



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

VICTOR HUGO DE AGUIAR ARRUDA

**DETECÇÃO DE POROS EM IMAGENS DE VESTÍGIOS DE IMPRESSÃO DIGITAL
POR MEIO DE REDES NEURAS CONVOLUCIONAIS**

Recife

2023

VICTOR HUGO DE AGUIAR ARRUDA

**DETECÇÃO DE POROS EM IMAGENS DE VESTÍGIOS DE IMPRESSÃO DIGITAL
POR MEIO DE REDES NEURAIAS CONVOLUCIONAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco como requisito parcial para obtenção do título de mestre em Engenharia Elétrica.
Área de Concentração: Comunicações.

Orientador: Prof. Dr. Juliano Bandeira Lima.

Recife

2023

Catálogo na fonte
Bibliotecário Gabriel Luz, CRB-4 / 2222

A779d Arruda, Victor Hugo de Aguiar.
 Detecção de poros em imagens de vestígios de impressão digital por meio
 de redes neurais convolucionais / Victor Hugo de Aguiar Arruda, 2023.
 79 f.: il.

 Orientador: Prof. Dr. Juliano Bandeira Lima.
 Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG.
Programa de Pós-Graduação em Engenharia Elétrica. Recife, 2023.
 Inclui referências.

 1. Engenharia elétrica. 2. Poros. 3. Vestígios de impressão digital. 4. Rede
 neural convolucional. 5. Análise em janelas. 6. Taxa de detecção real e falsa. I. Lima,
 Juliano Bandeira (Orientador). II. Título.

621.3 CDD (22. ed.)

UFPE

BCTG / 2023 - 231

VICTOR HUGO DE AGUIAR ARRUDA

**DETECÇÃO DE POROS EM IMAGENS DE VESTÍGIOS DE IMPRESSÃO DIGITAL
POR MEIO DE REDES NEURAIS CONVOLUCIONAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para a obtenção do título de mestre em Engenharia Elétrica. Área de concentração: Comunicações.

Aprovada em: 25-04-2023

BANCA EXAMINADORA

Prof. Dr. Juliano Bandeira Lima
(Orientador e Examinador Interno)
Universidade Federal de Pernambuco

Prof. Dr. João Marcelo Xavier Natário Teixeira
(Examinador Interno)
Universidade Federal de Pernambuco

Prof. Dr. Francisco Madeiro Bernardino Júnior
(Examinador Externo)
Universidade de Pernambuco

Recife

2023

AGRADECIMENTOS

Agradeço em primeiro lugar a Deus que me deu forças, saúde e vigor nessa jornada.

Aos meus pais, Ginaldo e Dilma, por terem investido em minha educação, apesar da escassez de recursos.

À minha esposa Nina, pela compreensão nos momentos de ausência e carinho e força nos momentos em que tudo parecia impossível.

Às minhas filhas, e jóias da minha vida, Clarice e Helena, cujos nascimentos me deram o maior título que hoje possuo: Pai.

Ao professor Juliano Bandeira, pela confiança, incentivo e orientação prestados neste trabalho.

Aos meus irmãos, Yann e Sofia, por fazerem parte da minha vida.

Aos colegas da Polícia Federal, pela troca de experiências que muito me ajudou no direcionamento deste trabalho.

Ao meu caçorro e terceiro filho Jack, por estar sempre quieto, mas em meus pés, nos momentos em que estava isolado estudando.

RESUMO

Neste trabalho são apresentadas e avaliadas técnicas que permitem a detecção de poros em imagens de impressão digital. A primeira avaliação se dá em relação às técnicas que utilizam ferramentas de filtragem, enquanto a segunda utiliza as redes neurais convolucionais (CNNs) para o processo de extração e detecção dos poros. Traçados os comparativos entre as técnicas existentes, foi implementada a topologia de rede neural proposta por (ALI; WANG; AHMAD, 2021). A escolha foi motivada pelo fato de o método em questão garantir os melhores resultados comparativamente aos outros, levando-se em conta as métricas denominadas de taxa de detecção real (R_T) e taxa de detecção falsa (R_F). Tais métricas são utilizadas para verificação da taxa de acerto e de erro da rede ao detectar um conjunto de poro na imagem de impressão digital, respectivamente. Como é de se imaginar, a rede será melhor quanto maior for a taxa de acerto e menor a taxa de erro. A primeira contribuição deste trabalho consiste na utilização de um banco de dados, que até onde se sabe na literatura pesquisada, nunca foi utilizado. Esse banco é composto por imagens reais obtidas através de fotografia forense em imagens de vestígios de impressão digital encontrados em local de crime. Além disso, o referido banco de dados utiliza imagens com resolução de 96 dpi, que é muito mais baixa que a resolução de 800 dpi estabelecida como mínima para esse processo de detecção de poros. Os resultados obtidos após as implementações anteriormente descritas revelam que, inicialmente, as taxas de R_T e de R_F não se mostraram muito satisfatórias, atingindo valores de $R_F = 77,9\%$ e $R_T = 56,5\%$, respectivamente. Nessa esteira, outra contribuição deste trabalho consiste na realização de uma análise em janelas, que corresponde à restrição da investigação em subimagens da imagem principal. Assim, foi verificado que para uma janela de tamanho 500×200 , o algoritmo apresentou resultados satisfatórios, atingindo um R_T de $80,8\%$ e um R_F de $4,5\%$, respectivamente. Ao final de tudo, a utilização do método proposto por (ALI; WANG; AHMAD, 2021), com os devidos ajustes, mostrou-se eficiente no processo de detecção de poros em imagens de vestígios de impressão digital encontrados em local de crime, fato que contribui significativamente no âmbito das Ciências Forenses, já que com a marcação de no mínimo 20 poros é possível individualizar uma pessoa.

Palavras-chave: poros; vestígios de impressão digital; rede neural convolucional; análise em janelas; taxa de detecção real e falsa.

ABSTRACT

In this work, techniques that allow the detection of pores in fingerprint images are presented and evaluated. The first evaluation takes place in relation to techniques that use filtering tools, while the second uses convolutional neural networks (CNNs) for the detection process and detection of pores. After comparing the existing techniques, the neural network topology proposed by (ALI; WANG; AHMAD, 2021) was implemented. The choice was motivated by the fact that the method in question guarantees the best results compared to the others, taking into account the metrics called true detection rate (R_T) and false detection rate (R_F). Such metrics are used to verify the success and error rate of the network when detecting a set of pores in the fingerprint image, respectively. As you can imagine, the network will be better the higher the hit rate and the lower the error rate. The first contribution of this work is the use of a database, which, as far as is known in the researched literature, has never been used. This database is composed of real images obtained through forensic photography on images of fingerprint traces found at a crime scene. In addition, said database uses images with a resolution of 96 dpi, which is much lower than the resolution of 800 dpi established as a minimum for this pore detection process. The results obtained after the previously described implementations reveal that, initially, the rates of R_T and R_F were not very satisfactory, reaching values of $R_F = 77,9\%$ and $R_T = 56,5\%$, respectively. In this vein, another contribution of this work consists in carrying out a window analysis, which corresponds to the restriction of the investigation in subimages of the main image. Thus, it was verified that for a window of size 500×200 , the algorithm presented satisfactory results, reaching an R_T of $80,8\%$ and an R_F of $4,5\%$, respectively. In the end, the use of the method proposed by (ALI; WANG; AHMAD, 2021), with the necessary adjustments, proved to be efficient in the process of detecting pores in images of fingerprint traces found at a crime scene, a fact that significantly contributes to the scope of Forensic Sciences, since with the marking of at least 20 pores it is possible to individualize a person.

Keywords: pores; fingerprint traces; convolutional neural network; analyze in windows; true and false detection rate.

LISTA DE FIGURAS

Figura 1 – Categorização das impressões digitais.	16
Figura 2 – Hierarquia dos níveis de características.	18
Figura 3 – Tipos de poros: fechado (marcado na cor verde); aberto em um dos lados (marcados como 1, na cor amarela); aberto nos dois lados (marcado como 2, na cor amarela); aberto, que se parece com o início e o fim de duas cristas diferentes (marcado como 3, na cor amarela).	19
Figura 4 – Esquema de uma estrutura híbrida de detecção de poros.	37
Figura 5 – Topologia de CNN utilizada para obtenção do mapa de intensidade de poros, X_θ	45
Figura 6 – Etapa de pós-processamento.	46
Figura 7 – Esquema do método LOOCV.	50
Figura 8 – Esquema de deslocamento de um <i>patch</i> 350×350 durante a fase de treinamento da rede, com sobreposição.	51
Figura 9 – Esquema de deslocamento de um <i>patch</i> 400×400 durante a fase de teste da rede, sem sobreposição.	52
Figura 10 – Função de perdas (<i>loss functions</i>) para as imagens de 1 a 8 do banco de teste.	59
Figura 11 – Fluxograma da simulação do algoritmo para cada uma das 8 imagens do conjunto de teste.	60
Figura 12 – Representação dos poros da imagem 1 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.	62
Figura 13 – Representação dos poros da imagem 2 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.	63
Figura 14 – Representação dos poros da imagem 3 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.	64

Figura 15 – Representação dos poros da imagem 4 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.	65
Figura 16 – Representação dos poros da imagem 5 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.	66
Figura 17 – Representação dos poros da imagem 6 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.	67
Figura 18 – Representação dos poros da imagem 7 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.	68
Figura 19 – Representação dos poros da imagem 8 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.	69
Figura 20 – Análise detalhada das subimagens de 1 e 8.	71

LISTA DE TABELAS

Tabela 1 – Arquitetura da CNND do método proposto por (LABATI et al., 2017). . . .	31
Tabela 2 – Arquitetura da CNNR do método proposto por (LABATI et al., 2017). . . .	32
Tabela 3 – Arquitetura da CNN do DeepPore.	34
Tabela 4 – Arquitetura da CNN do método proposto por (DAHIA; SEGUNDO, 2018). . . .	35
Tabela 5 – Arquitetura da CNN do DeepResPore.	36
Tabela 6 – Tabela comparativa entre os variados esquemas de detecção de poros existentes. . . .	44
Tabela 7 – Desempenho do esquema proposto após escolha do tamanho do <i>patch</i> e otimização dos hiperparâmetros pelo <i>Grid Search</i>	58
Tabela 8 – Desempenho do esquema proposto após aplicação da análise em janelas. . . .	61

SUMÁRIO

1	INTRODUÇÃO	12
1.1	OBJETIVOS	13
1.1.1	Objetivo geral	13
1.1.2	Objetivos específicos	13
1.2	ESTRUTURA DA DISSERTAÇÃO	14
2	CONTEXTUALIZAÇÃO	15
2.1	INTRODUÇÃO	15
2.2	AS IMPRESSÕES DIGITAIS	15
2.2.1	Níveis de características das impressões digitais	17
2.2.2	Poroscopia Forense	17
2.3	BREVE APANHADO SOBRE AS REDES NEURAIS ARTIFICIAIS	20
2.3.1	Redes Neurais Convolucionais	20
2.3.2	A operação de convolução	21
2.3.3	Redes Neurais Residuais	22
3	MÉTODOS PARA DETECÇÃO DE POROS	24
3.1	INTRODUÇÃO	24
3.2	MÉTODOS DE DETECÇÃO DE POROS BASEADOS EM FILTROS	25
3.2.1	Método proposto por Jain, Chen e Demirkus	25
3.2.2	Diferença de Gaussianas Adaptado (DoG adaptado)	26
3.2.3	Modelo de Poro Anisotrópico Dinâmico (DAPM)	27
3.2.4	Método proposto por Teixeira e Leite	28
3.2.5	Filtragem Dinâmica de Poros (DPF)	29
3.3	MÉTODOS DE DETECÇÃO DE POROS BASEADOS EM ANNS	31
3.3.1	Método proposto por Labati <i>et al.</i>	31
3.3.2	DeepPore	33
3.3.3	Método proposto por Dahia e Segundo	34
3.3.4	DeepResPore	35
3.3.5	Método proposto por Ali, Wang e Ahmad	36

3.3.5.1	Parte 1: extração dos poros	38
3.3.5.2	Parte 2: pós-processamento	39
3.3.5.2.1	Passo 1: detecção dos centroides	39
3.3.5.2.2	Passo 2: refinamento	42
3.4	COMPARAÇÃO DOS RESULTADOS OBTIDOS PELA UTILIZAÇÃO DOS DIFERENTES MÉTODOS DE DETECÇÃO DE POROS	43
4	IMPLEMENTAÇÃO DO ALGORITMO DE DETECÇÃO DE POROS PROPOSTO POR ALI, WANG E AHMAD COM ADAPTAÇÕES . . .	47
4.1	INTRODUÇÃO	47
4.2	DETALHES DO ALGORITMO IMPLEMENTADO	48
4.2.1	Banco de dados	48
4.2.2	Método de generalização do modelo	49
4.2.3	Esquema de particionamento das imagens para treinamento e teste da CNN	50
4.2.4	Procedimento de aumento de dados	52
4.2.5	Função de perda e otimizador	53
4.2.6	<i>Grid Search</i>	54
5	RESULTADOS	56
5.1	Introdução	56
5.2	Resultados Obtidos	56
5.2.1	Resultados obtidos na primeira etapa	56
5.2.2	Resultados obtidos na segunda etapa	61
6	CONCLUSÕES	72
6.1	Sugestões para Trabalhos Futuros	73
	REFERÊNCIAS	75

1 INTRODUÇÃO

Nas últimas décadas, o número de pesquisas relacionadas ao processo de identificação de uma pessoa tem crescido. Isso se deve, entre outras coisas, ao avanço na busca por sistemas de biometria mais confiáveis e que ofereçam níveis de segurança maiores aos usuários (SEMEDO et al., 2021). De maneira geral, os sistemas biométricos podem ser de dois tipos: os que utilizam características fisiológicas, a exemplo do rosto, impressão digital, geometria da mão, íris e retina; e os que utilizam as características comportamentais, como a marcha, a assinatura e a fala (KILIÇ; ASKERZADE; KAYA, 2020).

No âmbito criminal, as impressões digitais foram usadas pela primeira vez em 1982 por Juan Vucetich (U. S. Department of Justice, 2009) para solucionar um crime. Desde então, esse tipo de prova é considerada uma ferramenta poderosa para identificar vítimas e criminosos. As técnicas de obtenção de impressões digitais em cena de crime evoluíram ao longo do tempo, permitindo a extração desses vestígios em vários tipos de superfícies, seja utilizando métodos químicos ou físicos (HAZARIKA; RUSSELL, 2012).

Atualmente, as impressões digitais obtidas em cenas de crime são comparadas com os registros mantidos em um banco de dados de impressões digitais, através de um Sistema Automatizado de Identificação de Impressões Digitais (*Automated Fingerprint Identification System – AFIS*). No entanto, para a comparação de impressões digitais, é necessária qualidade técnica suficiente. Isso significa uma quantidade suficiente de minúcias para realizar a comparação com os modelos de impressão digital (SOUZA et al., 2021).

Como evolução ao modelo de individualização comumente utilizado nos AFIS – que se vale apenas de características de nível 1 e 2 das impressões digitais para comparação, ou seja, utiliza basicamente as pontas de linhas e bifurcações das cristas de ficção, que são as regiões da impressão digital em que estão localizados os poros – é que surge a análise de nível 3. O que se busca analisar sob essa nova ótica é a possibilidade de individualizar uma pessoa mediante outros atributos, por exemplo, a localização e o formato dos poros das impressões digitais (JAIN; CHEN; DEMIRKUS, 2007). Isso traz como vantagem a possibilidade de estabelecer a identidade de alguém com vestígios de fragmentos da ordem de 1 cm. Por outro lado, surge o inconveniente de as imagens dos vestígios necessitarem ter no mínimo 800 dpi (ZHANG et al., 2011).

É na busca por detectar os poros nas impressões digitais, e utilizá-las como maneira de identificação, que surge a ideia deste trabalho. Para tanto, foram utilizadas as redes neurais convolucionais (CNNs), que são treinadas mediante apresentação de diversas imagens cujos poros estão previamente definidos, e, como o próprio nome sugere, utilizam operações de convolução para geração dos mapas de poros que, posteriormente, serão utilizados no processo de identificação.

1.1 OBJETIVOS

1.1.1 Objetivo geral

O objetivo geral deste trabalho é investigar o desempenho de uma topologia específica de rede neural convolucional aplicada à detecção de poros em imagens de impressões digitais, utilizando, para isso, um banco de dados (*dataset*) de imagens de vestígios de impressões digitais encontradas em local de crime.

1.1.2 Objetivos específicos

Os objetivos específicos desta proposta são:

1. Comparar os diferentes métodos de detecção de poros existentes na literatura, tanto aqueles que utilizam as ferramentas de filtragem, quanto os que utilizam as redes neurais convolucionais.
2. Avaliar a taxa de acerto e de erro do algoritmo de detecção de poros quando da utilização da topologia proposta por (ALI; WANG; AHMAD, 2021), levando em conta o banco de dados disponível, que é composto por imagens de vestígios de impressões digitais encontradas em local de crime.
3. Utilizar a análise em janelas para geração de subimagens, de maneira a obter melhores resultados em relação às métricas de R_T e R_F mensuradas.
4. Empregar o *Grid Search* para escolha otimizada dos hiperparâmetros da CNN.

1.2 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está estruturada em seis capítulos, sendo o primeiro reservado à introdução e objetivos.

No Capítulo 2 é feita uma contextualização sobre as impressões digitais, abordando aspectos como as características que são usadas no processo de identificação e a Poroscopia Forense. Também são abordados resumidamente os conceitos de redes neurais artificiais, redes neurais convolucionais e redes neurais residuais.

No Capítulo 3 são abordados vários métodos de detecção de poros, partindo dos métodos de detecção baseados em filtros, até os mais modernos que utilizam as CNNs.

As técnicas propostas nesta dissertação estão contempladas no Capítulo 4, em que são apresentados os detalhes da implementação, englobando o banco de dados, método de generalização do modelo, o esquema de particionamento das imagens para treinamento e teste, o procedimento de aumento de dados, função de perda, otimizador e a implementação do *Grid Search*.

O Capítulo 5 é dedicado aos resultados de simulação. São apresentados os resultados obtidos das taxa de acerto e de erro das 8 imagens do conjunto de teste, analisando tanto a imagem como um todo, quanto as subimagens destas (análise em janelas).

As conclusões da dissertação e sugestões para trabalhos futuros são apresentadas no Capítulo 6

2 CONTEXTUALIZAÇÃO

2.1 INTRODUÇÃO

Neste Capítulo, são abordados temas importantes para compreensão do trabalho como um todo. Na Seção 2.2, busca-se contextualizar o leitor sobre o tema das impressões digitais. Para isso, é feita uma análise sobre os conjuntos de classificação das impressões; o processo de identificação através de minúcias e poros; e, também, é abordado o impacto desse tema nas Ciências Forenses. Na Seção 2.3, é abordada brevemente a temática das redes neurais artificiais. Mais especificamente, são tratados temas como as redes neurais convolucionais; a operação de convolução; e as redes neurais residuais.

Esse arcabouço temático envolve assuntos aparentemente tão distintos, mas que encontram convergência. Isso porque, para esta dissertação, as redes neurais artificiais constituem a ferramenta computacional necessária à aplicação das operações de detecção de poros em imagens de vestígios de impressão digital encontradas em local de crime. Os detalhes de como é realizado o procedimento de detecção serão tratados nos Capítulos 3 e 4 deste trabalho.

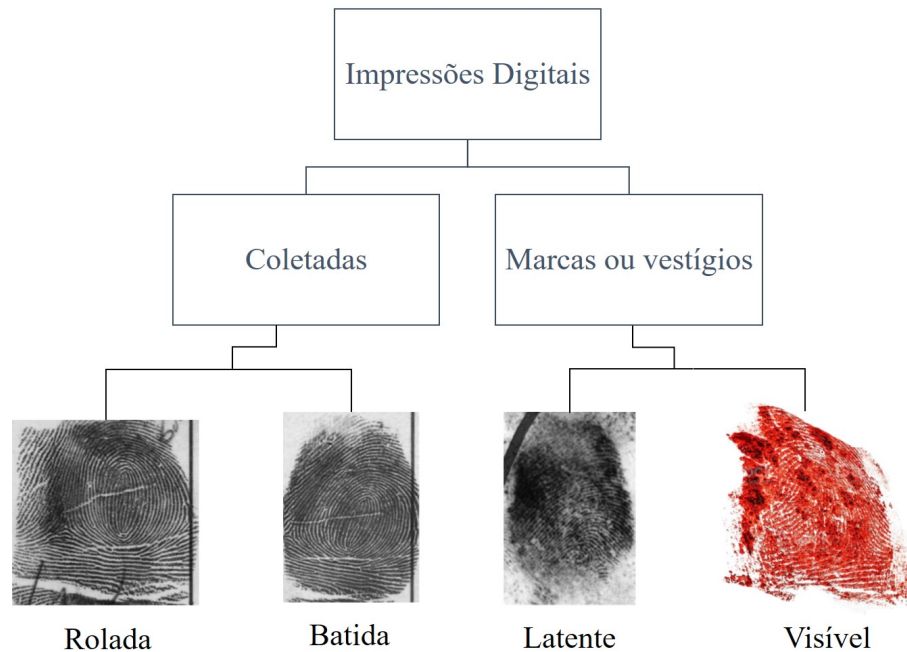
2.2 AS IMPRESSÕES DIGITAIS

Devido a preocupações com segurança, privacidade e fraude de identidade, a demanda por sistemas confiáveis de identificação pessoal aumentou. Nessa esteira, a impressão digital tornou-se uma das bioassinaturas mais usadas para reconhecimento pessoal devido à singularidade, permanência e conveniência de uso (MALTONI et al., 2009).

No campo criminal, o processo de identificação ganhou ainda mais notoriedade com a chegada dos Sistemas Automatizados de Identificação de Impressão Digital (*Automated Fingerprint Identification System* – AFIS), que executam o processo de busca e comparação de vestígios de impressão digital com o banco de dados existente (ALI; WANG; AHMAD, 2021).

Aprofundando a análise sobre essa temática, compete esclarecer que as impressões digitais podem ser categorizadas, basicamente, em dois conjuntos: as impressões coletadas, que abarcam tanto as impressões roladas quanto as impressões batidas; e as marcas ou vestígios de impressões digitais, que podem ser visíveis ou latentes (DVORNYCHENKO; GARRIS, 2006.;

Figura 1 – Categorização das impressões digitais.



Fonte: o autor, adaptado de (DVORNYCHENKO; GARRIS, 2006.).

KOMARINSKI, 2004; CHAMPOD et al., 2004). A Figura 1 ilustra essa categorização.

De maneira geral, as impressões digitais roladas são adquiridas rolando o dedo de uma extremidade a outra da unha. Essa forma de coleta contém o máximo de detalhes das cristas de fricção. Por outro lado, as impressões digitais batidas são impressões de toque, ou seja, contêm basicamente a região central da impressão e, desse modo, possuem menos detalhes que a rolada.

As latentes, por sua vez, são vestígios invisíveis de impressões digitais apostas em algum tipo de superfície, sendo necessário algum processo físico ou químico para ficarem aparentes (CHAMPOD et al., 2004). Esse tratamento é denominado de revelação de impressão digital latente. Por fim, existem também os vestígios de impressões digitais visíveis, que, como o próprio nome sugere, são visíveis sem a necessidade de tratamento específico (CHAMPOD et al., 2004). Podemos citar como exemplo as impressões digitais deixadas mediante marcas de sangue.

Esses dois últimos tipos de impressão digital têm importância particular para as ciências forenses, haja vista que tratam-se de vestígios deixados muitas vezes de maneira não intencional. Entretanto, fatores como baixa qualidade, área pequena e distorções provocadas por borrões, marcas de arrasto ou textura irregular das superfícies, afetam o processo de identificação dessas impressões (AGARWAL; BANSAL, 2021). Assim, para contornar esses problemas, é necessário

fazer uma análise pormenorizada de tais impressões, conforme será abordado na Seção 2.2.1.

2.2.1 Níveis de características das impressões digitais

Existem características nas impressões digitais que são hierarquizadas em 3 níveis. As de nível 1 referem-se aos aspectos morfológicos, como o fluxo de linhas e o tipo fundamental da impressão. Os tipos fundamentais podem ser subclassificados em 4 categorias: os Arcos; as Presilhas Internas; as Presilhas Externas; e os Verticilos.

No nível 2 estão presentes as minúcias, que consistem em pontos característicos singulares que permitem a identificação. Podem ser citados como exemplos de minúcias as bifurcações; terminações de cristas (também chamadas de pontas de linhas); ilhotas; encerros; entre outros. Já no nível 3 as características são mais detalhadas, observando-se, por exemplo, os poros; a largura e forma das cristas de fricção; as linhas incipientes; as linhas alodactilares; as cicatrizes; entre outros (ALSHDADI et al., 2020). Os diferentes níveis de hierarquia dessas características estão ilustrados na Figura 2 (JAIN; CHEN; DEMIRKUS, 2007).

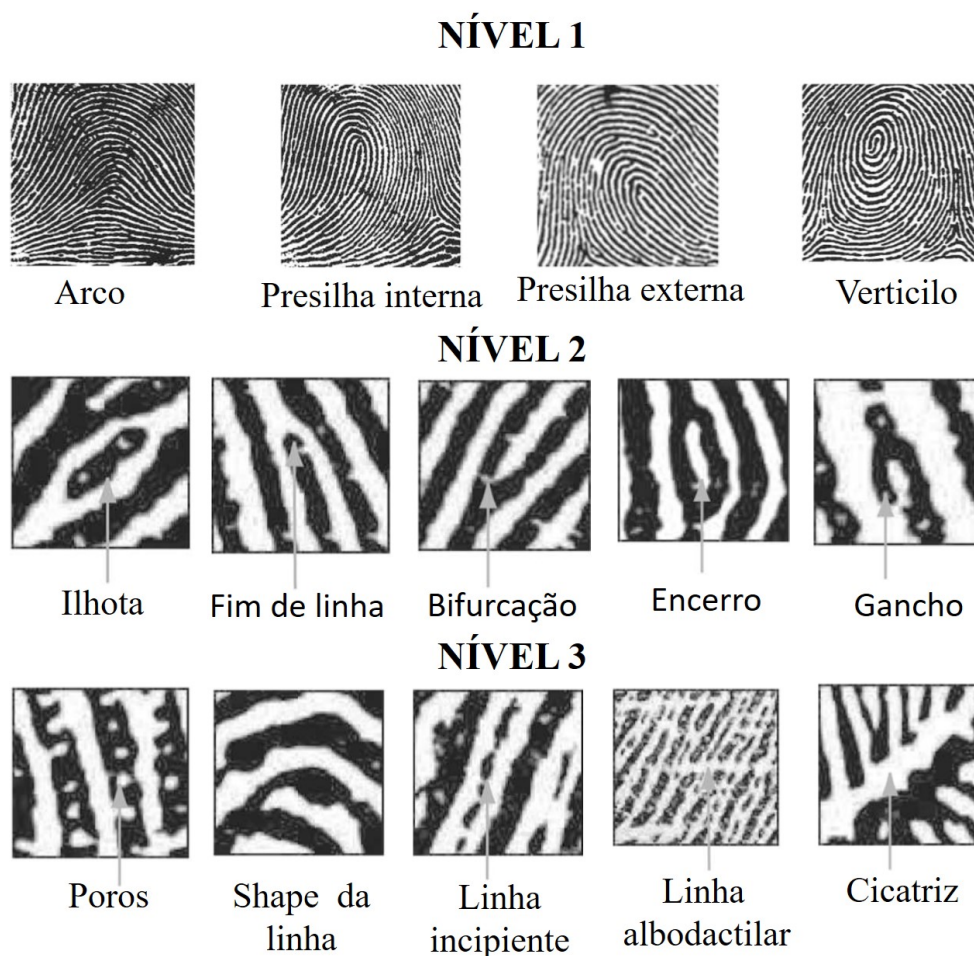
Nas aplicações forenses, os examinadores de vestígios latentes ou visíveis frequentemente aproveitam as características de nível 3 para auxiliar na identificação, apesar dos AFIS usualmente executarem o algoritmo de busca e *match* utilizando apenas as características de nível 1 e 2. Isso acontece porque normalmente o padrão de resolução de impressão digital do *Federal Bureau of Investigation* (FBI) para o AFIS é de 500 *dots per inch* (dpi), que não é o mais adequado para capturar características de nível 3, como os poros (JAIN; CHEN; DEMIRKUS, 2007).

Com os avanços na tecnologia de detecção de impressões digitais, muitos sensores estão equipados com capacidade de digitalização em resolução dupla (500 dpi/1.000 dpi). Isso tem tornado cada vez mais viável o uso da identificação através das minúcias de nível 3, sobretudo através dos poros, já que as características de nível 3 são, geralmente, observáveis em imagens de impressão digital com resolução superior a 800 dpi (ZHANG et al., 2011).

2.2.2 Poroscopia Forense

Entre as características de nível 3, ganha destaque o processo de identificação através dos poros, denominado de Poroscopia Forense. Este método mostrou ser eficaz para o reconhecimento biométrico (ZHAO et al., 2009), dadas as propriedades de imutabilidade e singularidade em termos de localização das coordenadas dos poros; da forma dos poros (se poro aberto ou fechado);

Figura 2 – Hierarquia dos níveis de características.



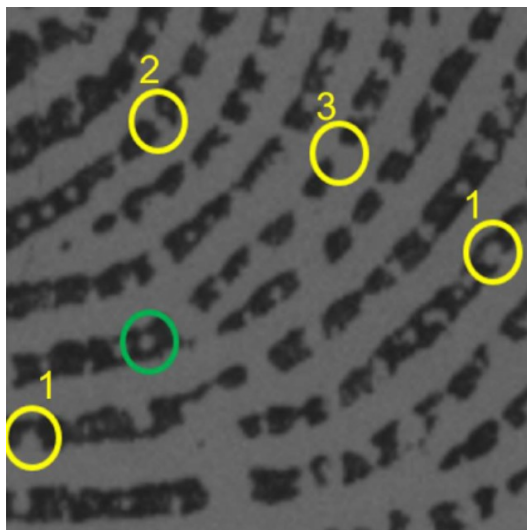
Fonte: o autor, adaptado de (JAIN; CHEN; DEMIRKUS, 2007).

e do tamanho deles (JAIN; CHEN; DEMIRKUS, 2007; CHAMPOD et al., 2004).

Conforme citado anteriormente, os poros podem ser classificados em dois tipos: em poros fechados, ou seja, aqueles que encontram-se dentro de uma crista; ou poros abertos, ou seja, aqueles que encontram-se parcialmente dentro de uma crista (ASHBAUGH, 1999). A Figura 3 ilustra esses tipos de poros, sendo que o círculo de cor verde representa os fechados e os círculos amarelos os abertos. De qualquer forma, segundo (JAIN; CHEN; DEMIRKUS, 2007) cerca de 20 poros são suficientes para estabelecer a identidade de uma pessoa.

O que também se observa é que em um pequeno segmento da imagem da impressão digital existe grande quantidade de poros. Por exemplo, ao longo de 1 cm de comprimento de uma crista, existem cerca de 9 a 18 poros (JAIN; CHEN; DEMIRKUS, 2007). Assim, apenas um pequeno segmento da imagem da impressão digital é necessário para ter um número suficiente

Figura 3 – Tipos de poros: fechado (marcado na cor verde); aberto em um dos lados (marcados como 1, na cor amarela); aberto nos dois lados (marcado como 2, na cor amarela); aberto, que se parece com o início e o fim de duas cristas diferentes (marcado como 3, na cor amarela).



Fonte: (ALI; WANG; AHMAD, 2021).

de poros para estabelecer a identificação.

Em contraste, para se estabelecer a identificação utilizando apenas características de nível 2 é necessário, em média, 12 pontos característicos, tomando como base a prática forense da Polícia Federal do Brasil. Sendo que, de acordo com (GARRIS, 2000) o número médio de minúcias encontradas são 27 para os vestígios latentes, ao passo que são encontradas 106 para as impressões roladas (PAULINO; FENG; JAIN, 2013). Ou seja, é muito comum serem encontrados vestígios com quantidade insuficiente de minúcias de nível 2 para a identificação.

Assim, em uma mesma área pode não haver minúcias de nível 2 suficientes, mas pode haver de nível 3 suficientes para estabelecer a identificação (ANAND; KANHANGAD, 2021). Esse fato, atrelado à imutabilidade e singularidade dos poros é uma razão convincente para usá-los no processo de identificação.

Apesar das vantagens da Poroscopia Forense, vale destacar que é imprescindível que as coordenadas dos poros sejam detectadas com precisão. No Capítulo 3 serão abordados diversos métodos de detecção de poros presentes na literatura.

2.3 BREVE APANHADO SOBRE AS REDES NEURAIIS ARTIFICIAIS

O Deep Learning é uma das vertentes do Machine Learning, cujos algoritmos são inspirados na funcionalidade e estrutura do cérebro humano. Tais estruturas são conhecidas como Redes Neurais Artificiais (*Artificial Neural Networks* – ANNs) que são formadas por uma série de nós de processamento interconectados (neurônios) que transformam uma coleção de entradas em uma coleção de saídas. A transformação é definida basicamente por duas variáveis: os pesos que são aplicados a cada entrada; e um viés, também chamado de *bias* (BACANIN et al., 2021).

O valor dos pesos e vieses são atualizados durante o processo de aprendizado, de forma a reduzir a taxa de erro de classificação, que é medida pela função de perda (*loss function*). O ajuste dos pesos e do valor do *bias* é chamado de treinamento da rede neural, que pertence ao tipo de aprendizado supervisionado, ou seja, aquele que aprende com os dados rotulados (BACANIN et al., 2021). Tal procedimento de aprendizado supervisionado permite que as redes neurais aprendam com exemplos e adaptem sua estrutura interna para adquirir a capacidade de generalização (HAYKIN, 2011).

Por outro lado, a maioria das redes neurais artificiais descritas na literatura consiste em camadas de neurônios que processam dados na forma de sinais unidimensionais (LABATI et al., 2017). Assim, em muitos casos, uma etapa de extração de características calcula os sinais de entrada unidimensionais (LABATI et al., 2016) a partir de dados com maior dimensionalidade (por exemplo, uma imagem). A etapa de extração de características requer um conhecimento prévio do problema, de maneira a reduzir de forma eficiente a dimensionalidade dos dados de entrada, mantendo as informações mais significativas (LECUN; BENGIO; HINTON, 2015).

2.3.1 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (*Convolutional Neural Networks* – CNNs), também conhecidas como ConvNets, são uma forma particular de redes neurais artificiais, cujas camadas possuem uma estrutura que permite processar dados diretamente na forma de arranjos multidimensionais, como imagens, por exemplo (LECUN; BENGIO; HINTON, 2015). Assim, entre as vantagens das CNNs em relação às redes neurais capazes de processar apenas sinais de entrada unidimensionais, está o fato de que as CNNs não requerem uma etapa preliminar de extração de recursos porque uma camada convolucional genérica pode extrair os recursos visuais buscados.

Por esta razão, pesquisadores usam CNNs em diferentes cenários de aplicação, como

classificação de objetos (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) ou processamento de linguagem natural (KALCHBRENNER; GREFFENSTETTE; BLUNSOM, 2014). Na biometria, existem métodos de reconhecimento baseados em CNNs para diferentes características, tais como: face (DING; TAO, 2015), íris (LIU et al., 2016) e impressão digital (LABATI et al., 2017; JANG et al., 2017; DAHIA; SEGUNDO, 2018; ANAND; KANHANGAD, 2019; ALI; WANG; AHMAD, 2021).

Em particular, vale destacar duas propriedades interessantes das ConvNets. A primeira é que os padrões que elas aprendem são invariantes a deslocamento (*shift invariant*) ou invariantes a espaço (*space invariant*). Assim, depois de aprender um certo padrão no canto inferior direito de uma imagem, uma ConvNet pode reconhecê-lo em qualquer lugar, por exemplo, no canto superior esquerdo. Isso faz com que esse tipo de rede precise de menos amostras de treinamento para aprender representações que têm poder de generalização (CHOLLET, 2018).

A segunda propriedade é que elas podem aprender hierarquias espaciais de padrões. Isso quer dizer que uma primeira camada de convolução aprenderá pequenos padrões locais, como arestas, por exemplo. Já numa segunda camada de convolução, por sua vez, aprenderá padrões maiores, feitos das feições das primeiras camadas, e assim por diante. Isso permite que as CNNs aprendam com eficiência conceitos visuais cada vez mais complexos e abstratos (CHOLLET, 2018).

2.3.2 A operação de convolução

Já que as CNNs constituem uma importante ferramenta computacional para aprendizado de máquina, vale mencionar o porquê de se chamarem redes convolucionais. Na verdade, esse nome está atrelado ao fato de que esse tipo de rede utiliza operações de convolução sobre tensores $3D$ (chamados de mapas de recursos) que, por serem $3D$, possuem dois eixos espaciais (altura e largura) e um eixo de profundidade (também chamado de eixo de canais).

A operação de convolução, por sua vez, transforma o mapa de recursos de entrada em uma mapa de recursos de saída que ainda é um tensor $3D$ (CHOLLET, 2018). Entretanto, a profundidade da saída da operação de convolução não representa mais o eixo dos canais como na entrada. Na saída, a profundidade representa as respostas ao mapa de recursos quando da aplicação dos filtros convolucionais. Assim, por exemplo, caso o mapa de recursos seja uma imagem em tons de cinza de tamanho 30×30 , onde seja utilizado um *kernel* de convolução

de tamanho 3×3 com 32 filtros convolucionais, o resultado será 32 canais de saída contendo imagens de tamanho 28×28 , considerando um *same padding* ou zero *padding* e *stride* igual a 1. Ao final será produzida uma saída de tamanho $(28, 28, 32)$, em detrimento de uma entrada de tamanho $(30, 30, 1)$.

Assim, logo percebe-se que as operações de convolução são definidas por dois parâmetros principais. O primeiro é o tamanho do *kernel* do filtro, cujos tamanhos são tipicamente 3×3 , ou 5×5 , nas CNNs. O segundo é a profundidade de saída do mapa de recursos, que é dado pelo número de filtros convolucionais utilizados. Em geral a profundidade inicia em 32 e aumenta para valores de 64, 128 e 256 (CHOLLET, 2018).

2.3.3 Redes Neurais Residuais

As CNNs, conforme abordado na Seção 2.3.1, têm uma estrutura em camadas e cada camada consiste em filtros convolucionais. Enquanto as primeiras camadas de filtros convolucionais representam características de baixo nível, como bordas e contrastes de cores, as camadas mais profundas tentam capturar formas mais complexas (SERMANET et al., 2014). Mesmo assim, embora as CNNs possam frequentemente ter um desempenho superior na classificação quando comparado a outros tipos de rede, elas são difíceis de treinar, por dois motivos principais.

O primeiro é o Problema da Dissipação do Gradiente (*The Vanish Gradient Problem*), em que à medida que mais camadas são adicionadas à rede, os gradientes da função de perda se aproximam de zero, dificultando o treinamento da rede. O segundo é a dificuldade de treinar a rede, já que o modelo fica com mais parâmetros (SRIVASTAVA; GREFF; SCHMIDHUBER, 2015). Portanto, a conclusão que se chega é que CNNs são mais difíceis de treinar, embora tenham melhor desempenho de classificação.

Uma maneira eficaz de resolver os problemas de treinamento citados anteriormente consiste em utilizar as Redes Neurais Residuais (*Residual Neural Networks* – RNN) (HE et al., 2016), também conhecidas como ResNets. Estas são uma das espécies de arquitetura de CNNs existentes, cuja principal diferença em relação às ConvNets normais é que têm links de atalho paralelos às suas camadas convolucionais normais. Ao contrário das camadas de convolução, esses links de atalho estão sempre ativos e os gradientes podem se propagar facilmente de volta, resultando em um treinamento mais rápido (KILIÇ; ASKERZADE; KAYA, 2020).

As Seções 3.3.4 e 3.3.5 descrevem alguns métodos para detecção de poros em imagens

de impressões digitais que utilizam esse tipo de arquitetura de rede, e que merecem destaque devido aos bons resultados que apresentaram.

3 MÉTODOS PARA DETECÇÃO DE POROS

3.1 INTRODUÇÃO

Os poros são minúcias de nível 3 que além de estarem presentes em grande número nas impressões digitais, possuem outras duas características importantes que são a estabilidade e a singularidade. Para o campo criminal, as três características citadas são importantes, entretanto o fato dos poros estarem presentes em grande número merece atenção especial, já que, com pequenas áreas de impressão digital disponíveis, é possível determinar o autor do fragmento de impressão. Todavia, a utilização de poros como forma de biometria é relativamente recente, existindo basicamente duas formas de detecção: através de algoritmos baseados em filtros; e de algoritmos baseados em CNNs.

A ideia básica dos algoritmos baseados em filtros consiste em utilizar alguma ferramenta de filtragem, como, por exemplo, filtros de Gabor combinados com *wavelet* do tipo Chapéu Mexicano (JAIN; CHEN; DEMIRKUS, 2007); Gaussianas (ZHAO et al., 2010a; ZHAO et al., 2010b); superfície de decisão (TEIXEIRA; LEITE, 2014); ou ainda utilizar técnicas de binarização de imagens com ferramentas geométricas de detecção de poros (LEMES et al., 2014).

Avanços recentes na detecção de objetos utilizando ANNs trouxeram melhorias quanto à forma de detecção de poros nas impressões digitais. (LABATI et al., 2017), por exemplo, utilizaram duas CNNs em seu método, sendo uma para o processo de extração de poros e a outra para a etapa de refinamento. Já (JANG et al., 2017), por exemplo, utilizaram uma CNN treinada para estimar o mapa de intensidade de poros, seguido por um algoritmo denominado de *Pore Intensity Refinement* (PIR), que busca localizar no mapa de poros máximos locais e não apenas globais.

Já (DAHIA; SEGUNDO, 2018) utilizaram o conceito de *Fully Convolutional Neural Network* (FCN) que é amplamente utilizado em tarefas de segmentação de imagens, permitindo que as imagens de treinamento e teste não tenham, necessariamente, o mesmo tamanho. Já para (ANAND; KANHANGAD, 2019) e para (ALI; WANG; AHMAD, 2021) foram utilizadas as RNNs que, conforme citado na Seção 2.3.3, diminuem o Problema de Dissipação do Gradiente.

Outros autores cujos trabalhos não serão tratados nesta dissertação também trouxeram

diferentes abordagens no que diz respeito ao uso das redes neurais. (WANG et al., 2017) utilizaram a rede U-Net para segmentar os poros e sulcos, obtendo bons resultados. (SU et al., 2017), por sua vez, usaram um modelo de CNN treinado para reconhecer poros em um pequeno bloco de imagem. Enquanto isso, (LIU et al., 2020) propuseram um modelo de CNN combinado com operações lógicas.

3.2 MÉTODOS DE DETECÇÃO DE POROS BASEADOS EM FILTROS

O objetivo desta seção é analisar alguns métodos de detecção de poros baseados em filtros de imagens, cujos trabalhos são recorrentemente citados na literatura. Além disso, serão traçadas comparações entre os métodos descritos e elencados pontos positivos e negativos de cada um deles.

3.2.1 Método proposto por Jain, Chen e Demirkus

O método proposto por (JAIN; CHEN; DEMIRKUS, 2007) pode ser dividido em duas etapas e baseia-se no fato de que os poros encontram-se localizados nas cristas de fricção das impressões digitais.

Na primeira etapa, são utilizados filtros de Gabor (HONG; WAN; JAIN, 1998) para realçar as cristas das impressões digitais. Esses filtros possuem a seguinte forma

$$G(x, y : \theta, f) = \exp \left\{ -\frac{1}{2} \left[\frac{x_{\theta}^2}{\delta_x^2} + \frac{y_{\theta}^2}{\delta_y^2} \right] \cos(2\pi f x_{\theta}) \right\}, \quad (3.1)$$

em que θ e f são a orientação e a frequência do filtro, respectivamente; δ_x e δ_y são os desvios-padrões da Gaussiana envolvente ao longo dos eixos x e y , respectivamente; e (x_{θ}, y_{θ}) representam a posição de um ponto (x, y) que sofreu uma rotação no sentido horário por um ângulo $(90^{\circ} - \theta)$. Os quatro parâmetros $(\theta, f, \delta_x, \delta_y)$ do filtro de Gabor são determinados empiricamente com base na frequência de crista e orientação da imagem da impressão digital.

A segunda etapa consiste em realçar os poros das impressões digitais. Para isso é utilizado um filtro passa-banda para capturar as respostas de alta frequência, que nesse caso são marcadas pela transição abrupta da cor branca dos poros para a preta das cristas. Para essa função foi escolhida a wavelet de Chapéu Mexicano (BURRUS; GUO, 1998), que aplicada em uma imagem de entrada $f(x, y) \in \mathbb{R}^2$ tem a resposta em frequência w , dada por

$$w(s, a, b) = \frac{1}{\sqrt{s}} \iint_{\mathbb{R}^2} f(x, y) \phi \left(\frac{x-a}{s}, \frac{y-b}{s} \right) dx dy, \quad (3.2)$$

em que s é o fator de escala ($= 1,32$) e (a, b) e f são os parâmetros de deslocamento.

Por fim, a resposta da *wavelet* é normalizada usando normalização min-max, que deixa os valores dos *pixels* da imagem dentro da faixa de $(0 - 255)$, evidenciando ainda mais as regiões dos poros.

Uma desvantagem desse método consiste no fato de que os poros são tratados como isotrópicos e estáticos (ou de escala unitária). Assim, os poros são considerados basicamente estruturas circulares e a frequência de crista/vale na impressão digital é considerada constante, o que não é necessariamente verdade. Assim, como evolução, surge o método o abordado na Seção 3.2.2.

3.2.2 Diferença de Gaussianas Adaptado (DoG adaptado)

O método de extração de poros denominado de Diferença de Gaussianas (*Difference of Gaussians* – DoG) foi proposto inicialmente por (PARSONS et al., 2008). A ideia do método consiste basicamente em usar um filtro passa-banda para detectar círculos nas impressões digitais. Em outras palavras, o método pressupõe que os poros aparecem como objetos circulares. O problema desse método é que os autores não consideraram a variação das escalas de poros e vales nas imagens de impressão digital, mas simplesmente usaram uma escala unitária em seu modelo. A fim de contornar o problema de escala, (ZHAO et al., 2010b) propuseram o método DoG adaptado.

Considerando que a escala de poros e vales geralmente não é uniforme em uma imagem de impressão digital e que diferentes impressões digitais podem ter diferentes frequências de crista/vale, uma imagem da impressão digital foi particionada em vários blocos e as escalas de filtros gaussianos foram estimadas para cada bloco. Tomando uma imagem de bloco I_B como exemplo, suponha que o período médio de crista sobre este bloco seja p , então p é uma boa medida da escala nesse bloco. Assim, são definidos os desvios-padrão dos dois filtros Gaussianos como k_1p e k_2p , respectivamente (em que, $0 < k_1 < k_2$ são duas constantes). As saídas dos filtros são, então, calculadas por

$$\begin{cases} F_1 = G_{k_1p} * I_B, \\ F_2 = G_{k_2p} * I_B, \end{cases} \quad (3.3)$$

$$\begin{cases} G_\sigma(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x^2+y^2)/2\sigma^2} - m_G, \\ |x|, |y| \leq 3\sigma, \end{cases} \quad (3.4)$$

em que "*"denota convolução e m_G é usado para normalizar o filtro Gaussiano para ser de média zero. Observe que as configurações de k_1 e k_2 devem levar em consideração as larguras das cristas e vales e o tamanho dos poros. Nos experimentos de (ZHAO et al., 2010b), os valores de k_1 e k_2 foram escolhidos empiricamente. As saídas de filtragem F_1 e F_2 são normalizadas para $[0, 1]$ e binarizadas, resultando em B_1 e B_2 . O filtro gaussiano de pequena escala G_{k_1p} realçará os poros e os vales, enquanto o filtro de grande escala G_{k_2p} realçará apenas os vales. Portanto, subtraindo B_2 de B_1 , obtêm-se resultado final da extração de poros.

Apesar de uma evolução para o método DoG convencional, este método possui como limitação o fato de os poros serem modelados como estruturas isotrópicas, ou seja, de conformação circular, quando é sabido que isso não é bem verdade. Para compensar este problema, (ZHAO et al., 2010a) propuseram um modelo de poro anisotrópico dinâmico, conforme pormenorizado na Seção 3.2.3.

3.2.3 Modelo de Poro Anisotrópico Dinâmico (DAPM)

O método de detecção de poros denominado de Modelo de Poro Anisotrópico Dinâmico (*Dynamic Anisotropic Pore Model* – DAPM), foi proposto por (ZHAO et al., 2010a). A discussão inicial desse método é em torno da anisotropia dos poros. Para isso, foram investigadas as aparências espaciais de centenas de poros em imagens de impressões digitais. Basicamente, foram catalogados três tipos de poros: poros fechados; abertos em um lado; abertos em dois lados. O que se percebeu no estudo é que, conforme esperado, os poros fechados e, especialmente, os abertos, não são isotrópicos.

Essas observações mostram claramente que modelar os poros em isotrópicos e estáticos (ou de escala unitária) não é razoável. Para representar os poros com mais precisão, é necessário um modelo de poros que tenha dois parâmetros para ajuste: um de escala; e um de orientação. Ao aplicar este modelo a um poro real, esses dois parâmetros são determinados de acordo com as características locais da crista (ou seja, orientação e frequência da crista). Portanto, o DAPM é definido como

$$\begin{cases} P_0(i, j) = e^{(-j^2/2\sigma^2)} \cos\left(\frac{\pi}{3\sigma}i\right), \\ -3\sigma \leq i, j \leq 3\sigma, \end{cases} \quad (3.5)$$

$$\left\{ \begin{array}{l} P_{\sigma}(i, j) = ROT(P_0, \theta) = e^{-(\hat{j}^2/2\sigma^2)} \cos\left(\frac{\pi}{3\sigma}\hat{i}\right), \\ \hat{i} = i \cos(\theta) - j(\theta), \\ \hat{j} = i(\theta) + j \cos(\theta), \\ -3\sigma \leq i, j \leq 3\sigma, \end{array} \right. \quad (3.6)$$

em que (3.5) é o modelo de referência (o modelo de grau zero) e (3.6) é um modelo rotacionado. Aqui, σ é o parâmetro de escala que é usado para controlar o tamanho dos poros. Ele pode ser determinado pela frequência local das cristas das impressões digitais. Por outro lado, θ é o parâmetro de orientação que é usado para controlar a direção do modelo do poro. Ele pode ser estimado levando em conta a orientação das cristas.

Em resumo, a ideia do método consiste basicamente em particionar a imagem de impressão digital em blocos. De acordo com a orientação e a frequência da crista dominante do bloco, obtém-se modelo de poro apropriado. Por fim, tais poros são extraídos.

3.2.4 Método proposto por Teixeira e Leite

O método proposto por (TEIXEIRA; LEITE, 2014) é baseado principalmente na análise do perfil de distribuição espacial dos poros. O objetivo é criar uma superfície de decisão local que possa ser usada para classificar o conjunto de poros de uma imagem, levando em consideração informações de distância. Como essa análise depende apenas das coordenadas cartesianas dos poros, qualquer método de detecção de poros existente, como, por exemplo, (JAIN; CHEN; DEMIRKUS, 2007; ZHAO et al., 2010b; ZHAO et al., 2010a; LEMES et al., 2014)) pode ser usado. Assim, o método em tela é uma espécie de refinamento dos outros métodos de detecção.

O método consiste basicamente na criação de uma superfície de decisão e baseia-se em duas premissas: a primeira de que a distância entre poros em uma mesma crista é proporcional à largura da crista (ASHBAUGH, 1999; CUMMINS; MIDLO, 1943); a segunda afirma que existem algumas variações regionais na largura das cristas epidérmicas (CUMMINS; MIDLO, 1943). Por conseguinte, a superfície de decisão mede a largura, tomando uma região em que a largura da crista se mantenha quase que constante, e estima a distância entre os poros.

Sob o ponto de vista prático, a primeira etapa do método consiste em dividir uma imagem de impressão digital em blocos não sobrepostos de tamanho $W \times W$. A etapa seguinte consiste em estimar a largura média ($\mathbf{x}$) das cristas em cada um desses blocos. Isso é importante para definir a função de decisão local que classifica pares de poros adjacentes, considerando uma

função de decisão linear simples aplicada a $(\mathbf{x})$, como segue

$$\begin{cases} \hat{w}(d) = \hat{w}_1, & \text{se } \mathbf{w}^T \mathbf{x} - d < 0, \\ \hat{w}_0, & \text{caso contrário,} \end{cases} \quad (3.7)$$

em que d representa a distância entre os poros, $\mathbf{w}$ é o vetor de peso, x é a largura média das cristas em cada bloco, $\hat{w}_1$ e $\hat{w}_0$ são os rótulos das classes correspondentes a verdadeiros positivos e falsos positivos, respectivamente. Os valores do vetor peso $\mathbf{w}$ em (3.7) podem ser calculados explorando a relação entre a distância intraporos e a largura da crista. Nos experimentos o autor setou $\mathbf{w} = \begin{pmatrix} 8,0 \\ -1,93 \end{pmatrix}$.

A última etapa consiste em realizar um processo iterativo que faz uma varredura dentro da janela $W \times W$ e mede as distâncias locais entre as coordenadas cartesianas daquilo que seriam poros. Isso, atrelado aos critérios de distância previamente setados, faz com que sejam distinguidos com mais precisão poros de não poros, melhorando assim a taxa de falsos positivos, que é uma das vantagens desse método.

3.2.5 Filtragem Dinâmica de Poros (DPF)

Alguns métodos analisados até agora possuem como limitadores o fato de alguns deles serem não adaptativos (JAIN; CHEN; DEMIRKUS, 2007; ZHAO et al., 2010b), fato que foi superado com a chegada de métodos adaptativos (ZHAO et al., 2010a; TEIXEIRA; LEITE, 2014). Entretanto, mesmo os adaptativos possuem inconvenientes, como o alto custo computacional para a estimativa de orientação e frequência na representação dos poros, além do fato de serem inadequados em imagens de baixa qualidade.

Para contornar esses inconvenientes (LEMES et al., 2014) projetaram um novo método de detecção de poros, denominado de Filtragem Dinâmica de Poros (*Dynamic Pore Filtering*—DPF). Este método não depende de etapas de processamento de maior custo computacional e se adapta a diferentes tamanhos e formas de poros.

O método em tela é dividido em cinco etapas. Na primeira etapa acontece a extração da região de interesse e a equalização da imagem. Na etapa seguinte, é aplicado um limiar de binarização global T_G que é estimado pelo método de Otsu (OTSU, 1979).

Na terceira etapa é feita uma estimativa do tamanho dos poros. Para isso, é realizada uma busca para cada *pixel* com valor acima de T_G em cada uma das quatro direções: para cima; para

baixo; à esquerda; e à direita. Mas essa busca não acontece indefinidamente, ou seja, existe uma distância máxima de pesquisa para encontrar esses vizinhos, que foi empiricamente definido pelo autor do método como de 20 *pixels*. Este limite é baseado na análise do tamanho dos poros de várias imagens de impressão digital.

Acontece que, se mais de uma direção atingir esse valor máximo, o *pixel* atual é descartado, pois é considerado um ponto de vale. Poros fechados sempre alcançam um vizinho em todas as direções, e poros abertos podem ter uma direção sem um vizinho. Um raio de poro global r_G é então calculado usando

$$r_G = \frac{1}{4|P|} \sum_{p \in P} (d_{up}^p + d_{down}^p + d_{left}^p + d_{right}^p), \quad (3.8)$$

em que P são os *pixels* com valores setados acima de T_G e com pelo menos 3 vizinhos conhecidos. d_θ^p é a distância do corrente *pixel* p para o vizinho na direção θ .

A quarta etapa consiste na classificação dos poros. Para isso, para cada *pixel* $p \in P$, é definido um limiar de binarização local T_L como a média dos *pixels* em uma máscara de tamanho $6r_G \times 6r_G$ centrado em p .

Além disso, um raio de poro local r_L é calculado usando

$$r_L = \frac{1}{4|P'|} \sum_{p \in P'} (d_{up}^p + d_{down}^p + d_{left}^p + d_{right}^p), \quad (3.9)$$

onde P' é o conjunto de *pixels* em P que também estão dentro da máscara definida acima. Então, cada *pixel* é classificado como um poro se satisfizer três condições trazidas pelo autor.

A etapa final, por sua vez, consiste na estimativa de localização de poros. O que acontece aqui é transformar todos os *pixels* que compõem um poro em um único ponto centrado no centro de massa do conjunto. Essa etapa foi implementada apenas para melhorar a representação gráfica dos resultados do esquema de detecção de poros.

A desvantagem desse método é que ele utiliza apenas relações geométricas dos poros para classificação. Assim, se eventual poro tiver tamanho muito maior ou muito menor dentro da mesma região local, o algoritmo não irá interpretá-los como poros. Assim, para mitigar problemas dessa ordem, que são típicos problemas de classificação, passaram a ser utilizadas as ANNs, conforme passaremos a tratar nas seções seguintes.

Tabela 1 – Arquitetura da CNND do método proposto por (LABATI et al., 2017).

#	Camada	Tamanho	Filtros
1	Conv.	5×5	5
	<i>Max-pooling</i>	3×3	-
2	Conv.	5×5	15
	<i>Max-pooling</i>	3×3	-
3	Conv.	5×5	1

Fonte: o autor, adaptado de (LABATI et al., 2017).

3.3 MÉTODOS DE DETECÇÃO DE POROS BASEADOS EM ANNS

O objetivo desta seção é analisar alguns métodos de detecção de poros baseados em aprendizado de máquina, enquanto que os outros até agora apresentados se limitam apenas ao uso de ferramentas de filtragem. Além disso, serão traçadas comparações entre os trabalhos que são recorrentemente citados na literatura e elencados pontos positivos e negativos de cada um deles.

3.3.1 Método proposto por Labati *et al.*

O método proposto por (LABATI et al., 2017) divide-se basicamente em duas etapas. Na primeira etapa, é utilizada uma CNN para realizar o processo de detecção dos poros e que foi batizada pelo autor como CNND (*Convolutional Neural Network Detection*). Nessa etapa o algoritmo recebe como entrada uma imagem em escala de cinza I e retorna a imagem I_{CNN} , na qual os poros são realçados e o padrão de crista das impressões digitais é removido. O objetivo da rede é computar uma imagem I_{CNN} em que a intensidade dos centros dos poros seja igual a 1 e a dos outros *pixels* igual a -1 .

Em seguida, é calculado um mapa binário dos poros B utilizando o algoritmo de Otsu (OTSU, 1979). Finalmente, são calculadas as coordenadas de cada poro $P_i = (x_i, y_i)$, bem como os centroides desses pontos de cada região de B . Regiões com área igual ou menor a um limiar t_a determinado empiricamente pelo autor são descartadas.

Diferentemente das configurações de CNN mais comumente usadas na literatura para problemas de classificação, a CNND não usa uma camada *softmax* final totalmente conectada. Usa apenas duas camadas convolucionais seguidas de camadas de *max-pooling*. Ao final, utiliza apenas uma camada convolucional de filtro unitário. A arquitetura da CNND está mostrada na Tabela 1.

Tabela 2 – Arquitetura da CNNR do método proposto por (LABATI et al., 2017).

#	Camada	Tamanho	Filtros
1	Conv.	$a_1 \times a_1$	10
	<i>Max-pooling</i>	3×3	-
2	Conv.	$a_2 \times a_2$	25
	<i>Max-pooling</i>	3×3	-
3	Conv.	$a_3 \times a_3$	250
	ReLU	-	-
4	Conv.	$a_4 \times a_4$	2
	<i>Softmax</i>	-	-

Fonte: o autor, adaptado de (LABATI et al., 2017).

A segunda etapa do método consiste em um processo de refinamento, que foi batizado pelo autor de CNNR (*Convolutional Neural Network Refinement*) e cujo objetivo consiste em remover possíveis erros da matriz de poros candidatos. Para tanto, a CNNR busca poros estimados erroneamente. Para cada poro da matriz de poros, a CNNR analisa os valores locais da imagem e retorna um valor binário 1 para as coordenadas consideradas como poros verdadeiros e 0 para poros falsos. Para acelerar a convergência do processo de aprendizado, foi inserida uma etapa de pré-processamento simples para cada imagem antes de usar a CNNR. Foi realizada uma normalização min-max seguida da subtração da intensidade média da imagem.

A CNNR é composta de oito camadas, incluindo quatro camadas convolucionais; duas camadas de *max-pooling*; uma camada cuja função de ativação é a Unidade Retificada Linear (*Rectified Linear Unit* – ReLU); e uma camada de classificação *softmax*. A primeira camada convolucional tem 10 filtros, cujo tamanho varia devido aos diferentes experimentos realizados pelo autor. Essa camada convolucional é seguida por uma camada de *max-pooling* com um *kernel* igual a 3×3 pixels. A segunda camada convolucional, por sua vez, tem 25 filtros e é seguida por uma camada de *max-pooling* com um *kernel* igual a 3×3 pixels. A terceira camada convolucional, entretanto, tem 250 filtros e é seguida por uma camada ReLU. Já a última camada convolucional tem 2 filtros e é seguida de um classificador *softmax* totalmente conectado retornando um valor binário representando se o poro pode ser considerado como estimado corretamente. A arquitetura da CNNR está mostrada na Tabela 2.

O algoritmo proposto foi avaliado em três conjuntos de imagens diferentes, quais sejam: imagens adquiridas através de leitor biométrico; imagens adquiridas sem toque através de fotografia; imagens de impressão digital latentes. Essa, inclusive, é uma das maiores contribuições de (LABATI et al., 2017), já que os outros trabalhos que serão tratados nas próximas seções

limitam a análise apenas a impressões digitais obtidas através de leitor biométrico.

Para simulação realizada utilizando as imagens adquiridas através de leitor biométrico, foi empregado o conjunto de dados que contém as 30 imagens rotuladas pertencentes ao *PolyU High-Resolution-Fingerprint* (HRF) (KUMAR, 2009). Os rótulos representam as coordenadas cartesianas de 12.767 poros. As imagens têm resolução de 1.200 dpi, têm tamanho 320×240 pixels, e representam a região central do dedo.

Já para a simulação realizada utilizando as imagens de impressão digital latentes, foram empregadas 36 impressões digitais de 22 dedos que foram adquiridas em laboratório mediante procedimento forense apropriado. Essas imagens têm uma resolução de 1.200 dpi, com tamanho 320×240 pixels. Foram rotuladas as coordenadas dos poros de cada imagem, totalizando 1.785 poros.

Para imagens obtidas através de leitor biométrico, foi utilizada apenas a CNND. Os resultados aferidos estão apostos na Tabela 6. Já para as imagens de impressões digitais latentes foram utilizadas a CNND e a CNNR. Os resultados também estão apostos na Tabela 6.

3.3.2 DeepPore

O método apresentado por (LABATI et al., 2017) utiliza o método de limiarização binária de (OTSU, 1979) como uma maneira de refinar um mapa de intensidade de poros. No entanto, é difícil detectar com precisão os poros desse modo em todo o mapa, já que a intensidade de poros é diferente para cada poro. Como resultado, a precisão desse método, que é baseada em aprendizado, torna-se semelhante às técnicas que utilizam modelos anisotrópicos para classificação dos poros. Para superar esse problema na extração de poros, (JANG et al., 2017) propuseram um método de extração profunda de poros baseada em CNN batizado como DeepPore.

A representação do pico de intensidade de cada poro varia dependendo da espessura da crista a que este pertence; ou se o poro é do tipo aberto ou fechado. Assim, torna-se difícil detectar os poros com precisão utilizando um limiar de binarização em todo o mapa. Para contornar esse problema, o autor passou a utilizar o *Pore Intensity Refinement* (PIR) que é uma maneira peculiar de pós-processamento para detectar os poros. Aqui os autores desenvolveram um algoritmo que busca identificar máximos locais do mapa de poros e não apenas os máximos globais, estimando, portanto, os poros com maior assertividade.

Tabela 3 – Arquitetura da CNN do DeepPore.

#	Camada	Tamanho	Filtros
1	Conv.	3×3	64
	BatchNorm	-	-
	ReLU	-	-
2	Conv.	3×3	64
	BatchNorm	-	-
	ReLU	-	-
$\vdots$	$\vdots$	$\vdots$	$\vdots$
9	Conv.	3×3	64
	BatchNorm	-	-
	ReLU	-	-
10	Conv.	3×3	64

Fonte: o autor, adaptado de (JANG et al., 2017).

Outra contribuição do DeepPore é a utilização de uma rede mais profunda e, desse modo, por ter mais camadas, tem um grande campo receptivo (SIMONYAN; ZISSERMAN, 2015). A Tabela 3 ilustra a arquitetura da rede utilizada. Note que o processo de extração de poros usa dez camadas. Cada camada tem 64 filtros com tamanho de 3×3 . Portanto, 64 mapas de características são gerados para cada camada. As camadas 1 a 9 realizam convolução, normalização de lote (*batch normalization* ou *BatchNorm* – BN) e ReLU. Na última camada, por sua vez, apenas a convolução prossegue.

O algoritmo em tela foi avaliado usando o PolyU HRF (KUMAR, 2009). Os resultados aferidos estão apostos na Tabela 6.

3.3.3 Método proposto por Dahia e Segundo

O método proposto por (DAHIA; SEGUNDO, 2018) utiliza o conceito de *Fully Convolutional Neural Network* (FCN) que é amplamente utilizado em tarefas de segmentação de imagens devido à sua capacidade de classificar *pixels* de imagem (LONG; SHELHAMER; DARREL, 2015). A FCN permite lidar com imagens de entrada de qualquer tamanho, sem exigir que todas as imagens de treinamento ou teste tenham o mesmo tamanho, economizando a sobrecarga de operações de fixação de tamanho. Também mantém o número de parâmetros do modelo relativamente baixo.

A Arquitetura da FCN utilizada pelo autor consiste em quatro camadas convolucionais. Nas três primeiras camadas, o tamanho do *kernel* é 3×3 e são realizadas operações de ReLU,

Tabela 4 – Arquitetura da CNN do método proposto por (DAHIA; SEGUNDO, 2018).

#	Camada	Tamanho	Filtros
1	Conv.	3×3	32
	ReLU	-	-
	BatchNorm	-	-
	<i>Max-pooling</i>	3×3	-
2	Conv.	3×3	64
	ReLU	-	-
	BatchNorm	-	-
	<i>Max-pooling</i>	3×3	-
3	Conv.	3×3	128
	ReLU	-	-
	BatchNorm	-	-
	<i>Max-pooling</i>	3×3	-
4	Conv.	5×5	1
	BatchNorm	-	-
	<i>Sigmoid</i>	-	-

Fonte: o autor, adaptado de (DAHIA; SEGUNDO, 2018).

normalização de lote (*Batch Normalization* – BN) e *max-pooling* 3×3 . O número de filtros é 32, 64 e 128 da primeira à terceira camada, respectivamente. A camada de saída é precedida por *dropout*, possui *kernel* de tamanho 5×5 e um único filtro e ativação sigmoide. A Tabela 4 apresenta a arquitetura detalhada do método apresentado.

O algoritmo em tela foi avaliado usando o PolyU HRF (KUMAR, 2009). Os resultados aferidos estão apostos na Tabela 6.

3.3.4 DeepResPore

O método proposto por (ANAND; KANHANGAD, 2019) foi batizado pelos autores como DeepResPore. Ele consiste em uma metodologia automatizada de detecção de poros que emprega um modelo de CNN baseado em aprendizado residual, ou seja, uma ResNet.

A metodologia proposta nesse método envolve duas etapas principais. Primeiramente, um mapa de intensidade de poros é gerado a partir da entrada de imagem de impressão digital. Em segundo lugar, o mapa de intensidade de poros é processado para obter estimativas das coordenadas dos poros.

A rede proposta contém 18 camadas com 8 blocos residuais. Assim, há um total de 8 conexões de atalho no DeepResPore. A entrada para DeepResPore é um *patch* de impressão digital de tamanho 80×80 *pixels* e a saída é um mapa de intensidade de poro correspondente do

Tabela 5 – Arquitetura da CNN do DeepResPore.

#	Camada	Tamanho	Filtros
1	Conv.	7×7	64
	BatchNorm	-	-
	ReLU	-	-
2	Conv.	3×3	64
	BatchNorm	-	-
	ReLU	-	-
$\vdots$	$\vdots$	$\vdots$	$\vdots$
6	Conv.	3×3	128
	BatchNorm	-	-
	ReLU	-	-
$\vdots$	$\vdots$	$\vdots$	$\vdots$
10	Conv.	3×3	256
	BatchNorm	-	-
	ReLU	-	-
$\vdots$	$\vdots$	$\vdots$	$\vdots$
14	Conv.	3×3	512
	BatchNorm	-	-
	ReLU	-	-
$\vdots$	$\vdots$	$\vdots$	$\vdots$
18	Conv.	3×3	1

Fonte: o autor, adaptado de (ANAND; KANHANGAD, 2019).

mesmo tamanho. A Tabela 5 apresenta a arquitetura detalhada do DeepResPore.

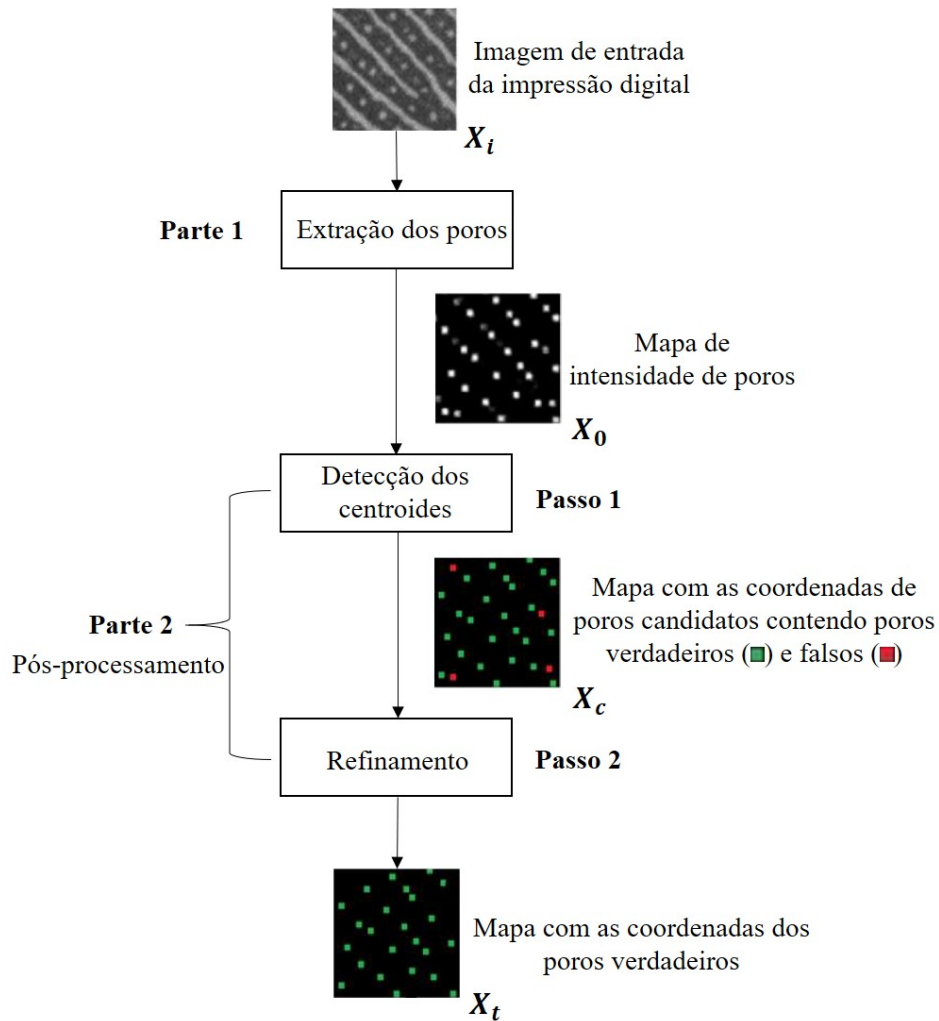
Para obter um mapa de características do mesmo tamanho que o da entrada não foram realizadas operações de *pooling* na rede proposta. Além disso, todas as camadas convolucionais, exceto a última, realizam a convolução seguida da BN e ReLU. No final da rede, foi introduzida uma camada convolucional contendo um único filtro para fornecer um mapa de intensidade de poros.

O algoritmo em tela foi avaliado usando o PolyU HRF (KUMAR, 2009). Os resultados aferidos estão apostos na Tabela 6.

3.3.5 Método proposto por Ali, Wang e Ahmad

O método proposto por (ALI; WANG; AHMAD, 2021) consiste em um esquema híbrido de detecção de poros. A primeira parte do método consiste do uso de uma CNN para realizar a extração dos poros, enquanto que a segunda parte é uma etapa de pós-processamento que

Figura 4 – Esquema de uma estrutura híbrida de detecção de poros.



Fonte: o autor, adaptado de (ALI; WANG; AHMAD, 2021).

consiste em: determinar os centroides dos poros verdadeiros; e remover os poros falsos que tenham sido detectados. A Figura 4 mostra o diagrama de blocos do método proposto.

Sendo um pouco mais específico, a primeira parte do método – que está representada no diagrama da Figura 4 como a etapa de extração dos poros – utiliza uma imagem de entrada X_i em escala de cinza que passa por uma CNN, cujo esquema de detecção de poros extrai os recursos em vários níveis e constrói um único mapa de intensidade de poros X_0 em escala de cinza.

A segunda parte do método consiste do pós-processamento, que também é dividido em dois passos. No primeiro passo acontece a detecção do centroide do poro, que recebe o mapa de intensidade de poros X_0 como entrada e o transforma em um mapa de poros binário X_c , onde o valor de 1 indica que o *pixel* em questão pertence a uma das regiões de poro e 0

indica que não. Por sua vez, o segundo passo do pós-processamento consiste no refinamento que também se subdivide em outras duas etapas que serão detalhadas adiante. Assim, ao final do pós-processamento, alguns dos poros candidatos falsamente detectados são descartados para produzir um mapa de poros verdadeiro X_t .

3.3.5.1 Parte 1: extração dos poros

A topologia da CNN utilizada para a extração de poros tem links de atalhos paralelos às suas camadas convolucionais, ou seja, é do tipo ResNet, assim como a utilizada do DeepResPore. Ela foi projetada em 4 estágios, como mostra a Figura 5, onde os estágios de 1 a 3, produzem, respectivamente, níveis hierárquicos de saída baixo, médio e alto. Cada um dos 3 estágios de extração de recursos consiste em duas camadas de convolução e no quarto estágio ocorre uma fusão concatenativa global dos 3 primeiros estágios.

A fim de alcançar diversidade suficiente entre as características de diferentes canais e ainda manter a complexidade baixa, foram aplicados 8 *kernels* diferentes de tamanho 3×3 em um *patch* de entrada X_{ip} da imagem de entrada X_i de maneira a obter 8 canais distintos na primeira camada do estágio 1. Cada um desses 8 mapas de recursos resultantes da primeira camada deste estágio é operado por outro *kernel* 3×3 . Assim, a segunda camada desta etapa também emprega 8 diferentes *kernels* para produzir outros 8 mapas distintos. O resíduo aditivo proveniente da fusão dos mapas correspondentes das duas camadas, em seguida, produz os mapas de recursos de saída X_1 do primeiro estágio.

Buscando alcançar ainda mais diversidade nos recursos extraídos, o número de filtros foi duplicado. Na primeira camada do estágio 2, foram aplicados 2 *kernels* 3×3 em cada um dos 8 mapas de X_1 resultantes do primeiro estágio e 1 *kernel* 3×3 na segunda camada para cada um dos mapas extraídos da primeira camada desta etapa. Assim, cada uma das duas camadas do segundo estágio usa 16 núcleos de tamanho 3×3 diferentes. Da mesma forma, no terceiro estágio, 2 *kernels* são aplicados para cada um dos 16 mapas de X_2 e 1 *kernel* 3×3 para cada um dos mapas extraídos de sua primeira camada. Assim, cada uma das duas camadas do estágio 3 emprega um total de 32 *kernels* 3×3 diferentes.

Nos três primeiros estágios, a BN (IOFFE; SZEGEDY, 2015) e a ReLU (NAIR; HINTON, 2010) são executadas após cada operação de convolução para normalizar a distribuição dos recursos produzidos pela camada anterior. Já no estágio 4 acontece a concatenação dos diversos

canais gerados nos estágios anteriores dando origem à X_{ci} sobre o qual é aplicada uma convolução pontual ($X_{ci} = (X_1 X_2 X_3)^T$) com *kernel* de tamanho $1 \times 1 \times 56$ *pixels*, produzindo um único mapa de recursos X_{c0} .

Assim, cada valor de *pixel* do *patch* de saída X_{c0} deve ser indicativo de sua probabilidade de pertencer a um poro, ou seja, temos um problema de classificação em que a probabilidade de um *pixel* pertencer a um poro é determinada. Dessa feita, já no último passo do estágio 4, é aplicada uma função sigmoide (MITCHELL, 1997) em X_{c0} para produzir X_{op} , ajustando a faixa de probabilidade de determinado *pixel* pertencer a um poro entre 0 e 1. Feito todo esse processo, finalmente a entrada X_i é convertida no mapa de poros X_0 .

3.3.5.2 Parte 2: pós-processamento

A fase seguinte à detecção de poros é o pós-processamento em que é apresentado um método para determinar com precisão as coordenadas de todos os poros verdadeiros do mapa de intensidade X_0 . Este pós-processamento consiste em duas etapas, conforme mostrado no diagrama de blocos da Figura 6. O passo 1 consiste em um módulo de binarização de imagem que transforma o mapa de intensidade de poros X_0 em um mapa binário X_b acrescido de um módulo que partindo de X_b obtém um mapa de centroide de poro candidato X_c . O passo 2 consiste em um processo de refinamento de poros que identifica os verdadeiros e remove os falsos levando a um mapa final de poros X_t contendo os centroides dos poros verdadeiros.

3.3.5.2.1 Passo 1: detecção dos centroides

Uma vez que o mapa de intensidade de poros X_0 fornece apenas uma medida da probabilidade de seus *pixels* individuais pertencerem a um poro, é necessário, para diferenciar mais claramente se X_0 é ou não um *pixel* pertencente a uma região de poros, atribuir-lhe um valor binário distinto. Portanto, para obter um mapa binário X_b precisamos de um limiar adequado T para que um *pixel* com um valor igual ou superior a T possa ser classificado como um candidato pertencente a uma região de poros, atribuindo-lhes um valor de 1. Por outro lado, se um *pixel* em X_0 tiver um valor menor que T , é atribuído um valor de 0 e, portanto, considerado como não pertencente a uma região de poros.

Por outro lado, deve-se notar que os valores de intensidade em X_θ são diretamente relacionados com as distribuições de intensidade dos vários poros dentro de uma imagem de impressão digital X_i . Mesmo que a maioria dos poros na imagem de entrada X_i tenha a mesma forma e intensidade padrão, há sempre alguns poros cujas formas e valores de intensidade típicos podem diferir e, portanto, os padrões de intensidade de tais poros podem ser diferentes de um para outro. A distribuição dos valores de *pixel* em X_θ é necessariamente afetada por tal não uniformidade das formas dos poros e valores de intensidade em X_i .

Assim, se um limiar fixo T for escolhido para todo o mapa X_θ para convertê-lo em um mapa binário X_b , pode resultar numa falsa classificação de alguns dos *pixels* de X_θ . Logo, um método para determinar o limiar T deve levar em conta a natureza não idêntica de distribuição de poros ao longo do mapa de intensidade de poros X_θ . Para contornar o problema da distribuição não idêntica da forma e intensidade dos poros na imagem X_i foi proposto um limiar adaptativo T_l para binarizar o mapa de intensidade de poros X_θ , ou seja, é adaptativo de uma janela para outra janela de X_θ . Quando este limiar T_l é usado para a janela l de X_θ , então X_θ é convertido em um mapa binário X_b .

O algoritmo Niblack (NIBLACK, 1998) foi usado com sucesso para determinar um limiar de adaptação local T_l para identificar a região de um objeto em imagens de documentos. Neste algoritmo, a média m_l e o desvio-padrão σ_l na janela l são usados para determinar um valor local para o limiar T_l a ser usado para todos os *pixels* dentro da janela l , da seguinte maneira

$$T_l = m_l + \beta \sigma_l, \quad (3.10)$$

em que os parâmetros m_l e σ_l calculados como

$$m_l = \frac{1}{N_l} \sum_{k=1}^{N_l} p_k, \quad (3.11)$$

$$\sigma_l = \sqrt{\frac{\sum_{k=1}^{N_l} (p_k - m_l)^2}{N_l}}, \quad (3.12)$$

em que p_k é o valor em tons de cinza do k -ésimo *pixel* da l -ésima janela de X_θ , N_l é o número de *pixels* na janela l -ésima, e β , chamado de fator Niblack, é um parâmetro que controla o efeito do desvio-padrão local na determinação do limite de um objeto. Em (NIBLACK, 1998), o valor de β foi empiricamente determinado para ser 0,2. O uso deste valor fixo de β (ou seja, um valor não adaptativo) na determinação do limiar T_l leva a uma determinação bem sucedida das regiões de

objetos em imagens que são puramente documentos de texto. Em (RAIS; HANIF; TAJ, 2004), o fator Niblack também foi adaptado localmente selecionando seu valor pela expressão

$$\beta_l = -0,3 \times \frac{(m_g \sigma_g) - (m_l \sigma_l)}{\max((m_g \sigma_g), (m_l \sigma_l))}, \quad (3.13)$$

em que m_g e σ_g são a média e o desvio-padrão de todo mapa, respectivamente. Assim, o fator Niblack β_l é feito para mudar de janela para janela, mas permanece fixo dentro de uma mesma janela.

Com o objetivo de transformar o mapa de intensidade de poros X_θ no mapa binário X_b , foi empregado o limiar adaptativo dado por (3.10) com o fator Niblack adaptativo β_l de (3.13). Por uma questão de simplicidade, foram usadas janelas de tamanho fixo $N \times N$ para todo o mapa de intensidade de poros X_θ , ou seja, $N_l = N$.

Para escolha do tamanho da janela (N) levou-se em consideração o fato de que a intensidade e as formas dos poros em uma imagem típica de impressão digital variam mais globalmente do que localmente. Portanto, escolher um tamanho de janela grande falha em capturar essa variação global na imagem da impressão digital. Por outro lado, a escolha de uma janela de tamanho pequeno resultaria em problema de classificar as regiões de ruído da impressão digital como poros. Além disso, um tamanho de janela pequeno afetaria negativamente o tempo de processamento. Assim, o tamanho da janela precisa ser escolhido adequadamente.

A etapa seguinte consiste em determinar os centroides dos poros candidatos. Nessa etapa, um único poro pertence a uma região específica na imagem de impressão digital, cujos *pixels* pertencentes a esta região não podem ser desconectados. Portanto, normalmente uma região do mapa binário X_b que tem, digamos, N_p *pixels* com os valores de 1 e estão conectados, pode ser considerada uma região de poro único.

No entanto, esta definição de região de poros não é absolutamente precisa, pois, por exemplo, se N_p for muito pequeno, a região pode ou não ser considerada uma região porosa, tendo em vista o fato de que uma região conectada de X_b com muito poucos *pixels* poderia ser apenas ruído. Por outro lado, se N_p for muito grande, pode ser uma situação em que a região em questão pode ser uma região correspondente a dois ou mais poros fundidos.

Em trabalhos anteriores, o centroide de um poro foi calculado usando apenas as coordenadas dos *pixels* que pertencem à região dos poros. No entanto, o sistema visual humano determina a localização do centroide de um poro não apenas pelo centro geométrico, mas é

influenciado pela intensidade da variação dos *pixels* na região em destaque. Daí, o cálculo do centroide do poro usando apenas as coordenadas geométricas não ser uma representação muito precisa da localização do poro.

Consequentemente, considerando que estamos lidando com um único poro, a localização dos poros é representada pelo seu centroide ponderado. É, então, calculado o peso do centroide usando a geometria euclidiana da região do poro definida como um conjunto de *pixels* que têm um valor de 1 e estão conectados no mapa binário X_b , bem como as informações de intensidade de poros dos *pixels* subjacentes da própria imagem da impressão digital. Assim, o centroide ponderado (X, Y) de um poro candidato é calculado como

$$(X, Y) = \left(\frac{\sum_{\nu=1}^{\nu=N_p} (X_\nu P_\nu)}{\sum_{\nu=1}^{\nu=N_p} P_\nu}, \frac{\sum_{\nu=1}^{\nu=N_p} (Y_\nu P_\nu)}{\sum_{\nu=1}^{\nu=N_p} P_\nu} \right). \quad (3.14)$$

Contudo, o que acontece quando a CNN detecta uma região grande no mapa binário de poros? Nesse caso, podemos estar diante de uma situação em que dois poros estão próximos. Em princípio seria de esperar que uma imagem de impressão digital tivesse tantos poros quanto a quantidade de centroides encontrados no mapa X_c . No entanto, na prática, cada centroide em X_c não necessariamente representa um poro verdadeiro. Nessa esteira, é iniciado o segundo passo do processo de pós-processamento, que consiste no processo de remoção de poros de X_c através do critério da distância euclidiana.

3.3.5.2.2 Passo 2: refinamento

Para distinguir os poros verdadeiros dos falsos, primeiro é necessário apresentar algumas características sobre os poros verdadeiros que ajudarão a distinguir um poro verdadeiro de um falso. A primeira característica é que um poro verdadeiro tem uma distância de seu vizinho mais próximo que varia de poro para poro. (ALI; WANG; AHMAD, 2021) apontam que em nenhuma imagem de impressão digital do banco PolyU HRF (KUMAR, 2009) há um par de poros com uma distância euclidiana menor que 4,12 *pixels*. A distância mínima é denotada por d_{min} . Portanto, em um mapa de centroide ponderado X_c , se um par de poros vizinhos é detectado com distância euclidiana menor que d_{min} ambos os poros do par não podem ser considerados poros verdadeiros, simultaneamente.

A segunda característica é que, em diferentes janelas de uma imagem de impressão digital, pode haver diferenças significativas nas intensidades médias de seus poros, bem como diferenças no desvio-padrão das intensidades de poros dentro da janela. Com base nessa observação, a média e valores de desvio-padrão das intensidades de poros em uma determinada janela em X_c têm valores que variam de janela para janela (W). Portanto, se a intensidade de um poro em qualquer janela tem um valor que é muito menor do que a intensidade média dos poros na janela, então tal poro não pode ser considerado um verdadeiro poro e, portanto, seu centroide deve ser removido de X_c .

Geralmente, a intensidade de um poro falso é menor que a de um poro verdadeiro em uma determinada vizinhança da imagem da impressão digital. O desvio-padrão das intensidades de poros é uma indicação do grau pela qual a intensidade de um poro verdadeiro pode ser menor que a intensidade média na vizinhança. Assim, para remover um poro falso, exige-se que a intensidade do poro seja menor que um valor limite dado por $t = m - c\sigma$, onde m e σ são a média e o padrão-desvio das intensidades dos poros dentro de uma determinada vizinhança da imagem da impressão digital e c é um parâmetro que precisa ser determinado adequadamente.

Foi adotado o valor de c empiricamente para todo o conjunto de dados da base de (KUMAR, 2009). As imagens do conjunto de dados são divididas em $W \times W$ janelas. Primeiramente, para uma determinada imagem no conjunto de dados, é determinado um valor de c de modo que a precisão na remoção dos poros falsos seja maior que 96%. Então, um valor $C_{avg} = 1,12$, que é a média dos valores de c , assim determinado para as imagens individuais, é usado para remover poros de cada uma das janelas de uma imagem de metodos.

Removendo todos os falsos poros de X_c , usando o critério do d_{min} e o critério da média e desvio-padrão da intensidade dos poros na janela, o resultado final é um mapa de poros verdadeiros, chamado de X_r . O algoritmo em tela foi avaliado usando o PolyU HRF (KUMAR, 2009). Os resultados aferidos estão apostos na Tabela 6.

3.4 COMPARAÇÃO DOS RESULTADOS OBTIDOS PELA UTILIZAÇÃO DOS DIFERENTES MÉTODOS DE DETECÇÃO DE POROS

Nesta seção, são comparados em termos numéricos os resultados dos experimentos realizados em cada um dos métodos de detecção de poros trazidos anteriormente. Para tanto, elaborou-se a Tabela 6.

Tabela 6 – Tabela comparativa entre os variados esquemas de detecção de poros existentes.

Esquema	R_T	R_F
Jain	75,90%	23,00%
DoG adaptado	80,80%	22,20%
DAPM	84,80%	17,60%
Teixeira	86,10%	8,60%
DPF	83,50%	9,90%
Labati <i>et al.</i> : imagens de leitor	84,69%	15,31%
Labati <i>et al.</i> : imagens de impressões latentes	52,70%	24,00%
DeepPore	93,09%	8,64%
Dahia	91,95%	8,88%
DeepResPore	94,49%	8,50%
Ali	96,69%	4,18%

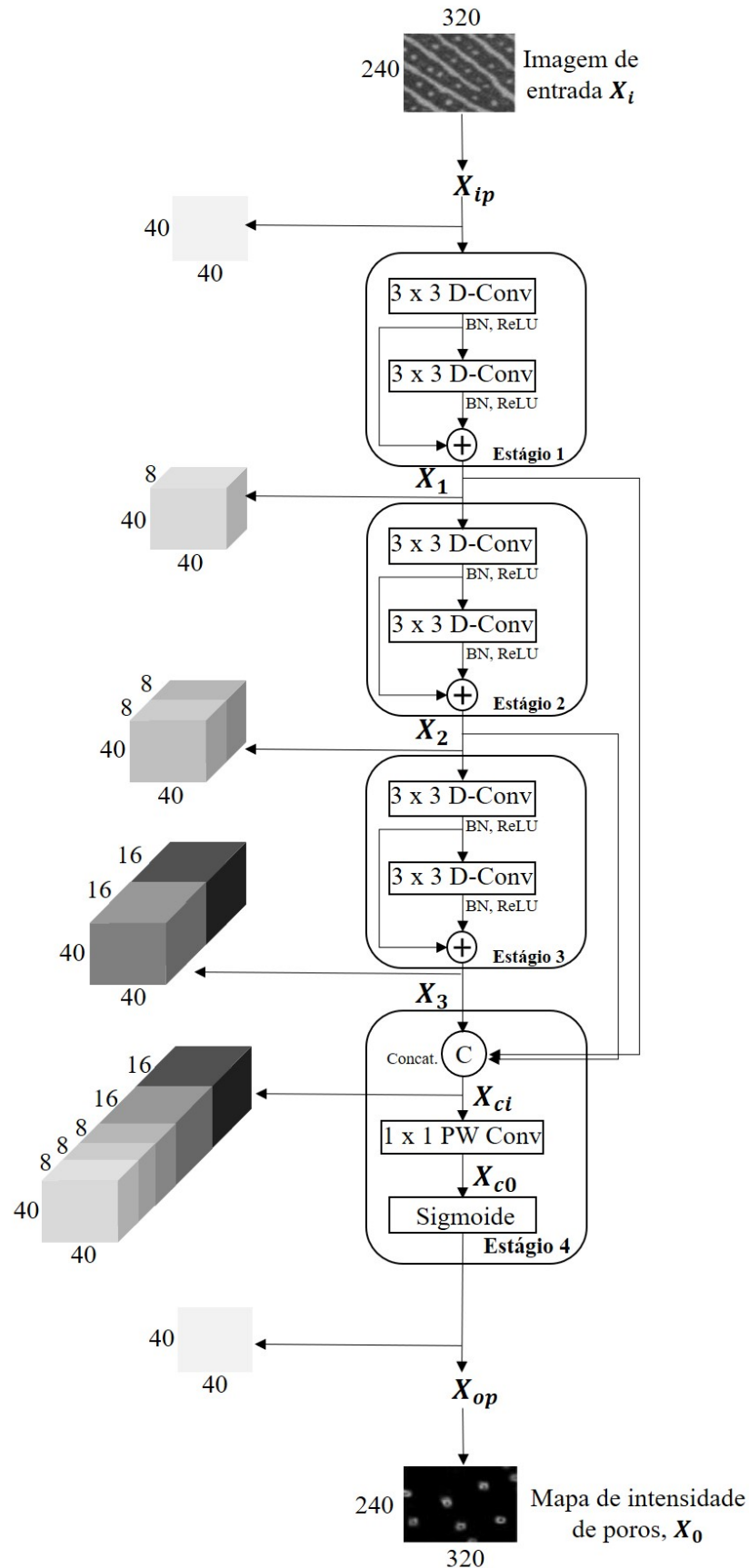
Fonte: o autor.

Para fins de comparação, são usadas as seguintes figuras de mérito para a avaliação de desempenho da rede:

- Taxa de detecção real (R_T), indica a razão entre o número de poros verdadeiros detectados e o número de poros reais presentes na imagem de impressão digital (LABATI et al., 2017).
- A taxa de detecção falsa (R_F), indica a razão entre o número de poros falsamente detectados e o número total de poros detectados (LABATI et al., 2017).

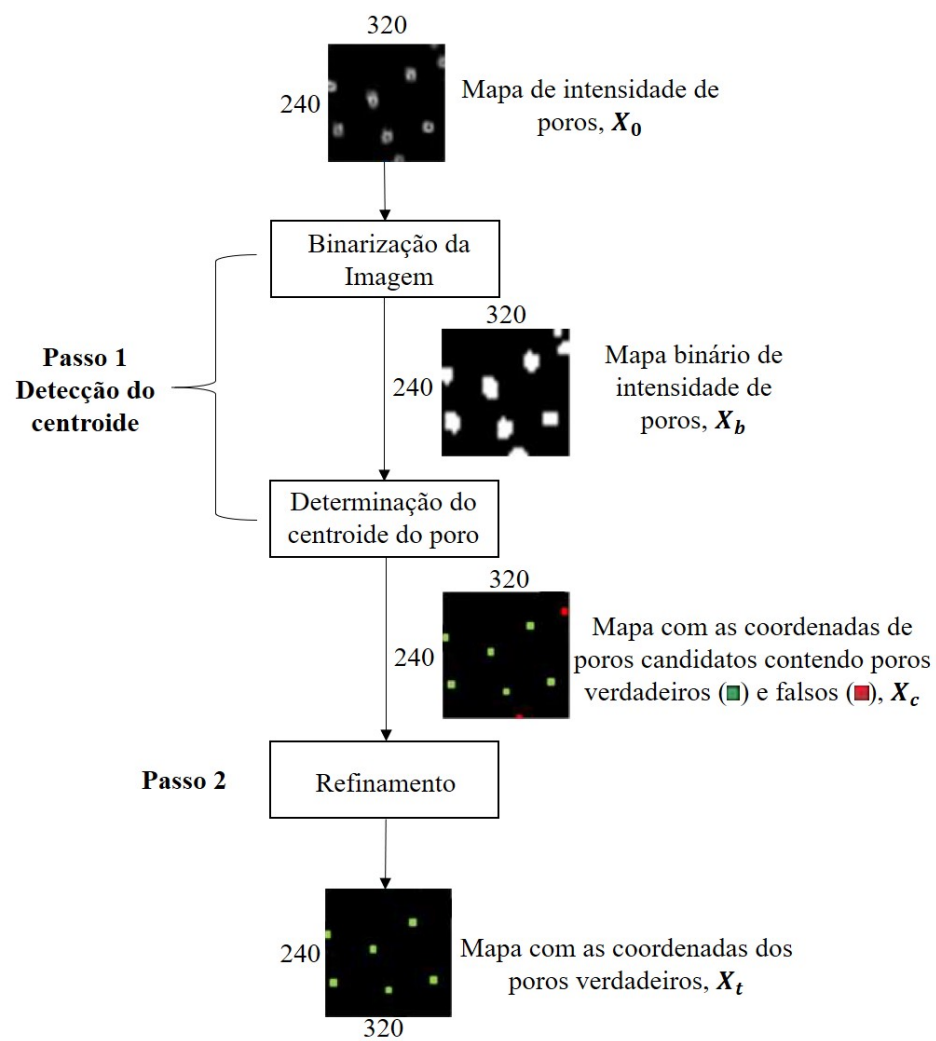
Os valores desejáveis para R_T e R_F são 1 e 0, respectivamente. Alta taxa de detecção verdadeira e baixa taxa de detecção falsa são indicativo de um bom desempenho do sistema de detecção de poros.

Figura 5 – Topologia de CNN utilizada para obtenção do mapa de intensidade de poros, X_0 .



Fonte: o autor, adaptado de (ALI; WANG; AHMAD, 2021).

Figura 6 – Etapa de pós-processamento.



Fonte: o autor, adaptado de (ALI; WANG; AHMAD, 2021).

4 IMPLEMENTAÇÃO DO ALGORITMO DE DETECÇÃO DE POROS PROPOSTO POR ALI, WANG E AHMAD COM ADAPTAÇÕES

4.1 INTRODUÇÃO

Os Capítulos 2 e 3 desta dissertação expuseram os referenciais teóricos mínimos para o entendimento do que será proposto adiante. Especificamente no Capítulo 2, foi feito um apanhado geral sobre as temáticas das impressões digitais e das redes neurais artificiais. Já no Capítulo 3, foi elaborado um breve resumo dos métodos mais relevantes de detecção de poros em imagens de impressões digitais presentes na literatura.

Ainda no Capítulo 3, foi confeccionada a Tabela 6 que traça um comparativo entre as métricas de R_T e R_F dos métodos de detecção de poros apresentados. Da análise comparativa, observou-se que o método proposto por (ALI; WANG; AHMAD, 2021) foi o que apresentou melhor resultado, com $R_T = 96,69\%$ e $R_F = 4,18\%$. Além disso, o trabalho supracitado foi o que apresentou maior riqueza de detalhes e parâmetros para fins de reprodutibilidade.

Esses motivos, aliados com a ausência de um trabalho específico na literatura que utilize como banco de dados imagens de vestígios de impressões digitais encontrados em local de crime, levou o autor desta dissertação a implementar a topologia proposta por (ALI; WANG; AHMAD, 2021) e verificar os resultados.

De todo modo, cabe aqui uma ressalva para o artigo publicado por (LABATI et al., 2017), que utilizou um banco de dados que o autor chamou de forense. Entretanto, essa base, que é composta de 36 imagens de impressões digitais de 22 dedos, não se aproxima da que foi usada nesta dissertação e que será melhor descrita na Seção 4.2.1. Na verdade, consiste apenas de um conjunto de dados cujas impressões digitais foram inseridas em objetos em ambiente de laboratório, e, por esse motivo, estão livres de distorções provocadas por tremores; borrões; sobreposição; ou diferença de pressão entre a superfície e o dedo. As imagens que compõem a base dados de (LABATI et al., 2017) têm resolução de 1.200 dpi, com tamanho de 320×240 pixels, em que foram rotuladas as coordenadas dos centroides dos poros de cada imagem, totalizando 1.785 poros.

4.2 DETALHES DO ALGORITMO IMPLEMENTADO

Nesta seção, serão abordados detalhes importantes da simulação do algoritmo. Destacam-se: caracterização do banco de dados; método de generalização do modelo; particionamento das imagens em *patches*, para fim de treinamento e teste; procedimento de aumento de dados; detalhes sobre a função de perda e otimizador; e o *Grid Search*.

4.2.1 Banco de dados

Todos os trabalhos constantes na Seção 3.3 utilizaram o mesmo banco de dados para realizar as simulações de seus algoritmos, qual seja, imagens de impressão digital coletadas em leitor biométrico de alta resolução denominado PolyU HRF (KUMAR, 2009), que dispõe de 30 imagens de impressão digital com tamanho 240×320 *pixels* e resolução de 1.200 dpi, totalizando 12.767 coordenadas cartesianas de centroides estimados de poros.

Por outro lado, neste trabalho, busca-se analisar o comportamento do algoritmo de detecção de poros proposto por (ALI; WANG; AHMAD, 2021) em imagens de vestígios de impressão digital encontrados em local de crime. Assim, o próprio autor confeccionou um banco de dados que é composto por 8 imagens de impressões digitais obtidas através da perícia papiloscópica e fotografia forense em objetos relacionados a cenas de crime. Essas imagens têm o tamanho de 1.000×1.000 *pixels* e resolução de 96 dpi, sendo tais imagens utilizadas tanto no conjunto de teste quanto no conjunto de treinamento da rede neural. Essas 8 imagens possuem um total de 1.200 poros, cujas coordenadas dos centroides dos poros também foram marcadas manualmente pelo autor desta dissertação.

Esse banco possui ainda outras 5 imagens de impressão digital obtidas através de leitor biométrico, com tamanho 1.000×1.000 *pixels* e resolução de 96 dpi, sendo tais imagens utilizadas apenas no conjunto de treinamento da rede neural. Essas 5 imagens possuem um total de 2.641 poros, cujas coordenadas dos centroides dos poros também foram marcadas manualmente e têm como objetivo principal aumentar a quantidade de dados disponível para aprendizado da rede.

Vale destacar que apesar de a Seção 2.2.1 ressaltar que a resolução das imagens para compor um banco de dados capaz de detectar poros precisar de ao menos 800 dpi, a base de dados utilizada nesta dissertação tem resolução de 96 dpi, de maneira que também será avaliado o desempenho do algoritmo proposto por (ALI; WANG; AHMAD, 2021) para uma base de resolução mais baixa.

4.2.2 Método de generalização do modelo

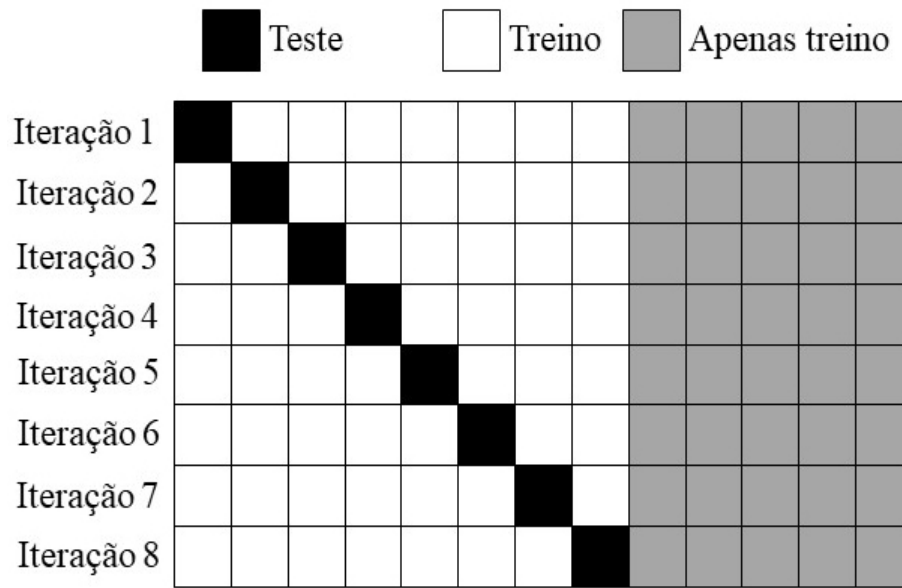
De acordo com (ALI; WANG; AHMAD, 2021), as imagens de impressão digital são divididas em conjuntos de treinamento e teste, usando, para isso, o método de validação cruzada *K-fold*, com $k = 5$ (DUDA; HART; STORK, 2012). Esse é um método usado para avaliar a capacidade de generalização de um modelo, a partir de um conjunto de dados (KOHAVI, 1995) e consiste em separar aleatoriamente o conjunto de dados em k grupos, de tamanho aproximadamente igual, em que, o primeiro conjunto será o conjunto de teste e os $k - 1$ restantes serão utilizados no treinamento. Posteriormente, o segundo conjunto será utilizado no teste e os conjuntos 1, 3 até $k - 1$ serão utilizados para treinamento. Esse procedimento se repete até que todos os conjuntos sejam utilizados no teste (TAMILARASI; RANI, 2020).

Dessa feita, (ALI; WANG; AHMAD, 2021) realizaram experimentos com as imagens do banco PolyU HRF, utilizando 5 conjuntos diferentes de treinamento e teste, em que cada conjunto contém 6 imagens. Além disso, nenhuma das 6 imagens de impressões digitais são repetidas em um mesmo conjunto, tampouco são utilizadas a mesma imagem em conjuntos diferentes. Assim, passam a existir 5 modelos de treinamento diferentes, cada um dos quais avaliados usando um conjunto de teste exclusivo. Dos resultados das 5 avaliações, é então calculada a média para obter o desempenho da rede.

Entretanto, essa metodologia proposta para verificação de desempenho da rede não é sumariamente aplicável a esta dissertação, já que o banco de dados aqui utilizado é diferente e pouco volumoso. Assim, foi utilizado um caso particular do método *K-fold* conhecido como método *Leave-one-out* ou *Leave-one-out Cross Validation* (LOOCV) que divide os dados em conjuntos de treinamento e teste, em que cada amostra é usada uma vez como um conjunto de teste (*singleton*) enquanto as amostras restantes formam o conjunto de treinamento (PEDREGOSA et al., 2011). Também vale destacar que o método LOOCV é imparcial e confiável, desde que exista distribuição semelhante nos dados de treinamento e teste (ER; WANG, 2016). Por outro lado, a confiabilidade vem ao custo de mais esforço computacional, que é contornado devido ao baixo número de imagens que compõem a base de dados (PEDREGOSA et al., 2011).

Para exemplificar o método LOOCV, caso a imagem 1 do subconjunto de imagens obtidas em local de crime seja escolhida para teste, as imagens 2, 3, 4, 5, 6, 7 e 8 e as 5 imagens do subconjunto de imagens obtidas pelo leitor biométrico serão utilizadas para treinamento. Nos mesmos moldes, caso a imagem 2 seja utilizada para teste, as imagens 1, 3, 4, 5, 6, 7 e 8 serão

Figura 7 – Esquema do método LOOCV.



Fonte: o autor.

escolhidas para treinamento, juntamente com as demais 5 imagens do outro subconjunto. A Figura 7 ilustra o método LOOCV para um caso particular do método *K-fold* em que $k = 8$.

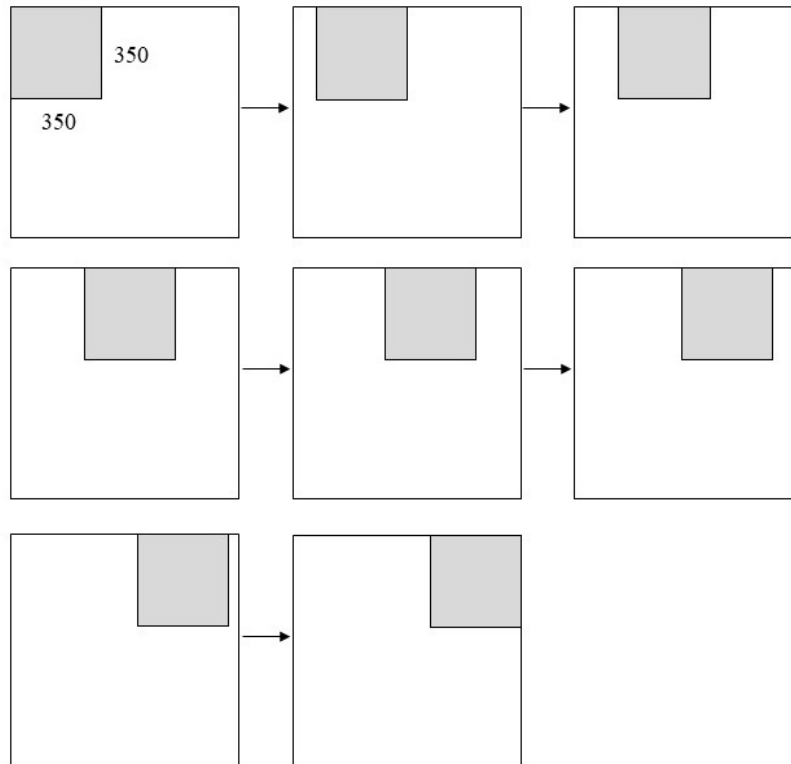
4.2.3 Esquema de particionamento das imagens para treinamento e teste da CNN

Segundo (ALI; WANG; AHMAD, 2021), as imagens de impressão digital no conjunto de treinamento, bem como seus rótulos correspondentes, são particionados em $p \times p$ *patches*, onde p varia de 10 até 100. Para isso é usado um processo de extração de *patches* com sobreposição. Então, um conjunto de 142.800 *patches* é obtido através da extração aleatória de 5.950 *patches* em cada imagem de impressão digital de treinamento. Já durante a fase de teste, é utilizado o mesmo tamanho do *patch* do treinamento com a ressalva de que as imagens que compõem o *patch* não possuem sobreposição e o *patch* não extrapola a fronteira da imagem.

De maneira análoga, nas implementações realizadas neste trabalho, o conjunto de imagens de treinamento, bem como seus rótulos correspondentes, foram particionados em $p \times p$ *patches*, de tamanho 100×100 , 150×150 , 200×200 , 250×250 , 300×300 , 350×350 e 400×400 , saltando de 100 em 100 *pixels*. A Figura 8 ilustra o tamanho dos *patches* utilizados durante o treinamento e o deslocamento de cada um deles.

Essa forma de deslocamento do *patch* funciona como uma espécie de aumento de dados para os conjuntos de treinamento, já que, ao final da varredura de cada imagem, serão produzidos

Figura 8 – Esquema de deslocamento de um *patch* 350×350 durante a fase de treinamento da rede, com sobreposição.

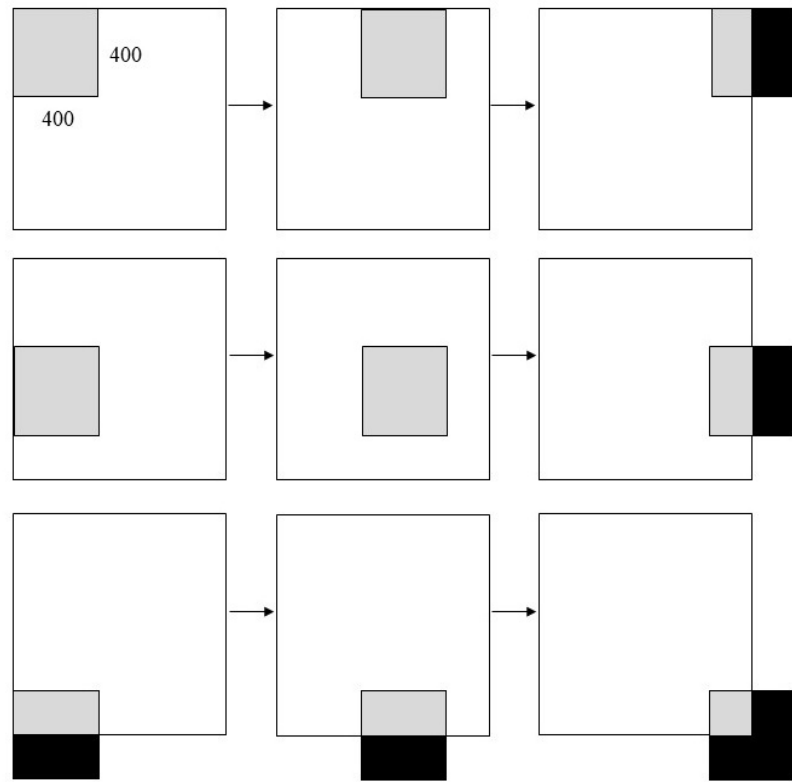


Fonte: o autor.

um total de 539 *patches* de treinamento. Como o conjunto de treinamento é composto por 12 imagens, o número total dos *patches* de treinamento será de 6.468. Entretanto, vale ressaltar que, para tamanho de *patch* cujos valores sejam múltiplos de 100, não existe possibilidade de problema de fronteira, ou seja, não há risco de ultrapassar a borda da imagem. Contudo, no caso do exemplo mostrado na Figura 8 o escolhido foi de 350×350 . Assim o *patch* se deslocará de 100 em 100 *pixels* e, ao final, dará o passo residual para completar a imagem. No caso do *patch* em tela, o deslocamento será: 0 – 350; 100 – 450; 200 – 550; 300 – 650; 400 – 750; 500 – 850; 600 – 950; 650 – 1000.

Durante o processo de teste, não há sobreposição das imagens dos *patches*, haja vista que o passo de deslocamento é o tamanho do próprio *patch*. Isso garante que não sejam realizados testes para os mesmos pedaços da imagem, removendo, assim, eventual tipo de viés no teste da rede. Por outro lado, por conta desse deslocamento, pode ocorrer um problema de fronteira. Assim, para mitigar esse problema – diferentemente da solução implementada por (ALI; WANG; AHMAD, 2021), que optou por replicar a parte do *patch* que ultrapassa a fronteira da imagem

Figura 9 – Esquema de deslocamento de um *patch* 400×400 durante a fase de teste da rede, sem sobreposição.



Fonte: o autor.

com as mesmas informações obtidas na parte de dentro da imagem – foi escolhida como solução ignorar qualquer informação referente ao pedaço que ultrapassa a fronteira. Assim, a rede não enxerga nenhum poro durante o processo de teste, sem interferir na taxa de acerto ou de erro da rede neural. A Figura 9 representa um esquema de teste utilizando um *patch* de tamanho 400×400 *pixels* em deslocamento utilizando saltos também de 400 *pixels*.

4.2.4 Procedimento de aumento de dados

Durante o processo de treinamento feito por (ALI; WANG; AHMAD, 2021), são aplicados procedimentos de aumento de dados (*data augmentation*) dos *patches* por meio de suas rotações por 90° e 180° , flipping e desfoque Gaussiano. Isso é utilizado para enriquecer o conjunto de dados de treinamento e aumentar o desempenho da rede (HUSSAIN et al., 2017).

Em contrapartida, nesta dissertação, a utilização do procedimento de aumento de dados através do desfoque Gaussiano não se mostrou satisfatório, uma vez que o subconjunto de 8 imagens composto por aquelas obtidas em cenas de crime possui distorções provocadas por

tremores; borrões; sobreposição; diferença de pressão entre a superfície e o dedo; iluminação irregular; entre outras características que já contribuem negativamente na qualidade das imagens. Quanto aos demais métodos de aumento de dados, foram aplicados o flipping horizontal, flipping vertical, rotação por 90° e por 180° , gerando, ao final do processo de treinamento, o *patch* original e mais 4 *patches* de aumento de dados, totalizando 32.340 *patches*.

4.2.5 Função de perda e otimizador

Tanto em (ALI; WANG; AHMAD, 2021), quanto neste trabalho, foram utilizados núcleos de convolução com suporte espacial de $s \times s$, inicializados aleatoriamente com uma distribuição Gaussiana tendo uma média zero e uma variância de $\sqrt{\frac{2}{s^2n}}$, em que n é a largura da janela. Uma função de perda (*loss function*) de entropia cruzada binária é otimizada para treinar os parâmetros da rede. A função de perda de entropia cruzada binária é dada por (MURPHY, 2014)

$$L = -\frac{1}{N_B} \sum_{i=1}^{N_B} (y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)), \quad (4.1)$$

em que y_i representa o i -ésimo *patch* rotulado em um lote de tamanho N_B e $\hat{y}_i$ representa o *patch* previsto correspondente, que é usado para treinar a rede.

A função de perda, por sua vez, mede a qualidade da saída de uma rede. Essa função calcula uma pontuação de distância entre o que é previsto ($\hat{y}_i$) e o que de fato a rede calcula (y_i), verificando, portanto, quão bem a rede se saiu. O truque fundamental no aprendizado profundo é usar essa pontuação como um sinal de feedback para ajustar o valor dos pesos da rede em uma direção que diminua a pontuação de perda. Quem faz esse ajuste é o otimizador, que implementa o que é chamado de algoritmo *Backpropagation* (CHOLLET, 2018).

Inicialmente, são atribuídos valores aleatórios aos pesos da rede, de modo que esta simplesmente implementa uma série de transformações aleatórias. Naturalmente, sua saída está longe do que deveria ser idealmente, e a pontuação de perda é, portanto, muito alta. Assim, o *loop* de treinamento, que, repetido um número suficiente de vezes (normalmente dezenas de iterações em milhares de exemplos), produz valores de peso que minimizam a função de perda. Uma rede com perda mínima é aquela para quais as saídas estão o mais próximo possível dos alvos: uma rede treinada (CHOLLET, 2018).

Além disso, foi empregado tanto no trabalho de (ALI; WANG; AHMAD, 2021), quanto neste, a função de otimização chamada de Gradiente Descendente Estocástico (*Stochastic*

Gradient Descent – SGD) com o esquema de Aceleração de Gradiente de Nesterov (*Nesterov Acceleration Gradient* – NAG) que foi proposto por (NESTEROV, 1983) para minimizar funções convexas suaves. A cada iteração, o método de aceleração NAG é usado para avaliar o ponto do gradiente e fornece uma correção maior e mais oportuna da velocidade de aprendizado. O esquema NAG é uma das abordagens mais eficazes entre as técnicas de otimização de primeira ordem (GU; TIAN; CHEN, 2019). Desse modo, o gradiente controla tanto a rapidez com que o viés muda durante o aprendizado, quanto a rapidez com que os pesos inseridos no neurônio mudam.

Outro fato que se constata é que as camadas iniciais aprendem mais lentamente que as camadas posteriores. Assim, o gradiente tende a diminuir à medida que ocorre o deslocamento para trás das camadas ocultas. Esse fenômeno é conhecido como o Problema da Dissipação do Gradiente (*The Vanishing Gradient Problem*) (CHOLLET, 2018).

Por fim, compete destacar o conceito de *momentum* que também merece atenção, pois está relacionado a duas questões importantes e que estão interligadas: a velocidade de convergência da CNN e o problema dos mínimos locais. Sem o *momentum*, o otimizador poderia ficar preso em um mínimo local, assim os pesos da rede não seriam corretamente atualizados em busca de um mínimo global. Na prática, isso representa perda de eficiência do algoritmo tanto na questão de não atingir os valores ótimos, quanto no tempo de processamento da operação (CHOLLET, 2018). Neste trabalho, o parâmetro de *momentum* e a taxa de aprendizado inicial são definidos como 0,9 e 0,1, respectivamente, e a taxa de aprendizado é diminuída em 0,01 a cada 10 épocas até que a taxa se torne 0,01. A rede é treinada com um tamanho de lote (*batch size*) de 34 e um número máximo de 100 épocas.

4.2.6 *Grid Search*

O *Grid Search* busca otimizar os valores dos hiperparâmetros da rede. Esses hiperparâmetros são parâmetros que não são aprendidos diretamente nos estimadores da CNN. É possível e recomendado pesquisar no espaço de hiperparâmetros o melhor *score* de validação cruzada.

Dois abordagens genéricas para pesquisa de parâmetros são fornecidas no *scikit-learn*: o *GridSearchCV*, que considera exaustivamente todas as combinações de parâmetros, e o *RandomizedSearchCV*, que pode amostrar um determinado número de candidatos de um espaço de parâmetros com uma distribuição especificada. Ambas têm *halvings* sucessivos (*HalvingGridSe-*

archCV e *HalvingRandomSearchCV*) que podem ser muito mais rápidos em encontrar uma boa combinação de parâmetros (PEDREGOSA et al., 2011).

A vantagem da aplicação do *Grid Search* é que todas as combinações possíveis de valores de parâmetros são avaliadas e a melhor combinação é mantida. Por outro lado, o tempo de processamento para obtenção desses resultados é demasiado longo. Assim, para contornar esse problema de tempo, foram estipuladas faixas estreitas de variação dos valores dos hiperparâmetros, conforme detalhado na Seção 5.2.1.

5 RESULTADOS

5.1 INTRODUÇÃO

Neste capítulo, são apresentados os resultados obtidos a partir da implementação do método proposto por (ALI; WANG; AHMAD, 2021). Para isso, foi levada em consideração a necessidade das adaptações propostas no Capítulo 4, inclusive com a utilização do *Grid Search* para otimização da escolha dos hiperparâmetros da CNN.

Além disso, os resultados aqui apresentados foram obtidos utilizando o banco de dados criado pelo autor desta dissertação, que, conforme descrito na Seção 4.2.1, é composto por 13 imagens, sendo 8 imagens de impressão digital obtidas através da perícia papiloscópica e fotografia forense em objetos relacionados a cenas de crime; e outras 5 imagens de impressão digital coletadas em leitor biométrico.

Os experimentos foram realizados no ambiente *Colab-Pro* utilizando a linguagem *Python* (versão 3.7.13). Foi utilizada uma GPU com 26,7 GB de memória RAM, sistema de interface de gerenciamento NVIDIA-SMI 460.32.03 (utilizado para monitorar e obter informações sobre dispositivos NVIDIA-GPU instalados no sistema) e processador Intel Xeon de 2.20GHz.

5.2 RESULTADOS OBTIDOS

A análise dos resultados foi dividida em duas etapas. Na primeira etapa, foram realizadas simulações buscando definir o tamanho do *patch* das imagens para fins de treinamento e teste da CNN, bem como otimizar a escolha dos hiperparâmetros da rede que retornassem os melhores valores das métricas R_T e R_F . Além disso, foram plotadas as *loss functions* de cada uma dessas 8 imagens, visando detectar possível *underfitting* ou *overfitting*. Na segunda etapa, buscou-se analisar as métricas de R_T e R_F em relação a subimagens obtidas a partir de recortes em janelas de cada uma das 8 imagens do conjunto de teste.

5.2.1 Resultados obtidos na primeira etapa

Resultados experimentais preliminares foram sugestivos para a fixação de uma faixa em que os valores dos hiperparâmetros da CNN pudessem variar quando da aplicação do *Grid Search*. Desse modo, foi estabelecido que os hiperparâmetros N , W e β variassem do seguinte

modo: N de 311 até 411, com passo de 2; W de 800 a 1.000, com passo de 25; e β de -6 a 0 com passo de 0,1, onde N é o tamanho da janela usada na etapa de binarização (3.3.5.2.1), W o tamanho da janela usada na etapa de refinamento (3.3.5.2.2) e β é o fator de Niblack (3.3.5.2.2).

Assim, para cada valor de *patch* (100×100 ; 150×150 ; 200×200 ; 250×250 ; 300×300 ; 350×350 ou 400×400) utilizado durante o treinamento e teste das imagens, foram encontrados valores de N , W e β que produzem os melhores valores de R_T e R_F . A Figura 11 ilustra o fluxograma da simulação do algoritmo que é realizado para cada uma das 8 imagens que compõem a base de dados do conjunto de teste. A Tabela 7, por sua vez, apresenta um resumo dos valores computados para cada imagem de teste.

A análise da Tabela 7 permite inferir que, embora utilizados tamanho de *patch* e hiperparâmetros que produzam os melhores valores de R_T e de R_F , os resultados obtidos quando comparados com os trazidos na Tabela 6, estão muito aquém, quando utilizados outros algoritmos de detecção de poros. Isso pode ser sugestivo para duas situações: ou a quantidade de imagens utilizadas no banco de dados não é suficiente para garantir a avaliação do modelo com precisão; ou a qualidade e resolução das imagens que compõem o *dataset* não estão adequadas.

Nessa esteira, foram plotadas as funções de perdas para cada uma das 8 imagens do conjunto de teste da CNN, cujos valores do tamanho de *patch* e hiperparâmetros estão registrados na Tabela 7. A análise dessas curvas permite observar se a rede neural está em eventual *underfitting* ou *overfitting*. A Figura 10 ilustra as *loss functions* dessas 8 imagens.

A partir da análise da Figura 10, permite-se inferir que o erro do treinamento e do teste diminuem à medida que as épocas aumentam. Sabendo que o *underfitting* ocorre quando o erro de treinamento não reduz, ou seja, o modelo não é capaz de determinar uma relação nos dados de treinamento, e que, por outro lado, o *overfitting* ocorre quando o erro de treinamento diminui, porém o de teste continua alto, ou seja, a rede não é capaz de generalizar para exemplos não vistos, conclui-se que a CNN não está nem em *underfitting* nem em *overfitting*. Logo, é possível dizer que a CNN tem dados de treinamento sólidos e é capaz de generalizar a detecção de poros para um novo conjunto de dados, e não somente para dados já conhecidos.

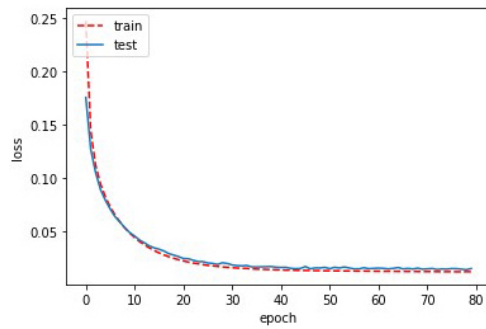
Dessa forma, resta averiguar o comportamento do algoritmo de (ALI; WANG; AHMAD, 2021) frente à qualidade das imagens que compõem o *dataset* deste trabalho. Para isso foi realizada uma análise em janelas, feita através de recortes em regiões específicas das imagens do banco de teste. A Seção 5.2.2 aprofunda essa análise.

Tabela 7 – Desempenho do esquema proposto após escolha do tamanho do *patch* e otimização dos hiperparâmetros pelo *Grid Search*.

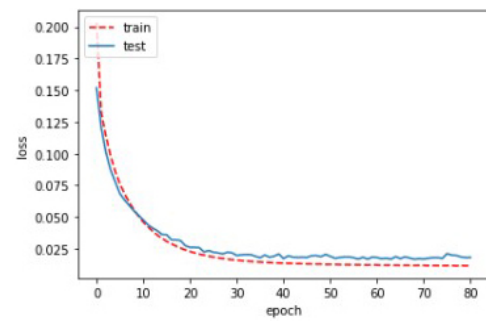
Imagem teste	Tamanho do <i>patch</i>	Hiperparâmetros	R_T	R_F
Imagem 1	150×150	$N = 409$ $W = 925$ $\beta = -5,2$	70,5%	68,4%
Imagem 2	100×100	$N = 405$ $W = 1000$ $\beta = -4,3$	58,0%	66,3%
Imagem 3	150×150	$N = 405$ $W = 950$ $\beta = -4,2$	62,7%	70,0%
Imagem 4	100×100	$N = 405$ $W = 1000$ $\beta = -5,1$	53,3%	76,7%
Imagem 5	100×100	$N = 405$ $W = 1000$ $\beta = -5,1$	77,9%	58,1%
Imagem 6	150×150	$N = 405$ $W = 1000$ $\beta = -5,1$	47,2%	60,7%
Imagem 7	150×150	$N = 401$ $W = 1000$ $\beta = -5,0$	62,8%	56,5%
Imagem 8	150×150	$N = 391$ $W = 1000$ $\beta = -5,0$	71,1%	60,5%

Fonte: o autor (2022).

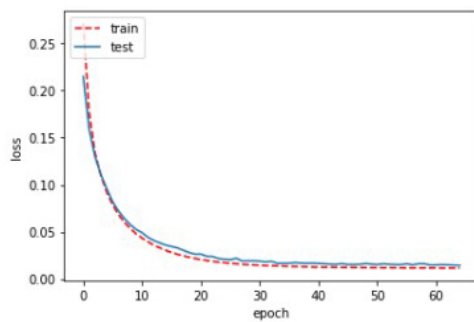
Figura 10 – Função de perdas (*loss functions*) para as imagens de 1 a 8 do banco de teste.



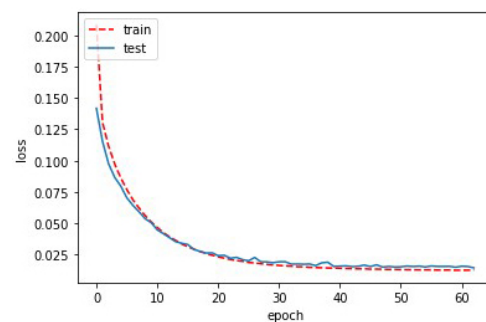
(a) Imagem 1.



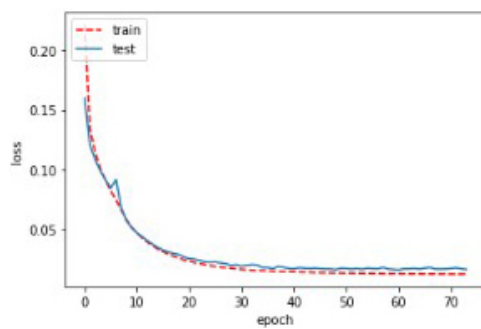
(b) Imagem 2.



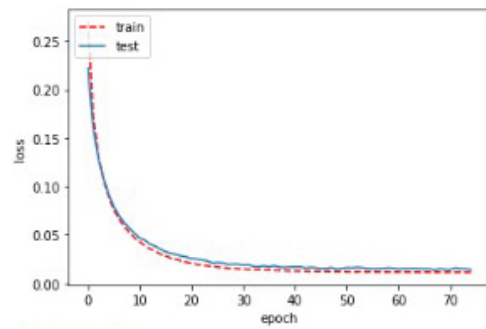
(c) Imagem 3.



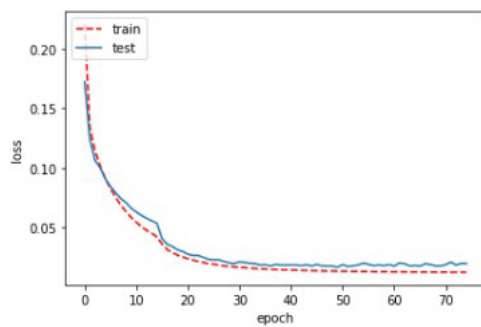
(d) Imagem 4.



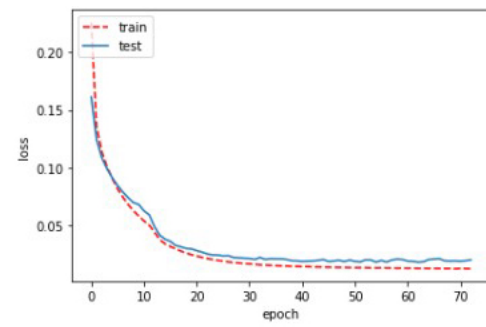
(e) Imagem 5.



(f) Imagem 6.

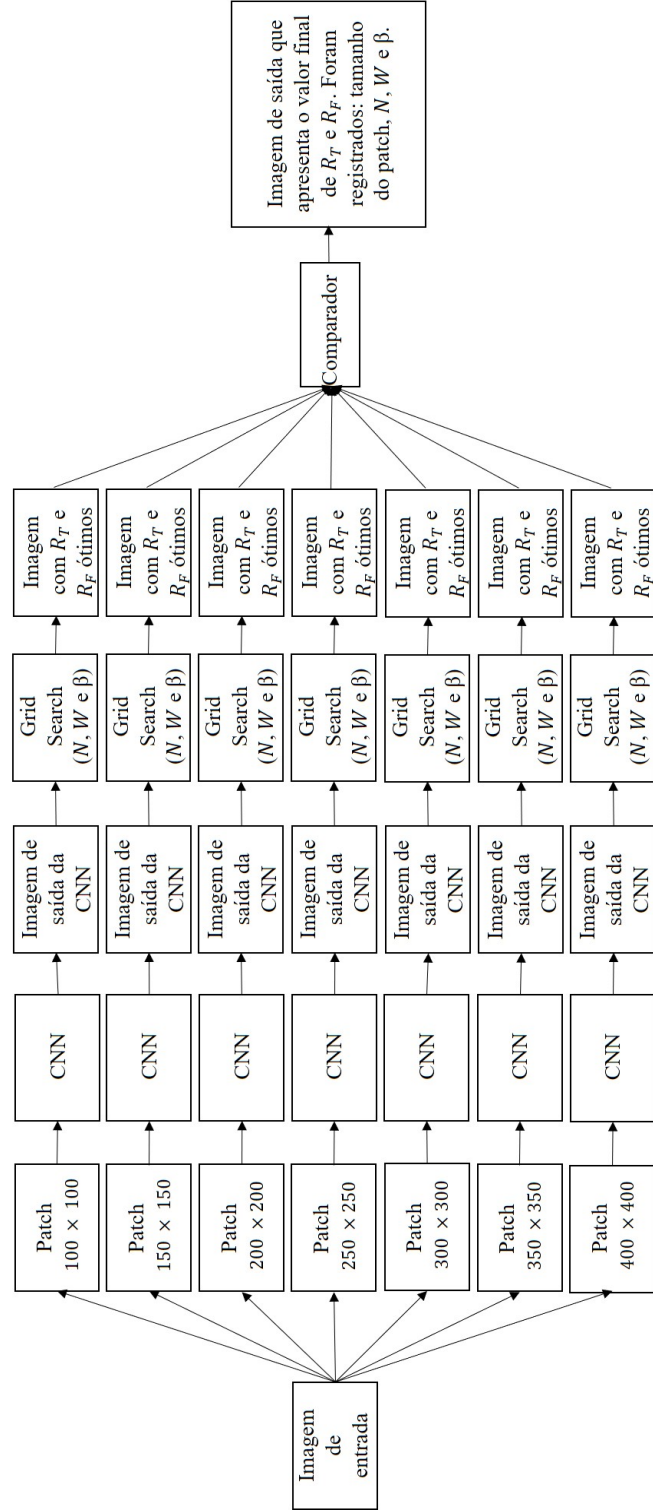


(g) Imagem 7.



(h) Imagem 8.

Figura 11 – Fluxograma da simulação do algoritmo para cada uma das 8 imagens do conjunto de teste.



Fonte: o autor (2022).

Tabela 8 – Desempenho do esquema proposto após aplicação da análise em janelas.

Imagem	Janela	Poros reais	Poros detectados	Poros verdadeiros detectados	Poros falsamente detectados	R_T	R_F
Imagem 1	400×400	22	34	20	14	90,9%	41,2%
Imagem 2	200×400	34	40	26	14	76,5%	35,0%
Imagem 3	400×200	46	47	37	10	80,4%	21,3%
Imagem 4	400×200	26	27	21	6	80,8%	22,2%
Imagem 5	400×200	27	24	22	2	81,5%	8,3%
Imagem 6	600×200	48	33	27	6	56,3%	18,2%
Imagem 7	500×200	26	22	21	1	80,8%	4,5%
Imagem 8	400×400	34	39	26	13	76,5%	33,3%

Fonte: o autor.

5.2.2 Resultados obtidos na segunda etapa

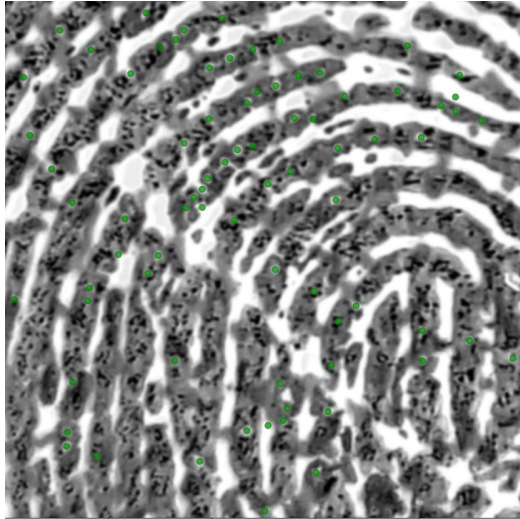
Apesar de globalmente o conjunto de imagens do banco de teste possuir problemas como: iluminação irregular; ruído provocado pelo ambiente de captação das fotografias; distorções nas impressões digitais provocadas por borrões, marcas de arrasto ou textura irregular das superfícies onde estas encontram-se depositadas, determinadas subregiões dessas imagens mantiveram-se preservadas.

Por esse motivo, a análise persistiu em verificar o comportamento do algoritmo de detecção de poros proposto por (ALI; WANG; AHMAD, 2021) nessas sub-regiões específicas. O tamanho de cada janela foi escolhido levando-se em conta a necessidade de haver no mínimo 20 poros reais dentro dessa subimagem, o que garante a quantidade mínima de minúcias para o processo de identificação, conforme detalhado na Seção 2.2.2.

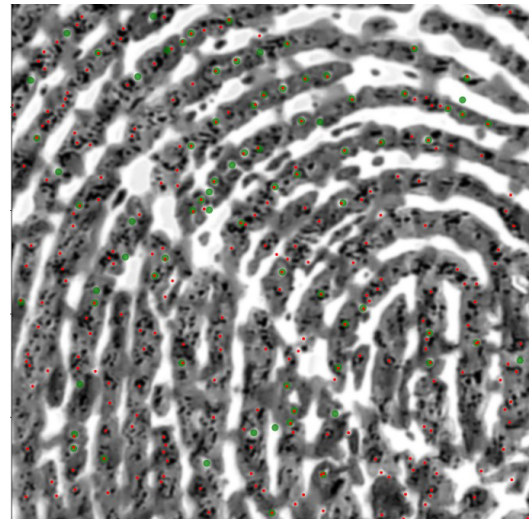
A Tabela 8 ilustra o resultado de R_T e R_F obtido depois da análise em janela. As Figuras 12, 13, 14, 15, 16, 17, 18 e 19 ilustram as imagens de impressão digital 1 a 8 do conjunto de teste, cujos poros foram: marcados pelo especialista humano; detectados pela CNN antes do refinamento; detectados pela CNN depois do refinamento; detectados pela CNN com a análise em janela, nessa ordem.

A partir da análise da Tabela 8, infere-se que os valores das métricas de R_T e R_F apresentam melhoras significativas quando comparado aos resultados registrados na Tabela 7. Isso indica que o método de detecção de poros proposto por (ALI; WANG; AHMAD, 2021), levando-se em conta as adaptações propostas, é aplicável às imagens forenses, ainda que de maneira mais assertiva, em janelas, em que pese a baixa qualidade e resolução das imagens.

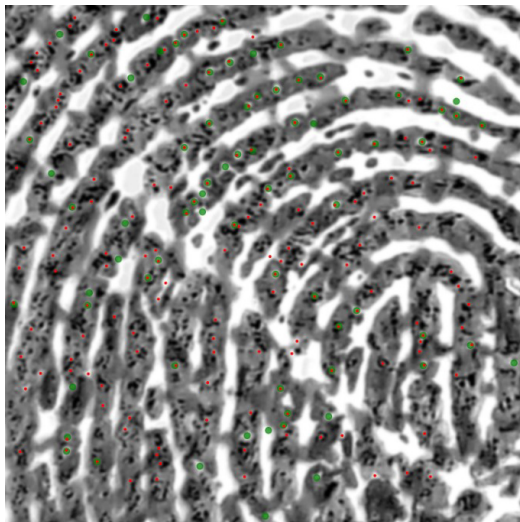
Figura 12 – Representação dos poros da imagem 1 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.



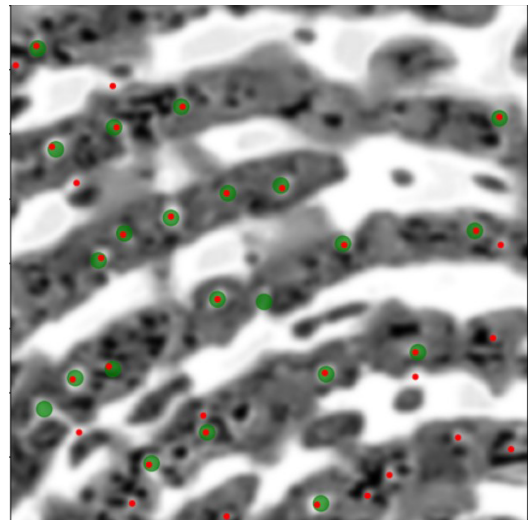
(a) Imagem original da impressão digital 1 com a marcação manual dos poros.



(b) Imagem 1 com a detecção dos poros pela CNN antes da etapa de refinamento.



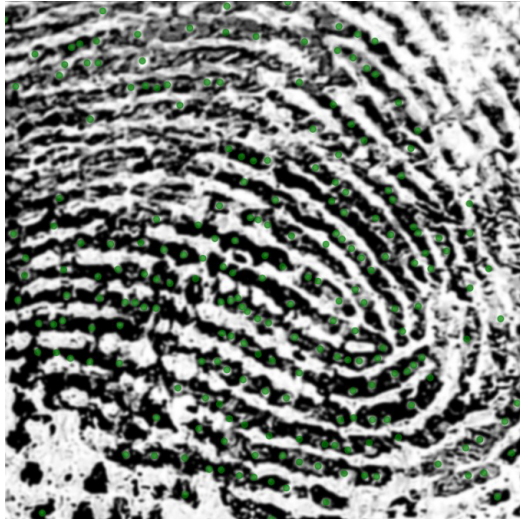
(c) Imagem 1 com a detecção dos poros pela CNN da etapa de refinamento.



(d) Subimagem de 1 (400×400 pixels) na região de interesse para avaliação de R_T e R_F .

Fonte: o autor.

Figura 13 – Representação dos poros da imagem 2 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.



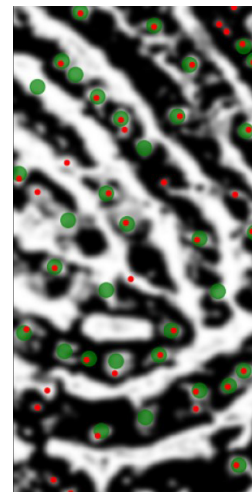
(a) Imagem original da impressão digital 2 com a marcação manual dos poros.



(b) Imagem 2 com a detecção dos poros pela CNN antes da etapa de refinamento.



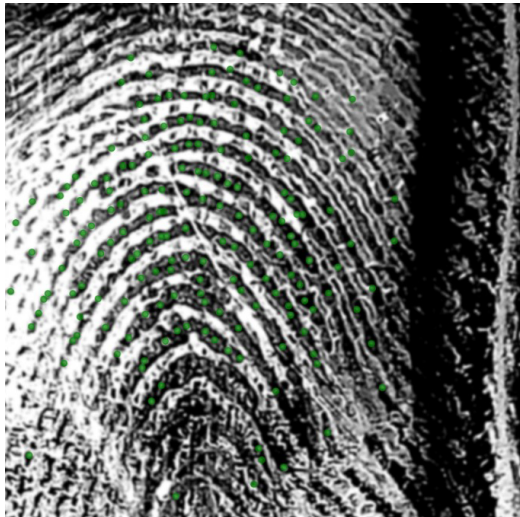
(c) Imagem 2 com a detecção dos poros pela CNN da etapa de refinamento.



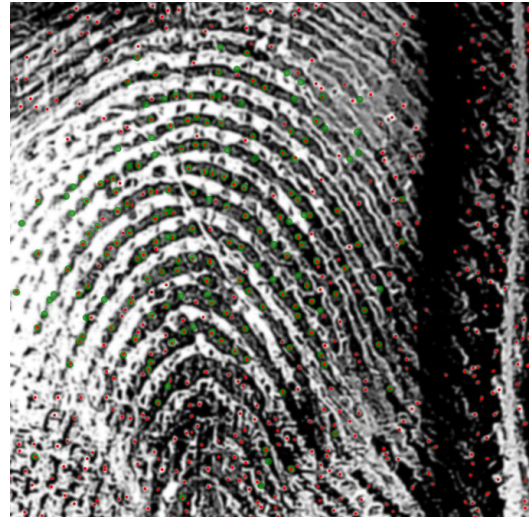
(d) Subimagem de 2 (200×400 pixels) na região de interesse para avaliação de R_T e R_F .

Fonte: o autor.

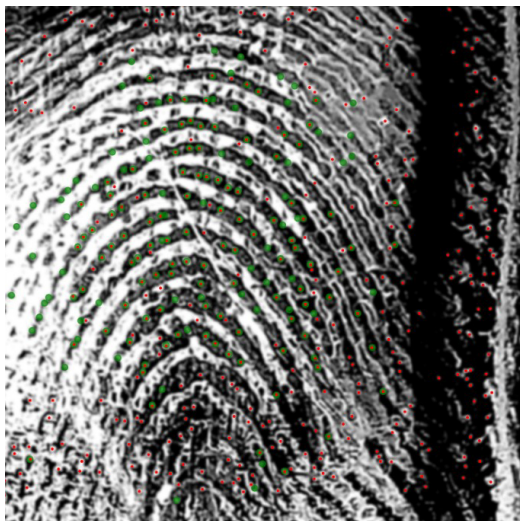
Figura 14 – Representação dos poros da imagem 3 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.



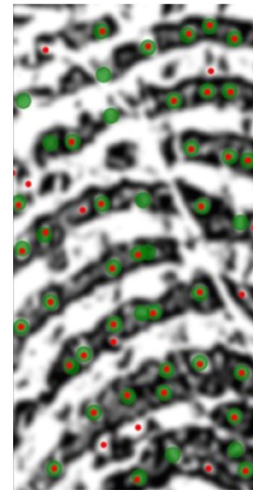
(a) Imagem original da impressão digital 3 com a marcação manual dos poros.



(b) Imagem 3 com a detecção dos poros pela CNN antes da etapa de refinamento.



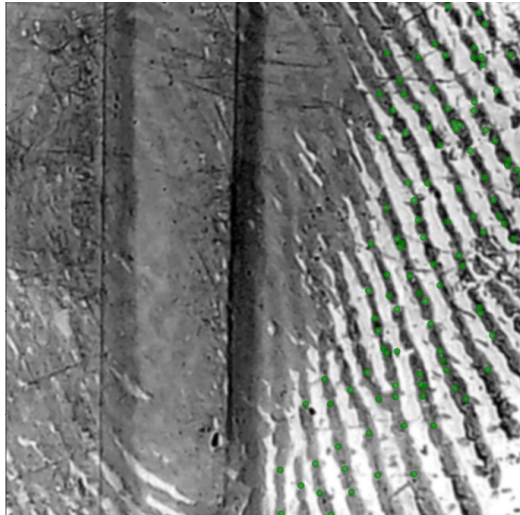
(c) Imagem 3 com a detecção dos poros pela CNN da etapa de refinamento.



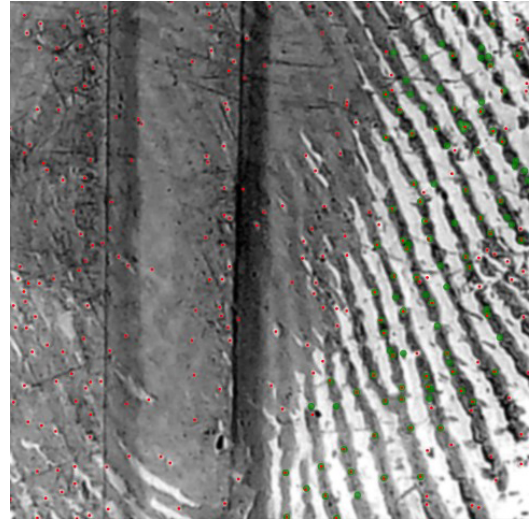
(d) Subimagem de 3 (200×400 pixels) na região de interesse para avaliação de R_T e R_F .

Fonte: o autor.

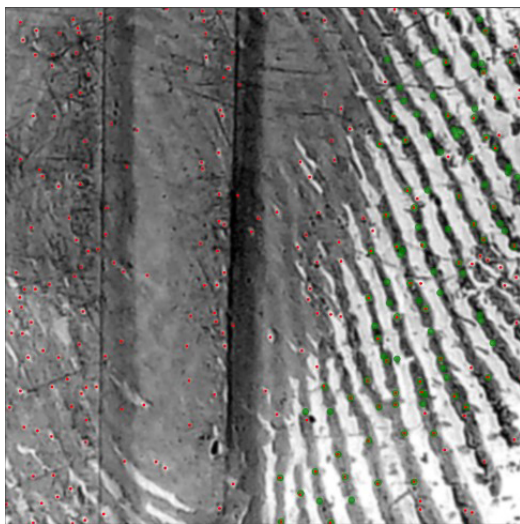
Figura 15 – Representação dos poros da imagem 4 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.



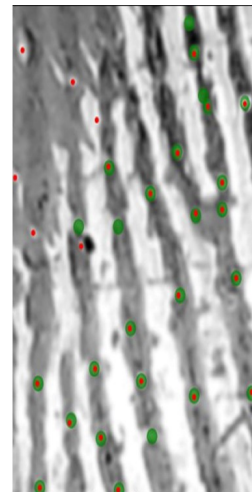
(a) Imagem original da impressão digital 4 com a marcação manual dos poros.



(b) Imagem 4 com a detecção dos poros pela CNN antes da etapa de refinamento.



(c) Imagem 4 com a detecção dos poros pela CNN da etapa de refinamento.



(d) Subimagem de 4 (200×400 pixels) na região de interesse para avaliação de R_T e R_F .

Fonte: o autor.

Figura 16 – Representação dos poros da imagem 5 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.



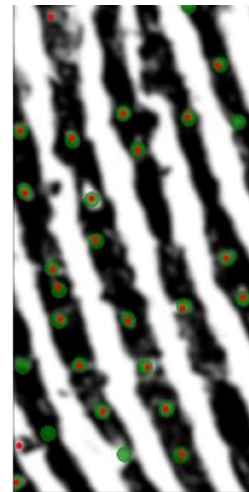
(a) Imagem original da impressão digital 5 com a marcação manual dos poros.



(b) Imagem 5 com a detecção dos poros pela CNN antes da etapa de refinamento.



(c) Imagem 5 com a detecção dos poros pela CNN da etapa de refinamento.



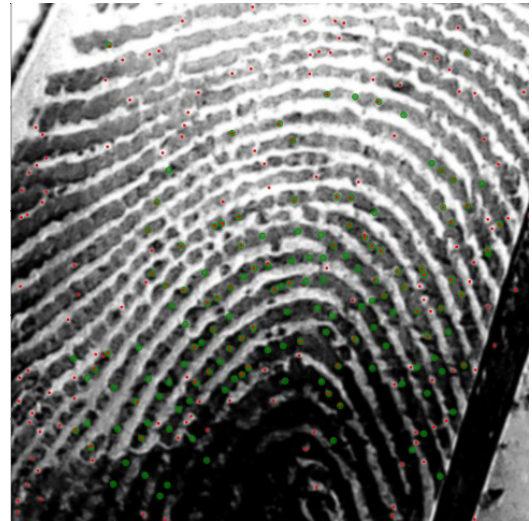
(d) Subimagem de 5 (200×400 pixels) na região de interesse para avaliação de R_T e R_F .

Fonte: o autor.

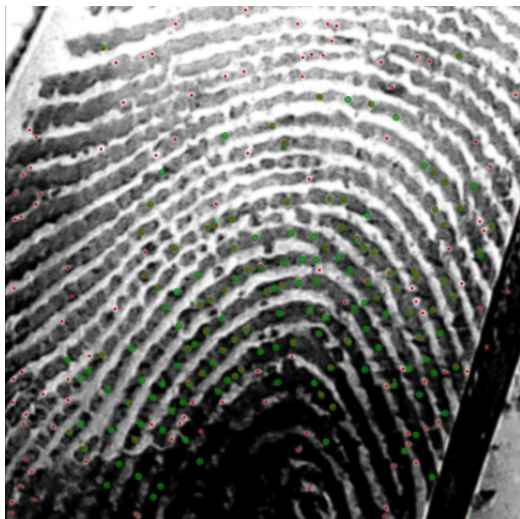
Figura 17 – Representação dos poros da imagem 6 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.



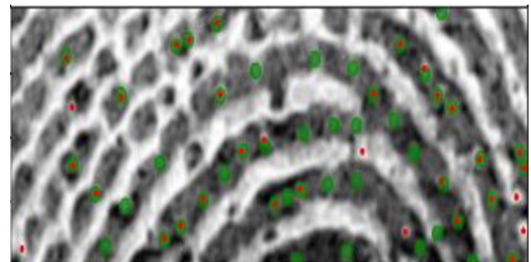
(a) Imagem original da impressão digital 6 com a marcação manual dos poros.



(b) Imagem 6 com a detecção dos poros pela CNN antes da etapa de refinamento.



(c) Imagem 6 com a detecção dos poros pela CNN da etapa de refinamento.



(d) Subimagem de 6 ($600 \times 200 \text{ pixels}$) na região de interesse para avaliação de R_T e R_F .

Fonte: o autor.

Figura 18 – Representação dos poros da imagem 7 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.



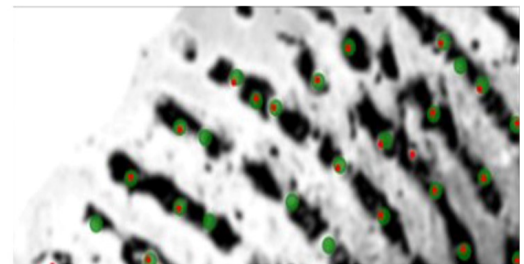
(a) Imagem original da impressão digital 7 com a marcação manual dos poros.



(b) Imagem 7 com a detecção dos poros pela CNN antes da etapa de refinamento.



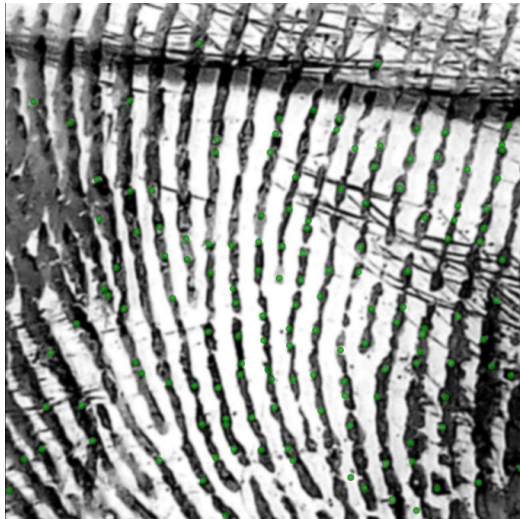
(c) Imagem 7 com a detecção dos poros pela CNN da etapa de refinamento.



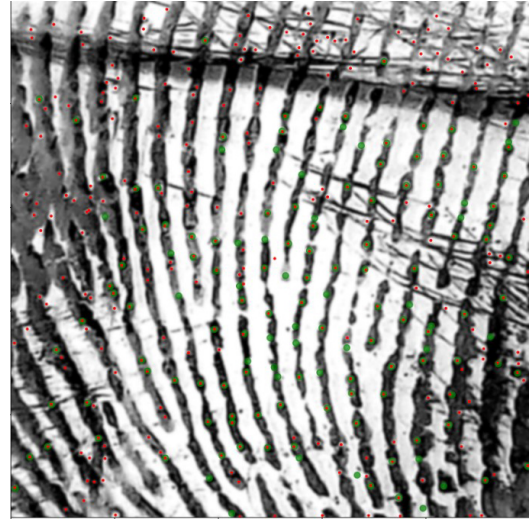
(d) Subimagem de 7 (500×200 pixels) na região de interesse para avaliação de R_T e R_F .

Fonte: o autor.

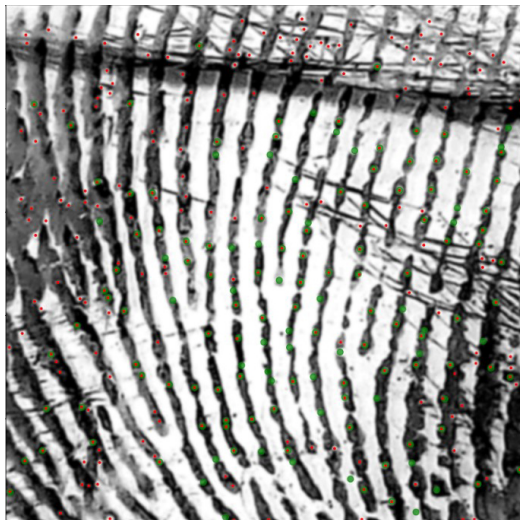
Figura 19 – Representação dos poros da imagem 8 do banco de teste, em que os marcados manualmente pelo autor estão na cor verde e os detectados pela CNN na cor vermelha.



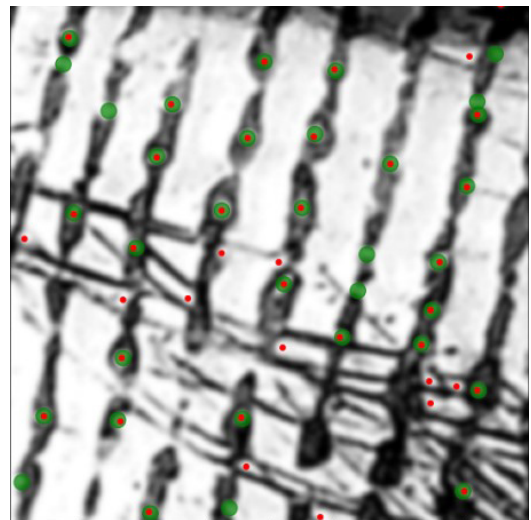
(a) Imagem original da impressão digital 8 com a marcação manual dos poros.



(b) Imagem 8 com a detecção dos poros pela CNN antes da etapa de refinamento.



(c) Imagem 8 com a detecção dos poros pela CNN da etapa de refinamento.



(d) Subimagem de 8 (400×400 pixels) na região de interesse para avaliação de R_T e R_F .

Fonte: o autor.

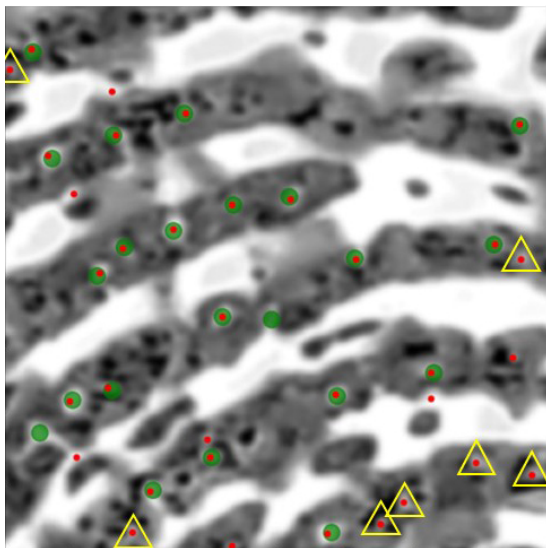
Uma outra comparação que merece ser feita é quanto aos resultados obtidos quando (LABATI et al., 2017) utilizaram as imagens latentes do banco criado em ambiente de laboratório. Esse é o resultado que mais se aproxima do obtido nesta dissertação e por isso merece ser detalhado. Os resultados obtidos foram de $R_T = 52,70\%$ e $R_F = 24,00\%$, ao passo que para as simulações em janelas o R_T e o R_F médios deste trabalho ficaram em $77,96\%$ e $23,00\%$, respectivamente. Isso representa um aumento 48% para a taxa de acerto da rede, enquanto a taxa de erro da rede se manteve praticamente constante. Isso quer dizer que comparativamente a rede passou a detectar mais poros corretamente, mantendo a mesma taxa de erro na detecção.

Além disso, quando da marcação manual de poros feita para a base forense utilizada, o especialista humano optou por não registrar os centroides dos poros cujo grau de certeza fosse questionável. Entretanto, isso não implica necessariamente que não existam poros nas regiões não marcadas. A Figura 20, por exemplo, ilustra a marcação manual de poros na Figura ??, cujos poros o especialista humano não registrou, mas que foram detectadas pela CNN e são passíveis de marcação. Isso elevaria a taxa de poros verdadeiros detectados, aumentando, portanto, o R_T calculado. Essas regiões estão marcadas por um triângulo amarelo.

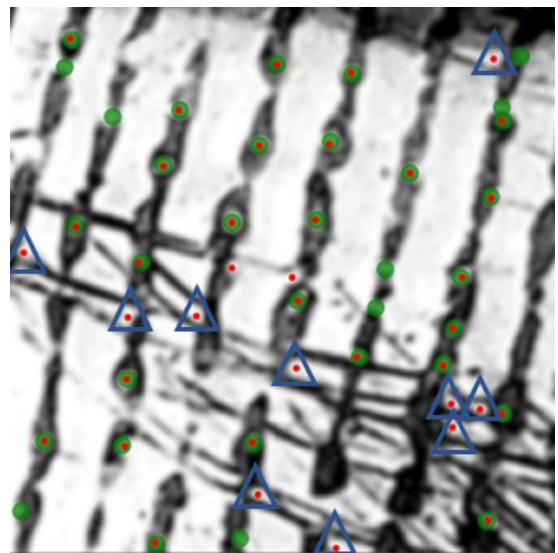
Por outro lado, a Figura 20 ilustra a marcação manual de poros na Figura ??, mas que o especialista humano também não registrou e que claramente não são poros. Na verdade, essa marcação errada pela rede neural foi influenciada pela presença de ranhuras no anteparo onde a impressão digital foi depositada. Ou seja, essa mesma impressão digital, caso depositada em um anteparo liso e sem rugosidades, muito provavelmente não influenciaria negativamente a detecção pela CNN. Existem diversas regiões na imagem cujos poros foram falsamente detectados elevando equivocadamente o R_F calculado. Essas regiões estão marcadas por um triângulo azul.

De todo modo, o fato é que nas 8 imagens do conjunto de teste o algoritmo se mostrou viável à sua principal funcionalidade: detecção de poros reais voltada ao processo pericial forense de individualização, já que nas janelas analisadas foi possível detectar ao menos 20 poros, número suficiente de acordo com a literatura existente para garantir a individualização de uma pessoa.

Figura 20 – Análise detalhada das subimagens de 1 e 8.



(a) Subimagem de 1, com destaque para as regiões em amarelo que representam prováveis poros detectados pela CNN, mas não detectados pelo especialista.



(b) Subimagem de 8, com destaque para as regiões em azul que representam poros falsamente detectados pela CNN, devido a ranhuras no anteparo.

Fonte: o autor.

6 CONCLUSÕES

A detecção de poros em imagens de impressão digital através das CNNs é algo relativamente recente, mas que tem apresentado bons resultados. Para este trabalho foi utilizada uma topologia de rede neural convolucional proposta por (ALI; WANG; AHMAD, 2021), que tem links de atalhos paralelos às camadas convolucionais, criando caminhos de retorno para propagação do Gradiente (*backpropagation gradient*). Além disso, a rede detecta recursos de nível alto, médio e baixo, utilizando para tanto três estágios convolucionais em cascata. A esse tipo peculiar de CNN é dado o nome de Rede Neural Residual (RNN).

O trabalho proposto por (ALI; WANG; AHMAD, 2021) provou ser bastante eficiente quando aplicado ao banco de dados do PolyU HRF, cujas imagens possuem resolução de 1.200 dpi e são obtidas através de leitores biométricos específicos. Entretanto, o desafio desta dissertação consiste, entre outras coisas, em avaliar esse modelo de rede para imagens de vestígios de impressão digital obtidas em local de crime e cujas imagens tenham resolução de 96 dpi.

Assim, como primeira contribuição deste trabalho, foi utilizada a base de dados criada pelo próprio autor, onde as coordenadas dos centroides dos poros foram catalogadas manualmente. Parte da base é composta de 8 imagens de vestígios, que são utilizadas tanto na etapa de treinamento, quanto de teste da rede neural. A outra parte da base é composta de 5 imagens obtidas através de leitor biométrico, mas que foram utilizadas apenas durante o treinamento. O que se buscou avaliar utilizando essa nova base foi a eficiência da taxa de detecção de falsos positivos (R_T) e falsos negativos (R_F), associado à utilização do *Grid Search* para escolha e otimização dos hiperparâmetros da rede neural, procedimento que, até onde se sabe, nunca foi feito em outros tipos de trabalhos.

Realizados os experimentos descritos na Seção 5.2.1, os resultados alcançados retornaram valores de $R_T = 77,9\%$, para a imagem 5, onde foram utilizados os valores otimizados dos hiperparâmetros em 150×150 para o tamanho do *patch*, $N = 409$, $W = 925$ e $\beta = -5,2$. Enquanto isso, o valor alcançado por (ALI; WANG; AHMAD, 2021) foi de 96,69%, cujos hiperparâmetros utilizados foram de 40×40 para o tamanho do *patch*, $N = 40$ e $W = 60$, sem que tenha havido qualquer mecanismo de otimização da escolha dos hiperparâmetros da rede. Os valores de R_F , por sua vez, ficaram em 56,5% para a imagem 7, valor muito superior aos

registrados por (ALI; WANG; AHMAD, 2021), cujo máximo registrado foi de 10,2%. Isso é um indicativo de que a rede poderia não estar detectando corretamente os poros nas imagens.

Algumas hipóteses podem explicar os resultados obtidos. O tamanho da base de dados trazida pelo autor, que pelo fato de possuir poucas imagens, poderia ser inexpressiva e a rede neural poderia estar em *underfitting*, ou seja, o modelo não se mostrou capaz de determinar uma relação nos dados de treinamento. Ou ainda, a rede poderia não estar sendo capaz de generalizar o teste para exemplos não vistos, hipótese conhecida como *overfitting*. Fato é que foram traçadas as *loss functions* das 8 imagens de teste e foi descartado possível causa de *underfitting* ou *overfitting*.

Desse modo, atribui-se como possibilidade para explicar os resultados, que em certa medida não estão satisfatórios, à qualidade das imagens do banco. Isso porque a resolução de 96 dpi não é adequada para este tipo de análise, em que pese a base de dados disponível estar nessa qualidade devido às imagens terem sido obtidas através de câmeras fotográficas de dispositivos celulares. Além disso, as imagens possuem diversos problemas como: iluminação irregular; ruído provocado pelo ambiente de captação das fotografias; distorções nas impressões digitais provocadas por borrões, marcas de arrasto ou textura irregular das superfícies.

Em que pese tais pontos negativos, a dissertação prosseguiu no sentido de analisar não mais as imagens como um todo, mas analisar subimagens do conjunto de teste, consoante descrito na Seção 5.2.2. Essa análise em janelas foi escolhida de maneira que as subimagens apresentassem ao menos 20 poros reais, número suficiente para garantir a identificação. Como consequência dessa análise, foram encontrados valores de $R_T = 90,9\%$ e de $R_F = 4,5\%$, utilizando para isso janelas de tamanho 400×400 e 500×200 , para as imagens 1 e 7, respectivamente. Essa análise em janelas é outra contribuição desta dissertação.

Ao final, a utilização do método proposto nesse trabalho mostrou-se eficiente no processo de detecção de poros em imagens de vestígios de impressão digital encontrados em local de crime, desde que analisadas regiões específicas das imagens, ou seja, em janelas. Isso é um fato que contribui significativamente no âmbito das Ciências Forenses, já que com a marcação de no mínimo 20 poros é possível estabelecer a identidade de uma pessoa, não havendo, portanto, necessidade de analisar a imagem como um todo para identificação.

6.1 SUGESTÕES PARA TRABALHOS FUTUROS

Para trabalhos futuros podem ser citados:

- Avaliação de outros mecanismos de otimização dos hiperparâmetros da rede neural, como, por exemplo, o *Randomized Search*, ou algoritmos que utilizem inteligência de enxames.
- Investigação do efeito do aumento ou diminuição das camadas convolucionais da topologia de CNN usada para detecção de poros;
- Extensão da pesquisa para outras topologias de redes neurais aplicadas à detecção de poros em imagens de vestígios de impressões digitais;
- Estudo da complexidade computacional da técnica proposta.

REFERÊNCIAS

- AGARWAL, D.; BANSAL, A. A utility of pores as level 3 features in latent fingerprint identification. **Multimedia Tools and Applications**, v. 80, p. 23605–23624, 2021. Citado na página 16.
- ALI, M.; WANG, C.; AHMAD, M. O. An efficient convolutional neural network for fingerprint pore detection. **IEEE Transactions on Biometrics, Behavior, and Identity Science**, v. 3, p. 332–346, 2021. Citado 24 vezes nas páginas 5, 6, 13, 15, 19, 21, 24, 36, 37, 42, 45, 46, 47, 48, 49, 50, 51, 52, 53, 56, 57, 61, 72 e 73.
- ALSHDADI, A. et al. Exploiting level 1 and level 3 features of fingerprints for liveness detection. **Biomedical Signal Processing and Control**, v. 61, p. 169135 – 169155, 2020. Citado na página 17.
- ANAND, V.; KANHANGAD, V. Pore detection in high-resolution fingerprint images using deep residual network. In: **Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2019. Citado 4 vezes nas páginas 21, 24, 35 e 36.
- ANAND, V.; KANHANGAD, V. Cross-sensor pore detection in high-resolution fingerprint images. **IEEE Sensors Journal**, v. 22, p. 555 – 564, 2021. Citado na página 19.
- ASHBAUGH, D. R. **Quantitative-Qualitative Friction Ridge Analysis**. [S.l.]: CRC Press, 1999. ISBN 978-1420048810. Citado 2 vezes nas páginas 18 e 28.
- BACANIN, N. et al. Artificial neural networks hidden unit and weight connection optimization by quasi-reflection-based learning artificial bee colony algorithm. **IEEE Access**, v. 9, p. 169135–169155, 2021. Citado na página 20.
- BURRUS, R. G. C.; GUO, H. **Introduction to Wavelets and Wavelet Transforms**. [S.l.]: Prentice Hall, 1998. Citado na página 25.
- CHAMPOD, C. et al. **Fingerprints and Other Ridge Skin Impressions**. [S.l.]: CRC PRESS, 2004. Citado 3 vezes nas páginas 15, 16 e 18.
- CHOLLET, F. **Deep Learning With Python**. [S.l.]: Manning, Publications Co, 2018. ISBN 978-1617294433. Citado 4 vezes nas páginas 21, 22, 53 e 54.
- CUMMINS, H.; MIDLO, C. **Finger Prints, Palms and Soles**. [S.l.]: Dover Publications, 1943. Citado na página 28.
- DAHIA, G.; SEGUNDO, M. P. Improving fingerprint pore detection with a small fcn. In: **Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2018. Citado 5 vezes nas páginas 9, 21, 24, 34 e 35.
- DING, C.; TAO, D. Robust face recognition via multimodal deep face representation. **IEEE Transactions on Multimedia**, v. 17, p. 2049 – 2058, 2015. Citado na página 21.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification**. [S.l.]: John Wiley Sons, 2012. ISBN 978-0-471-05669-0. Citado na página 49.

DVORNYCHENKO, V. N.; GARRIS, M. D. **Summary of NIST Latent Fingerprint Testing Workshop**. [S.l.]: NISTIR 7377, 2006. Citado 2 vezes nas páginas 15 e 16.

ER, M. J.; WANG, N. An efficient leave-one-out cross-validation-based extreme learning machine (eloo-elm) with minimal user intervention. **IEEE TRANSACTIONS ON CYBERNETICS**, v. 46, p. 1939–1951, 2016. Citado na página 49.

GARRIS, R. M. M. D. **NIST special database 27 fingerprint minutiae from latent and matching Tenprint images**. 2000. <https://doi.org/10.6028/NIST.IR.6534>. Citado na página 19.

GU, P.; TIAN, S.; CHEN, Y. Q. Iterative learning control based on nesterov accelerated gradient method. **IEEE Access**, v. 7, p. 115836 – 115842, 2019. Citado na página 54.

HAYKIN, S. **Neural Networks and Learning Machines**. [S.l.]: Pearson Education, 2011. Citado na página 20.

HAZARIKA, P.; RUSSELL, D. A. Advances in fingerprint analysis. **Angewandte Chemie**, v. 51, p. 3493–3722, 2012. Citado na página 12.

HE, K. et al. Deep residual learning for image recognition. In: **IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Las Vegas, NV, USA: [s.n.], 2016. Citado na página 22.

HONG, L.; WAN, Y.; JAIN, A. Fingerprint image enhancement: algorithm and performance evaluation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 20, p. 777 – 789, 1998. Citado na página 25.

HUSSAIN, Z. et al. Differential data augmentation techniques for medical imaging classification tasks. In: **AMIA Annual Symposium Proceedings**. Stanford, EUA: [s.n.], 2017. p. 979–984. Citado na página 52.

IOFFE, S.; SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: **32nd International Conference on Machine Learning**. [S.l.: s.n.], 2015. p. 448–456. Citado na página 38.

JAIN, A. K.; CHEN, Y.; DEMIRKUS, M. Pores and ridges: High-resolution fingerprint matching using level 3 features. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 29, p. 15 – 27, 2007. Citado 7 vezes nas páginas 12, 17, 18, 24, 25, 28 e 29.

JANG, H. U. et al. Deeppore: Fingerprint pore extraction using deep convolutional neural networks. **IEEE Signal Processing Letters**, v. 24, p. 1808 – 1812, 2017. Citado 4 vezes nas páginas 21, 24, 33 e 34.

KALCHBRENNER, N.; GREFFENSTETTE, E.; BLUNSON, P. A convolutional neural network for modelling sentences. In: **52nd Annual Meeting of the Association for Computational Linguistics (ACL)**. [S.l.: s.n.], 2014. p. 1–11. Citado na página 21.

KILIÇ, S.; ASKERZADE, I.; KAYA, Y. Using resnet transfer deep learning methods in person identification according to physical actions. **IEEE Access**, v. 8, p. 220364 – 220373, 2020. Citado 2 vezes nas páginas 12 e 22.

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: **International joint Conference on artificial intelligence**. Stanford, EUA: [s.n.], 1995. p. 1137–1145. Citado na página 49.

- KOMARINSKI, P. **Automated Fingerprint Identification Systems (AFIS)**. [S.l.]: Elsevier Academic Press, 2004. ISBN 9780124183513. Citado 2 vezes nas páginas 15 e 16.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: **Advances in Neural Information Processing Systems**. [S.l.: s.n.], 2012. Citado na página 21.
- KUMAR, A. **PolyU HRF Database, The Hong Kong Polytechnic University**. 2009. Citado 7 vezes nas páginas 33, 34, 35, 36, 42, 43 e 48.
- LABATI, R. D. et al. Computational intelligence for biometric applications: a survey. **International Journal of Computing**, v. 1, p. 40 – 49, 2016. Citado na página 20.
- LABATI, R. D. et al. A novel pore extraction method for heterogeneous fingerprint images using convolutional neural networks. **Pattern Recognition Letters**, v. 113, p. 58–66, 2017. Citado 10 vezes nas páginas 9, 20, 21, 24, 31, 32, 33, 44, 47 e 70.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, p. 436–444, 2015. Citado na página 20.
- LEMES, R. P. et al. Dynamic pore filtering for keypoint detection applied to newborn authentication. In: **22nd International Conference on Pattern Recognition**. Stockholm, Sweden: [s.n.], 2014. Citado 3 vezes nas páginas 24, 28 e 29.
- LIU, F. et al. Fingerprint pore matching using deep features. **Pattern Recognition**, v. 102, p. 107 – 208, 2020. Citado na página 25.
- LIU, N. et al. Accurate iris segmentation in noncooperative environments using fully convolutional networks. In: **International Conference on Biometrics (ICB)**. Halmstad, Sweden: [s.n.], 2016. Citado na página 21.
- LONG, J.; SHELHAMER, E.; DARREL, T. Fully convolutional networks for semantic segmentation. In: **Proceedings of the IEEE conference on computer vision an**). [S.l.: s.n.], 2015. p. 3431–3440. Citado na página 34.
- MALTONI, D. et al. **Handbook of Fingerprint Recognition**. [S.l.]: Pringer-Verlag, 2009. Citado na página 15.
- MITCHELL, T. M. **Machine Learning**. [S.l.]: McGraw-Hill Science, 1997. ISBN 978-0070428072. Citado na página 39.
- MURPHY, K. P. **The Machine Learning: A Probabilistic Perspective**. [S.l.]: MIT Press, 2014. Citado na página 53.
- NAIR, V.; HINTON, G. E. Rectified linear units improve restricted boltzmann machines. In: **Proceedings of the 27th International Conference on Machine Learning**. Haifa, Israel: [s.n.], 2010. p. 807–814. Citado na página 38.
- NESTEROV, Y. A method of solving a convex programming problem with convergence rate $o(1/k^2)$. **Soviet mathematics doklady**, v. 27, p. 372–376, 1983. Citado na página 54.
- NIBLACK, W. **Introduction to Digital Image Processing**. [S.l.]: London, England: Pearson Education, 1998. ISBN 978-0134806747. Citado na página 40.

- OTSU, N. A threshold selection method from gray-level histograms. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 9, p. 62 – 66, 1979. Citado 3 vezes nas páginas 29, 31 e 33.
- PARSONS, N. R. et al. Rotationally invariant statistics for examining the evidence from the pores in fingerprints. **Law, Probability and Risk**, v. 7, p. 1 – 14, 2008. Citado na página 26.
- PAULINO, A. A.; FENG, J.; JAIN, A. K. Latent fingerprint matching using descriptor-based hough transform. **IEEE Transactions on Information Forensics and Security**, v. 8, p. 31 – 45, 2013. Citado na página 19.
- PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011. Citado 2 vezes nas páginas 49 e 55.
- RAIS, N. B.; HANIF, M. S.; TAJ, I. A. Adaptive thresholding technique for document image analysis. In: **8th International Multitopic Conference, Proceedings of INMIC**. [S.l.: s.n.], 2004. p. 61–66. Citado na página 41.
- SEMEDO, D. et al. Contact-free airport borders with biometrics-on-the-move. In: **IEEE International Workshop on Biometrics and Forensics (IWBF)**. Rome, Italy: [s.n.], 2021. Citado na página 12.
- SERMANET, P. et al. Overfeat: Integrated recognition, localization and detection using convolutional networks. In: **Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2014. Citado na página 22.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. In: **Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2015. Citado na página 34.
- SOUZA, M. A. et al. Raman spectroscopy of fingerprints and chemometric analysis for forensic sex determination in humans. **Forensic Chemistry**, v. 27, p. 100395, 2021. Citado na página 12.
- SRIVASTAVA, R. K.; GREFF, K.; SCHMIDHUBER, J. Training very deep networks. In: **Adv. neural Inf. Process. Syst.** [S.l.: s.n.], 2015. p. 2377–2385. Citado na página 22.
- SU, H. et al. A deep learning approach towards pore extraction for high-resolution fingerprint recognition. In: **IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. New Orleans, LA, USA: [s.n.], 2017. Citado na página 25.
- TAMILARASI, P.; RANI, R. U. Diagnosis of crime rate against women using k-fold cross validation through machine learning. In: **International Conference on Computing Methodologies and Communication (ICCMC)**. [S.l.: s.n.], 2020. p. 1034 – 1038. Citado na página 49.
- TEIXEIRA, R. F. S.; LEITE, N. J. Improving pore extraction in high resolution fingerprint images using spatial analysis. In: **IEEE International Conference on Image Processing (ICIP)**. Paris, France: [s.n.], 2014. Citado 3 vezes nas páginas 24, 28 e 29.
- U. S. Department of Justice. **The fingerprint - Sourcebook**. [S.l.]: Washington, D.C.: National Institute of Justice, 2009. Citado na página 12.
- WANG, H. et al. Fingerprint pore extraction using u-net based fully convolutional network. In: **Chinese Conference on Biometric Recognition**. [S.l.: s.n.], 2017. Citado na página 25.

ZHANG, D. et al. Selecting a reference high resolution for fingerprint recognition using minutiae and pores. **IEEE Transactions on Instrumentation and Measurement**, v. 60, p. 863 – 871, 2011. Citado 2 vezes nas páginas 12 e 17.

ZHAO, Q. et al. Adaptive fingerprint pore modeling and extraction. **Pattern Recognition**, v. 43, p. 2833–2844, 2010. Citado 4 vezes nas páginas 24, 27, 28 e 29.

ZHAO, Q. et al. High resolution partial fingerprint alignment using pore–valley descriptors. **Pattern Recognition**, v. 43, p. 1050–1061, 2010. Citado 5 vezes nas páginas 24, 26, 27, 28 e 29.

ZHAO, Q. et al. Direct pore matching for fingerprint recognition. In: **International Conference on Biometrics**. Springer: [s.n.], 2009. p. 597–606. Citado na página 17.