduos colaboradores da etapa anterior. Para tanto, alguns elementos denominados

de materiais base foram utilizados: espátula plástica, cuba de silicone flexível, desmoldante, borracha de silicone líquida e o catalisador como reagente químico acelerando a secagem.

Ambas as IDs, originais e réplicas, devem ser obtidas utilizando o mesmo dispositivo de posicionamento do dedo, descrito na Seção 4.1.3.1, desenvolvido para essa finalidade.

A. Construção do dedo artificial

Definimos algumas expressões utilizadas na construção dos dedos artificiais, para um entendimento adequado dos procedimentos envolvidos. Desta forma teremos:

- Moldeira: base em material acrílico, que será preenchida com massa de *biscuit* para executar a moldagem;
- Moldagem: é o ato de moldar o dedo. A moldagem inclui o posicionamento correto, a preparação e manuseio dos materiais. O dedo será pressionado na moldeira para obtermos o molde;
- Molde: é a reprodução do dedo em negativo com seu formato, suas medidas e características correspondentes ao dedo original;
- Dedo artificial: é a réplica do dedo derivada do preenchimento do molde, com o material definido na construção do artefato;
- Desmoldante: substância que forma uma camada que impede a aderência entre os materiais, permitindo a desmoldagem.

Moldes

Os moldes, que formaram a matriz dos dedos artificiais, foram construídos em uma moldeira de acrílico de *design* circular, com raio de 40 mm de diâmetro, altura de 13 mm e 1 mm de espessura. Em cada moldeira foi acrescentada 30g de massa *biscuit*, e retirados os excessos com espátula plástica, deixando seu conteúdo uniforme para

moldagem. O procedimento de moldagem ocorre pressionando cada dedo sobre essa massa, durante cinco segundos, tempo necessário para a estabilização do material. A Figura 14(a) ilustra uma moldeira de acrílico, e a Figura 14(b), um molde de biscuit construído sobre a moldeira.

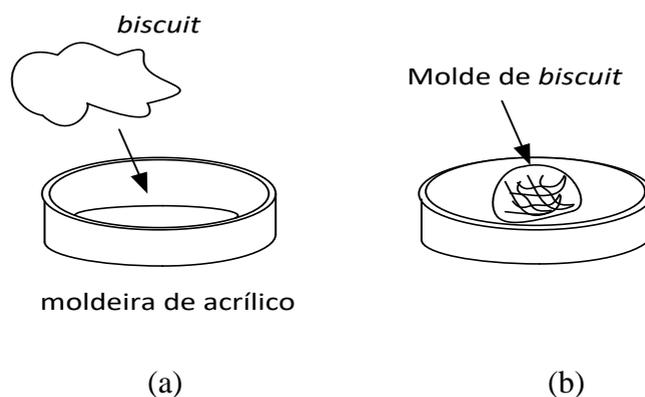


Figura 14. (a) moldeira de acrílico, (b) molde de *biscuit* construído sobre a moldeira.

O tempo de preparação da massa e a moldagem é de 10 a 15 minutos, mas sua secagem final é lenta, sem utilizar catalisadores, necessitando de pelo menos quinze dias.

B. Dedos artificiais e a obtenção de réplicas de IDs

O material escolhido na construção dos dedos artificiais foi a borracha de silicone líquida. Para essa etapa é importante aplicar o desmoldante sobre o molde antes de preenchê-lo com o material escolhido. Uma fina camada de desmoldante, deve ser aplicada com um pincel número 2, repetidamente a cada 10 minutos, para potencializar o efeito que impede a aderência entre os materiais.

Em seguida, é realizada uma mistura de material, borracha de silicone líquida com 4% de catalisador, por 20 segundos, e preenchidos lentamente os cinco moldes, um de cada dedo da mão esquerda. No ensaio realizado, a secagem em borracha de silicone foi de oito horas para a desmoldagem e 24 horas para a cura completa. A Figura 15

ilustra a construção de um dedo artificial, sendo (a) molde de *biscuit* com introdução de borracha de silicone líquida, e (b), dedo artificial resultante.

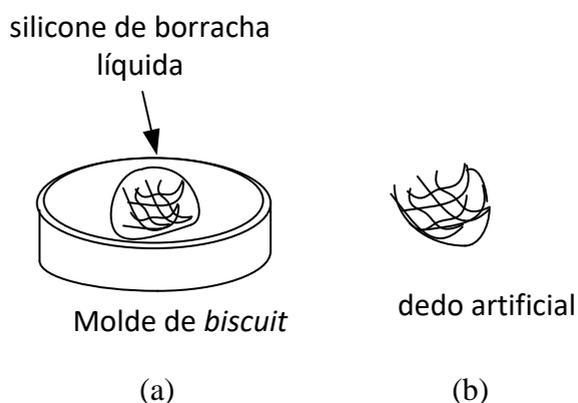


Figura 15. (a) molde de *biscuit*, (b) dedo artificial.

Após a construção dos dedos artificiais, são obtidas as respectivas amostras de IDs, para a composição do banco de dados próprio com réplicas de imagens de IDs.

Nota-se que apesar da intenção inicial da aplicação da abordagem do dedo artificial para a obtenção das imagens réplicas das IDs originais, não foi possível obter resultados significativos. Algumas construções de dedos artificiais mostraram que a captura de características de impressões digitais dos indivíduos não foi satisfatória para a construção de um banco de dados de réplicas de IDs. Em função disso, foi adotada a segunda abordagem para a obtenção dessas réplicas, a de simulação incorporando texturas sobre as IDs originais, abordagem SIMTEX. A descrição prossegue com a metodologia de pesquisa, que será detalhado até o final do capítulo.

4.2. Plano de Pesquisa e Metodologia

A seguir é descrito o plano de pesquisa para o trabalho de dissertação, acompanhado das metodologias adotadas. As etapas desenvolvidas podem ser resumidas nas atividades, cujo diagrama de bloco é mostrado na Figura 16.

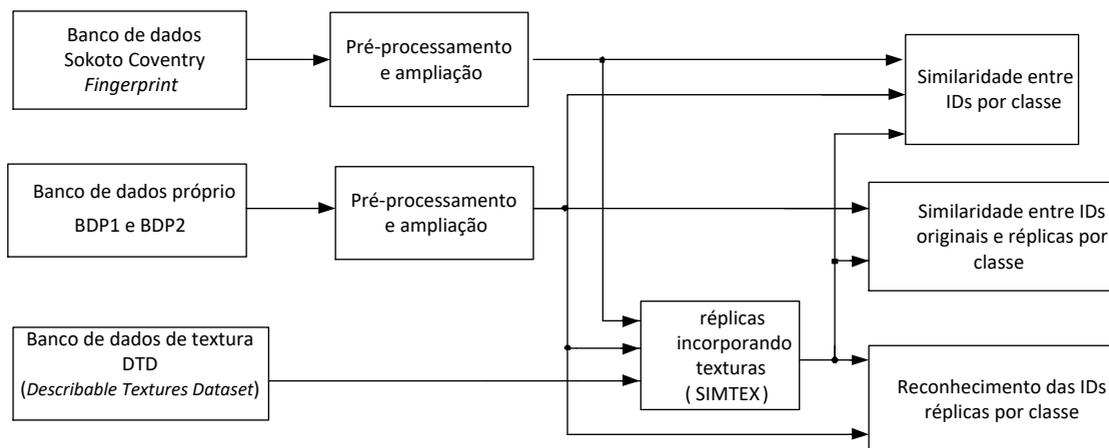


Figura 16. Diagrama de blocos – Plano de Pesquisa.

Conforme descrito na seção anterior, para o desenvolvimento desta pesquisa de dissertação faz uso de bancos de dados disponíveis, o *Sokoto Coventry*, o DTD e de bancos de dados próprio desenvolvido. O plano de pesquisa é constituído das etapas:

- 1) Obtenção das imagens dos bancos de dados *Sokoto Coventry* e DTD no sítio disponível;
- 2) Desenvolvimento de banco de dado próprio, com amostras de IDs originais;
- 3) Pré-processamento e ampliação dos dados utilizando Pix2pix, para o banco de dados *Sokoto Coventry*;
- 4) Pré-processamento e ampliação dos dados utilizando Pix2pix, para o banco de dados próprio;
- 5) Desenvolvimento de IDs réplicas incorporando texturas (SIMTEX);
- 6) Cálculo de similaridade entre IDs por classes;
- 7) Cálculo de similaridade entre IDs originais e réplicas por classes e
- 8) Reconhecimento de IDs réplicas por classes.

4.2.1. Pré-processamentos

Nesta seção são descritos os pré-processamentos realizados sobre as imagens de IDs consideradas. Na primeira etapa é descrita a obtenção da região de interesse, ROI (*Region of Interest*). Em seguida, descreve-se sobre o método de equalização de histograma CLAHE, e sobre o operador de Sobel para realce de bordas.

A. Obtenção da região de interesse ROI

Nesta etapa, temos uma avaliação visual da imagem, verificando os traços comuns das IDs, o formato e nitidez presentes em cada uma delas. Nos ensaios realizados, a partir da imagem de 3096x3096 pixels, foi obtida a região de interesse, ROI, de 1500x1500 na imagem capturada e depois redimensionado para o tamanho de 256x256 pixels com o software GIMP (The GIMP Development Team, 2019). Para fins ilustrativos, a Figura 17 mostra a limitação da região de interesse da imagem, ROI.

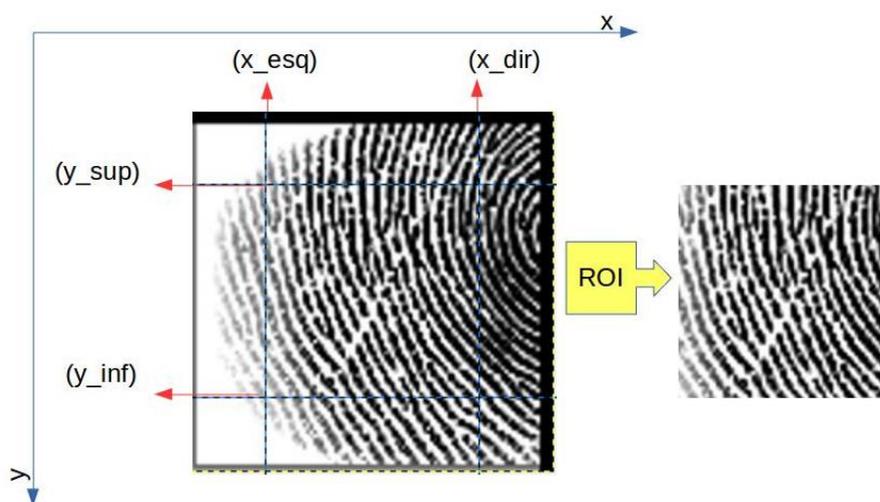


Figura 17. Definição da região de interesse, ROI.

Tendo a necessidade em definir de forma mais específica a área de impressão digital, selecionamos subimagens que serão destacadas exclusivamente para o treinamento e nos testes do experimento. Inicialmente utilizamos o GIMP (The GIMP

Development Team, 2019) como recurso de edição de imagens, mas o desenvolvimento de um algoritmo foi necessário para esta finalidade, a fim de obter melhor precisão e agilidade no procedimento. A exemplificação da rotina consta no Algoritmo 3.

Algoritmo 3 – Definição da região de interesse na imagem (ROI).

```
- Imagementrada
- Ler roi_larg # largura do ROI.
- Ler roi_alt # altura do ROI.
- (larg, alt) = obter tamanho (Imagementrada) # obter a largura, altura da
img
- x_esq = (larg/2 - roi_larg/2) # x_esquerdo
- y_sup = (alt/2 - roi_alt/2) # y_superior
- x_dir = (larg/2 + roi_larg/2) # x_direita
- y_inf = (alt/2 + roi_alt/2) # y_inferior
- Chamar função recorta (x_esq, y_sup, x_dir, y_inf) #função recorta img
- Redimensionar imagem.
- Armazenar o resultado na variável roi_img.
```

Depois incluímos algumas rotinas em Python com openCv e biblioteca PIL (*Python Imaging Library*) de Clark, A. (2015), para uma melhor confiabilidade nos procedimentos durante o pré-processamento.

Para uma melhor uniformidade entre as amostras de IDs, são aplicados: 1) método de equalização de histograma e 2) operador de Sobel, conforme descrições seguintes.

B. Equalização de histograma CLAHE

Foi utilizada uma abordagem para melhorar o detector de impressões digitais apresentadas, incorporando o método CLAHE, uma equalização de histograma adaptativo por contraste limitado, técnica aplicada por Zuiderveld (1994) e também em trabalhos de Pizer et al. (1987), para ajustar o nível de cinza, onde o contraste é limitado. Para isso, a imagem é dividida em subimagens e cada uma recebe o ajuste do histograma. Pressupõe-se o aprimoramento da imagem, seus detalhes, o contraste,

inclusive as bordas. Faremos uma comparação que será baseada na quantificação de possível melhoria, medindo a precisão de cada método. O Algoritmo 4, detalha a rotina de equalização de histograma.

Algoritmo 4 – Equalização de histograma adaptativo por contraste limitado.

- Imagem_{entrada} é convertida para escala de cinza, como imagem_{saída}.
- Ler imagem_{saída}.
- Chamar a função CLAHE.
- Armazenar o resultado na variável CLAHE_img1.
- Exibir ambos: imagem original e a CLAHE_img1.

A expressão do algoritmo com a função CLAHE, trabalha com uma distribuição uniforme, na qual o brilho será redistribuído para alterar o contraste e será transformada em uma imagem semelhante, para realçar os seus detalhes. Essa técnica de pré-processamento possibilita diferenciar os principais atributos de interesse nas imagens. Temos com isso, a intenção de potencializar a discriminação durante o processo de aprendizagem e posteriormente na classificação.

C. Realce de bordas com Operador Sobel

Estabelecer a borda no processamento de imagens é uma tarefa que pode ser desempenhada por operadores de borda baseados em gradiente, fornecendo informações importantes na análise de características fundamentais no campo de reconhecimento de imagens, mas existem diversas técnicas que podem ser aplicadas. Neste sentido, incluímos para fins experimentais a detecção de borda baseada no operador Sobel, iniciados por Sobel & Feldman (1968), na intenção de determinar as alterações nos níveis de intensidade, e encontrar esta transição, que possibilita descobrir as diferentes regiões da imagem que podem representar uma possível borda. Se considerarmos $f(x, y)$

a imagem de entrada, podemos calcular o gradiente conforme as equações 7 e 8. Onde, ∇_x é o gradiente na direção de x e ∇_y o gradiente na direção de y e denotado por ∇f na equação 9, a magnitude do gradiente:

$$\nabla_x = \frac{\partial f(x,y)}{\partial x} = \frac{f(x+dx,y)-f(x,y)}{dx} \quad (7)$$

$$\nabla_y = \frac{\partial f(x,y)}{\partial y} = \frac{f(x,y+dy)-f(x,y)}{dy} \quad (8)$$

$$\nabla f = \sqrt{\nabla_x^2 + \nabla_y^2} = [\nabla_x^2 + \nabla_y^2]^{\frac{1}{2}} \quad (9)$$

As implementações com o operador de Sobel determinam a estimativa do gradiente central da imagem, aplicando a convolução com duas máscaras 3x3 mostradas na Figura 18 com (a) componente-x e (b) componente-y,

$$\nabla_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad \nabla_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

(a) (b)

Figura 18. (a) componente-x e (b) componente-y – adaptado de Sobel & Feldman (1968).

O realce de bordas com operador de Sobel possui a vantagem de fornecer a estimativa com base no pixel central, (i, j). Os limites das máscaras de Sobel calculam a intensidade do gradiente da imagem, obtendo o aumento ou suavização da parte que provavelmente representa a borda. No processo de localização da borda, a procura está em realçar as informações consideradas relevantes, preservando a estrutura mais significativa.

4.2.2. Metodologia para ampliação utilizando Pix2pix

Descreve-se aqui a metodologia para a ampliação do banco de dados utilizando a rede Pix2pix. Essa metodologia inclui uma fase de treinamento e posterior fase de ampliação, propriamente dita.

Temos no treinamento o processo de aprendizagem para gerar imagens de IDs sintetizadas. Considera-se os detalhes da associação das entradas e saídas de acordo com Isola et al. (2017) no Pix2pix, utilizando z como sendo um vetor aleatório de entrada ao Gerador, e imagens x , de IDs originais, como condições para gerar as imagens sintéticas. A saída do gerador, imagem y , é inserida na entrada do Discriminador.

O Gerador deve ser otimizado para que seus resultados se aproximem cada vez mais da imagem alvo, durante a geração do conjunto sintetizado, de novas IDs similares às originais, reconhecíveis nas respectivas classes.

Temos na Figura 19, a ilustração da estrutura utilizada para a ampliação do banco de imagens de IDs, com imagens sintetizadas. No diagrama consta a aplicação da detecção de borda baseada no operador Sobel, ferramenta importante na adequação dos pares de imagens de entrada, ID original e vetor aleatório, durante o treinamento.

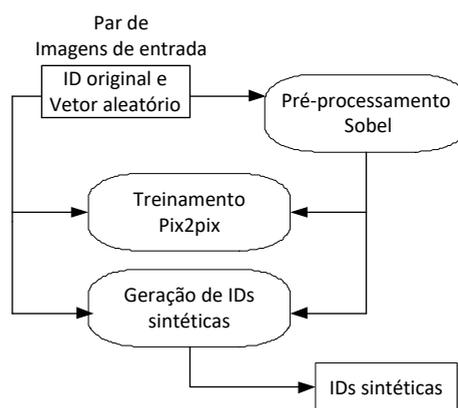


Figura 19. Diagrama ilustrativo do treinamento da rede Pix2pix e geração das IDs sintetizadas.

Seguimos as recomendações de (Kingma & Ba 2014) com o otimizador Adam (*Adaptive moment estimation*), conhecido como estimativa do momento adaptativo, que utiliza o método de otimização estocástica pela variação do gradiente. São valores muito utilizados na lr (*learning rate*) sendo uma taxa de aprendizagem inicial, representada por α (*alpha*), com os hiperparâmetros β_1 (*beta 1*), β_2 (*beta 2*) e ϵ (*épsilon*) para controlar a

taxa de decaimento exponencial, onde respectivamente temos $\alpha=0,001$, $\beta_1 =0,9$, $\beta_2=0,999$ e $\varepsilon=10^{-8}$. São valores que também foram incorporados na rede Pix2pix por Isola et al. (2017) com algumas adequações em $\alpha=0,0002$ e $\beta_1=0,5$.

4.2.3. Metodologia SIMTEX para a obtenção de réplicas de IDs

Como mencionado, não foi possível a abordagem da utilização do dedo artificial para obtenção de réplicas de IDs. Nesse sentido, foi adotada a alternativa de obtenção de réplicas de IDs, incorporando texturas sobre as imagens de IDs originais.

A abordagem desenvolvida denominada SIMTEX, associa uma imagem de textura do banco DTD, como a entrada z da rede Pix2pix, além de uma imagem x , que é uma ID original do banco de dados próprio, para a obtenção da imagem y pela rede G .

A partir da imagem de 3096x3096 pixels, foi aplicado o método ROI que define uma região de interesse de 1500x1500 pixels e depois redimensionada para o tamanho de 256x256 pixels com o software GIMP (The GIMP *Development Team*, 2019). Uma equalização adaptativa de histograma limitada pelo contraste foi utilizada para melhorar o reconhecimento de impressões digitais, incorporando o método CLAHE aplicado por (Zuiderveld 1994) e (Pizer et al. 1987), para ajustar o nível de cinza, onde o contraste é limitado.

A Figura 20 ilustra a estrutura da metodologia SIMTEX para a geração de IDs acrescidas de texturas. A imagem de ID original (x), mostrada no canto esquerdo inferior, e a imagem de textura (z), no canto esquerdo superior, consistem nas entradas do sistema SIMTEX. As duas imagens passam por um pré-processamento, obtenção da região de interesse ROI, seguido da equalização de histograma, antes da aplicação da rede Pix2pix para a geração de IDs com textura, na saída (y).

Denominamos “transferência de aprendizagem” a associação da textura às IDs originais, considerando-se que a rede Pix2pix aplica o modelo aprendido da textura (z),

para a imagem da ID original (x). Como resultados teremos sínteses de IDs com textura e assim, podemos também considerar que obtivemos uma transferência de estilo. Para a obtenção de resultados que conservem as originalidades das IDs, mantendo a sua reconhecibilidade, verifica-se a necessidade de controlar o treinamento da rede pela limitação do número de épocas de treinamento.

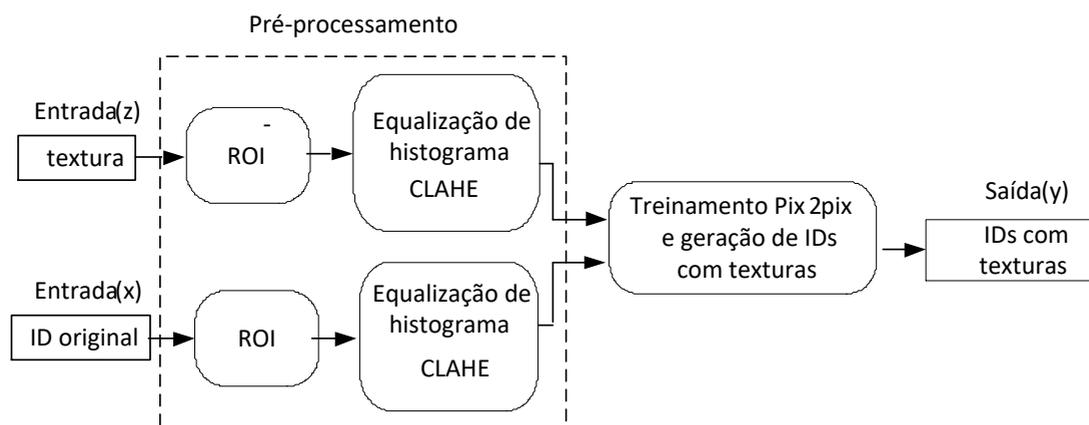


Figura 20. Estrutura da metodologia SIMTEX para geração de IDs com textura pela rede Pix2pix.

Teremos a obtenção de novas IDs, fazendo uma mistura de imagens de texturas sobre as imagens de IDs originais descritas. Essa forma de obtenção de amostras resulta em réplicas de IDs, como alternativa à abordagem de dedo artificial. Nessa formação da base de dados de impressões digitais com texturas, foi necessário nos ensaios optarmos pela incorporação do conjunto de dados DTD de Cimpoi et al. (2014), do banco de dados biométrico *Sokoto Coventry Fingerprint Dataset* (Shehu et al. 2018) e dos bancos de dados próprios, BDP1 e BDP2.

4.2.4. Similaridade entre IDs

A similaridade entre IDs deve ser verificada entre amostras de ID de dedos de um mesmo indivíduo, e entre uma amostra de ID original e uma respectiva réplica. Uma

das ferramentas para o cálculo da similaridade entre as amostras é a rede Xception. Os resultados de similaridade pela rede Xception são valores próximos a um, quando demonstrarem maiores semelhanças, e zero, caso contrário. Dentre outras ferramentas de cálculo de similaridade destacamos a métrica SSIM (*Structure Similarity Index Measure*) (Wang et al. 2004) e o cálculo pela distância Euclidiana, ou RMSE (*Root Mean Square Error*), descritos a seguir. Destacamos ainda, que mensuramos a proximidade das réplicas que foram geradas, quando comparadas com as originais.

4.2.4.1. SSIM

A métrica SSIM (Wang et al. 2004) avalia a similaridade entre as imagens em função do sistema visual humano ser sensível a mudanças nas estruturas locais, normalizadas pela luminância e pelo contraste. As estruturas da imagem de referência e a imagem modificada são comparadas para o cálculo de SSIM. A Figura 21 corresponde a um diagrama em que dadas as imagens original e estimada, os valores de luminância, contraste e estrutura são obtidos, comparados e combinados, para a obtenção da métrica SSIM.

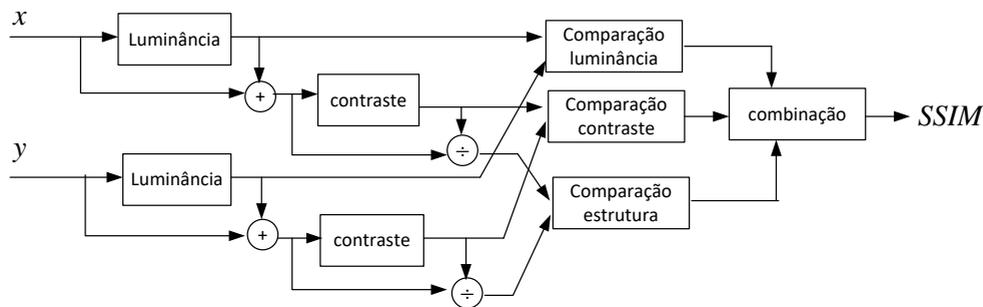


Figura 21. Diagrama ilustrativo do SSIM, adaptado de Wang et al. (2004).

As equações utilizadas para calcular as comparações de luminância (l), contraste (c) e estrutura (s) são apresentadas pelas equações 10, 11 e 12 (Wang et al. 2004),

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (10)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y+c_2}{\sigma_x^2+\sigma_y^2+c_2} \quad (11)$$

$$s(x, y) = \frac{\sigma_{xy}+c_3}{\sigma_x\sigma_y+c_3} \quad (12)$$

em que x e y são as imagens original e estimada, μ_x e μ_y são as médias de x e y , respectivamente. As variáveis σ_x e σ_y são as variâncias de x e y , respectivamente, σ_{xy} é a covariância entre as duas imagens, e c_1 , c_2 e c_3 são pequenas constantes obtidas com base no intervalo dos valores dos pixels em imagens, que para 8 bits é dado por [0,255]. Para calcular o índice SSIM com base na luminância, contraste e estrutura é utilizada a equação 13 (Wang et al. 2004),

$$SSIM(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (13)$$

em que α , β e γ são parâmetros que definem a ordem de importância dos três componentes. Assumindo-se que $\alpha = \beta = \gamma = 1$ e $c_3 = c_2/2$, tem-se a equação 14 (Wang et al. 2004):

$$SSIM(x, y) = \frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \quad (14)$$

O valor de SSIM varia entre -1 e 1, porém, são mais comuns valores entre 0 e 1. Quanto mais próximo de 1, mais similares são as imagens comparadas.

4.2.4.2. Distância Euclidiana

A distância Euclidiana ou Norma L2, pode ser utilizada para a obtenção da similaridade entre as imagens. É uma métrica muito utilizada quando precisamos representar a distância entre dois pontos em um espaço de n-dimensional.

Considerando a distância entre os pontos em n-dimensões tendo $a = (x_1, x_2, \dots, x_n)$ e $b = (y_1, y_2, \dots, y_n)$, podemos definir a distância Euclidiana entre os pontos a e b na equação 15 de acordo com (Gonzalez & Woods, 2010) e (Filho & Neto, 1999).

$$distância(a, b) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (15)$$

Temos representado na equação o cálculo da soma dos valores quadrados e finalizamos o cálculo obtendo a raiz quadrada, que nos fornecerá a distância mais curta do ponto a para o ponto b.

Para verificar a similaridade entre imagens de IDs, foi utilizada nos experimentos o cálculo da distância média pela raiz quadrada do erro médio quadrático, RMSE (*Root Mean Square Error*), considerando todos os pares de pixels das imagens. com os valores de pixels em nível de cinza no intervalo [0,255]. O resultado foi normalizado dividindo o RMSE por 255. A distância, ou erro normalizado, é dado $\frac{RMSE}{255}$. Como a similaridade é o complemento do erro, $similaridade = 1 - \frac{RMSE}{255}$.

4.2.5 Reconhecimento de IDs

O reconhecimento de IDs deve ser realizado para a verificação de quão reconhecível são as IDs obtidas pelos métodos descritos de construção de réplicas, com o método utilizando dedos artificiais ou método SIMTEX, em relação às IDs originais. Esse teste de reconhecimento é importante para a análise dos resultados obtidos com os métodos de replicação em estudo.

Como ferramenta de reconhecimento de réplicas de IDs a rede Xception é utilizada, sendo o treinamento realizado com amostras de IDs originais.

4.3. Avaliação do desempenho

Na verificação do desempenho, Prabhakar et al.(2003) empregam três métricas no sistema biométrico que podem ser utilizadas para avaliar sua precisão:

- FAR (*False Accept Rate*), uma medida da taxa de aceitação falsa. Isto seria a possibilidade de aceitação de um usuário impostor.

- FRR (*False Reject Rate*), a taxa de rejeição falsa, é um indicador ao recusar incorretamente um acesso válido.
- ERR (*Equal Error Rate*), uma taxa de erro igual que demonstra a FAR e FRR com igualdade. São indicadores de falhas no algoritmo e deve direcionar o ERR ao menor índice possível.

Na representação dos dados graficamente, podemos empregar a curva ROC (*Receiver Operating Characteristic*), uma curva característica de operação do receptor, na qual utilizamos o FAR contra FRR ou $(1 - FRR)$ com seus valores de limiar t . A medida FAR deve receber uma atenção significativa, pois reflete a possibilidade de impostores autorizados indevidamente. Para o presente trabalho de dissertação foi utilizada a taxa de verdadeiros positivos (*True Positive Rate*, TPR), no lugar de FRR para a construção da curva ROC. Desta forma, temos uma representação gráfica com a TPR quando o classificador acertou a predição e a FAR com a taxa de falsos positivos, considerando quando o classificador não acertou.

4.4. Considerações finais

Neste capítulo foi descrita a metodologia para o desenvolvimento do Trabalho de Dissertação. Os resultados obtidos serão apresentados no próximo capítulo, detalhando os experimentos abordados.

CAPÍTULO 5 – Experimentos e Resultados

Neste capítulo são descritos os experimentos realizados e os resultados. No Experimento 1, foi utilizada a rede Pix2pix para a geração de imagens sintéticas de IDs a partir de imagens originais. No Experimento 2, a geração de imagens sintéticas com o banco *Sokoto*. Já no Experimento 3, consta o ensaio relacionado à incorporação de textura com o banco de dados de BDP1 e no Experimento 4, incorporamos texturas às imagens de IDs, com o banco de dados biométrico Sokoto. No Experimento 5, foi associado textura no banco BDP2. As imagens sintéticas geradas pela rede Pix2pix foram comparadas com as imagens originais com a rede Xception, para a validação.

5.1. Experimento 1 - Geração de IDs sintéticas com Pix2pix a partir do banco BDP1

Primeiramente, teremos o processo de expansão do banco BDP1, para obtermos novas amostras sintéticas, para testar a metodologia Pix2pix, descrita na Seção 4.2.2.

Como descrito no Capítulo 4, o banco BDP1 contém um conjunto original de IDs de 3 indivíduos, com 10 imagens de cada dedo da mão esquerda, com o número total de 50 IDs por indivíduo. Identificamos as IDs originais dos indivíduos por P01, P02 e P03. Por outro lado, as IDs sintetizadas pelo Pix2pix foram denotadas por G01, G02 e G03. A Figura 22 ilustra a ID original e as imagens intermediárias de sintetização durante a aplicação da rede Pix2pix, temos a original e após 3, 14 e 168 épocas, da esquerda para a direita, respectivamente.



Figura 22. Imagens de IDs original e após 3, 14 e 168 épocas de aplicação da rede Pix2pix, da esquerda para a direita, respectivamente.

Medidas de similaridade das IDs originais com as sintetizadas

Aferimos os resultados das imagens sintetizadas pela rede Pix2pix em 200 épocas. No conjunto, constam 150 amostras sintetizadas. Para a utilização da rede Xception na verificação da similaridade, foi feito o treinamento da rede em 100 épocas. Além disso, foram incluídos os cálculos com a métrica SSIM e a RMSE. Temos na Tabela 4 a média e o desvio padrão, da similaridade das IDs originais, por classes, de IDs de pessoas representadas por P01, P02 e P03, comparadas com as IDs sintetizadas representadas por G01, G02 e G03. Devemos considerar uma maior similaridade nos resultados quando os valores estiverem próximos a 1 na métrica RMSE, no índice de similaridade estrutural SSIM e Xception.

Tabela 4. Medidas de similaridade das IDs originais com as respectivas IDs sintetizadas

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
P01xG01	0,9798	0,0105	0,9518	0,0815	0,9846	0,0171
P02xG02	0,9750	0,0094	0,9512	0,0182	0,9817	0,0139
P03xG03	0,9817	0,0035	0,9528	0,0073	0,7779	0,4146

Nota-se que para os dois primeiros conjuntos, a métrica Xception obteve melhores resultados (em negrito), seguida do RMSE e SSIM. Ao observarmos os valores estimados no conjunto P03xG03, temos uma redução da média para 0,7779 e o desvio padrão aumentou para 0,4146 na métrica Xception. Isso se explica pelo fato da métrica Xception adotar um cálculo da similaridade em função da reconhecibilidade. Quando uma amostra de ID for reconhecida como sendo de uma outra classe, essa amostra acaba distorcendo o valor médio de similaridade global entre as amostras. De uma forma ampla, altos índices de similaridade são alcançados com a métrica RMSE e com a métrica SSIM em todos os conjuntos aferidos. No cálculo da similaridade para as métricas RMSE e SSIM, os cálculos são feitos entre pares de IDs pré-estabelecidos, não havendo a distorção notada para Xception.

5.2. Experimento 2 - Geração de IDs sintéticas com Pix2pix a partir do banco Sokoto

No Experimento 2 foi utilizado o banco de dados biométrico *Sokoto Coventry Fingerprint Dataset* (Shehu et al. 2018), para a síntese de novas amostras utilizando Pix2pix, conforme Seção 4.2.2.

Neste experimento seguimos com alguns ajustes no pré-processamento para determinar a região de interesse, redefinindo o ROI aplicado para 150x150 pixels. Em seguida as imagens foram redimensionadas para 256x256 pixels, permanecendo a compatibilidade com a rede. Mantemos a detecção de borda baseada no operador Sobel, aplicada como artifício necessário na formação dos pares durante o treinamento com a rede Geradora.

Os conjuntos de IDs é relativo a 3 pessoas, com 10 imagens de cada dedo da mão esquerda, totalizando 50 IDs por indivíduo. O número de IDs relativo a apenas 3 pessoas, é devido ao interesse de compararmos os resultados com os obtidos com o BDP1 do Experimento 1, em que constam IDs de apenas 3 pessoas. Para identificarmos as IDs originais selecionadas do banco Sokoto, utilizamos as notações S01, S02 e S03, e para as imagens sintetizadas, SG1, SG2 e SG3.

Medidas de similaridade das IDs originais Sokoto com as sintetizadas

Foram obtidos a média e o desvio padrão, calculando-se a similaridade entre as IDs originais e as sintetizadas, conforme Tabela 5, em conjuntos separados por indivíduos. São mostrados, portanto, os resultados nos conjuntos S01xSG01, S02xSG02, e S03xSG03, para os três indivíduos, respectivamente.

Tabela 5. Proporção de similaridade dos conjuntos Sokoto por ID gerado

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
S01xSG01	0,9651	0,0007	0,9096	0,0227	0,9966	0,0066
S02xSG02	0,9664	0,0006	0,9281	0,0228	0,9910	0,0195
S03xSG03	0,9660	0,0008	0,9224	0,0204	0,9951	0,0099

Para a utilização da rede Xception, a mesma foi treinada em 100 épocas, como no Experimento 1. Em todos os conjuntos os valores estão próximos a 1 com desvio padrão pequeno. Considerando-se a similaridade com as três métricas, temos uma pequena vantagem com a Xception, seguida de RMSE e SSIM. diferença entre as médias, mas mantendo seus valores uniformes e desvio padrão próximo a zero. Quanto ao desvio padrão, existe uma vantagem para a métrica RMSE, cujo valor é mais próximo de zero.

Os valores que resultaram do Experimento 2, no treinamento em 100 épocas com a rede Xception, alcançou alta similaridade durante as medições, que foram refletidas da mesma forma quando comparadas com a métrica SSIM e RMSE.

5.3. Experimento 3 – Síntese de Réplicas utilizando SIMTEX para o banco BDP1

O Experimento 3 refere-se à síntese de réplicas de IDs originais, associando-se texturas nas amostras do banco de dados BDP1, utilizando SIMTEX, descrito na subseção 4.2.3.

Utilizamos o conjunto de dados com texturas descritíveis, DTD (Cimpoi et al. 2014), de onde foram escolhidas 03 imagens de texturas das classes *blotchy*, *bumpy*, *cracked*, *crosshatched* e *fibrous* do conjunto DTD.

Nos parâmetros para o treinamento da rede Pix2pix, foram necessárias algumas adequações. Percebemos que para cada classe de texturas possui imagens com características próprias, que necessita de parâmetros específicos. Neste contexto, utilizamos variações da taxa de aprendizagem (*learning rate*, *lr*) entre 0,0001 a 0,02, e constatamos que $lr=0,0002$ apresenta os melhores resultados para cada classe de textura.

A Figura 23 é uma ilustração de amostras de réplicas obtidas, com as classes de texturas selecionadas. Cada coluna corresponde a uma classe de textura, sendo *blotchy*, *bumpy*, *cracked*, *crosshatched* e *fibrous*, da esquerda para a direita. Em cada coluna, são mostrados de cima para baixo, as 3 imagens de textura, da respectiva classe, seguidas do resultado de treinamento, réplica, em 200 épocas.



Figura 23. Réplica obtida pelo SIMTEX, em 200 épocas, a partir de uma ID original, em 5 texturas *blotchy*, *bumpy*, *cracked*, *crosshatched* e *fibrous*, respectivamente, da esquerda para direita.

Foram obtidas 150 réplicas das IDs originais do banco BDP1, para cada uma das cinco classes de textura consideradas, e o número de épocas utilizadas para a síntese utilizando SIMTEX foi de 200 épocas.

Medidas de similaridade das réplicas de IDs com as originais BDP1

Na sequência, apresentamos as Tabelas 6, 7, 8, 9 e 10, a medida de similaridade entre as réplicas de IDs com as originais, nas métricas RMSE, SSIM e rede Xception.

Os resultados são da similaridade entre as IDs originais das pessoas 01, 02 e 03, do BDP1, com as réplicas utilizando-se diferentes texturas nas classes *blotchy* (*BL*),

bumpy(BU), *cracked*(CR), *crosshatched*(CS) e *fibrous*(FB), do conjunto DTD. Como exemplo, temos P01 como imagens de IDs originais da Pessoa 01, e BL01 as respectivas réplicas de IDs da Pessoa 01 com a textura BL, que representa *blotchy* na Tabela 6.

Tabela 6. Medidas de similaridade das réplicas com as IDs originais para a textura *blotchy*

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
P01xBL01	0,9678	0,0038	0,9741	0,0038	0,9996	0,0006
P02xBL02	0,9642	0,0023	0,9538	0,0231	0,9878	0,0271
P03xBL03	0,9675	0,0017	0,9640	0,0096	1,0000	0,0000

Os valores da média na Tabela 6 foram significativamente elevados em todos os conjuntos e métricas, estando próximos a 1 e com desvio padrão próximo a zero, indicando alta similaridade entre as réplicas com texturas sintetizadas, quando comparadas com as IDs originais. Os maiores valores de similaridade e os menores valores de desvio padrão estão em negrito. Nota-se que a métrica Xception obteve os maiores valores de similaridade e menores valores de desvio padrão, com exceção do desvio padrão na segunda linha, relativo a P02xBL02, em que o desvio padrão para o RMSE foi menor.

Na sequência temos a Tabela 7 com a textura *bumpy*. Os maiores valores de similaridade e os menores valores de desvio padrão estão em negrito.

Tabela 7. Medidas de similaridade das réplicas com as IDs originais para a textura *bumpy*

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
P01xBU01	0,9720	0,0024	0,9798	0,0034	0,9759	0,0530
P02xBU02	0,9684	0,0043	0,9580	0,0242	0,8939	0,1856
P03xBU03	0,9688	0,0015	0,9704	0,0095	0,9997	0,0004

Nos resultados da Tabela 7, temos uma uniformidade para as métricas RMSE e SSIM. Já no Xception, existe uma distorção na linha P02xBU02, tanto na similaridade como no desvio padrão, em função de um possível diferença no reconhecimento de uma ID, o que não acontece nas outras métricas, devido ao pré-estabelecimento dos pares de amostras de IDs comparadas.

Temos nos resultados da Tabela 8, para as comparações das IDs originais com as réplicas obtidas com a textura *cracked*. As métricas RMSE e SSIM apresentam valores próximos. Nota-se que como no caso anterior, a métrica Xception resultou em uma distorção nos resultados de similaridade e desvio padrão, na linha 2, quando os melhores resultados foram obtidos para a métrica RMSE. Nos outros casos, a rede Xception obteve melhores resultados.

Tabela 8. Medidas de similaridade das réplicas com as IDs originais para a textura *cracked*

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
P01xCR01	0,9612	0,0025	0,9708	0,0077	1,0000	0,0000
P02xCR02	0,9603	0,0013	0,9442	0,0337	0,8367	0,3488
P03xCR03	0,9613	0,0015	0,9570	0,0124	0,9998	0,0002

Os resultados na Tabela 9, mostram as medidas de similaridade entre as IDs originais em comparação com as réplicas obtidas com a textura *crosshatched*. Os resultados das medidas demonstram um equilíbrio entre as métricas, apresentando médias proporcionalmente distribuídas e com desvio padrão próximos a zero, demonstrando poucas variações.

Tabela 9. Medidas de similaridade das réplicas com as IDs originais para a textura *crosshatched*

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
P01xCS01	0,9748	0,0026	0,9635	0,0035	0,9468	0,1175
P02xCS02	0,9693	0,0025	0,9384	0,0298	0,9178	0,1247
P03xSS03	0,9724	0,0022	0,9523	0,0117	0,9738	0,0312

Por último, seguimos com a Tabela 10 com as medidas de similaridade entre as IDs originais e as réplicas com a textura *fibrous*. Novamente, nota-se na linha 2 uma distorção sobre os valores de similaridade e desvio padrão para a métrica Xception, caso em que a métrica RMSE obteve melhores resultados. Nas demais linhas a métrica Xception obteve melhores resultados de similaridade. Em relação ao desvio padrão, os resultados foram melhores para a métrica RMSE, para todas as linhas.

Tabela 10. Medidas de similaridade das réplicas com as IDs originais para a textura *fibrous*

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
P01xFB01	0,9609	0,0013	0,9524	0,0049	0,9940	0,0110
P02xFB02	0,9596	0,0015	0,9251	0,0222	0,7025	0,4017
P03xFB03	0,9613	0,0014	0,9398	0,0117	0,9932	0,0146

5.4. Experimento 4 – Síntese de Réplicas utilizando SIMTEX para o banco Sokoto

O Experimento 4 refere-se à construção de réplicas de IDs, utilizando SIMTEX, para amostras originais do banco Sokoto.

O experimento com o banco de dados Sokoto foi realizado utilizando SIMTEX, como no caso anterior. Como no Experimento 3, foram utilizadas as amostras de textura do banco DTD, para as classes *blotchy*, *bumpy*, *cracked*, *crosshatched* e *fibrous*.

Foram obtidas 150 réplicas das IDs originais do banco Sokoto, para cada uma das cinco classes de textura consideradas, e o número de épocas utilizadas para a síntese utilizando SIMTEX foi de 200 épocas.

A Figura 24 é uma ilustração do resultado de construção das réplicas, para diferentes classes de texturas. Na parte superior da Figura 24 contém as descrições das classes, seguidas logo abaixo das 3 imagens utilizadas de cada classe. Na parte inferior, constam exemplos de resultados das réplicas de IDs geradas do indivíduo em 200 épocas.

Medidas de similaridade das réplicas de IDs com as originais Sokoto

Os valores de similaridade entre as IDs originais do banco Sokoto e as réplicas construídas com o SIMTEX, constam nas Tabelas 11, 12, 13, 14 e 15. Como foram utilizadas as mesmas classes de texturas do Experimento 3, *blotchy (BL)*, *bumpy(BU)*,

cracked(CR), *crosshatched(CS)* e *fibrous(FB)*, do banco DTD, as notações dos conjuntos são similares.



Figura 24. Textura incorporada nos conjuntos por pessoa – Sokoto.

Consideramos as três pessoas, cujas IDs foram utilizadas, representadas por S01, S02 e S03. Então, na primeira coluna da Tabela 11 temos o conjunto S01xSBL01, correspondente às IDs da pessoa 1, e as respectivas réplicas utilizando a textura *BL*, *blotchy*.

Na Tabela 11 constata-se melhores valores de similaridade para a métrica Xception, nas duas primeiras linhas, porém, na terceira, ocorre distorção, o mesmo

ocorrendo para o valor do desvio padrão que fica com um valor alto. Nesse caso, a métrica RMSE obtém melhor resultado para a similaridade. Quanto ao desvio padrão, os melhores resultados, próximos de zero, são obtidos em todos os casos pela métrica RMSE. Nota-se também que os valores de similaridade para a métrica SSIM foram baixos em relação à métrica RMSE.

Tabela 11. Medidas de similaridade das réplicas com as IDs originais para a textura *blotchy*

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
S01xSBL01	0,9593	0,0002	0,7457	0,0355	0,9692	0,0660
S02xSBL02	0,9593	0,0002	0,7737	0,0264	0,9722	0,0419
S03xSBL03	0,9584	0,0005	0,8232	0,0339	0,8150	0,3899

A Tabela 12 apresenta os valores de similaridade obtidos comparando as IDs originais do banco Sokoto com as réplicas de IDs com a textura *bumpy*. Nota-se que os valores de similaridade são maiores para a métrica RMSE, com valores baixos de desvio padrão. No caso das métricas Xception e SSIM, os resultados de similaridade foram inferiores.

Tabela 12. Medidas de similaridade das réplicas com as IDs originais para a textura *bumpy*

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
S01xSBU01	0,9596	0,0006	0,7236	0,0325	0,8431	0,3121
S02xSBU02	0,9602	0,0005	0,7403	0,0236	0,7478	0,3062
S03xSBU03	0,9592	0,0007	0,8029	0,0270	0,8647	0,2662

Os resultados da Tabela 13, referentes às medidas de similaridade entre as IDs originais da Sokoto comparadas com as réplicas de IDs obtidas com a textura *cracked*, demonstram uma redução das médias e a elevação significativa dos desvios padrão, para a métrica Xception. Os melhores valores ficam para a métrica RMSE. Esses resultados são muito parecidos com os da Tabela 12, anterior.

Tabela 13. Medidas de similaridade das réplicas com as IDs originais para a textura *cracked*

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
S01xSCR01	0,9583	0,0006	0,7381	0,0386	0,4101	0,2958
S02xSCR02	0,9585	0,0005	0,7702	0,0274	0,5679	0,4732
S03xSCR03	0,9585	0,0009	0,8030	0,0316	0,4173	0,3488

Na Tabela 14 são mostradas as similaridades entre as medidas de IDs originais do banco Sokoto com as réplicas utilizando a textura *crosshatched*. Como nos casos anteriores, a métrica RMSE apresenta melhores resultados, tanto da similaridade como do desvio padrão. Mas, quando nos atentamos aos valores aferidos na métrica SSIM e no Xception, há uma discrepância nos grupos, resultando em uma queda mais expressiva na média. No caso do SSIM o desvio padrão é baixo, porém, para o Xception o desvio padrão é maior.

Tabela 14. Medidas de similaridade das réplicas com as IDs originais para a textura *crosshatched*

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
S01xSCS01	0,9588	0,0003	0,7502	0,0396	0,8069	0,1522
S02xSCS02	0,9588	0,0004	0,7815	0,0296	0,6899	0,3184
S03xSCS03	0,9587	0,0003	0,8162	0,0345	0,8301	0,1712

Por último, encontram-se na Tabela 15 os resultados de similaridade entre as IDs originais do banco Sokoto e as réplicas utilizando a textura *fibrous*.

Tabela 15. Medidas de similaridade das réplicas com as IDs originais para a textura *fibrous*

Conjuntos	RMSE		SSIM		Xception	
	Média	Desvio	Média	Desvio	Média	Desvio
S01xSFB01	0,9596	0,0006	0,7303	0,0409	0,9446	0,0836
S02xSFB02	0,9593	0,0002	0,7588	0,0313	0,7536	0,3230
S03xSFB03	0,9599	0,0004	0,8104	0,0342	0,9027	0,1889

Como nos casos anteriores, a métrica RMSE obteve os melhores resultados, mostrados em negrito. Observamos uma grande estabilidade nos valores aferidos com as métricas RMSE, desvio padrão próximo a zero e uma média sem oscilações.

5.5. Experimento 5 – Síntese de Réplicas utilizando SIMTEX para o banco BDP2

No Experimento 5 é utilizado o banco BDP2, constituído de 300 imagens de IDs de 3 indivíduos, sendo coletadas 20 imagens de cada dedo da mão esquerda, conforme descrito no Capítulo 4, Seção 4.1.3.2.

São realizados os experimentos: Experimento 5.1, de construção de réplicas utilizando SIMTEX, Experimento 5.2, de reconhecimento dessas réplicas com a rede Xception, treinado para reconhecer as IDs originais do BDP2, além do Experimento 5.3, de cálculo de similaridades entre as IDs originais e as réplicas.

Na subseção 5.5.1 é descrito o Experimento 5.1, de construção de réplicas, para o banco de dados BDP2. Já na subseção 5.5.2, é descrito o Experimento 5.2, no qual as réplicas de IDs obtidas, no Experimento 5.1, passam pelo reconhecimento, utilizando a rede Xception, e posteriormente, na subseção 5.5.3, é descrito o Experimento 5.3, no qual são calculadas as similaridades entre as imagens originais e as geradas pelo SIMTEX, utilizando RMSE, SSIM e a rede Xception, para mensurar a semelhança entre as imagens que foram geradas com textura, quando comparadas com as originais.

5.5.1. Experimento 5.1 – Construção de réplicas de IDs do banco BDP2, associando textura com SIMTEX

Neste experimento foram utilizadas as mesmas classes de textura aplicadas nos Experimentos 3 e 4, com exceção da textura *cracked*, pois obteve baixa similaridade no Experimento 4, do banco DTD. Foram utilizadas, portanto, as classes *blotchy*, *bumpy*, *crosshatched* e *fibrous* e o banco de dados de BDP2.

Na Figura 25, estão algumas amostras que resultaram do experimento.

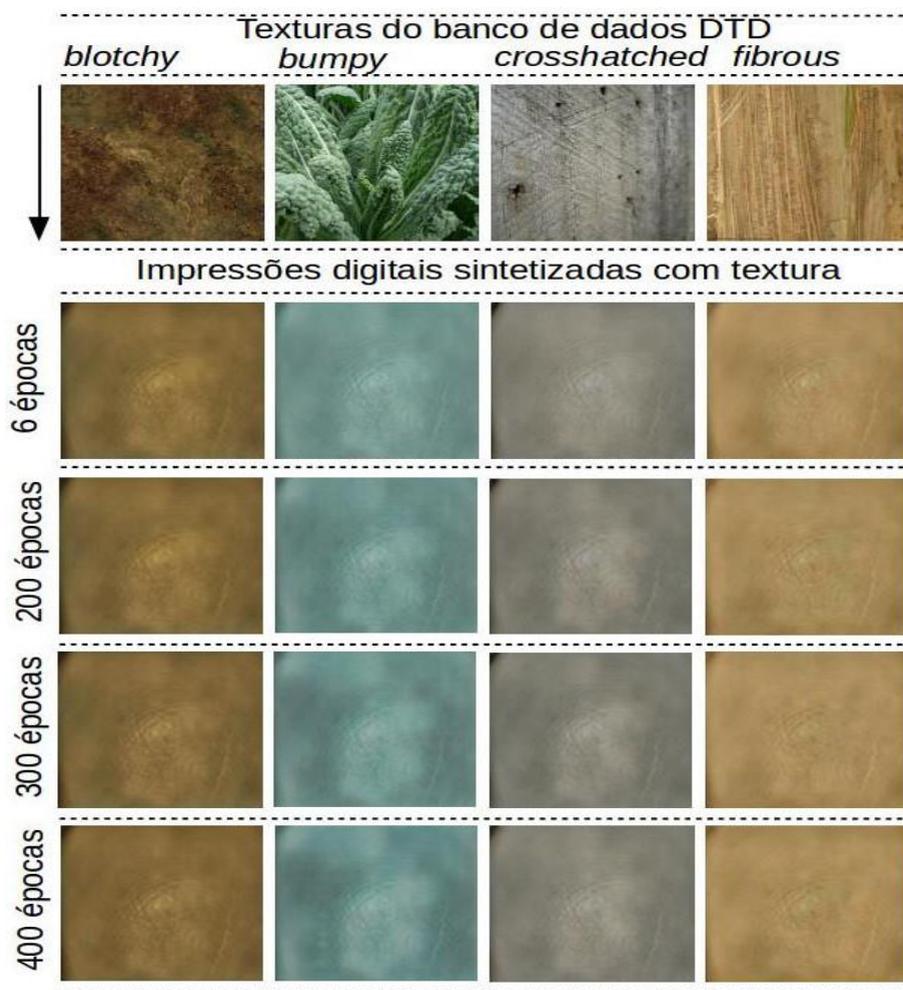


Figura 25. Impressões digitais que foram sintetizadas em 6, 200, 300 e 400 épocas com textura.

Na parte superior da Figura 25, na primeira linha de imagens, são mostradas as 4 imagens de texturas obtidas do banco DTD, da esquerda para a direita, *blotchy*, *bumpy*, *crosshatched* e *fibrous*. Em seguida, são mostradas a partir da segunda linha, as imagens de IDs com texturas obtidas, variando nas colunas as texturas. O número de épocas utilizadas para treinamento, foram 6, 200, 300 e 400 épocas, e as respectivas réplicas obtidas, são mostradas nas linhas 2, 3, 4 e 5, respectivamente.

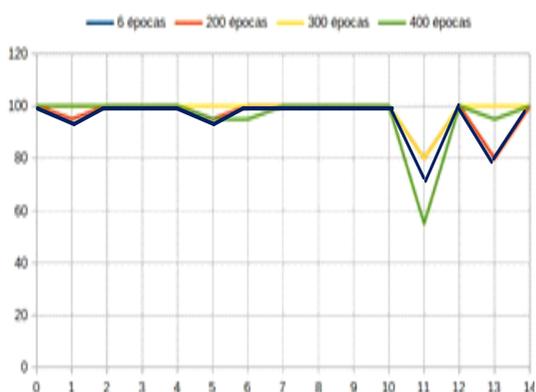
O Experimento 5.1, teve o intuito de verificar a influência do número de épocas para a construção de réplicas utilizando SIMTEX, e a sua interferência na reconhecibilidade. Nesse sentido, foram obtidos conjuntos de réplicas de IDs, com 6, 200, 300 e 400 épocas de iteração do SIMTEX, para cada uma das classes de textura consideradas. O número de 100 réplicas de IDs foi obtido para cada indivíduo, sendo 20 correspondentes a cada dedo da mão esquerda, para cada um dos quatro casos de treinamento, variando o número de épocas, descritos no parágrafo anterior.

5.5.2. Experimento 5.2 - Testes de reconhecimento das réplicas de IDs obtidas no Experimento 5.1

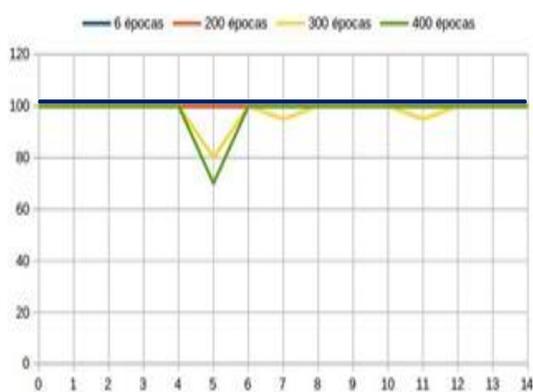
No Experimento 5.2 é realizado o reconhecimento das réplicas de IDs obtidas no Experimento 5.1, utilizando o treinamento da rede Xception, conforme descrito na Seção 4.2.5, com as amostras originais do BDP2. A intenção de se aplicar o teste de reconhecimento é a validação da metodologia SIMTEX, pois se as réplicas de IDs construídas são reconhecíveis, significa que a abordagem é viável. Mesmo que a similaridade entre a réplica e a ID original seja alta, pelas métricas descritas, a réplica pode não ser reconhecível. Como descrito na Seção 4.3, quando se aplica um sistema de reconhecimento de ID, pode resultar em erros de reconhecimento. Para efeitos de análise dos resultados do reconhecimento foi convencionado preliminarmente, neste trabalho, que o reconhecimento de 80% para cada classe seria razoável para validar a construção de uma réplica.

Na Figura 26 é mostrado o resultado dos testes de reconhecimento, de IDs com texturas sendo: (a) *blotchy*, (b) *bumpy*, (c) *crosshatched* e (d) *fibrous*. Os testes de reconhecimento foram realizados, treinando a rede Xception com as amostras de IDs originais dos 5 dedos da mão esquerda de 3 indivíduos, contido no banco de dados de BDP2 e aplicando os testes de reconhecimento para as réplicas geradas pelo SIMTEX, no Experimento 5.1. A coordenada horizontal corresponde às 15 classes de IDs, identificadas de 0 a 14, sendo 0 a 4, dos dedos do indivíduo 1, de 5 a 9, do indivíduo 2 e

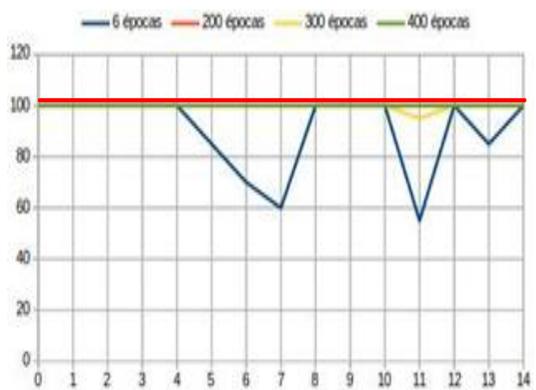
de 10 a 14, do indivíduo 3. A coordenada vertical corresponde às porcentagens de reconhecimento por classe. Para cada quantidade de épocas de treinamento aplicadas no SIMTEX, é mostrado um gráfico de porcentagem de reconhecimento, sendo cor azul (6 épocas), vermelha (200 épocas), amarela (300 épocas) e verde (400 épocas).



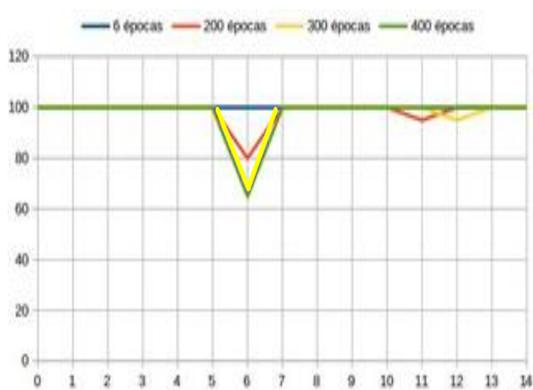
(a)



(b)



(c)



(d)

Figura 26. (a) blotchy, (b) bumpy, (c) crosshatched e (d) fibrous – Reconhecimento de IDs com texturas em 6, 200, 300 e 400 épocas.

Os valores obtidos de reconhecimento foram significativamente elevados na maioria das classes, com exceção das classes 5, 6, 7 e 11, cuja porcentagem de

reconhecimento foi abaixo de 80%, para alguns casos descritos a seguir. Essa porcentagem de reconhecimento foi baixa provavelmente pela ID original apresentar poucas características de identificação da classe. Na Figura 26(a) o reconhecimento foi de 75% para 6 épocas, e 55% para 400 épocas, ambos na classe 11. Na Figura 26(b) o reconhecimento foi de 70% na classe 5 em 400 épocas. Na Figura 26(c) o reconhecimento foi de 70%, 60% e 55%, para as classes 6, 7 e 11, respectivamente, em 6 épocas. Já na Fig. 26(d) o reconhecimento foi de 65% para 300 e 400 épocas, na classe 6. Observa-se que as especificidades de cada classe de textura, associada às características das IDs, não proporcionam resultados idênticos. Por isso, diferenças podem ser vistas nos gráficos para as diversas texturas da Figura 26.

A Figura 27 ilustra os casos citados, num gráfico onde a coordenada horizontal representa o número de épocas e a vertical, o número de classes em que a porcentagem foi abaixo de 80%.

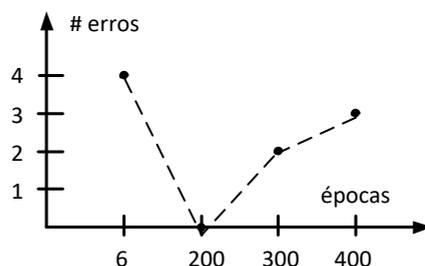


Figura 27. Quantidade de erros de reconhecimento para diferentes números de épocas de treinamento.

Os resultados de reconhecimento mostram uma certa tendência de que, quando a quantidade de épocas é insuficiente, no caso, 6 épocas, a quantidade de erros é alta. Esse número reduz para um número de épocas do treinamento suficiente de 200 épocas, porém volta a aumentar com números de épocas maiores.

Para os resultados de reconhecimento realizados foram calculadas as medidas de desempenho, ou sejam as taxas de Verdadeiros Positivos (*True Positive Rate*, TPR), a taxa de aceitação falsa (*False Accept Rate*, FAR) e a taxa de rejeição falsa (*False Reject Rate*, FRR) conforme descrito na Seção 4.3, para cada uma das réplicas construídas. As Tabelas 16, 17, 18 e 19, referem-se aos resultados do cálculo para as réplicas obtidas para as texturas *blotchy*, *bumpy*, *crosshatched* e *fibrous*, respectivamente.

Tabela 16. Medidas de desempenho das réplicas obtidas com a textura *blotchy*

	<i>Textura blotchy</i>			
Num. de épocas	6	200	300	400
TPR	0,9633	0,9633	0,9866	0,9600
FAR	0,0033	0,0033	0	0
FRR	0,0366	0,0366	0,0133	0,0400

Tabela 17. Medidas de desempenho das réplicas obtidas com a textura *bumpy*

	<i>Textura bumpy</i>			
Num. de épocas	6	200	300	400
TPR	1	1	0,9766	0,9800
FAR	0	0	0,0066	0
FRR	0	0	0,0233	0,0200

Tabela 18. Medidas de desempenho das réplicas obtidas com a textura *crosshatched*

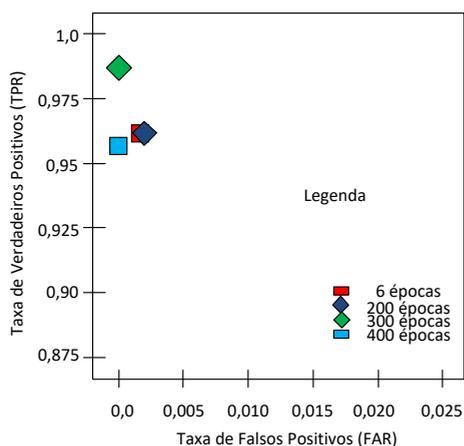
	<i>Textura crosshatched</i>			
Num. de épocas	6	200	300	400
TPR	0,9033	1	0,9966	1
FAR	0,0033	0	0	0
FRR	0,0966	0	0,0033	0

Tabela 19. Medidas de desempenho das réplicas obtidas com a textura *fibrous*

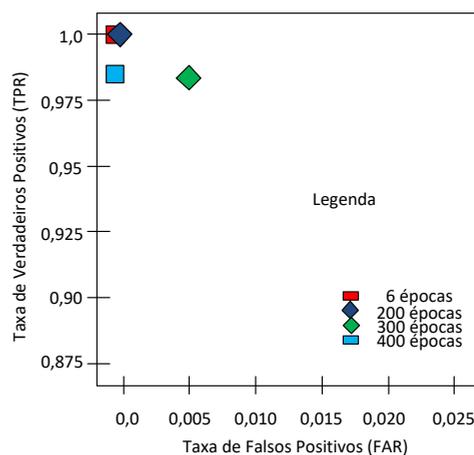
	<i>Textura fibrous</i>			
Num. de épocas	6	200	300	400
TPR	0,9500	0,9466	0,9466	0,9166
FAR	0,0033	0,0100	0,0033	0,0133
FRR	0,0500	0,0533	0,0533	0,0833

Em função dos resultados de desempenho mostrados nas tabelas descritas foram construídas as curvas ROC, conforme mostradas na Figura 28 para (a) textura *blotchy*, (b) *bumpy*, (c) *crosshatched* e (d) *fibrous*. Como podemos verificar, os pontos de

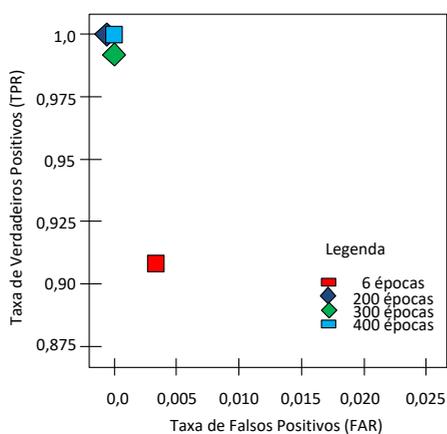
encontro das taxas FAR e TPR, ficaram próximo do ideal, ou seja, canto esquerdo superior onde TPR é igual a 1 e FAR é zero. Os detalhes dos resultados de reconhecimento das IDs que resultaram nessas medidas de desempenho, estão no Apêndice 1. Nota-se que foi considerado o índice de similaridade mínima de 0,8 para o reconhecimento das IDs réplicas, em comparação com as originais.



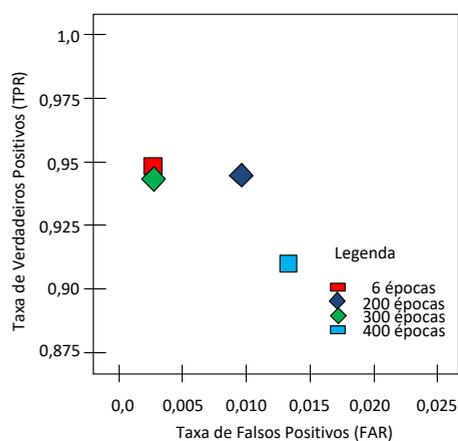
(a) Textura *blotchy*



(b) Textura *bumpy*



(c) Textura *crosshatched*



(d) Textura *fibrous*

Figura 28. Curva ROC para as réplicas obtidas com as texturas: (a) *blotchy*, (b) *bumpy*, (c) *crosshatched* e (d) *fibrous*.

5.5.3 Experimento 5.3 - Medidas de Similaridade das réplicas de IDs, obtidas no Experimento 5.1, com as originais do BDP2

As Tabelas 20, 21, 22 e 23, correspondem às medidas de similaridade média e o desvio padrão das réplicas de IDs obtidas com a aplicação da textura *blotchy*, *bumpy*, *crosshatched* e *fibrous*, respectivamente. Cada linha das tabelas corresponde a um número de épocas de treinamento, e as colunas, às métricas RMSE, SSIM e Xception.

Tabela 20. Similaridade média e desvio padrão para as réplicas usando textura *blotchy*

Épocas	<i>RMSE</i>		<i>SSIM</i>		<i>Xception</i>	
	Média	Desvio	Média	Desvio	Média	Desvio
6	0,9592	0,0006	0,9537	0,0019	0,9763	0,0465
200	0,9592	0,0006	0,9537	0,0019	0,9763	0,0465
300	0,9592	0,0006	0,9546	0,0025	0,9900	0,0268
400	0,9589	0,0006	0,9536	0,0024	0,9686	0,0859

Tabela 21. Similaridade média e desvio padrão para as réplicas usando textura *bumpy*

Épocas	<i>RMSE</i>		<i>SSIM</i>		<i>Xception</i>	
	Média	Desvio	Média	Desvio	Média	Desvio
6	0,9716	0,0045	0,9660	0,0174	0,9997	0,0007
200	0,9698	0,0023	0,9555	0,0195	0,9983	0,0042
300	0,9673	0,0026	0,9512	0,0172	0,9783	0,0495
400	0,9654	0,0020	0,9485	0,0188	0,9833	0,0505

Tabela 22. Similaridade média e desvio padrão para as réplicas usando textura *crosshatched*

Épocas	<i>RMSE</i>		<i>SSIM</i>		<i>Xception</i>	
	Média	Desvio	Média	Desvio	Média	Desvio
6	0,9713	0,0040	0,9548	0,0113	0,9410	0,0932
200	0,9653	0,0032	0,9763	0,0083	0,9998	0,0006
300	0,9645	0,0025	0,9764	0,0080	0,9992	0,0031
400	0,9679	0,0042	0,9756	0,0088	0,9999	0,0003

Tabela 23. Similaridade média e desvio padrão para as réplicas usando textura *fibrous*

Épocas	<i>RMSE</i>		<i>SSIM</i>		<i>Xception</i>	
	Média	Desvio	Média	Desvio	Média	Desvio
6	0,9666	0,0043	0,9646	0,0125	0,9996	0,0012
200	0,9655	0,0039	0,9632	0,0114	0,9838	0,0509
300	0,9618	0,0016	0,9586	0,0129	0,9777	0,0748
400	0,9642	0,0028	0,9602	0,0123	0,9836	0,0577

Os maiores valores de similaridade estão destacados em negrito, e os menores valores, em vermelho. Observa-se que os valores de similaridade são muito próximos para todas as quantidades de épocas de treinamento, e provavelmente por esta razão, os valores mínimos e máximos destacados não coincidem para ambas as métricas. Nota-se também que as medidas de similaridade média não refletem na reconhecibilidade das réplicas de IDs.

Os índices obtidos no reconhecimento foram significativamente elevados em todos os casos. Nos gráficos constatamos que os índices estão próximos a 100% com algumas oscilações, indicando alto índice de similaridade entre as réplicas de IDs e as originais. Porém, nota-se uma certa tendência de que quando as épocas do treinamento são aumentadas, temos uma possibilidade de diminuição na similaridade.

É necessário explicitar que devido às especificidades de cada classe de textura, e às características de cada classe de IDs, os resultados apresentam diferenças.

As Figuras 29, 30 e 31, mostram os gráficos de similaridade média das Tabelas 20, 21, 22 e 23, relativos aos valores obtidos pelas métricas RMSE, SSIM e Xception, respectivamente. Os gráficos são mostrados para as texturas *blotchy*, *bumpy*, *crosshatched* e *fibrous*, da esquerda para a direita e de cima para baixo, respectivamente.

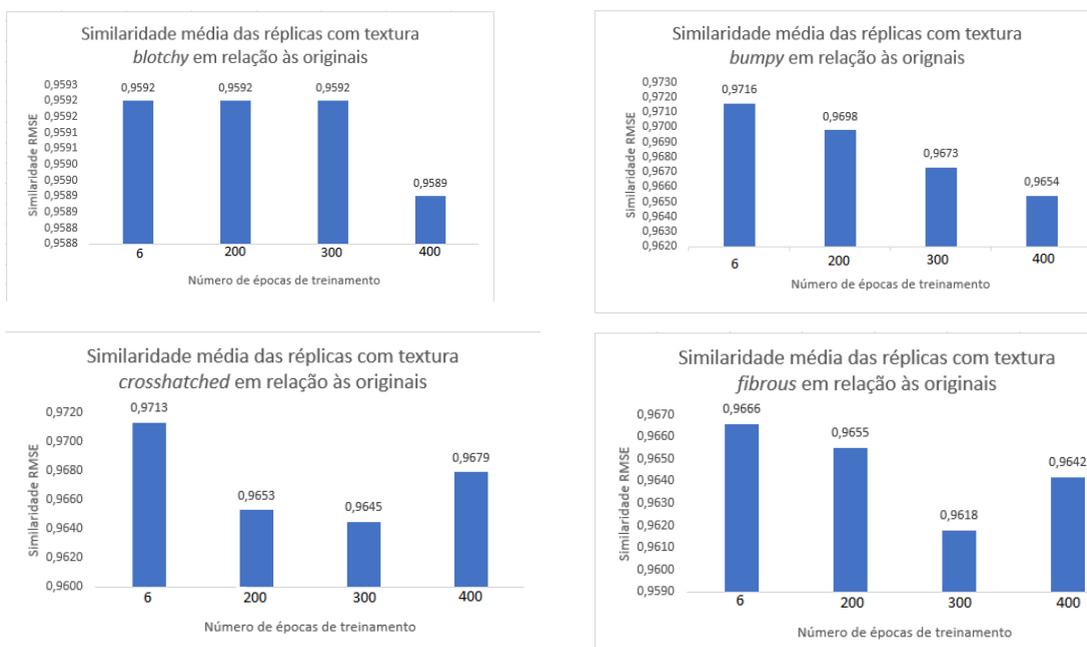


Figura 29. Gráficos de Similaridade média obtidas pela métrica RMSE das réplicas com texturas *blotchy*, *bumpy*, *crosshatched* e *fibrous*, da esquerda para a direita e de cima para baixo.

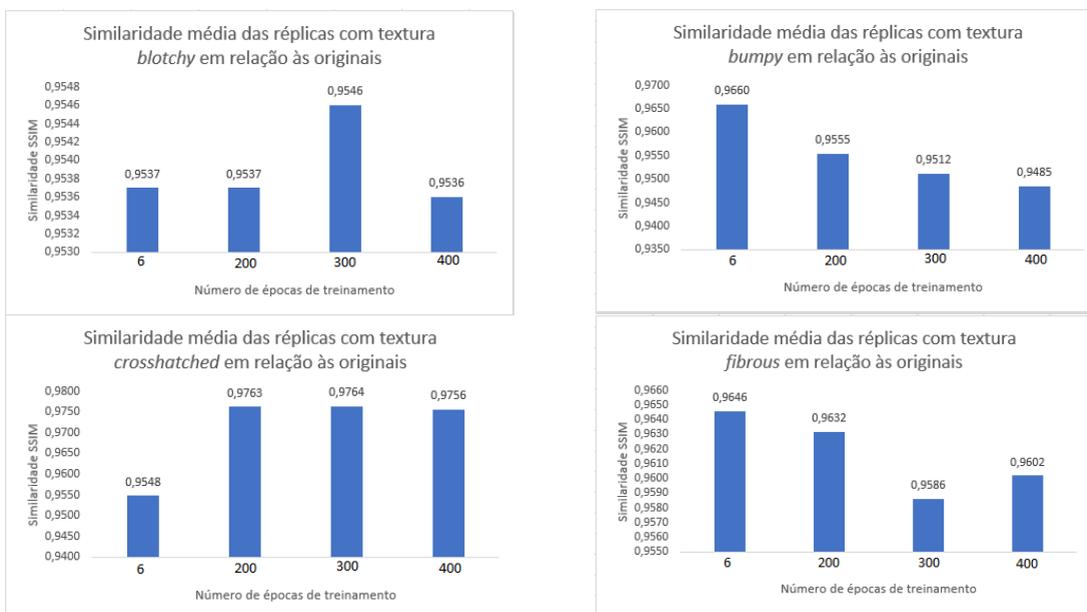


Figura 30. Gráficos de Similaridade média obtidas pela métrica SSIM das réplicas com texturas *blotchy*, *bumpy*, *crosshatched* e *fibrous*, da esquerda para a direita e de cima para baixo.

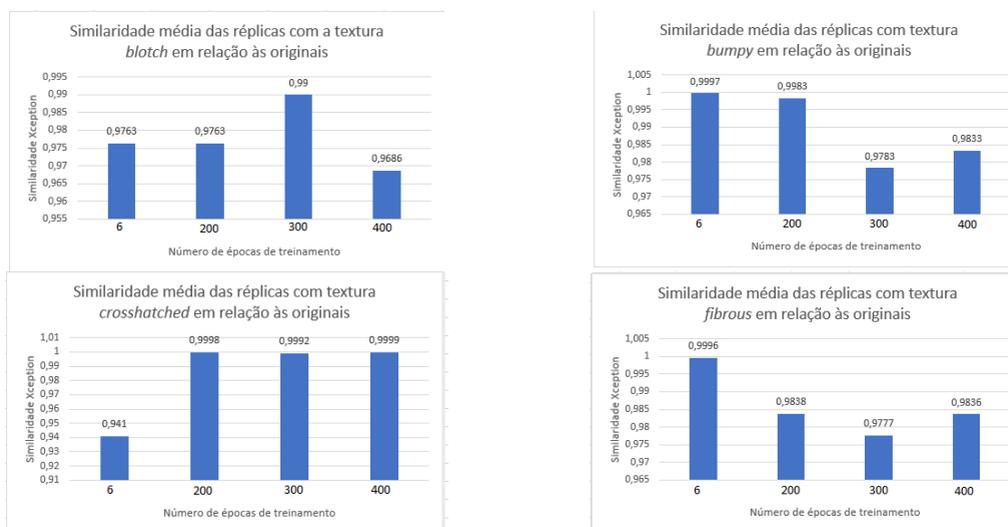


Figura 31. Gráficos de Similaridade média obtidas pela métrica Xception das réplicas com texturas *blotchy*, *bumpy*, *crosshatched* e *fibrous*, da esquerda para a direita e de cima para baixo.

5.6. Recursos computacionais

O trabalho desenvolvido utilizou-se de recursos computacionais de software necessários para a implementação dos programas, na linguagem de programação Python com TensorFlow, Keras, OpenCV, e sistema operacional Linux na distribuição Ubuntu 20.04, sendo máquina local para o pré-processamento e treinamento das redes neurais. Na modelagem dos diagramas utilizamos o (ArchiMate 4.6.0, 2021) e para desenho 3D o (TinkerCAD, 2021). O sistema computacional foi constituído do processador i7-9750H, com 6 núcleos físicos, 16GB de memória RAM e GPU GTX 1660 Ti 6GB.

5.7. Considerações finais

Neste capítulo foram apresentados os resultados sobre a geração de réplicas de IDs utilizando a Pix2pix, com e sem acréscimo de textura. Foram aplicados os testes de reconhecibilidade das réplicas com o Xception, e realizados os cálculos de similaridade entre as réplicas e as originais, utilizando Xception, métrica RMSE e SSIM. No capítulo seguinte, será apresentada a conclusão do trabalho de Dissertação.

CAPÍTULO 6 – Conclusões e Trabalhos Futuros

Descrevemos neste trabalho de Dissertação uma abordagem experimental denominada SIMTEX, que utiliza a rede GAN Pix2pix, para a obtenção de réplicas de IDs a partir das amostras de IDs originais, acrescentando texturas, no contexto do estudo de vulnerabilidade da segurança dos sistemas de identificação biométricos. Também, foi descrita a abordagem de se obter ampliação do banco de dados de IDs, sem acrescentar texturas ao repertório de amostras em bancos de dados de IDs. Os resultados mostram que a escolha das texturas é importante para manter a reconhecibilidade das amostras obtidas, além da quantidade de épocas de treinamento. Uma tendência verificada é que um certo número de épocas de treinamento é necessário para se chegar a réplicas de IDs com bom índice de reconhecimento, porém, a partir disso, quanto maior a quantidade de épocas de treinamento, a reconhecibilidade fica reduzida. Nota-se que a similaridade entre as amostras originais e as geradas, apesar de altas, a reconhecibilidade das amostras geradas pode ser baixa.

Foi demonstrado nos resultados, uma predisposição de melhor reconhecimento de réplicas de IDs, obtidas em 200 épocas, quando as texturas foram geradas com a rede Pix2pix, mas pode suceder texturas que não reflitam da mesma forma e temos que adequar essas peculiaridades da textura, com os aspectos específicos das IDs, na adoção da técnica SIMTEX descrita.

Apesar do objetivo do desenvolvimento da técnica SIMTEX ser de obter réplicas de IDs, como as imagens de IDs utilizadas para a obtenção dos resultados são imagens originais, as IDs obtidas são variações das IDs originais. Para a continuidade da pesquisa seria interessante o desenvolvimento de alguma técnica de obtenção de ID falsa. Uma possibilidade é a utilização do dispositivo de posicionamento para captura de ID desenvolvido, descrito na Seção 4.1.3.1, para o posicionamento de imagens de IDs, no lugar de dedos dos indivíduos. Assim, o resultado da captura poderia ser considerado

uma ID falsa. E estudos poderiam ser realizados comparando essas IDs falsas com as réplicas de IDs obtidas com o SIMTEX.

Principais contribuições

Abordamos o método SIMTEX, para a construção de réplicas de IDs, mantendo as características de reconhecibilidade dessas réplicas, incluindo pré-processamentos como a extração de região de interesse, ROI, o operador de Sobel para o realce de bordas e o método CLAHE de equalização de histogramas. O operador de Sobel e a equalização de histogramas melhoram a precisão na aplicação da rede Pix2pix. A abordagem SIMTEX serviu para substituir a construção de um dedo artificial, utilizando diversos materiais, como ferramenta para a obtenção de réplicas de IDs. A pesquisa desenvolvida teve como objetivo investigar a complexidade na prática de falsificação de IDs, no mundo real, para o aperfeiçoamento de sistemas de autenticação biométrica utilizando impressões digitais.

Durante a pesquisa, foi necessário desenvolver um protótipo denominado como dispositivo de posicionamento de captura biométrica, utilizando a câmera de um *smartphone*, permitindo uma padronização nos procedimentos de coleta da imagem, reduzindo possíveis distorções durante a aquisição das IDs.

Trabalhos futuros

Abordamos especificamente neste trabalho imagens de IDs, porém existe a possibilidade de estudos mais amplos, de forma genérica, visto que, atuamos na transferência de aprendizagem, no intuito de reproduzir as características específicas do elemento, associando outra textura, para as mais diversas finalidades. Poderíamos simular faces, mãos, íris, entre outros, na intenção de obter réplicas com as mesmas características das originais, as quais poderiam ser utilizadas com a finalidade de ampliar pesquisas direcionadas ao aperfeiçoamento de sistemas de autenticação biométrica, pois manter a autenticidade e a integralidade de forma confiável, são requisitos fundamentais e uma tarefa desafiadora.

Como um outro trabalho futuro, um estudo poderia ser realizado, medindo a similaridade das réplicas, que não foram reconhecidas pelo Xception. Isso poderia explicitar a correspondência entre a similaridade medida utilizando as métricas utilizadas e a reconhecibilidade, uma vez que, neste trabalho, os experimentos de similaridade mostram a média das similaridades e não a similaridade específica.

Além disso, o desenvolvimento de banco de dados de impressões digitais falsas, obtidas por meio de captura de fotos de impressões digitais originais, para posterior reconhecimento, fazendo a distinção das originais e aferindo sua similaridade. Os resultados de similaridade e reconhecimento dessas IDs falsas poderiam ser comparados com os resultados obtidos com o SIMTEX, para validar o método SIMTEX como uma forma de simular a obtenção de réplicas de IDs falsas.

Trabalhos publicados

Os seguintes trabalhos foram publicados como resultados desta dissertação:

- Anais do WCF (*Workshop de Computação do UNIFACCAMP*), Vol 7, pp 15-20, 2020, ISSN 2447-4703, XVI WCF, 1-2 outubro de 2020, com o título: Similaridades em Imagens Geradas por Redes Neurais Pix2Pix medidas pela rede Xception.
- WVC 2021 (*XVII Workshop de Visão Computacional*), 22-23 novembro de 2021, com o título: Pix2pix network for fingerprint texture image synthesis.

REFERÊNCIAS

- Abhyankar, A., & Schuckers, S. (2006) Fingerprint Liveness Detection using Local Ridge Frequencies and Multiresolution Texture Analysis Techniques. *International Conference on Image Processing*, Atlanta, USA, 8 a 11 outubro, pp.321-324.
- Amancio, D. R., Comin, C. H., Casanova, D., Travieso, G., Bruno, O. M., Rodrigues, F. A., & Costa, L. da F. (2014) A Systematic Comparison of Supervised Classifiers. *PLoS ONE*, vol.9, n.4, pp.e94137.
- Amirthalingam, G., & Radhamani., G. (2013) A Multimodal Approach for Face and Ear Biometric System. *International Journal of Computer Science Issues*, vol.10, n.5, pp.234-241.
- ArchiMate 4.6.0. (2021) [Software]. Archi – Open Source ArchiMate Modelling [Online]. <https://www.archimatetool.com/#>. Acessado em janeiro de 2021.
- Arjovsky, M., & Bottou, L. (2017a) Towards Principled Methods for Training Generative Adversarial Networks. *ArXiv:1701.04862*.
- Arjovsky, M., Chintala, S., & Bottou, L. (2017b) Wasserstein GAN. *ArXiv:1701.07875 [Stat.ML]*.
- Berg, E. V. D., Friedlander, M. P., & SIAM, J. (2007) *SPGL1: A Solver for Large-Scale Sparse Reconstruction* [Software Online]. <https://www.cs.ubc.ca/~mpf/spgl1/index.html>. Acessado em dezembro de 2020.
- Bishop, C. M. (2006) *Pattern Recognition and Machine Learning*. New York, Springer, pp. 199-200.
- Cao, L. J., Chua, K. S., Chong, W. K., Lee, H. P., & Gu, Q. M. (2003) A Comparison of PCA, KPCA and ICA for Dimensionality Reduction in Support Vector Machine. *Neurocomputing*, vol.55, pp.321-336.

- Catté, F., Lions, P.-L., Morel, J.-M., & Coll, T. (1992) Image Selective Smoothing and Edge Detection by Nonlinear Diffusion. *Society for Industrial and Applied Mathematics*, vol.29, n.1, pp.182-193.
- Caviedes, J., & Oberti, F. (2004) A New Sharpness Metric Based on Local Kurtosis, Edge and Energy Information. *Signal Processing: Image Communication*, vol.19, pp.147-161.
- Chalana, V., & Kim, Y. (1997) A Methodology for Evaluation of Boundary Detection Algorithms on Medical Images. *IEEE Transactions on Medical Imaging*, vol.16, n.5, pp.642-652.
- Chen, C., & Ross, A. (2018) A Multi-Task Convolutional Neural Network for Joint Iris Detection and Presentation Attack Detection. *IEEE Winter Applications of Computer Vision Workshops (WACVW)*, Lake Tahoe, USA, 15 de março, pp.44-51.
- Chiachia, G. (2014) *Convolutional Neural Networks using Random Filter Weights (CNN-RFW)* [Software Online]. <https://github.com/giovanichiachia/convnet-rfw>. Acessado em novembro de 2020.
- Chollet, F. (2017) Xception: Deep Learning with Depthwise Separable Convolutions. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 de julho, pp.1800-1807.
- Chugh, T., Cao, K., & Jain, A. K. (2018) Fingerprint Spoof Buster: Use of Minutiae-Centered Patches. *IEEE Transactions on Information Forensics and Security*, vol.13, n.9, pp.2190-2202.
- Chugh, T., & Jain, A. K. (2020) Fingerprint Spoof Detector Generalization. *IEEE Transactions on Information Forensics and Security*, vol.16, p.42-55.
- Cimpoi, M., Maji, S., Kokkinos, I., Mohamed, S., & Vedaldi, A. (2014) Describing Textures in the Wild. *IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 23-28 junho*.

- Clark, A. (2015). Pillow (PIL Fork) Documentation. readthedocs. Retrieved from <https://buildmedia.readthedocs.org/media/pdf/pillow/latest/pillow.pdf>
- Demuth, H., & Beale, M. (1993) *Neural Network Toolbox User's Guide* (Version 4). The MathWorks, Inc. [Online]. http://128.174.199.77/matlab_pdf/nnet.pdf. Acessado em dezembro de 2020.
- Derakhshani, R., Schuckers, S. A. C., Hornak, L. A., & O’Gorman, L. (2003) Determination of Vitality from a Non-Invasive Biomedical Measurement for Use in Fingerprint Scanners. *Pattern Recognition*, vol.36, pp.383-396.
- Dumoulin, V., & Visin, F. (2018) A Guide to Convolution Arithmetic for Deep Learning. *ArXiv:1603.07285 [Cs, Stat]*, pp.19-20.
- Feng, Y., Chen, Z., Wang, D., Chen, J., & Feng, Z. (2020) DeepWelding: A Deep Learning Enhanced Approach to GTAW using Multisource Sensing Images. *IEEE Transactions on Industrial Informatics*, vol.16, n. 1, pp.465-474.
- Filho, O. M., & Neto, H. V. (1999). *Processamento Digital de Imagens*. Brasport; pp. 27.
- Galbally, J., Marcel, S., & Fierrez, J. (2014) Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint, and Face Recognition. *IEEE Transactions on Image Processing*, vol.23, n.2, pp.710-724.
- Gauthier, J. (2015) Conditional Generative Adversarial Nets for Convolutional Face Generation. *Technical Report* [Online]. <http://www.foldl.me/uploads/papers/tr-cgans.pdf>. Acessado em setembro de 2020.
- Geethanjali, N., & Thamaraiselvi, K. (2013) Feature Level Fusion of Multimodal Biometrics and Two Tier Security in ATM System. *International Journal of Computer Applications*, vol.70, n.14, pp.17-23.
- Gil, A. C. (2002) *Como elaborar projetos de pesquisa* (4 ed.). São Paulo-SP, Ed. Atlas; pp.59-85.

- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014) Generative Adversarial Networks. *ArXiv:1406.2661 [Cs, Stat]*.
- Gonzalez, R. C., & Woods, R. C. (2010). *Processamento Digital de Imagens* (3rd ed.). Pearson Prentice Hall; pp. 46.
- Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. (2017) Improved Training of Wasserstein GANs. *ArXiv:1704.00028v3 [Cs, LG]*.
- Gupta, K. (2017) Advances in Multi Modal Biometric Systems: A Brief Review. *International Conference on Computing, Communication and Automation (ICCCA)*, Greater Noida, Uttar Pradeshi, India, 5-6 de maio, pp.262-267.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009) The WEKA Data Mining Software: An Update. *ACM SIGKDD Explorations Newsletter*, vol.11, n.1, pp.10-18.
- Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973) Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, vol.3, n.6, pp.610-621.
- He, K., Zhang, X., Ren, S., & Sun, J. (2015) Deep Residual Learning for Image Recognition. *ArXiv:1512.03385 [Cs]*.
- He, P., Li, H., & Wang, H. (2019) Detection of Fake Images Via The Ensemble of Deep Representations from Multi Color Spaces. *IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan, 22-25 de setembro, pp.2299-2303.
- Hinton, G., Sabour, S., & Frosst, N. (2018) Matrix Capsules with EM Routing. *6th International Conference on Learning Representations (ICLR)*, Vancouver, Canadá, 30 de abril a 3 de maio.

- Hongliang, J., Qingshan, L., Hanqing, L., & Xiaofeng, T. (2004) Face Detection using Improved LBP under Bayesian Framework. *Third International Conference on Image and Graphics (ICIG'04)*, Hong Kong, China, 18-20 de dezembro, pp.306-309.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weijun, W., Weyand, T., Andreetto, M., & Adam, H. (2017) MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *ArXiv:1704.04861 [Cs]*.
- Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003) A Practical Guide to Support Vector Classification. *Technical Report, Department of Computer Science and Information Engineering*, University of National Taiwan, Taipei, pp.1-12.
- Huang, G., Liu, Z., Maaten, L. v. d., & Weinberger, K. Q. (2017) Densely Connected Convolutional Networks. *ArXiv:1608.06993[Cs]*.
- Huang, X., & Belongie, S. (2017) Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization. *IEEE International Conference on Computer Vision (ICCV)*, Veneza, Itália, 22 a 29 de outubro, pp.1510-1519.
- Hussein, M. E., Spinoulas, L., Xiong, F., & Abd-Almageed, W. (2018) Fingerprint Presentation Attack Detection using A Novel Multi-Spectral Capture Device and Patch-Based Convolutional Neural Networks. *IEEE International Workshop on Information Forensics and Security (WIFS)*, Hong Kong, China, 11 a 13 de dezembro.
- Huttenlocher, D. P., Klanderman, G. A., & Rucklidge, W. J. (1993) Comparing Images using the Hausdorff Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.15, n. 9, pp.850-863.
- Ioffe, S., & Szegedy, C. (2015) Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning*, vol.37, Lille, França, 6 a 11 de julho, pp.448-456.

- Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017) Image-to-Image Translation with Conditional Adversarial Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, USA, 21 a 26 de julho, pp.5967-5976.
- Jain, A. K., Ross, A., & Prabhakar, S. (2004) An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, vol.14, n.1, pp.4-20.
- Jayan, T. J., & Aneesh, R. P. (2018) Image Quality Measures Based Face Spoofing Detection Algorithm for Online Social Media. *International CET Conference on Control, Communication, and Computing (IC4)*, Thiruvanthapuram, India, 5 a 7 de julho, pp.245-249.
- Jin, J., Dundar, A., & Culurciello, E. (2015) Flattened Convolutional Neural Networks for Feedforward Acceleration. *Workshop Contribution at ICLR*, San Diego, CA, USA, 7 a 9 de maio.
- Johnson, J., Alahi, A., & Fei-Fei, L. (2016) Perceptual Losses for Real-Time Style Transfer and Super-Resolution. *Computer Vision – ECCV*, vol.9906, Amsterdã, Holanda, 11 a 14 de outubro, pp.694-711.
- Kannala, J., & Rahtu, E. (2012) BSIF: Binarized Statistical Image Features. *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, Tsukuba, Japão, 11 a 15 de novembro, pp.1363-1366.
- Karras, T., Laine, S., & Aila, T. (2019) A Style-Based Generator Architecture for Generative Adversarial Networks. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 16 a 20 de junho, pp.4396-4405.
- Khutlang, R., & Nelwamondo, F. V. (2014) Novelty Detection-Based Internal Fingerprint Segmentation in Optical Coherence Tomography Images. *Second*

- International Symposium on Computing and Networking*, Shizuoka, Japão, 10 a 12 de dezembro, pp.556-559.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012) ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, vol.60, n.6, pp.84-90.
- Kulkarni, N. A., & Sankpal, L. J. (2017) Efficient Approach Determination for Fake Biometric Detection. *International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, Pune, India, 17 a 18 de agosto, pp.1-4.
- Lei, H., & Govindaraju, V. (2005) Speeding Up Multi-Class SVM Evaluation by PCA and Feature Selection. *SIAM International Conference on Data Mining*, Newport Beach, CA, USA, 21 a 23 de abril, pp.72-79.
- Li, C., & Wand, M. (2016) Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks. *Computer Vision (ECCV), 14th European Conference*, vol.9907, Amsterdã, Holanda, 8 a 16 de outubro, pp.702-716.
- Li, L., Bao, J., Zhang, T., Yang, H., Chen, D., Wen, F., & Guo, B. (2020) Face X-Ray for More General Face Forgery Detection. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, 14 a 19 de junho, pp.5000-5009.
- Liu, A., Lin, W., & Narwaria, M. (2012) Image Quality Assessment Based on Gradient Similarity. *IEEE Transactions on Image Processing*, vol.21, n. 4, 1500–1512.
- Lowe, D. G. (2004) Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, vol.60, pp.91-110.
- Ma, L., Shuai, R., Ran, X., Liu, W., & Ye, C. (2020) Combining DC-GAN with ResNet for Blood Cell Image Classification. *Medical & Biological Engineering & Computing*, vol.58, pp.1251-1264.

- Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013) Rectifier Nonlinearities Improve Neural Network Acoustic Models. *30 Th International Conference on Machine Learning*, vol.28, Atlanta, USA, 17 a 19 de junho.
- Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S. (2009) *Handbook of Fingerprint Recognition (second edition)*. London, Springer.
- Marconi, M. de A., & Lakatos, E. M. (2017) *Fundamentos de metodologia científica* (sétima edição). São Paulo-SP, Atlas; pp.198-216.
- Mirza, M., & Osindero, S. (2014) Conditional Generative Adversarial Nets. *ArXiv:1701.04862*.
- Mirzaalian, H., Hussein, M., & Abd-Almageed, W. (2019) On the Effectiveness of Laser Speckle Contrast Imaging and Deep Neural Networks for Detecting Known and Unknown Fingerprint Presentation Attacks. *International Conference on Biometrics (ICB)*, Creta, Grécia, 4 a 7 de junho.
- Mittal, A., Moorthy, A. K., & Bovik, A. C. (2011) *BRISQUE Software Release* [Software Online]. http://live.ece.utexas.edu/research/quality/BRISQUE_release.zip. Acessado em agosto de 2020.
- Mittal, A., Moorthy, A. K., & Bovik, A. C. (2012) No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Transactions on Image Processing*, vol.21, n.12, pp.4695-4708.
- Nair, V., & Hinton, G. E. (2010) Rectified Linear Units Improve Restricted Boltzmann Machines. *27th International Conference on Machine Learning*, Haifa, Israel, 21 a 24 de junho, pp.807-814.
- Nash, J. (1951) Non-Cooperative Games. *The Annals of Mathematics*, vol.54, n.2, pp.286-295.
- Nikam, S. B., & Agarwal, S. (2008) Local Binary Pattern and Wavelet-based Spoof Fingerprint Detection. *International Journal of Biometrics*, vol.1, n.2, pp.141-159.

- Nogueira, R. F., Lotufo, R. de A., & Machado, R. C. (2014) Evaluating Software-Based Fingerprint Liveness Detection using Convolutional Networks and Local Binary Patterns. *IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings*, Roma, Itália, 17 de outubro, pp.22-29.
- Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002) Multiresolution Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.24, n.7, pp.971-987.
- Otsu, N. (1979) A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, vol.9, n.1, pp.62-66.
- Pala, F., & Bhanu, B. (2017) On the Accuracy and Robustness of Deep Triplet Embedding for Fingerprint Liveness Detection. *IEEE International Conference on Image Processing (ICIP)*, Beijing, China, 17 a 20 de setembro, pp.116-120.
- Patil, A. A., & Dhole, S. A. (2016) Image Quality (IQ) Based Liveness Detection System for Multi-Biometric Detection. *International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, 26 a 27 de agosto, pp.1-5.
- Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., & Zuiderveld, K. (1987) Adaptive Histogram Equalization and Its Variations. *Computer Vision, Graphics, and Image Processing*, vol.39, n.3, pp.355-368.
- Prabhakar, S., Pankanti, S., & Jain, A. K. (2003) Biometric Recognition: Security and Privacy Concerns. *IEEE Security & Privacy*, vol.1, n.2, pp.33-42.
- Radford, A., Metz, L., & Chintala, S. (2016) Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *4th International Conference on Learning Representations, (ICLR)*, San Juan, Porto Rico, 2 a 4 de maio.

- Raghavendra, R., Raja, K. B., & Busch, C. (2015) Presentation Attack Detection for Face Recognition using Light Field Camera. *IEEE Transactions on Image Processing*, vol.24, n.3, pp.1060-1075.
- Rattani, A., Akhtar, Z., & Foresti, G. (2015) A Preliminary Study on Identifying Fabrication Material From Fake Fingerprint Images. *IEEE Symposium Series on Computational Intelligence*, Cape Town, África do Sul, 7 a 10 de dezembro, pp.362-366.
- Rattani, A., & Ross, A. (2014) Minimizing the Impact of Spoof Fabrication Material on Fingerprint Liveness Detector. *IEEE International Conference on Image Processing (ICIP)*, Paris, França, 27 a 30 de outubro, pp.4992-4996.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, Nevada, USA, 27 a 30 de junho, pp.779-788.
- Redmon, J., & Farhadi, A. (2017) YOLO9000: Better, Faster, Stronger. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, USA, 21 a 26 de julho, pp.6517-6525.
- Ronneberger, O., Fischer, P., & Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI*, vol.9351, Munique, Alemanha, 5 a 9 de outubro. pp.234-241.
- Rössler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019) FaceForensics++: Learning to Detect Manipulated Facial Images. *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Coreia do Sul, 27 de outubro a 3 de novembro. pp.1-11.

- Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011) ORB: An Efficient Alternative to SIFT or SURF. *2011 International Conference on Computer Vision*, Barcelona, Espanha, 6 a 13 de novembro, pp.2564-2571.
- Sabour, S., Frosst, N., & Hinton, G. E. (2017) Dynamic Routing Between Capsules. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, CA, USA, 4 a 9 de dezembro, pp.3859-3869.
- Samarasinghe, P., Lakmal, L. K. P., Weilkala, A. V., Wickramarachchi, W. A. N. P. C., & Niroshana, E. R. S. (2017) Sri Lanka Driving License Forgery Detection. *Fourth International Conference on Image Information Processing (ICIIP)*, Himachal Pradesh, India, 21 a 23 de dezembro, pp.1-6.
- Shankaranarayana, S. M., Ram, K., Mitra, K., & Sivaprakasam, M. (2017) Joint Optic Disc and Cup Segmentation Using Fully Convolutional and Adversarial Networks. *Fetal, Infant and Ophthalmic Medical Image Analysis. OMIA, FIFI, vol.10554*, Québec, Canadá, 6 de setembro, pp.168-176.
- Shehu, Y. I., Ruiz-Garcia, A., Palade, V., & James, A. (2018) Sokoto Coventry Fingerprint Dataset [Online]. *ArXiv:1807.10609*, 3. Acessado em agosto de 2020.
- Shibuya, N. (2017) *Up-Sampling with Transposed Convolution* [Online]. <https://naokishibuya.medium.com/up-sampling-with-transposed-convolution-9ae4f2df52d0>. Acessado em Novembro de 2020.
- Sifre, L., & Mallat, S. (2013) Rotation, Scaling and Deformation Invariant Scattering for Texture Discrimination. *2013 IEEE Conference on Computer Vision and Pattern Recognition*, Portland, Oregon, USA, 23 a 28 de junho, pp.1233-1240.
- Sifre, L., & Mallat, S. (2014) Rigid-Motion Scattering for Texture Classification. *ArXiv:1403.1687 [Cs]*.
- Silva, M. V., Marana, A. N., & Paulino, A. A. (2015) On the Importance of using High Resolution Images, Third Level Features and Sequence of Images for Fingerprint

- Spoof Detection. *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, Queensland, Australia, 19 a 24 de abril, pp.1807-1811.
- Simonyan, K., & Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. *ArXiv:1409.1556 [Cs]*.
- Singh, A., Jaswal, G., & Nigam, A. (2019) FDSNet: Finger Dorsal Image Spoof Detection Network using Light Field Camera. *IEEE 5th International Conference on Identity, Security, and Behavior Analysis (ISBA)*, Kobe, Japão, 13 a 14 de julho, pp.1-9.
- Sivasankari, N., & Muthukumar, A. (2016) A Review on Recent Techniques in Multimodal Biometrics. *International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, 4 a 9 de janeiro, pp.1-11.
- Sobel, I., & Feldman, G. (1968) A 3×3 Isotropic Gradient Operator for Image Processing. *Stanford Artificial Intelligence Project (SAIL)* [Online]. https://www.researchgate.net/publication/281104656_An_Isotropic_3x3_Image_Gradient_Operator?channel=doi&linkId=55d5876408ae43dd17de57a4&showFulltext=true. Acessado em dezembro de 2020.
- Springenberg, J. T., Dosovitskiy, A., Brox, T., & Riedmiller, M. (2015) Striving for Simplicity: The All Convolutional Net. *ArXiv:1412.6806 [Cs]*.
- Sun, K., Zhao, Y., Jiang, B., Cheng, T., Xiao, B., Liu, D., Mu, Y., Wang, X., Liu, W., & Wang, J. (2019) High-Resolution Representations for Labeling Pixels and Regions. *ArXiv:1904.04514 [Cs]*.
- Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2016) Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *ArXiv:1602.07261 [Cs]*.

- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2014) Going Deeper with Convolutions. *ArXiv:1409.4842 [Cs]*.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2015) Rethinking the Inception Architecture for Computer Vision. *ArXiv:1512.00567 [Cs]*.
- Tabassi, E., Wilson, C. L., & Watson, C. I. (2004) Fingerprint Image Quality. *NIST Interagency/Internal Report (NISTIR) – 7151* [Online].
<https://www.nist.gov/publications/fingerprint-image-quality>. Acessado em dezembro de 2020.
- Tan, T., Wang, X., Fang, Y., & Zhang, W. (2019) The Impact of Data Correlation on Identification of Computer-Generated Face Images. *14th Chinese Conference on Biometric Recognition, CCBR 2019, vol.11818*, Zhuzhou, China, 12 a 13 de outubro, p.155-162.
- Tariq, S., Lee, S., Kim, H., Shin, Y., & Woo, S. S. (2018) Detecting Both Machine and Human Created Fake Face Images In the Wild. *2nd International Workshop on Multimedia Privacy and Security - MPS '18*, Toronto, Canadá, outubro, pp.81-87.
- The GIMP Development Team. (2019). GIMP. Retrieved from <https://www.gimp.org>.
- Thuy, M. B. H., & Hoang, V. T. (2020)Fusing of Deep Learning, Transfer Learning and GAN for Breast Cancer Histopathological Image Classification. *Advanced Computational Methods for Knowledge Engineering (H.A. Lethi et al. Eds. : ICCSAMA 2019, AISC 1121)*. *Suiça, Springer Nature*, pp. 255–266.
- TinkerCAD. (2021) [Software]. Tinkercad | From Mind to Design in Minutes [Online].
<https://www.tinkercad.com/>. Acessado em outubro de 2020.
- Viola, P., & Jones, M. J. (2001) Rapid Object Detection using a Boosted Cascade of Simple Features. *Proceedings of the IEEE Computer Society Conference on*

- Computer Vision and Pattern Recognition - CVPR*, Kauai, HI, USA, 8 a 14 de dezembro, pp.I-511-I-518.
- Viola, P., & Jones, M. J. (2004) Robust Real-Time Face Detection. *International Journal of Computer Vision*, vol.57, n. 2, pp.137-154.
- Wang, G., Kang, W., Wu, Q., Wang, Z., & Gao, J. (2018) Generative Adversarial Network (GAN) Based Data Augmentation for Palmprint Recognition. *Digital Image Computing: Techniques and Applications (DICTA)*, Canberra, Australia, 10 a 13 de dezembro, pp.1-7.
- Wang, M., Liu, B., & Foroosh, H. (2017) Factorized Convolutional Neural Networks. *IEEE International Conference on Computer Vision Workshops (ICCVW)*, Veneza, Itália, 22 a 29 de outubro, pp.545-553.
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004) Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, vol.13, n.4, pp.600-612.
- Williams, G. W. (1976) Comparing the Joint Agreement of Several Raters with Another Rater. *Biometrics*, vol.32, n.3, pp.619-627.
- Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., & Ma, Y. (2009) Robust Face Recognition via Sparse Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.31, n.2, pp.210-227.
- Yambay, D., Ghiani, L., Denti, P., Marcialis, G. L., Roli, F., & Schuckers, S. (2012) LivDet 2011 – Fingerprint Liveness Detection Competition 2011. *5th IAPR International Conference on Biometrics (ICB)*, Nova Delhi, India, 29 de março a 1 de abril, pp.208-215.
- Zaghetto, C., Mendelson, M., Zaghetto, A., & Vidal, F. de B. (2017) Liveness Detection on Touchless Fingerprint Devices using Texture Descriptors and Artificial Neural

- Networks. *IEEE International Joint Conference on Biometrics (IJCB)*, Denver, CO, USA, 1 a 4 de outubro, pp.406-412.
- Zhang, H., Chen, Z., Zhang, C., Xi, J., & Le, X. (2019) Weld Defect Detection Based on Deep Learning Method. *IEEE 15th International Conference on Automation Science and Engineering (CASE)*, Vancouver, Canadá, 22 a 26 de agosto, pp.1574-1579.
- Zhang, L., Wang, P., Dang, F., & Zhang, S. (2019) A Simple and Robust Attentional Encoder-Decoder Model for License Plate Recognition. *Pattern Recognition and Computer Vision: Second Chinese Conference, PRCV, Part I, vol.11857*, Xi'an, China, 8 a 11 de novembro, pp.295-307.
- Zhang, L., Zhou, Z., & Li, H. (2012) Binary Gabor Pattern: An Efficient and Robust Descriptor for Texture Classification. *19th IEEE International Conference on Image Processing*, Orlando, FL, USA, 30 de setembro a 3 de outubro, pp.81-84.
- Zhang, Y., Shi, D., Zhan, X., Cao, D., Zhu, K., & Li, Z. (2019) Slim-ResCNN: A Deep Residual Convolutional Neural Network for Fingerprint Liveness Detection. *IEEE Access*, vol.7, pp.91476-91487.
- Zhao, Q., Zhang, D., Zhang, L., & Luo, N. (2010) Adaptive Fingerprint Pore Modeling and Extraction. *19th International Conference on Pattern Recognition (ICPR2008)*, Tampa, FL, USA, pp.2833-2844.
- Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017) Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. *IEEE International Conference on Computer Vision (ICCV)*, Veneza, Itália, 22 a 29 de outubro, pp. 2242–2251.
- Zuiderveld, K. (1994) Contrast Limited Adaptive Histogram Equalization. In *Graphics Gems: Vol. VIII.5 (Paul S. Heckbert eds.)*, Amsterdã, Elsevier, (pp. 474–485).

APÊNDICE 1 – Tabelas de Reconhecimento das réplicas do banco BDP2

Tabela A1 – Reconhecimento das Réplicas com textura *blotchy* obtidas com 6 épocas de treinamento.

		Classes Reconhecidas															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Não reconhecidas
Classes originais das réplicas	0	20															
	1		19														1
	2			20													
	3				20												
	4					20											
	5						19										1
	6							20									
	7								20								
	8									20							
	9										20						
	10											20					
	11												15				5
	12													20			
	13														16	1	3
	14															20	

TPR = 289/300 = 0,9633 FAR = 1/300 = 0,0033 FRR = 11/300 = 0,0366

Tabela A2 – Reconhecimento das Réplicas com textura *blotchy* obtidas com 200 épocas de treinamento.

		Classes Reconhecidas																	
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Não reconhecidas		
Classes originais das réplicas	0	20																	
	1		19															1	
	2			20															
	3				20														
	4					20													
	5						19												1
	6							20											
	7								20										
	8									20									
	9										20								
	10											20							
	11												15						5
	12													20					
	13														16	1			3
	14																20		
TPR = 289/300 = 0,9633 FAR = 1/300 = 0,0033 FRR = 11/300 = 0,0366																			

Tabela A3 – Reconhecimento das Réplicas com textura *blotchy* obtidas com 300 épocas de treinamento.

		Classes Reconhecidas															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Não reconhecidas
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						20										
	6							20									
	7								20								
	8									20							
	9										20						
	10											20					
	11												16				4
	12													20			
	13														20		
	14															20	

TPR = 296/300 = 0,9866 FAR = 0 FRR = 4/300 = 0,0133

Tabela A4 – Reconhecimento das Réplicas com textura *blotchy* obtidas com 400 épocas de treinamento.

		Classes Reconhecidas															Não reconhecidas				
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14					
Classes originais das réplicas	0	20																			
	1		20																		
	2			20																	
	3				20																
	4					20															
	5						19											1			
	6							19											1		
	7								20												
	8									20											
	9										20										
	10											20									
	11												11							9	
	12													20							
	13														19						1
	14																20				

TPR = 288/300 = 0,9600 FAR = 0 FRR = 12/300 = 0,0400

Tabela A5 – Reconhecimento das Réplicas com textura *bumpy* obtidas com 6 épocas de treinamento.

		Classes Reconhecidas															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Não reconhecidas
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						20										
	6							20									
	7								20								
	8									20							
	9										20						
	10											20					
	11												20				
	12													20			
	13														20		
	14															20	

TPR = 300/300 = 1 FAR = 0 FRR = 0

Tabela A6 – Reconhecimento das Réplicas com textura *bumpy* obtidas com 200 épocas de treinamento.

		Classes Reconhecidas															Não reconhecidas
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						20										
	6							20									
	7								20								
	8									20							
	9										20						
	10											20					
	11												20				
	12													20			
	13														20		
	14															20	

TPR = 300/300 = 1 FAR = 0 FRR = 0

Tabela A7 – Reconhecimento das Réplicas com textura *bumpy* obtidas com 300 épocas de treinamento.

		Classes Reconhecidas															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Não reconhecidas
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						15										5
	6							20									
	7								19	1							
	8									20							
	9										20						
	10											20					
	11												19	1			
	12													20			
	13														20		
	14															20	
		TPR = 293/300 = 0,9766 FAR = 2/300=0,0066 FRR = 7/300=0,0233															

Tabela A8 – Reconhecimento das Réplicas com textura *bumpy* obtidas com 400 épocas de treinamento.

		Classes Reconhecidas															Não reconhecidas
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						14										6
	6							20									
	7								20								
	8									20							
	9										20						
	10											20					
	11												20				
	12													20			
	13														20		
	14															20	

TPR = 294/300 = 0,9800 FAR = 0 FRR = 6/300=0,0200

Tabela A9 – Reconhecimento das Réplicas com textura *crosshatched* obtidas com 6 épocas de treinamento.

		Classes Reconhecidas															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Não reconhecidas
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						17										3
	6							14									6
	7								12								8
	8									20							
	9										20						
	10											20					
	11												11	1			8
	12													20			
	13														17		3
	14															20	

TPR = 271/300 = 0,9033 FAR = 1/300 = 0,0033 FRR = 29/300 = 0,0966

Tabela A10 – Reconhecimento das Réplicas com textura *crosshatched* obtidas com 200 épocas de treinamento.

		Classes Reconhecidas															Não reconhecidas
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						20										
	6							20									
	7								20								
	8									20							
	9										20						
	10											20					
	11												20				
	12													20			
	13														20		
	14															20	

TPR = 1 FAR = 0 FRR = 0

Tabela A11 – Reconhecimento das Réplicas com textura *crosshatched* obtidas com 300 épocas de treinamento.

		Classes Reconhecidas															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Não reconhecidas
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						20										
	6							20									
	7								20								
	8									20							
	9										20						
	10											20					
	11												19				1
	12													20			
	13														20		
	14															20	

TPR = 299/300 = 0,9966 FAR = 0 FRR = 1/300 = 0,0033

Tabela A12 – Reconhecimento das Réplicas com textura *crosshatched* obtidas com 400 épocas de treinamento.

		Classes Reconhecidas															Não reconhecidas
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						20										
	6							20									
	7								20								
	8									20							
	9										20						
	10											20					
	11												20				
	12													20			
	13														20		
	14															20	

TPR = 1 FAR = 0 FRR = 0

Tabela A13 – Reconhecimento das Réplicas com textura *fibrous* obtidas com 6 épocas de treinamento.

		Classes Reconhecidas															Não reconhecidas	
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14		
Classes originais das réplicas	0	20																
	1		20															
	2			20														
	3				20													
	4					20												
	5						20											
	6							19										1
	7								18									2
	8									19								1
	9										20							
	10											20						
	11												11			1		8
	12													18				2
	13														20			
	14																20	

TPR = 285/300 = 0,9500 FAR = 1/300 = 0,0033 FRR = 15/300 = 0,0500

Tabela A14 – Reconhecimento das Réplicas com textura *fibrous* obtidas com 200 épocas de treinamento.

		Classes Reconhecidas															Não reconhecidas
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						19										1
	6							16									4
	7								20								
	8									20							
	9										20						
	10											20					
	11												10	2		1	7
	12													19			1
	13														20		
	14															20	
		TPR = 284/300 = 0,9466 FAR = 3/300 = 0,0100 FRR = 16/300 = 0,0533															

Tabela A15 – Reconhecimento das Réplicas com textura *fibrous* obtidas com 300 épocas de treinamento.

		Classes Reconhecidas															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Não reconhecidas
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						19										1
	6							16									4
	7								20								0
	8									19							1
	9										20						
	10											20					
	11												11	1			8
	12													19			1
	13														20		
	14															20	
TPR = 284/300 = 0,9466 FAR = 1/300 = 0,0033 FRR = 16/300 = 0,0533																	

Tabela A16 – Reconhecimento das Réplicas com textura *fibrous* obtidas com 400 épocas de treinamento.

		Classes Reconhecidas															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Não reconhecidas
Classes originais das réplicas	0	20															
	1		20														
	2			20													
	3				20												
	4					20											
	5						19										1
	6							11									9
	7								19								1
	8									20							1
	9										20						
	10											20					
	11												7	4			9
	12													20			
	13														20		
	14															20	
		TPR = 275/300 = 0,9166 FAR = 4/300 = 0,0133 FRR = 25/300 = 0,0833															