

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

WESLEY AUGUSTO CATUZZO DE BONA

**CONSTRUÇÃO DE DIGITAIS ROLADAS DE RECÉM-NASCIDOS POR MEIO
DE TÉCNICAS AVANÇADAS DE REGISTRO DE IMAGEM**

PATO BRANCO

2024

WESLEY AUGUSTO CATUZZO DE BONA

**CONSTRUÇÃO DE DIGITAIS ROLADAS DE RECÉM-NASCIDOS POR MEIO
DE TÉCNICAS AVANÇADAS DE REGISTRO DE IMAGEM**

**CONSTRUCTION OF DIGITAL ROLLED NEWBORN MEANS OF ADVANCED
IMAGE REGISTRATION TECHNIQUES**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção
do título de Bacharel em Engenharia de
Computação do Curso de Bacharelado em
Engenharia de Computação da Universidade
Tecnológica Federal do Paraná.

Orientador: Prof. Dr. Jefferson Tales Oliva

Coorientador: Prof. Dr. Luiz Fernando Puttow
Southier

PATO BRANCO

2024



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

WESLEY AUGUSTO CATUZZO DE BONA

**CONSTRUÇÃO DE DIGITAIS ROLADAS DE RECÉM-NASCIDOS POR MEIO
DE TÉCNICAS AVANÇADAS DE REGISTRO DE IMAGEM**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção
do título de Bacharel em Engenharia de
Computação do Curso de Bacharelado em
Engenharia de Computação da Universidade
Tecnológica Federal do Paraná.

Data de aprovação: 17/junho/2024

Prof. Dr. Fábio Favarim
Universidade Tecnológica Federal do Paraná, Campus Pato Branco

Prof. Me. Vinicius Pegorini
Universidade Tecnológica Federal do Paraná, Campus Pato Branco

Prof. Dr. Jefferson Tales Oliva
Universidade Tecnológica Federal do Paraná, Campus Pato Branco

Prof. Dr. Luiz Fernando Puttow Southier
Universidade Tecnológica Federal do Paraná, Campus Pato Branco

PATO BRANCO

2024

AGRADECIMENTOS

Primeiramente, agradeço a Deus, fonte de toda sabedoria, força e inspiração. Sua presença constante me guiou nos momentos de incerteza, me fortaleceu nos desafios e me proporcionou serenidade e discernimento ao longo de toda esta jornada. Sem Sua proteção e bênçãos, este trabalho não seria possível.

À minha família, expresso minha eterna gratidão. Aos meus pais e meu irmão, por todo amor, apoio e compreensão em todos os momentos, e por me incentivarem a seguir em frente mesmo diante das dificuldades. Vocês são minha base e minha inspiração.

À minha namorada, pelo amor incondicional, paciência e motivação que me proporcionou. Seu apoio constante foi essencial para que eu pudesse me dedicar inteiramente a este projeto.

Aos professores e orientadores, meus sinceros agradecimentos. Aos professores do projeto e da banca, pela dedicação, conhecimento compartilhado e por todas as contribuições valiosas que enriqueceram este trabalho. Especialmente aos meus orientadores, prof Jefferson e Luiz, por sua orientação incansável, paciência e por acreditarem no meu potencial. Suas críticas construtivas e incentivo foram fundamentais para o desenvolvimento deste trabalho.

Aos meus amigos, pela camaradagem e suporte ao longo desta jornada. Agradeço por todos os momentos de descontração e por estarem sempre ao meu lado, proporcionando o equilíbrio necessário entre estudo e lazer.

Agradece-se à InfantID por seu inestimável apoio operacional ao grupo de pesquisa em biometria neonatal na UTFPR. Sua parceria tem sido fundamental para o avanço de exploração de tecnologias biométricas no atendimento neonatal. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior -Brasil (CAPES) - Código de Financiamento 001 e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ).

A todos que estiveram comigo, meu muito obrigado. Este trabalho é o reflexo da colaboração e do apoio que recebi de cada um de vocês.

RESUMO

A identificação biométrica é um método de reconhecimento de indivíduos com base em características únicas e mensuráveis de seus corpos. A principal vantagem dessa tecnologia é sua alta precisão e a dificuldade de falsificação, proporcionando uma forma segura de autenticação. A impressão digital é dos tipos de identificação biométrica que possui um papel fundamental no contexto de saúde e segurança. Levando em consideração os recém-nascidos como público-alvo, problemas como troca de bebês em hospitais, tráfico de crianças, problemas com a vacinação e mortes são comuns e podem ser evitados. A captura de uma digital pode ser realizada de duas formas: de forma pousada ou de forma rolada. Na digital rolada, mais detalhes são registrados, pois a área capturada do dedo é maior. Porém, para a realização da coleta biométrica em recém-nascidos, opta-se por fazer a captura de forma pousada por conta do tamanho do dedo, do comportamento da criança e dos resultados razoáveis obtidos. Sendo assim, este trabalho objetivou construir imagens de digitais roladas a partir de várias digitais pousadas, utilizando o método de registro de imagem e o método da soma. O primeiro método envolve três etapas principais: identificação dos pontos-chave, correspondência das características e distorção da imagem. Os resultados mostraram que, embora o registro de imagem clássico tenha apresentado algumas limitações, a técnica de soma de imagens revelou um aumento significativo na quantidade de minúcias extraídas. Isso demonstra que a técnica de soma contribuiu substancialmente para a identificação de características importantes das impressões digitais.

Palavras-chave: impressão digital; recém-nascido; redes neurais; registro de imagem; match.

ABSTRACT

Biometric identification is a method of recognizing individuals based on unique and measurable characteristics of their bodies. The main advantage of this technology is its high precision and difficulty in forgery, providing a secure form of authentication. The fingerprint is one of the types of biometric identification that plays a fundamental role in the context of health and safety. Taking newborns into account as the target audience, problems such as changing babies in hospitals, child trafficking, problems with vaccination and deaths are common and can be avoided. Capturing a fingerprint can be carried out in two ways: flat or rolling. In digital rolled, more details are recorded as the captured area of the finger is larger. However, to perform biometric collection on newborns, it is chosen to capture it in a flat manner due to the size of the finger, the child's behavior and the reasonable results obtained. Therefore, this work aimed to construct rolled fingerprint images from several fingerprints placed using the image registration method and the sum method. The first method involves three main steps: identifying key points, matching features and distorting the image. The results showed that, although classical image registration had some limitations, the image summing technique revealed a significant increase in the amount of minutiae extracted. This demonstrates that the summation technique contributed substantially to the identification of important characteristics of fingerprints.

Keywords: fingerprint; newborn; neural networks; image registration; match.

LISTA DE FIGURAS

Figura 1 – Classificação de digital	14
Figura 2 – Classificação das minúcias	15
Figura 3 – Captura pousada x Captura rolada	15
Figura 4 – Distribuição de pontos com limite definido	16
Figura 5 – Exemplos de curvas DET (esquerda), ROC (centro) e CMC (direita) . . .	17
Figura 6 – Detecção dos pontos-chave	18
Figura 7 – Correspondência dos pontos-chave	20
Figura 8 – Imagem após a distorção	21
Figura 9 – Rede Neural	22
Figura 10 – Rede Neural Convolucional	24
Figura 11 – Protocolo de pesquisa proposto	29
Figura 12 – Pontuação dos Frames	31
Figura 13 – Digital em contato e sem contato com o <i>scanner</i>	32
Figura 14 – Pousadas Identificadas	33
Figura 15 – Detecção dos pontos-chave utilizando o algoritmo KAZE.	35
Figura 16 – Correspondência das características utilizando um BFMatcher.	36
Figura 17 – Digital rolada construída a partir de três digitais pousadas.	37
Figura 18 – <i>Frames</i> antes e depois do pré-processamento	38
Figura 19 – <i>Frames</i> 1ª pousada pré-processados	38
Figura 20 – <i>Frames</i> 2ª pousada pré-processados	38
Figura 21 – <i>Frames</i> 3ª pousada pré-processados	39
Figura 22 – Registro dos <i>frames</i> da 1ª pousada pré-processados	39
Figura 23 – Registro dos <i>frames</i> da 2ª pousada pré-processados	40
Figura 24 – Registro dos <i>frames</i> da 3ª pousada pré-processados	40
Figura 25 – <i>Frames</i> da 1ª pousada pré-processados e segmentados	41
Figura 26 – <i>Frames</i> da 2ª pousada pré-processados e segmentados	41
Figura 27 – <i>Frames</i> da 3ª pousada pré-processados e segmentados	42
Figura 28 – Registro dos <i>frames</i> da 1ª pousada pré-processados e segmentados .	42
Figura 29 – Registro dos <i>frames</i> da 2ª pousada pré-processados e segmentados .	43
Figura 30 – Registro dos <i>frames</i> da 3ª pousada pré-processados e segmentados .	43

Figura 31 – <i>Frames</i> da 1ª pousada remapeados e segmentados	44
Figura 32 – <i>Frames</i> da 2ª pousada remapeados e segmentados	44
Figura 33 – <i>Frames</i> da 3ª pousada remapeados e segmentados	45
Figura 34 – Soma resultante dos <i>frames</i> da 1ª pousada	45
Figura 35 – Soma resultante dos <i>frames</i> da 2ª pousada	46
Figura 36 – Soma resultante dos <i>frames</i> da 3ª pousada	46

LISTA DE ABREVIATURAS E SIGLAS

Siglas

AKAZE	<i>Accelerated-KAZE</i>
BFMatcher	Combinador de Força Bruta, do inglês <i>Brutal Force Matcher</i>
BRIEF	<i>Binary Robust Independent Elementary Features</i>
CEP	Comitê de Ética em Pesquisa
CMC	Característica de Correspondência Cumulativa, do inglês <i>Cumulative Match Characteristic</i>
CropRoi	Recorte da Região de Interesse, do inglês <i>Crop the Region of Interest</i>
DET	Características de Detecção de Erros, do inglês <i>Detection Error Trade-off</i>
EER	Taxa de Erro Equivalente, do inglês <i>Equal Error Rate</i>
FAIS	Sistemas de Identificação Automática de Impressões Digitais, do inglês <i>Fingerprint Automatic Identification Systems</i>
FAR	Taxa de Falso Aceitação, do inglês <i>False Accepted Rate</i>
FAST	<i>Features from Accelerated Segment Test</i>
FED	<i>Fast Explicit Diffusion</i>
FRR	Taxa de Falsas Rejeições, do inglês <i>False Reject Rate</i>
ICP	<i>Iterative Closest Point</i>
IFViT	Representação Interpretável de Comprimento Fixo para Correspondência de Impressões Digitais via Vision Transformer, do inglês <i>Interpretable Fixed-Length Representation for Fingerprint Matching via Vision Transformer</i>
KNN	<i>K-Nearest Neighbors</i>
MRF	Campo Aleatório de Markov, do inglês <i>Markov Random Field</i>
NFIQ	NIST <i>Fingerprint Image Quality</i>
ORB	<i>Oriented Fast and Rotate Brief</i>

ppi	<i>pixels per inch</i>
RANSAC	<i>Random Sample Consensus</i>
RNC	Rede Neural Convolucional
ROC	Características Operacionais do Receptor, do inglês <i>Receiver Operating Characteristics</i>
SIFT	<i>Scale-invariant feature transform</i>
SURF	<i>Speeded Up Robust Features</i>
TAR	Taxa de Aceitação Verdadeira, do inglês <i>True Acceptance Rate</i>
TPIR	Taxa de Identificação Positiva Verdadeira, do inglês <i>True Positive Identification Rate</i>
TRR	Taxa de Rejeição Verdadeira, do inglês <i>True Reject Rate</i>
UTFPR	Universidade Tecnológica Federal do Paraná
ViT	<i>Vision Transformer</i>

SUMÁRIO

1	INTRODUÇÃO	11
1.1	Objetivo geral	12
1.2	Objetivos específicos	12
2	REFERENCIAL TEÓRICO	14
2.1	Reconhecimento utilizando impressões digitais	14
2.1.1	Tipos de captura de impressão digital	14
2.1.2	Extração e <i>match</i>	16
2.2	Registro de Imagem	17
2.2.1	Detecção de pontos-chave e descrição de características	17
<u>2.2.1.1</u>	<u>SIFT</u>	18
<u>2.2.1.2</u>	<u>SURF</u>	19
<u>2.2.1.3</u>	<u>ORB</u>	19
<u>2.2.1.4</u>	<u>KAZE</u>	19
2.2.2	Correspondência de características	20
2.2.3	Distorção de imagem	21
2.3	Redes neurais	21
2.3.1	<i>Deep learning</i>	22
<u>2.3.1.1</u>	<u>Redes Neurais Convolucionais</u>	23
3	TRABALHOS RELACIONADOS	25
4	MATERIAIS E MÉTODOS	27
4.1	Materiais	27
4.2	Método	28
4.2.1	Registro de imagem sem processamento	29
4.2.2	Registro de imagem com pré-processamento	31
4.2.3	Registro de imagem com pré-processamento e segmentação	33
4.2.4	Soma de imagens	33
5	RESULTADOS E DISCUSSÕES	35
5.1	Registro de imagem sem processamento	35
5.2	Registro de imagem com pré-processamento	37
5.3	Registro de imagem com pré-processamento e segmentação	41

5.4	Soma de imagens	44
5.5	Discussões	47
6	CONCLUSÃO	49
	REFERÊNCIAS	50

1 INTRODUÇÃO

Na década de 1960, a população mundial tinha em média 3,03 bilhões de pessoas, aumentando para 7,95 bilhões em 2022 (Mundial, 2023). A cada ano mais de 134 milhões de crianças nascem no mundo, totalizando mais de 656 milhões de menores de cinco anos (United Nations, 2023). Com esse aumento da população mundial, surgem muitas dificuldades no controle e identificação de crianças e recém-nascidos, segundo a UNICEF United Nations International Children's Emergency Fund (2023b), cerca de 13.400 crianças com essa idade morreram por dia em 2022 por conta de doenças ou subnutrição. Nesse contexto, outros problemas são agravados como a troca de bebês em hospitais, o sequestro e o tráfico infantil (Sedlak *et al.*, 2002), além de transtornos gerados em sistemas de saúde, como de vacinação, esse endurecido por falta de identificação adequada (WHO World Health Organization, 2023a). No mundo, a contagem de crianças que não receberam vacinas é cerca de 14,3 milhões em 2022 (WHO World Health Organization, 2023b). Além disso, 1,5 milhão de crianças morrem a cada ano devido a doenças evitáveis por vacinação (UNICEF United Nations International Children's Emergency Fund, 2023a).

Para amenizar os desafios apresentados anteriormente, diversos métodos de reconhecimento por biometria foram propostos, seja utilizando retina, sola do pé, a palma da mão, entre outras. Embora o reconhecimento facial seja amplamente adotado atualmente, a coleta de impressões digitais em dedos é a principal abordagem de identificação, pois as imagens resultantes são únicas para cada pessoa, tornando-se uma ferramenta eficaz (Li, 2009). Essa técnica permite a autenticação de indivíduos com base em padrões únicos denominados minúcias¹ presentes na superfície do dedo, tornando esse método notável pela sua capacidade de fornecer uma identificação precisa e confiável (Dargan; Kumar, 2020). No entanto, é importante destacar que a coleta de impressões digitais pode ser um desafio em certos cenários, como o registro de recém-nascidos. Nesses casos, as digitais são notavelmente pequenas, o que pode dificultar a identificação das minúcias necessárias para uma correspondência precisa. Além disso, bebês e crianças pequenas nem sempre são cooperativos durante o processo de coleta, além de suas mãos serem delicadas e, muitas vezes, podendo estar úmidas e/ou sujas, interferindo nos resultados do reconhecimento. Adicionalmente, tem-se a característica de que os sensores utilizados para essa coleta foram desenvolvido para adultos, cujas imagens de digitais geralmente possuem resolução de 500 *pixels per inch* (ppi), o que dificulta a captura dos detalhes das digitais de recém nascidos (Southier *et al.*, 2023). Sensores capazes de gerar imagens com maior resolução estão disponíveis, porém são mais caros.

Também, as taxas de verificação (*match*) com digitais adultas atingem 98.68%, enquanto para recém-nascidos (até seis meses de idade), o *match* é de 64.27% (Jain; Cao; Arora, 2014). Dessa forma, é necessário encontrar soluções que contribuam para a melhoria do *match* em

¹ Características específicas e únicas que permitem a identificação individual de uma pessoa a partir de suas impressões digitais.

bebês. Uma das formas de melhoria está em contornar dificuldades inerentes da captura de digitais em recém-nascidos.

A captura de impressões digitais nos adultos é feita de forma rolada, ou seja, a pessoa apoia a digital sobre o leitor do escâner e, durante a captura, realiza a rolagem do dedo de um lado para o outro, compreendendo dessa maneira toda a superfície do dedo. Esse procedimento de rolagem em adultos frequentemente gera resultados satisfatórios porque trata-se de digitais com minúcias bem formadas e com características bem definidas. Já no caso dos bebês, a captura é feita de forma pousada por conta do comportamento dos recém-nascidos e das irregularidades encontradas na digital. Para isso, o dedo da criança é apoiado sobre o leitor do escâner, onde a captura da imagem é realizada. Porém, o resultado possui limitações, por exemplo, a região capturada de forma pousada do dedo do recém-nascido possui menos características de identificação do que se fosse de forma rolada, além disso a baixa resolução do escâner pode ser outro fator (Southier *et al.*, 2023). Com base nessas comparações de procedimentos e também nas taxas apresentadas, percebe-se a importância de obter-se os registros das digitais roladas por conta da rica quantidade de detalhes e de informações, fundamentais na correspondência. Logo, espera-se que a obtenção das digitais roladas dos recém-nascidos a partir das digitais pousadas possa contribuir para a melhora da correspondência.

Para a obtenção dessas digitais roladas, existem os métodos de registro de imagem que são responsáveis por fazerem correspondências dessas imagens a partir de ângulos e posições diferentes. O *Speeded Up Robust Features* (SURF) é um exemplo de algoritmo que encontra as relações entre objetos da mesma cena (Bay; Tuytelaars; Gool, 2006). Em sistemas de captura de digitais infantis, pode ser realizada gravações de vídeos com vários (*frames*) da mesma digital, o que possibilita a aplicação do método de registro. Neste trabalho é aplicado registro de imagem para construir uma composição de digitais, utilizando os *frames* de alta resolução a fim de se obter a impressão digital rolada.

1.1 Objetivo geral

O objetivo desse trabalho é desenvolver um algoritmo que identifique, a partir de múltiplas imagens, quais são as melhores partes de uma digital e realizar o registro de imagem desses fragmentos. Essa abordagem visa aprimorar a qualidade das imagens resultantes, contribuindo assim para uma correspondência mais precisa das impressões digitais.

1.2 Objetivos específicos

- Coletar uma base de dados com vídeos de digitais de recém-nascidos;
- Realizar o pré-processamento dos vídeos em *frames*;

- Aplicar os métodos de registro de imagem para construir artificialmente imagens roladas (com mais detalhes) com base nos *frames* extraídos (imagens pousadas);
- Realizar testes e comparar os resultados de correspondência.

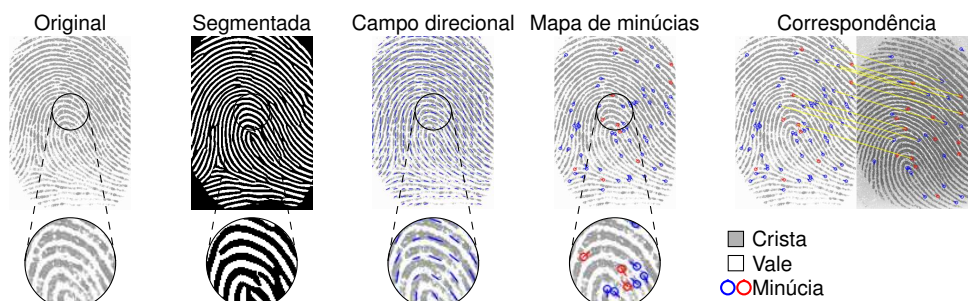
2 REFERENCIAL TEÓRICO

2.1 Reconhecimento utilizando impressões digitais

A tarefa fundamental no gerenciamento de identidade é estabelecer a associação entre um indivíduo e os seus dados pessoais (Jain; Ross; Nandakumar, 2011). O reconhecimento biométrico utilizando impressões digitais é um dos métodos que estabelece a identidade da pessoa, de forma única e imutável, com base em suas características físicas ou comportamentais.

Uma impressão digital consiste em cristas papilares e vales que formam um padrão de textura orientado na ponta do dedo. As cristas são as regiões mais acentuadas da pele, enquanto que os vales são as regiões que separam uma crista de outra. Além disso, as cristas ainda podem ter alguns formatos diferentes ao longo do seu desenho, o que caracterizam as chamadas minúcias (Li; Jain, 2015). Uma categoria crucial de impressões digitais são as segmentadas, as quais são obtidas através da segmentação, um processo que divide uma imagem em seus componentes constituintes. Normalmente, uma imagem de impressão digital capturada consiste em dois componentes distintos: o primeiro plano e o fundo (Bazen; Gerez, 2001). Essas características e classificações são mostradas na Figura 1.

Figura 1 – Classificação de digital



Fonte: Southier *et al.* (2023).

As minúcias são classificadas a partir de três parâmetros: posição, orientação e tipo. A posição e a orientação são referentes à localização e à direção da minúcia na digital, respectivamente. O tipo define a classificação das minúcias entre bifurcação, ilhota, encerro, ponta de linha, entre outras (Natosafe, 2022), como mostrado da Figura 2.

2.1.1 Tipos de captura de impressão digital

A aquisição de uma impressão digital costuma ocorrer por meio de um dispositivo de entrada de dados, como um escâner, que converte a imagem para o formato digital. Porém, essa captura pode ocorrer de forma pousada ou rolada, as quais referem-se ao posicionamento dos dedos durante o processo de coleta e varia de acordo com a necessidade, como alguma

Figura 2 – Classificação das minúcias

Representação	Descrição
	Terminação
	Lago
	Bifurcação
	Espora
	Ponto ou Ilha
	Ponte

Fonte: NascimentoFilho *et al.* (2019).

limitação do equipamento, condição física ou comportamental do indivíduo (Julião Medeiros, 2023). A principal diferença entre esses dois métodos está na forma com que a impressão é capturada, como mostrada na Figura 3.

A impressão digital rolada é obtida durante a rolagem do dedo de uma extremidade à outra (*e.g.* esquerda para direita) sobre uma superfície que contenha tinta ou outro material que possa ajudar no procedimento de captura. Nesse cenário de rolagem, é possível obter todas as informações e características da impressão digital, como as cristas, sulcos e detalhes minuciosos devido a possibilidade de captura de imagem de toda a área de interesse do dedo. Essa abordagem é comum em procedimentos de identificação criminal, no qual um indivíduo é solicitado a rolar os dedos em um papel ou superfície coberta de tinta para criar uma impressão digital completa (Jain; Ross; Nandakumar, 2011).

A impressão digital pousada é obtida quando a pessoa simplesmente pressiona o dedo contra uma superfície, como um escâner. Nesse caso, não há rolagem do dedo sobre a superfície e o resultados terão uma menor área coletada do dedo. O dispositivo de captura é equipado com sensores preparados para a digitalização de impressão digital na posição em que o dedo é colocado. Essa forma captura comumente utilizada em sistemas de autenticação (Matos, 2000).

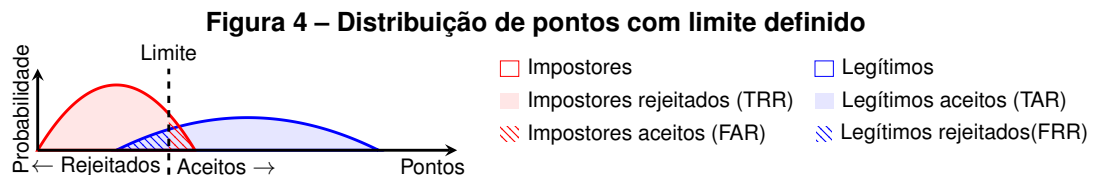
Figura 3 – Captura pousada x Captura rolada

Fonte: Julião Medeiros (2023).

2.1.2 Extração e *match*

De acordo com Southier *et al.* (2023), para extrair características, desenvolver modelos e realizar correspondência, utilizam-se Sistemas de Identificação Automática de Impressões Digitais, do inglês *Fingerprint Automatic Identification Systems* (FAIS). Esses sistemas, em sua maioria, são representados por kits de desenvolvimento de software comerciais, como o Verifinger (Neurotechnology, 2023). Para avaliar a qualidade das impressões digitais coletadas, três métricas de qualidade de imagem são amplamente empregadas. O NIST *Fingerprint Image Quality* (NFIQ) é capaz de categorizar as impressões digitais em cinco níveis, sendo o nível um para representar figuras de melhor qualidade e o nível cinco, para as de pior qualidade. Sua versão aprimorada, o NFIQ 2, oferece vantagens em termos de precisão e adota uma escala de pontuação que varia de 0 a 100, em que pontuações mais elevadas indicam melhor qualidade (Tabassi *et al.*, 2021). Além disso, o VERIQ é uma ferramenta proprietária projetada para funcionar em conjunto com o Verifinger (Neurotechnology, 2023).

Os sistemas apresentados nessa seção determinam a pontuação biométrica comparando o modelo armazenado com a impressão digital do usuário (Li, 2009). Caso a pontuação biométrica ultrapasse um limite predefinido, o usuário é autorizado. No entanto, um sistema biométrico perfeito, que sempre aceitaria usuários legítimos e rejeitaria impostores, é improvável (Southier *et al.*, 2023). Como ilustrado na Figura 4, a aceitação ou rejeição de usuários legítimos e impostores depende do limite estabelecido e do escore biométrico.



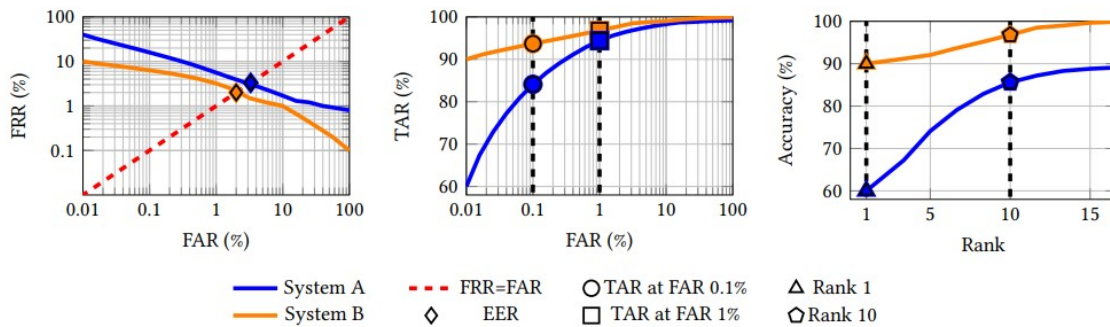
Fonte: Southier *et al.* (2023).

A porcentagem de usuários legítimos que são erroneamente rejeitados é quantificada pela Taxa de Falsas Rejeições, do inglês *False Reject Rate* (FRR). Seu complemento ($1 - \text{FRR}$), que representa a porcentagem de usuários legítimos corretamente autorizados, é conhecido como Taxa de Aceitação Verdadeira, do inglês *True Acceptance Rate* (TAR). A porcentagem de impostores rejeitados é denominada Taxa de Rejeição Verdadeira, do inglês *True Reject Rate* (TRR), e a porcentagem de impostores autorizados é a Taxa de Falso Aceitação, do inglês *False Accepted Rate* (FAR), sendo $\text{TRR} = 1 - \text{FAR}$. Essas quatro medidas estão interligadas e dependem do limite selecionado. A variação desse limite é comumente representada por duas curvas para ilustrar o desempenho biométrico do sistema: a curva de Características Operacionais do Receptor, do inglês *Receiver Operating Characteristics* (ROC) e a curva de Características de Detecção de Erros, do inglês *Detection Error Trade-off* (DET) (Li, 2009; Jain; Kumar, 2012).

Quanto à identificação, cada biometria é comparada com todas as amostras da galeria em tempo real. As pontuações resultantes são classificadas e ordenadas. A Taxa de Identifi-

cação Positiva Verdadeira, do inglês *True Positive Identification Rate* (TPIR) é calculada para determinar a classificação em que ocorre uma correspondência verdadeira e para mensurar a probabilidade de reconhecer corretamente a identidade nas K primeiras classificações. A Característica de Correspondência Cumulativa, do inglês *Cumulative Match Characteristic* (CMC) é utilizada para apresentar os resultados de identificação. A Figura 5 apresenta exemplos das curvas DET, ROC e CMC para dois sistemas biométricos. A curva DET representa as taxas de erro (FAR e FRR) nos dois eixos usando uma escala logarítmica. A Taxa de Erro Equivalente, do inglês *Equal Error Rate* (EER) de um sistema biométrico é o ponto em que FAR=FRR. A curva ROC mostra a Taxa de Aceitação Verdadeira (TAR) em vários valores de FARs. Normalmente, os valores de TAR em FAR = 0,1% e 1% são usados para a comparação entre diferentes sistemas. A curva CMC é traçada para mostrar o TPIR em relação às classificações.

Figura 5 – Exemplos de curvas DET (esquerda), ROC (centro) e CMC (direita)



Fonte: Southier et al. (2023).

2.2 Registro de Imagem

O registro de imagem é um processo fundamental na visão computacional que consiste sobrepor diversas imagens de uma cena em um sistema de coordenadas unificado, podendo ser originadas a partir de momentos distintos (registro multitemporal), por sensores variados (registro multimodal) ou de perspectivas diferentes (Emna Kamoun, 2022). Esse processo é dividido em três etapas: detecção de pontos-chave e descrição de características (Seção 2.2.1); correspondência dessas características (Seção 2.2.2); e distorção de imagem (Seção 2.2.3).

2.2.1 Detecção de pontos-chave e descrição de características

Um ponto-chave, também conhecido como ponto de interesse, é um elemento que identifica o que é significativo e único em uma imagem, como cantos, bordas, entre outros aspectos. Cada ponto-chave é associado a um descritor, que é um vetor de características que contém informações cruciais. A robustez do descritor é fundamental, pois ele deve resistir a diferentes

Figura 6 – Detecção dos pontos-chave



Fonte: Emna Kamoun (2022).

transformações na imagem, como mudanças na localização, escala, brilho e outras variações (Emna Kamoun, 2022). Na Figura 6, é possível visualizar os pontos de interesse identificados na moldura da Mona Lisa e algumas outras na própria pintura.

A descrição de características é o processo de calcular a abstração das informações e criar um local de decisão em cada ponto da imagem. Técnicas consideradas eficientes para a detecção dessas características precisam ser robustas na análise de escala, ruído, rotação, iluminação e transformação (Shan, 2010). Além disso, características ideais devem ser altamente distintivas, de modo que uma única característica seja combinada corretamente com alta precisão (Shan, 2010). Diversos algoritmos são empregados para realizar a detecção de pontos-chave e a descrição de suas características, os quais são apresentados a seguir.

2.2.1.1 SIFT

Scale-invariant feature transform (SIFT) é um algoritmo de detecção de pontos-chave, utilizado para identificar objetos parecidos em múltiplas imagens. O SIFT tem quatro fases computacionais que incluem: construção do espaço de escala, detecção de extremos em escala espacial, localização de pontos-chave, atribuição de orientação e definição de descritores de pontos-chave (Vaghela; Naina *et al.*, 2014). A etapa inicial do processo, consiste em identificar potenciais pontos de interesse determinando a localização e a escala. Na etapa seguinte, a seleção é baseada em sua estabilidade para garantir a resistência às distorções. Durante a fase de orientação, o algoritmo SIFT calcula a direção dos gradientes nas proximidades dos pontos-

chave. Por fim, são atribuídas uma ou mais orientações para cada um deles. Embora seja lento comparado a outros algoritmos que realizam a mesma tarefa, o SIFT é considerado adequado para detecção de objetos em imagens com alta resolução (Vaghela; Naina *et al.*, 2014).

2.2.1.2 SURF

SURF é um descritor de características que emprega duas fases distintas: a extração e a construção do descritor. A fase inicial envolve estabelecer uma orientação reproduzível com base em dados provenientes de uma área circular ao redor do ponto de interesse. Posteriormente, cria-se uma região quadrada alinhada à orientação escolhida e dela é extraída o descritor (Bay; Tuytelaars; Gool, 2006). O SURF otimiza o processo de detecção em relação ao método SIFT, baseado em propriedades semelhantes com uma complexidade reduzida, aprimorando a qualidade dos pontos detectados (Bay; Tuytelaars; Gool, 2006). A matriz Hessiana fornece informações sobre como as derivadas parciais de uma função variam à medida que se desloca pelo espaço de entrada e é usada junto com descritores de baixa dimensionalidade para aumentar significativamente a velocidade de correspondência (Vaghela; Naina *et al.*, 2014). Primeiramente, a filtragem gaussiana é aplicada na imagem. Em seguida, a matriz Hessiana é obtida pelo método da diferença gaussiana. Os pontos-chave são obtidos a partir das comparações entre as características da matriz (Li *et al.*, 2017). Em relação ao desempenho, para dispositivos de baixo consumo ou sistemas em tempo real (*e.g.* odometria visual), o algoritmo SURF possui um menor custo computacional (Bay; Tuytelaars; Gool, 2006).

2.2.1.3 ORB

Oriented Fast and Rotate Brief (ORB) é um descritor binário de alta velocidade que resulta da fusão do detector de pontos-chave *Features from Accelerated Segment Test* (FAST) com o descritor *Binary Robust Independent Elementary Features* (BRIEF). O FAST é eficiente e encontra pontos-chave de canto razoáveis, ademais muitos detectores de pontos-chave incluem um operador de orientação (SIFT e SURF são dois exemplos proeminentes), mas o FAST não (Rublee *et al.*, 2011). Já o BRIEF possui desempenho semelhante ao SIFT em muitos aspectos, incluindo resistência à variação de luz, desfoque e distorção de perspectiva. Esse método é notável por sua invariância a rotação e robustez contra o ruído e representa uma alternativa aos dois algoritmos apresentados anteriormente (Rublee *et al.*, 2011).

2.2.1.4 KAZE

KAZE, palavra em japonês que significa vento, é um algoritmo de detecção e descrição de características 2D em múltiplas escalas, em espaços de escala não linear (Alcantarilla; Bartoli; Davison, 2012). Diferentemente do métodos anteriores que envolviam diferentes níveis

de escala, dependendo do espaço gaussiano, o KAZE é baseado em espaços de escala não linear. No entanto, o desfoque gaussiano não leva em consideração as fronteiras naturais dos objetos e tende a suavizar tanto os detalhes quanto o ruído de forma equivalente, resultando na diminuição da precisão e na dificuldade de distinguir as localizações com clareza. Assim, o KAZE realiza uma filtragem de difusão linear, o que o capacita a ajustar o desfoque de forma adaptável aos dados da imagem, reduzindo o ruído sem comprometer os contornos do objeto. Isso resulta em uma precisão e clareza de localização superiores. (Alcantarilla; Bartoli; Davison, 2012). Comparado com os algoritmos anteriores, houve aumento moderado do custo computacional. A versão acelerada do algoritmo, a *Accelerated-KAZE* (AKAZE) (Alcantarilla; Nuevo; Bartoli, 2013), é baseado no determinante da Matriz Hessiana e seus espaços de escala não linear são construídos usando uma estrutura de filtragem chamada *Fast Explicit Diffusion* (FED) (Tareen; Saleem, 2018). Essa estrutura foi projetada para melhorar o desempenho computacional, especialmente em termos de velocidade de processamento e tende a ser mais sensível a pontos-chave de baixo contraste em comparação com o KAZE.

2.2.2 Correspondência de características

Uma vez identificados os pares de pontos-chave, formado em ambas as imagens, é necessário vincular esses pontos, ou seja, determinar quais deles correspondem ao mesmo local. Um método viável para essa tarefa é utilizar um Combinador de Força Bruta, do inglês *Brutal Force Matcher* (BFMatcher), algoritmo responsável por avaliar a menor distância entre os descritores de cada par de pontos-chave e, em seguida, atribuir a cada um desses pontos às suas k melhores correspondências. BFMatcher explora todas as opções e as melhores correspondências são então encontradas (Appana *et al.*, 2021). Na Figura 7, é possível observar a correspondência de características entre imagens de tamanhos diferentes, evidenciada pela conexão de traços.

Figura 7 – Correspondência dos pontos-chave



Fonte: Emna Kamoun (2022).

2.2.3 Distorção de imagem

A distorção de imagem é o ato de deformar uma imagem de origem em uma imagem de destino de acordo com um mapeamento entre os espaços de origem e de destino (Heckbert, 1989), após a concordância de um conjunto de pontos. Essa etapa é usada no processamento de imagens principalmente para a correção de distorções geométricas introduzidas por sistemas de imagem imperfeitos (Gonzales; Wintz, 1987). Tratando-se de duas imagens em uma superfície plana no espaço, essa distorção é representada por uma homografia, uma transformação geométrica que possui 8 parâmetros livres representada por uma matriz 3×3 .

A Figura 8 ilustra o efeito da distorção da imagem. Pode-se notar que, ao combinar as imagens, algumas deformações surgiram ao redor dela, devido às diferenças nos ângulos e tamanhos das imagens individuais.

Figura 8 – Imagem após a distorção

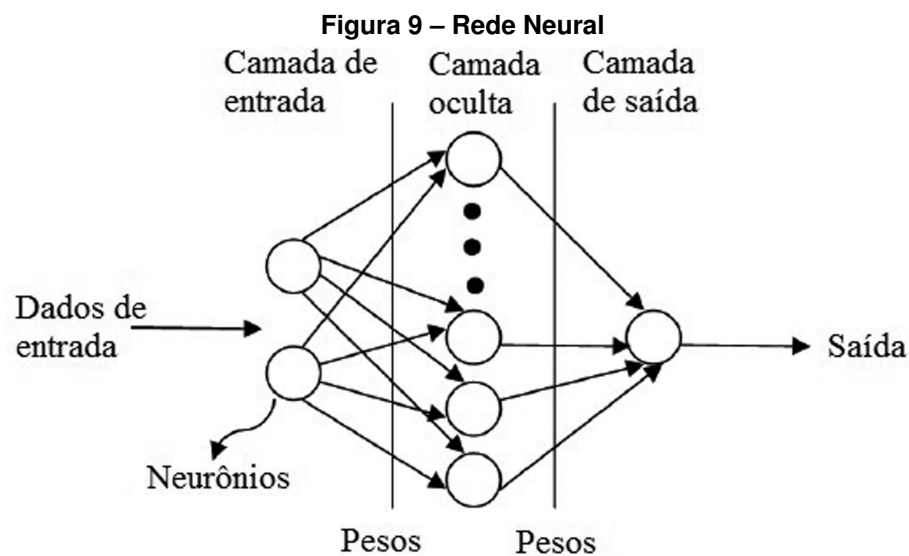


Fonte: Emna Kamoun (2022).

2.3 Redes neurais

Redes Neurais Artificiais são abordagens computacionais que se baseiam em um modelo matemático inspirado na estrutura neural humana e têm a capacidade de adquirir conhecimento por meio de experiência. Redes neurais profundas podem conter várias (de centenas a milhares) unidades de processamento interconectadas, formando camadas. A camada de entrada é onde dados (*e.g.* conjunto de treinamento) são apresentados a rede. As camadas intermediárias, também conhecidas como camadas ocultas, é onde realiza-se o processamento das informações. A camada de saída fornece o resultado do processamento da entrada (Wu;

Feng, 2018). As Redes Neurais são compostas por nós (ou neurônios artificiais) conectados entre si (Jenkins; Tanguay, 1995) e organizados em camadas. Cada neurônio artificial tem uma função de ativação, que determina se e em que medida ele deve ser ativado com base na entrada recebida, permitindo que a rede neural aprenda e modele relações complexas nos dados. A conexão entre cada par de neurônios é representado por um valor de peso determinado durante o treinamento da rede, ele é ajustado iterativamente para minimizar a diferença entre a saída predita pela rede e a saída desejada, esse valor de peso é equivalente a memória de rede neural artificial (Bulsari, 1993), como mostrada na Figura 9.



Fonte: Barros et al. (2020).

A característica fundamental das redes neurais é sua capacidade de aprendizado a partir de técnicas de reforço, supervisionado ou não supervisionado, resultando em melhorias no seu próprio desempenho (DE LEON F. DE CARVALHO, 2009). Esse aprimoramento é alcançado por meio de um processo iterativo de ajustes nos pesos da rede, chamado treinamento. O aprendizado se concretiza quando a rede neural alcança uma solução generalizada para o problema. Outro aspecto significativo é a forma como uma rede neural interage com o ambiente, recebendo ou não uma resposta desejada para o padrão de entrada do agente externo. (DE LEON F. DE CARVALHO, 2009).

2.3.1 *Deep learning*

O aprendizado profundo, ou *deep learning*, é uma subcategoria do campo de aprendizado de máquina e consiste, fundamentalmente, na construção de redes neurais compostas por várias camadas ocultas, como apresentada na Figura 9. O aspecto principal do aprendizado profundo é que essas camadas não são projetadas por engenheiros humanos, elas absorvem informações de várias fontes de dados durante o treinamento e são apresentados à rede neural em lotes. Cada camada da rede processa esses dados, aplicando transformações matemáticas

que são ajustadas iterativamente para minimizar a diferença entre as previsões da rede e as saídas desejadas. Além disso uma vez que a rede tenha sido treinada ela pode ser usada para analisar esses dados em tempo real, sem a necessidade de intervenção humana (LeCun; Bengio; Hinton, 2015), isso é possível porque as operações realizadas em cada camada da rede são geralmente simples o suficiente para serem calculadas rapidamente. Métodos de aprendizado profundo permitem que modelos computacionais compostos de múltiplas camadas de processamento aprendam representações de dados com vários níveis de abstração (LeCun; Bengio; Hinton, 2015).

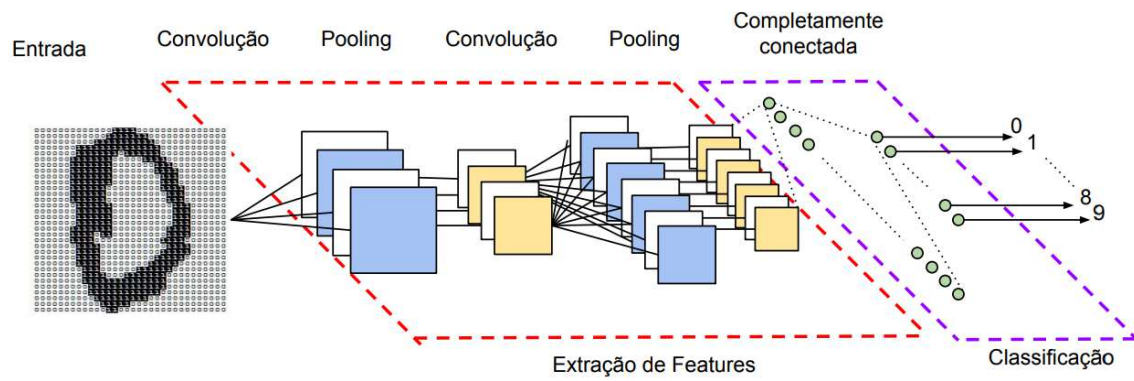
Vários são os motivos para se utilizar aprendizado profundo como sua flexibilidade e capacidade de treinamento em uma ampla variedade de dados e aplicações; maior precisão na análise de dados; melhora no reconhecimento de padrões; e aprendizado automático de recursos.

Na literatura foram propostos diversos métodos de aprendizado profundo, como as redes neurais convolucionais, utilizadas principalmente em aplicativos de visão computacional e processamento de imagem (Albawi; Mohammed; Al-Zawi, 2017a).

2.3.1.1 Redes Neurais Convolucionais

Rede Neural Convolucional (RNC) é uma conhecida arquitetura de aprendizado profundo inspirada no mecanismo natural de percepção visual das criaturas vivas (Gu *et al.*, 2018). O termo "convolução" deriva da operação matemática linear que ocorre entre matrizes, e seu objetivo é apreender características relevantes a partir dos dados de entrada. A camada de convolução é constituída de vários *kernel*s convolucionais também conhecidos como filtros. As RNCs superaram o desempenho dos métodos clássicos em diferentes campos especialmente no reconhecimento de padrões (Albawi; Mohammed; Al-Zawi, 2017b). Na literatura encontra-se diversas variações de arquiteturas de RNCs que foram desenvolvidas ou reformuladas acarretando em avanços significativos para a área de aprendizado profundo.

A Figura 10 ilustra que a RNC é composta por dois estágios distintos: extração de características e classificação. Ao receber uma imagem como entrada, composta por pixels organizados em matrizes, a rede passa por procedimentos de convolução e *pooling* durante o estágio de extração. Após a extração, as características obtidas são então utilizadas para realizar a classificação, por meio do aprendizado dos padrões presentes em cada uma das características.

Figura 10 – Rede Neural Convolucional

Fonte: Vargas, Paes e Vasconcelos (2016).

3 TRABALHOS RELACIONADOS

Na literatura, a fim de entender e explicar o estado da arte, a maioria dos artigos relacionados ao registro de imagem, aplicados a impressões digitais, possuem como público alvo os adultos. Além disso, segundo Kwon, Yun e Lee (2010), a construção de impressões digitais roladas convergem na maioria das vezes em mosaicos, considerando sensores de pequenas áreas. Logo, dado que diversas impressões do mesmo dedo em sensores de área reduzidas frequentemente compartilham uma sobreposição limitada. Essa área específica se revela crucial para aprimorar o desempenho do reconhecimento de impressões digitais (Kwon; Yun; Lee, 2010).

Em um estudo relacionado com registro de imagem, é explorado a utilização de algoritmos genéticos para otimizar o alinhamento de pares de imagens de impressões digitais. O procedimento é feito por meio do alinhamento entre duas regiões de interesse nas digitais utilizando operações de rotação e translação. Após atingir a região de interesse, o algoritmo genético é aplicado. Em testes de desempenho com outros algoritmos de alinhamento de imagem 2D e 3D, ambos essencialmente aleatórios, foi verificado que o algoritmo de força bruta 3D, ao tentar todas as possibilidades, sempre encontra a melhor correlação possível (Ammar; Tao, 2000). Ao comparar o desempenho do algoritmo genético com o algoritmo de busca 2D, eles verificam como a precisão do registro melhora ao utilizar o algoritmo genético. Já na comparação com o algoritmo de força bruta 3D, o objetivo deles é determinar com que frequência o algoritmo genético encontra a solução ideal e avaliar seu desempenho médio. Utilizaram o coeficiente de correlação como medida para a precisão do registro em todos os casos e dessa forma verificaram que o algoritmo genético é uma abordagem eficaz para o registro de imagens de impressões digitais (Ammar; Tao, 2000).

Um outro cenário que aparece na literatura, envolve a utilização de um algoritmo amplamente empregado na realização de alinhamento de modelos geométricos tridimensionais, o *Iterative Closest Point* (ICP). Esse algoritmo, comparado com o mosaico de imagem, pesquisa iterativamente os conjuntos de pontos mais próximos e calcula a transformação correspondente usando matrizes de covariância até que a mudança no erro entre os conjuntos de pontos esteja abaixo de um valor limite T (Moon *et al.*, 2004). Seu objetivo é encontrar a matriz de transformação que define a relação espacial entre as duas impressões digitais (Jain; Ross, 2002). Após os testes com os algoritmos ICP e a técnica de mosaico de imagem, foi percebido que ambos foram capazes de melhorar a precisão da verificação de impressões digitais, mas que o tamanho da imagem se torna um fator crítico na decisão (Moon *et al.*, 2004).

Uma outra abordagem na construção de impressões digitais roladas é realizada por meio da incorporação do método de registro de imagem baseado em um modelo de Campo Aleatório de Markov, do inglês *Markov Random Field* (MRF) (Kwon; Yun; Lee, 2010). A técnica proposta estabelece correspondências densas entre imagens de sequências de impressões digitais roladas, realizando distorções abrangentes na área da impressão digital para criar uma

representação rolada mais autêntica. Esse método pode gerar impressões digitais roladas conceitualmente mais precisas, preservando as propriedades geométricas da superfície do dedo, em oposição as impressões roladas a base de tinta e outros métodos existentes. Para validação da precisão dessa abordagem, foram conduzidos diversos experimentos comparativos, revelando diferenças substanciais entre os métodos de construção rolado. Os resultados evidenciam a eficiência do método proposto em diversos aspectos quando comparado a outras estratégias.

Em outro estudo é proposto uma nova abordagem para a correspondência de impressões digitais chamada Representação Interpretável de Comprimento Fixo para Correspondência de Impressões Digitais via Vision Transformer, do inglês *Interpretable Fixed-Length Representation for Fingerprint Matching via Vision Transformer* (IFViT), que funciona em dois estágios principais. No primeiro estágio, é usado um módulo de registro denso interpretável com uma rede siamesa baseada em *Vision Transformer* (ViT). Isso ajuda a captar dependências de longo alcance e o contexto geral em pares de impressões digitais, proporcionando correspondências detalhadas e interpretáveis nos pontos característicos das impressões digitais (Qiu *et al.*, 2024). Isso melhora o alinhamento e a clareza na correspondência das impressões. No segundo estágio, é considerado as representações locais e globais das impressões digitais alinhadas para extrair e corresponder representações de comprimento fixo de forma interpretável. Utiliza-se os ViTs treinados no primeiro estágio, acrescentando uma camada adicional e treinando novamente para gerar uma representação clara e detalhada de comprimento fixo, bem como correspondências densas nos pontos característicos (Qiu *et al.*, 2024). Experimentos extensivos com diversos bancos de dados de impressões digitais públicos mostram que a abordagem não só oferece um desempenho superior em registro e correspondência, mas também melhora significativamente a interpretabilidade das correspondências de impressões digitais (Qiu *et al.*, 2024).

4 MATERIAIS E MÉTODOS

4.1 Materiais

Os materiais empregados na realização deste trabalho, abrangendo ferramentas e tecnologias utilizadas, estão detalhadamente relacionados no Quadro 1.

A linguagem de programação Python foi escolhida para a execução deste projeto pois, além de se tratar de uma linguagem *Open Source* de alto nível com uma sintaxe mais simples, ela possui inúmeras bibliotecas dedicadas que foram utilizadas para executar diversas tarefas. Dentre essas tarefas pode-se citar o aprendizado de máquina e a identificação de pontos-chave, e entre as bibliotecas o Scikit-learn ¹, o Keras ² e o OpenCV ³.

Além disso, Python é a linguagem adotada pelo Google Colab, ferramenta que viabiliza a escrita compartilhada e execução de códigos em uma máquina virtual através do navegador, comumente utilizada na área de inteligência artificial. Além disso, o Colab possui grande parte das bibliotecas que são rapidamente instaláveis e configuráveis. A ferramenta permitiu executar os códigos das etapas de coleta no caso da leitura dos vídeos e na comunicação do escâner com o computador, além do pré e pós-processamento das imagens.

A seguir, são apresentadas breves descrições das bibliotecas do Python que foram empregadas na construção deste trabalho:

- NumPy(Harris *et al.*, 2020): biblioteca do Python que foi utilizada para manipulação de matrizes, permitiu a manipulação simples de matrizes durante o desenvolvimento dos códigos.
- Matplotlib(Hunter, 2007): biblioteca que permitiu gerar os gráficos para comparação de resultados.
- OpenCV(Bradski; Kaehler, 2008): biblioteca desenvolvida visando eficiência computacional e com especial ênfase em aplicações em tempo real. Permitiu a identificação e correspondência dos pontos-chave.
- Scikit-learn(Kramer, 2016): biblioteca que possibilitou a criação e utilização de diversos modelos de algoritmos de aprendizado de máquina e modelagem estatística, incluindo algoritmos de classificação, regressão e agrupamento.
- TensorFlow(Gulli; Pal, 2017): biblioteca robusta, porém acessível, para construção de modelos de aprendizado de máquina.

¹ <https://scikit-learn.org/stable/>

² <https://www.tensorflow.org/install?hl=pt-br>

³ <https://opencv.org/get-started/>

- MINDTCT(Ko, 2007): detector de minúcias de impressão digital contida em um arquivo ANSI/NIST 2000

Quadro 1 – Ferramentas empregadas no desenvolvimento do trabalho.

Tecnologias/Ferramentas	Finalidade
Google Colab	Máquinas virtuais do Google que possibilitam rápida escrita e execução de código em Python diretamente no navegador.
NumPy	Biblioteca para a linguagem Python, que suporta o processamento de matrizes e vetores.
Matplotlib	Biblioteca para a linguagem Python, que permite a geração de gráficos e histogramas.
OpenCV	Biblioteca para linguagem Python que permite o desenvolvimento e utilização de funções na área de visão computacional.
Python	Linguagem de programação.
Scikit-learn	Biblioteca para linguagem Python com recursos de aprendizado de máquina
TensorFlow	Biblioteca para linguagem Python com recursos de redes neurais
MINDTCT	Detectar minúcias de impressões digitais

Fonte: Autoria própria (2024).

A base de dados utilizada neste trabalho é composta por um vídeo de captura de digitais de forma rolada, disponibilizada pela Natosafe, do grupo Akiyama, empresa dedicada ao desenvolvimento de tecnologias de identificação biométrica para crianças. Já as coletas realizadas nos hospitais, pelo grupo de pesquisa Biometria neonatal da Universidade Tecnológica Federal do Paraná (UTFPR) serão utilizadas futuramente. Esses vídeos serão capturas feitas por escâneres com alta resolução que registram vários momentos (*frames*) das impressões digitais. Eles não estarão disponíveis ao público, somente para os membros do grupo de pesquisa que assinaram a liberação no Comitê de Ética em Pesquisa (CEP) ⁴, concordando que trata-se de uma pesquisa que envolve seres humanos.

4.2 Método

Esse trabalho objetiva aplicar métodos de registro de imagem para melhorar o *match* de impressões digitais de recém-nascidos. Mais especificamente, nessas imagens foi pretendida a construção de digitais roladas a partir de digitais pousadas, coletadas por meio de vários *frames* capturados de um mesmo dedo. Sendo assim, foi tomado como base para a realização

⁴ Estudo aprovado com Certificado de Apresentação de Apreciação Ética 73791023.7.0000.0177 na Plataforma Brasil.

desse projeto, estudos sobre registro de imagem em digitais de adultos (Kwon; Yun; Lee, 2010; Ammar; Tao, 2000; Moon *et al.*, 2004), método esse que combinado com outros algoritmos é também aplicado em estudos de outras áreas, como os de análise de imagens de retina.

A seguir são apresentadas as etapas que foram adotadas, para o desenvolvimento deste trabalho, tais como: coleta dos vídeos, separação dos *frames*, extração dos pontos-chave, correspondência de características e comparação dos *datasets*. A Figura 11 fornece uma visão geral destas etapas.

Figura 11 – Protocolo de pesquisa proposto



Fonte: Autoria própria (2024).

O trabalho foi dividido em quatro experimentos, o registro de imagem aplicado nas imagens sem pré-processamento, o registro aplicado nas imagens com pré-processamento, o registro aplicado nas imagens com pré-processamento e segmentação e, por fim, o ultimo experimento a composição das imagens utilizando um método de soma. Os três primeiros experimentos seguem, igualmente, com as etapas de coleta dos vídeos, separação dos *frames*, extração dos pontos-chave e correspondência das características. Já o experimento da soma mantém apenas os passos da coleta, separação e não utiliza a técnica de registro de imagem.

Estão apresentadas e descritas, nas seções a seguir, as etapas utilizadas no desenvolvimento desse trabalho. Cada seção fornecerá uma visão detalhada dos procedimentos e materiais adotados.

4.2.1 Registro de imagem sem processamento

Nesse primeiro experimento, coletas foram realizadas por escâneres concedidos pela empresa parceira Akiyama que capturaram em alta resolução várias partes do mesmo dedo. O vídeo utilizado nos experimentos tem duração de 1 minuto e 2 segundos e resolução de

4160x3104 pixels, refere-se a um recém-nascido do sexo masculino e foi capturado 24 horas após o nascimento da criança.

Após a coleta, o passo seguinte foi obter o conteúdo do vídeo e separá-lo em 628 *frames*. Essa separação foi realizada utilizando a linguagem Python e algumas funções como *VideoCapture* do OpenCV e *array* do NumPy que permitiram a manipulação e visualização de qualquer *frame* de forma fácil e rápida. Dessa forma, foi possível escolher manualmente alguns deles para a realização do registro de imagem

Nessa etapa iniciou-se o registro, independente da posição, os pontos-chave das imagens foram identificados e extraídos. Essa extração foi realizada por um dos algoritmos clássicos, como o KAZE, disponibilizado pela biblioteca OpenCV. Primeiramente, o algoritmo recebeu as imagens (*frames*) e as converteu para escala de cinza, já que o mesmo opera em imagens monocromáticas para simplificar o processamento e melhorar a precisão da detecção de características. As configurações do algoritmo incluíram o ajuste do *threshold* para balancear a sensibilidade e evitar a inclusão de ruídos, e a configuração da *diffusivity* para controlar a suavidade da imagem, melhorando a precisão em condições de ruído. Após isso, o KAZE utilizou um descritor para detectar e computar os pontos em comum por meio de vetores descritivos que representam as características locais das imagens. Finalmente, o algoritmo desenhou as marcações de pontos em ambas as figuras.

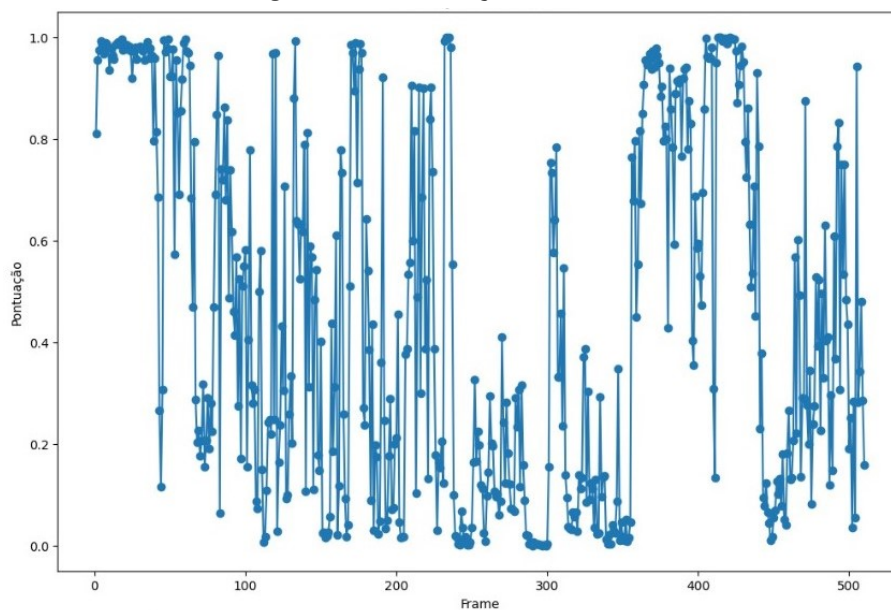
Após a etapa de extração, foi iniciada a etapa de correspondência das características. Nessa etapa, inicialmente, foi utilizado o combinador de força bruta com um classificador *K-Nearest Neighbors* (KNN). O combinador de força bruta foi configurado para comparar diretamente todos os descritores de pontos-chave de uma imagem com todos os descritores de outra imagem, permitindo encontrar as melhores correspondências com base na distância mínima, uma abordagem eficaz devido à sua simplicidade e precisão. O KNN foi utilizado com $k = 2$, significando que para cada descritor de ponto-chave, os dois descritores mais próximos foram identificados. Esta configuração foi escolhida para aplicar a técnica de razão de distância, que ajuda a filtrar correspondências ambíguas e melhorar a robustez da correspondência ao considerar a razão entre as distâncias dos dois vizinhos mais próximos. A razão de distância foi usada para validar as correspondências, onde apenas as correspondências com uma razão de distância inferior a 0,75 foram consideradas válidas, eliminando correspondências falsas positivas e aumentando a precisão geral do processo.

Os pontos correspondentes obtidos serviram de entrada para a matriz de transformação. Esta matriz foi calculada usando o método de *Random Sample Consensus* (RANSAC) para garantir a robustez contra valores significativamente diferentes do resto. A matriz de transformação foi então utilizada para alinhar as imagens, corrigindo distorções e permitindo a composição das imagens.

4.2.2 Registro de imagem com pré-processamento

Nesse segundo experimento, a mesma etapa de coleta foi reaproveitada, utilizando-se o mesmo vídeo de 1 minuto e 2 segundos. No entanto, algumas mudanças foram implementadas na etapa de separação dos *frames*. Utilizando novamente a linguagem Python e funções como *VideoCapture* do OpenCV e *array* do NumPy, procedeu-se à separação. Em seguida, foi feita a seleção dos *frames* de melhor qualidade, por meio da rede neural fornecida (InfantID, 2024), que atribui valores entre 0 (qualidade inferior) e 1 (qualidade superior) às imagens.

Figura 12 – Pontuação dos Frames



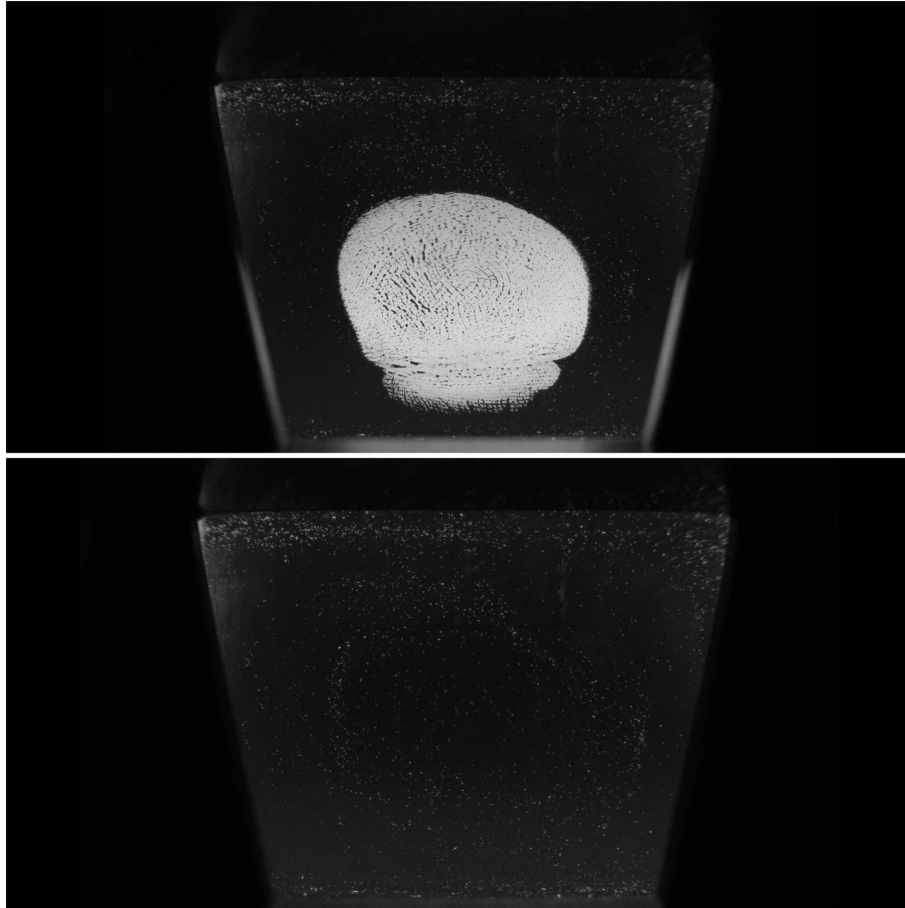
Fonte: Autoria própria (2024).

A Figura 12 apresenta a pontuação de cada *frame* no vídeo, além disso é importante destacar que antes da verificação de qualidade, entre a separação e a obtenção do *score*, o *frame* passa por um procedimento de Recorte da Região de Interesse, do inglês *Crop the Region of Interest* (CropRoi). Em seguida, o *frame* passa por um remapeamento de coordenadas que é realizado utilizando a função *cv2.remap* da biblioteca OpenCV, ela aplica uma transformação geométrica aos pixels da imagem. A função usa dois mapas (*map1* e *map2*) para definir como os pixels do *frame* original devem ser reposicionados na imagem remapeada. Esses mapas especificam as novas coordenadas para cada pixel, permitindo ajustes precisos conforme necessário. O remapeamento garante que o *frame* tenha a configuração correta para o processamento subsequente pela rede.

Após a atribuição de pontos, observou-se que era preciso identificar no gráfico quando o dedo estava em contato com o sensor (digital pousada) e quando o dedo era retirado. Em alguns testes utilizando métodos de derivadas e filtros, foram obtidos resultados considerados insatisfatórios. Por fim, a ideia foi utilizar um método simples de análise, que consiste em verificar a presença de um agrupamento considerável de pixels brancos no *frame*. A presença desse

agrupamento indica que o dedo está em contato com o sensor do escâner; caso contrário, não há contato do dedo, conforme ilustrado na Figura 13.

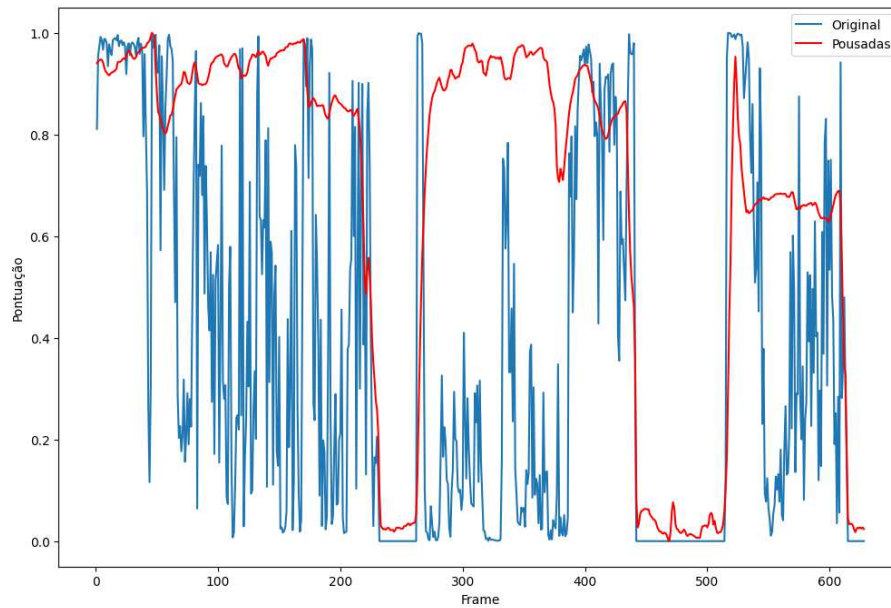
Figura 13 – Digital em contato e sem contato com o scanner



Fonte: Autoria própria (2024).

Com a análise na cor de pixels e aplicação de um filtro passa-baixa foi possível suavizar os dados, reduzir a influência de variações rápidas e manter as mudanças mais lentas. Isso foi útil para identificar as pousadas do dedo em todo o vídeo que antes possuía uma sinal ruidoso e altamente variável. Conforme apresentado na Figura 14, quando a digital está visível, ou seja, em contato com o sensor, a linha vermelha permanece em um nível de qualidade entre 0.6 e 1. Esse comportamento coincide com os *frames* de maior pontuação mostrados pela linha azul, indicando que esses são os momentos de melhor qualidade. Já no momento em que essa linha vermelha desce para a pontuação 0, significa que o dedo é retirado e somente volta a subir quando o dedo é novamente posicionado. É possível notar com clareza as três regiões de pousadas no gráfico, por exemplo do *frame* 0 ao 220, 280 ao 440 e 520 ao 600.

Por fim, o registro de imagem foi realizado com as mesmas configurações do primeiro experimento, seguindo as mesmas exigências nas etapas de extração de pontos-chave e correspondência de características.

Figura 14 – Pousadas Identificadas

Fonte: Autoria própria (2024).

4.2.3 Registro de imagem com pré-processamento e segmentação

Nesse terceiro experimento, todas as etapas anteriores, desde a coleta até o pré-processamento, foram realizadas novamente. Além disso, foi introduzida a técnica de segmentação que é um processo fundamental na análise e processamento de imagens de impressões digitais. Esse processo envolve a separação ou extração de regiões de interesse da imagem, geralmente a área contendo os padrões de cristas e vales que são utilizados para a identificação e verificação de indivíduos. Além disso, serve para a melhoria da qualidade da imagem, pois técnicas de filtragem e realce podem ser aplicadas para melhorar a nitidez e o contraste das cristas e vales, facilitando a detecção e extração de minúcias.

Para os testes, foi utilizado o segmentador SEGV4, disponibilizado por (InfantID, 2024), com o objetivo de melhorar a qualidade das digitais e o registro de imagem. Foram necessárias algumas configurações no algoritmo para realizar a segmentação, como a organização das imagens em uma pasta de origem, a escolha de uma pasta destino e a seleção da resolução adequada, que foi de 3000 ppi.

4.2.4 Soma de imagens

Esta etapa adicional de composição, utilizando máscaras e roi, foi realizada especificamente no experimento quatro. Inicialmente, utilizando as bibliotecas OpenCV, NumPy e Matplotlib, as imagens são carregadas e invertidas para escala de cinza para facilitar o processamento. Em seguida, são aplicadas operações de dilatação e erosão para a criação de máscaras que destacam certas características das imagens. Essas máscaras são então combinadas de

acordo com uma condição que garante que apenas os pixels que estão presentes em uma ou duas máscaras, mas não em todas as três, sejam incluídos na imagem resultante. Isso é feito para destacar regiões que são distintas em pelo menos uma ou duas das imagens originais. Por fim, a resultante é gerada somando as regiões das figuras originais com base nas máscaras. Pixels que não atendem à condição especificada são zerados. Além disso, pixels que estão presentes em todas as três imagens originais são definidos como brancos no resultado final.

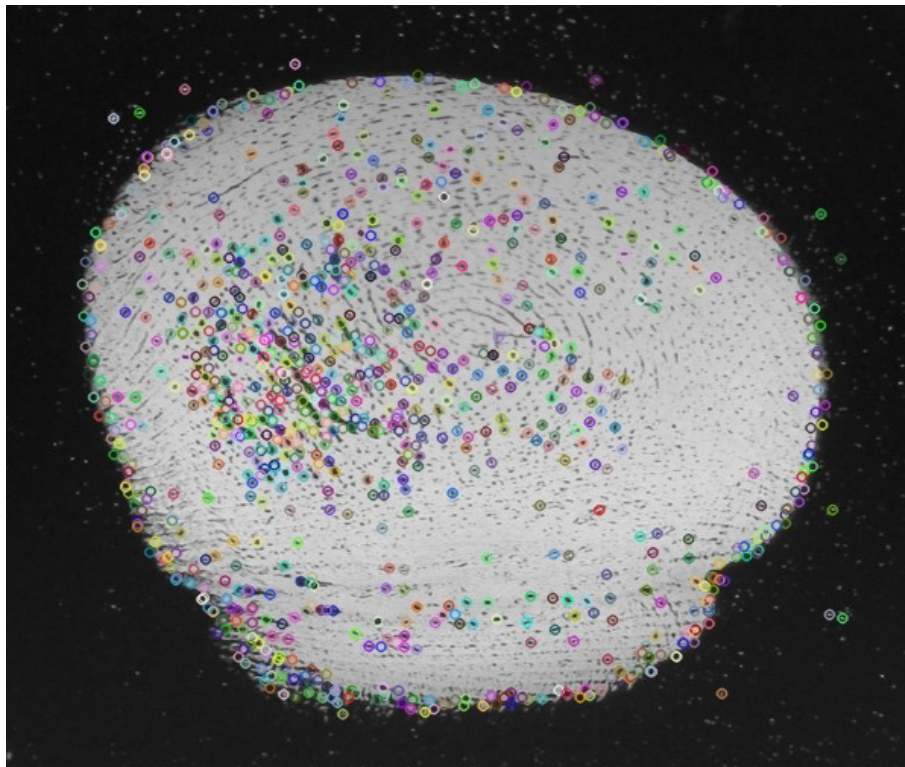
Por fim, a comparação das imagens foi realizada com base na quantidade de minúcias presentes na impressão digital de um único *frame* e no resultado obtido pela soma dos *frames*. Foi utilizada uma ferramenta de identificação de minúcias chamada MINDTCT, que recebe uma imagem como entrada e salva um arquivo contendo informações de posição X e Y, ângulo e qualidade das minúcias. A posição X e Y pode variar conforme a largura e altura da imagem, enquanto o ângulo pode variar entre 0 e 180 graus e a qualidade entre 1 e 100. O principal interesse do experimento é determinar a quantidade de minúcias, representada por cada linha do arquivo. O detector processa a imagem de entrada, converte-a para escala de cinza e então calcula as quatro variáveis mencionadas.

5 RESULTADOS E DISCUSSÕES

5.1 Registro de imagem sem processamento

Experimentos com a detecção de pontos-chave foram realizados, utilizando o algoritmo AKAZE, uma versão otimizada e acelerada do KAZE, apresentada na Seção 2.2.1.4. Para isso, foi utilizado um vídeo disponibilizado pela (InfantID, 2024) e a biblioteca OpenCV do Python. Na Figura 15, é exibido um *frame* do vídeo, mostrando o momento em que o dedo entra em contato com o sensor do *scanner*, capturando assim uma das regiões da digital. Essa imagem, sem nenhum pré-processamento, foi selecionada manualmente e inserida no algoritmo, que identifica e destaca vários pontos-chave. Esses pontos podem representar potenciais minúcias, que são características cruciais para a identificação. Contudo, observou-se um desafio adicional, pois alguns pontos foram marcados fora da região de interesse, além disso, a qualidade da própria impressão digital dificultou a diferenciação entre cristas e vales.

Figura 15 – Detecção dos pontos-chave utilizando o algoritmo KAZE.

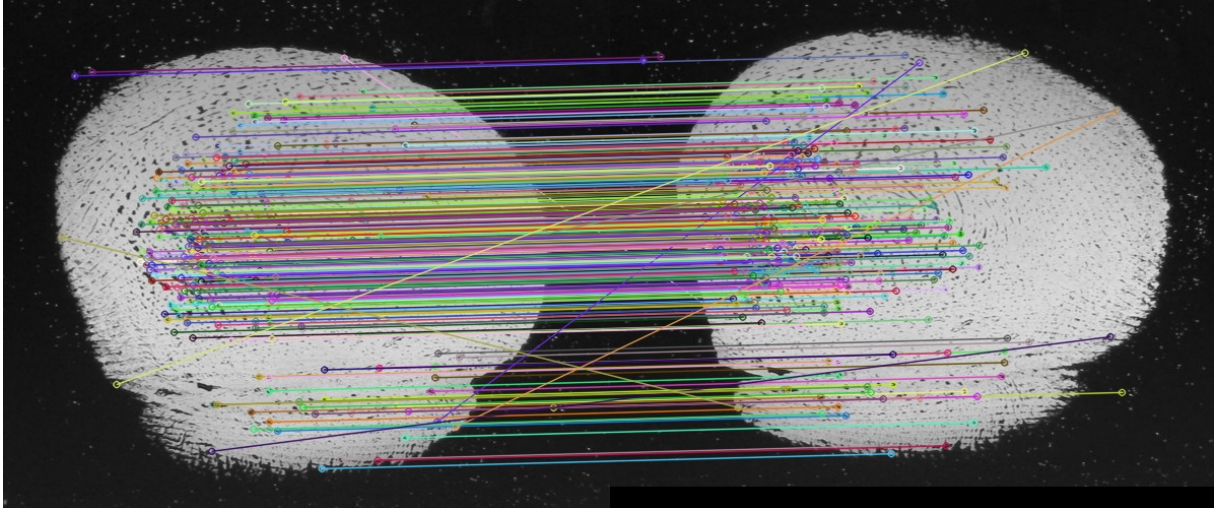


Fonte: Autoria própria (2024).

Após a detecção dos pontos-chave na digital, mais dois *frames* foram escolhidos aleatoriamente, cada um com uma posição e angulação diferente, como mostra a Figura 16. Esses novos *frames* também tiveram seus pontos-chave identificados e destacados. Então as imagens, juntas, foram submetidas ao BFMatcher, que avaliou a menor distância entre os descritores de cada par de pontos-chave e realizou as correspondências ligando os pontos de acordo com as retas traçadas. Observou-se uma grande quantidade de pontos conectados, inclusive fora das

áreas de interesse das digitais, os quais não contribuem para o processo. Além disso, é evidente a presença de linhas conectivas que cruzam, ligando pontos localizados em áreas superiores de uma impressão digital a pontos inferiores de outra. Este padrão é inesperado, uma vez que o objetivo era buscar pontos em comum entre as imagens.

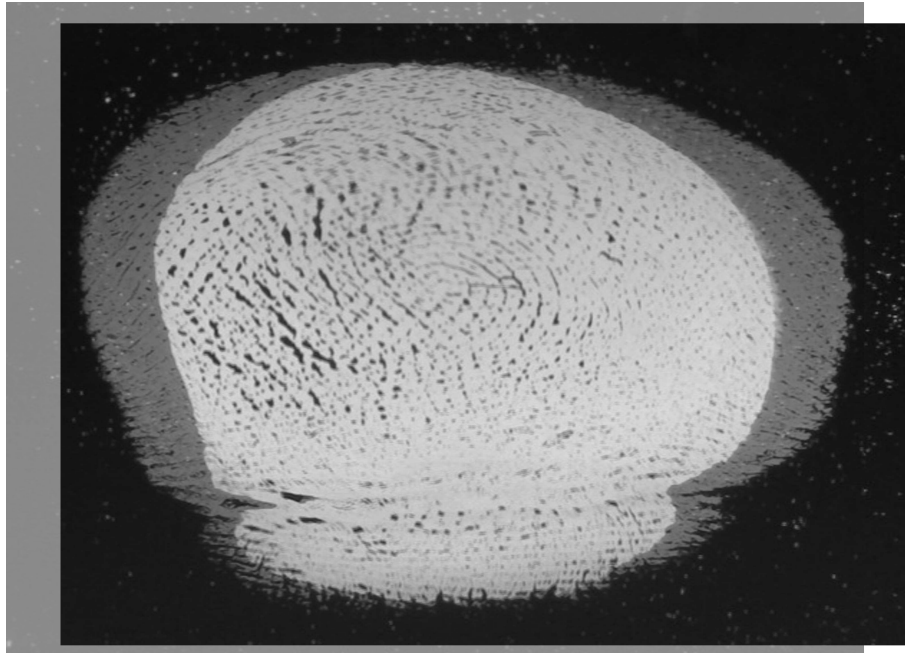
Figura 16 – Correspondência das características utilizando um BFMatcher.



Fonte: Autoria própria (2024).

Com base nas etapas anteriores de detecção dos pontos-chave e de correspondência das características, foi possível construir uma imagem da digital rolada a partir de três imagens com digitais pousadas. Porém, ainda é necessário realizar um pré-processamento na quantidade de informações presentes nas cenas, preservando as partes importantes como possíveis regiões de minúcias e eliminando os elementos dispensáveis, como os ruídos. Na Figura 17 é possível visualizar o resultado da sobreposição das digitais, nota-se que algumas partes da imagem continuam saturadas, o que ainda pode dificultar o reconhecimento de algumas características como vales e cristas. Essa saturação é observada desde a captura da imagem e é atribuída tanto ao tamanho da impressão digital quanto à pressão aplicada durante o contato com o sensor. Partes como o centro da composição estão menos prejudicadas mas mesmo assim ainda o resultado é considerado insatisfatório.

Figura 17 – Digital rolada construída a partir de três digitais pousadas.



Fonte: Autoria própria (2024).

5.2 Registro de imagem com pré-processamento

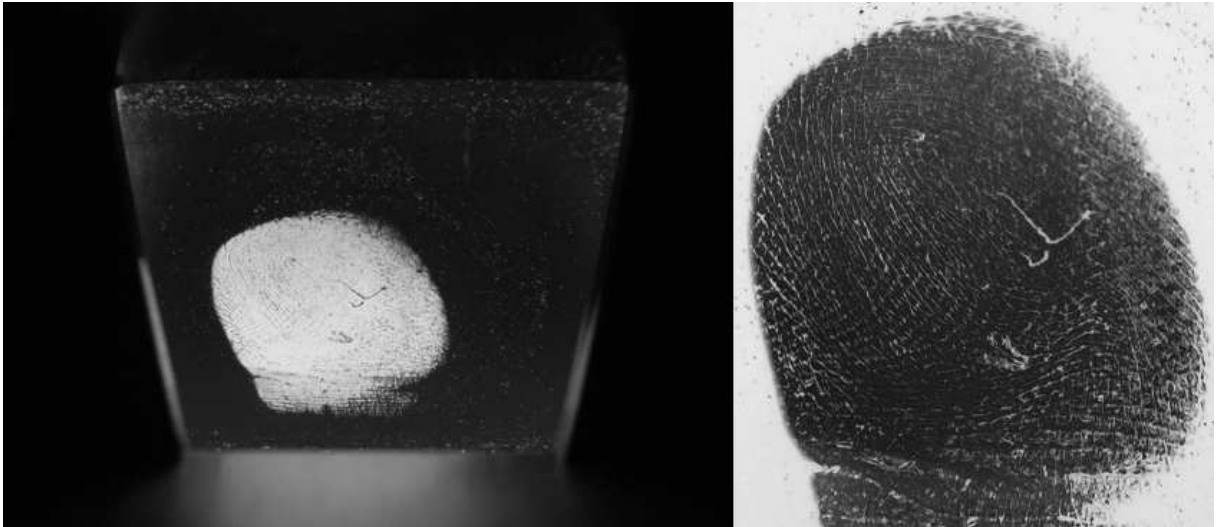
No segundo experimento, antes de executar os procedimentos de registro de imagem, 3 *frames* foram escolhidos aleatoriamente dentro da cada uma das 3 regiões de pousada, conforme ilustrado na Figura 14. Embora a escolha tenha sido aleatória, ela foi baseada nas pontuações atribuídas pela rede de qualidade, descritas na Seção 4.2.2 que incluiu o recorte da área de interesse, remapeamento de coordenadas e inversão das cores. O resultado dessas etapas podem ser visualizadas na Figura 18.

A imagem à esquerda retrata o estado antes do pré-processamento. Nela, observa-se um trapézio correspondente à abertura do sensor, com a impressão digital em branco e uma área preta significativa ao redor. A imagem à direita representa o estado depois do pré-processamento, houve uma inversão de cores, com a impressão digital em preto e o fundo em branco, ocorreu um achatamento vertical (remapeamento) e um recorte da área de interesse, permitindo uma visualização mais próxima.

Com os 3 *frames* selecionados em cada região de pousada e o pré-processamento concluído, como ilustrado nas Figuras 19, 20, 21 foi possível avançar para a técnica de registro.

Por fim, o registro de imagem foi realizado, ainda utilizando o método clássico com o AKAZE detectando os pontos-chave e o BFMatcher fazendo a correspondência das características. Os resultados são apresentados nas Figuras 22, 23, 24, nas quais podem ser observados que o registro de imagem gerou resultados insatisfatórios e mostrou que esse método não funcionou por conta do nível de detalhes, também dos cantos e bordas que as imagens apresentam.

Figura 18 – *Frames* antes e depois do pré-processamento



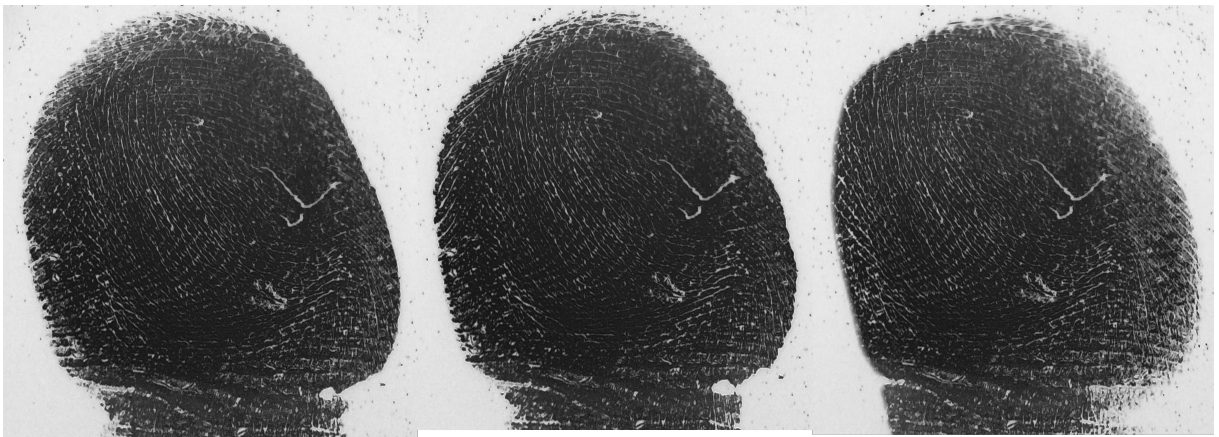
Fonte: Autoria própria (2024).

Figura 19 – *Frames* 1ª pousada pré-processados



Fonte: Autoria própria (2024).

Figura 20 – *Frames* 2ª pousada pré-processados



Fonte: Autoria própria (2024).

Figura 21 – *Frames* 3ª pousada pré-processados



Fonte: Autoria própria (2024).

Figura 22 – Registro dos *frames* da 1ª pousada pré-processados



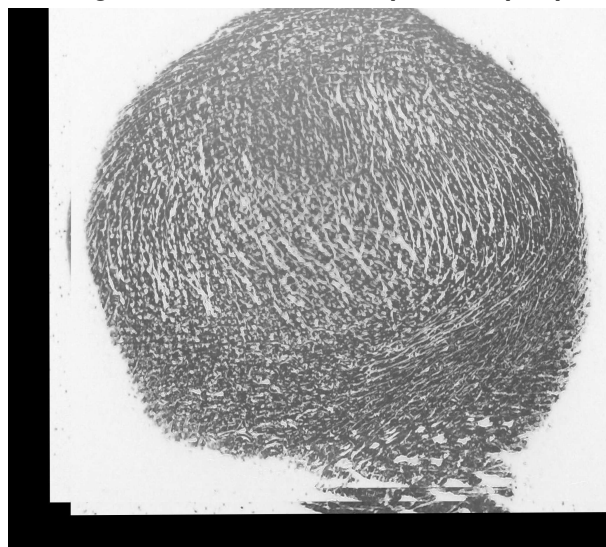
Fonte: Autoria própria (2024).

Figura 23 – Registro dos *frames* da 2ª pousada pré-processados



Fonte: Autoria própria (2024).

Figura 24 – Registro dos *frames* da 3ª pousada pré-processados



Fonte: Autoria própria (2024).

5.3 Registro de imagem com pré-processamento e segmentação

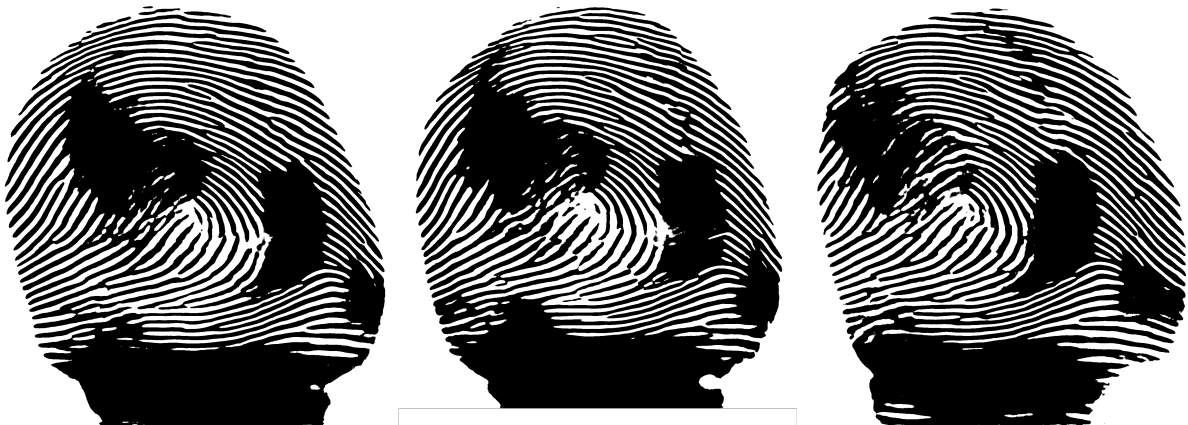
Nesse terceiro experimento, optou-se por realizar a segmentação das imagens resultantes do pré-processamento. Foi utilizado o segmentador SEGV4, que recebe como entrada as imagens e a resolução de 3000 ppi, e fornece como saída as imagens segmentadas, como apresentada nas Figuras 25, 26, 27.

Figura 25 – Frames da 1ª pousada pré-processados e segmentados



Fonte: Autoria própria (2024).

Figura 26 – Frames da 2ª pousada pré-processados e segmentados



Fonte: Autoria própria (2024).

Figura 27 – Frames da 3ª pousada pré-processados e segmentados



Fonte: Autoria própria (2024).

Novamente, conforme observado nas Figuras 28, 29, 30, mesmo deixando as características da digital mais notáveis (cristas e vales) em comparação com o experimento anterior, resultados foram gerados novamente como o registro de imagem e falharam. Apesar da sobreposição dos *frames*, o algoritmo de alguma forma mesclou, mas não de forma satisfatória. É possível ver como as imagens estão em planos diferentes e posições diferentes.

Figura 28 – Registro dos frames da 1ª pousada pré-processados e segmentados



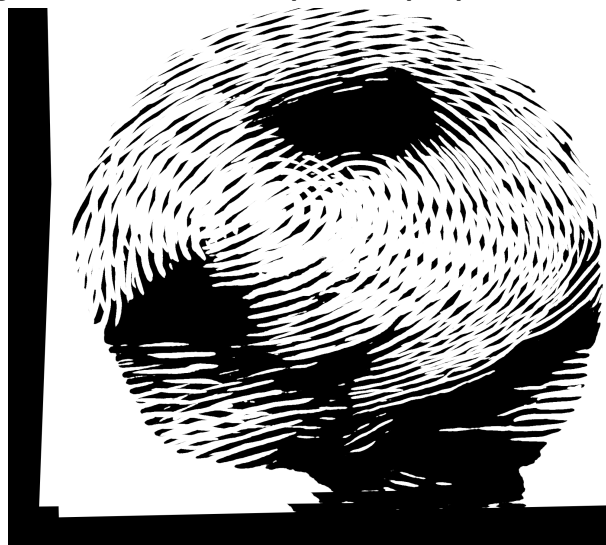
Fonte: Autoria própria (2024).

Figura 29 – Registro dos *frames* da 2ª pousada pré-processados e segmentados



Fonte: Autoria própria (2024).

Figura 30 – Registro dos *frames* da 3ª pousada pré-processados e segmentados



Fonte: Autoria própria (2024).

5.4 Soma de imagens

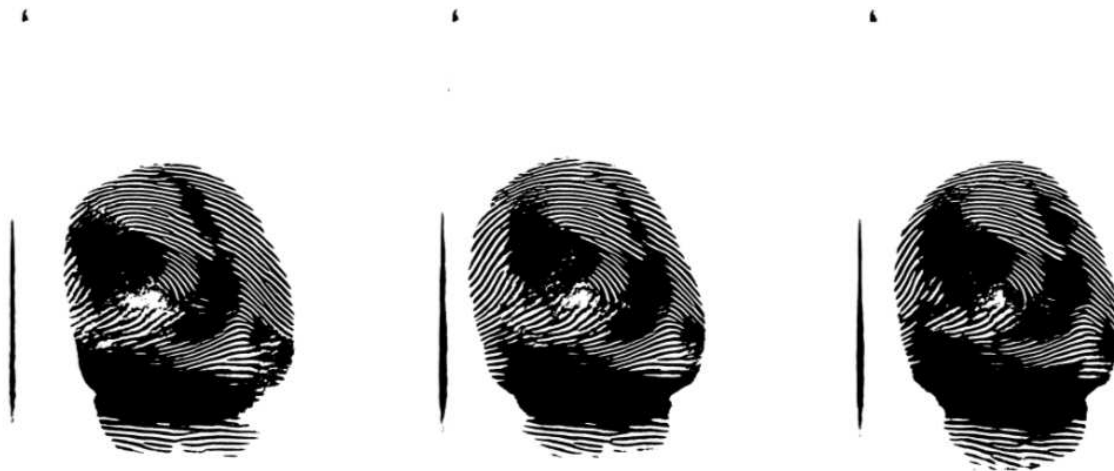
Nesse quarto experimento, foi utilizado um método de soma de imagens em *frames* aleatórios. Porém, dessa vez, durante o pré-processamento foi feito apenas o remapeamento, conforme ilustrado nas Figuras 31, 32, 33, nota-se que as impressões digitais não estão recortadas na região de interesse e que aparecem alguns riscos nas laterais por conta da abertura do sensor.

Figura 31 – *Frames* da 1ª pousada remapeados e segmentados



Fonte: Autoria própria (2024).

Figura 32 – *Frames* da 2ª pousada remapeados e segmentados



Fonte: Autoria própria (2024).

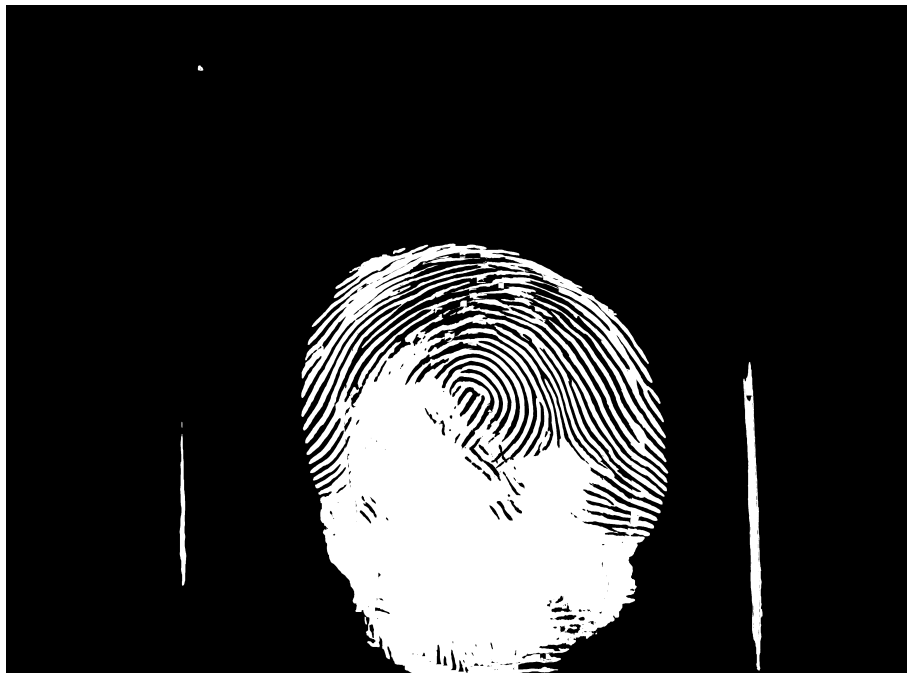
Figura 33 – Frames da 3ª pousada remapeados e segmentados



Fonte: Autoria própria (2024).

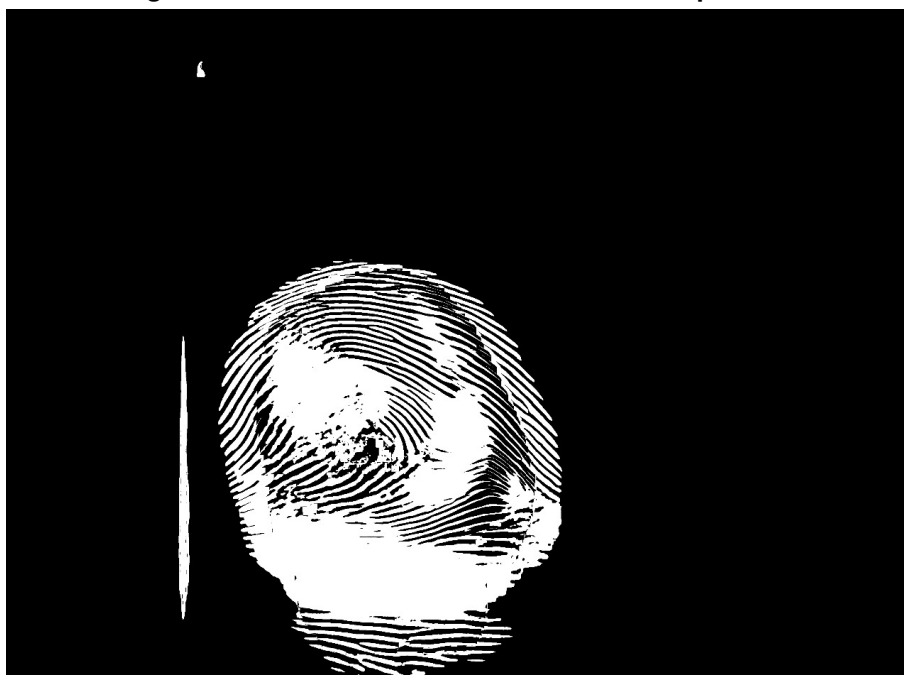
Após o remapeamento das imagens, procedeu-se à segmentação e à soma das mesmas. Contudo, logo após a conclusão deste processo, foram observados pontos brancos na impressão digital que não foram identificados anteriormente pelo segmentador. Essas áreas foram submetidas à soma, utilizando operações com as máscaras criadas a partir das imagens em escala de cinza invertidas, utilizando um kernel retangular de tamanho (25, 25). Essa tentativa de soma das bordas das três imagens resultou no surgimento de algumas ranhuras, além de um aumento na área de contato da impressão digital, como pode ser verificado nas Figuras 34, 35, 36.

Figura 34 – Soma resultante dos frames da 1ª pousada



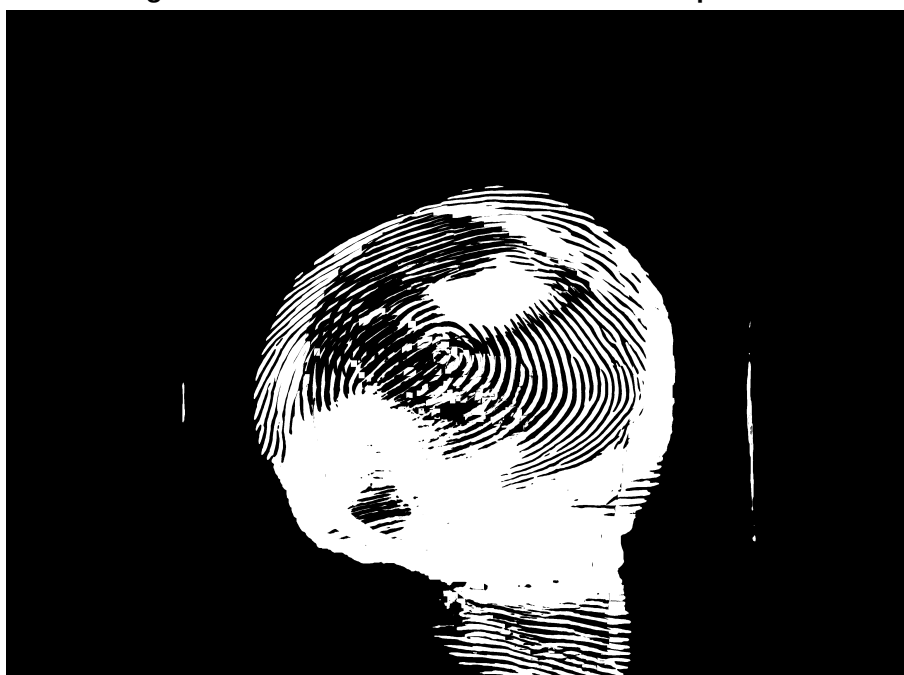
Fonte: Autoria própria (2024).

Figura 35 – Soma resultante dos *frames* da 2ª pousada



Fonte: Autoria própria (2024).

Figura 36 – Soma resultante dos *frames* da 3ª pousada



Fonte: Autoria própria (2024).

Ao analisar esses resultados, tornou-se crucial avaliar se a técnica proporcionou alguma vantagem ou melhoria. Para realizar essa avaliação, utilizou-se a ferramenta MINDTCT para examinar a quantidade de minúcias encontradas tanto nos *frames* individuais quanto nos resultados da soma. As quantidades de minúcias foram as seguintes conforme a Tabela 1.

Tabela 1 – Quantidade de minúcias extraídas

Pousada(s)	1º Frame	2º Frame	3º Frame	Soma das imagens
Primeira	41	42	57	68
Segunda	39	64	38	122
Terceira	55	41	64	184

Fonte: Autoria própria (2024).

Após análise da Tabela, constatou-se que em todas as ocasiões, a soma dos três *frames* revelou um número maior de minúcias. Isso evidenciou que, com a técnica empregada, foi possível obter minúcias que antes não haviam sido detectadas.

5.5 Discussões

Os resultados apresentados no Capítulo 5 mostram, com detalhes, os estudos e experimentos realizados desde o início do projeto até o momento. Apesar de muitos resultados negativos, eles foram de grande importância para as discussões de modelos e métodos que deveriam ser utilizados ou não no trabalho, moldando dessa forma, o assunto registro de imagem em digitais. Avalia-se que o estudo nesse contexto de registro de imagem aplicado em digitais de recém-nascidos que, apesar de ser embrionário, é amplamente explorável.

A escolha de uma técnica para a separação dos *frames* demandou tempo, porém acreditou-se que tenha sido feita uma ótima escolha ao analisar as cores dos pixels somado à utilização da rede de qualidade que confere uma pontuação a cada um deles. As estratégias se somaram e contribuíram para a escolha das melhores imagens. O vídeo utilizado, com duração de 1 minuto e 2 segundos e composto por mais de 600 *frames*, apresentando variações nas posições e colocações do dedo, contribuiu de forma positiva para a realização de experimentos com amostras mais bem distribuídas.

No primeiro experimento em que foram utilizadas os *frames* sem nenhum tipo de processamento, visualmente o registro de imagem não colaborou para um reconhecimento ou correspondência considerável das digitais e as características como cristas, vales e minúcias são dificilmente detectáveis. Porém, esses resultados desempenharam um importante papel para os experimentos seguintes, principalmente no passo de detecção de pontos-chaves que mostrou a necessidade de filtrar aquilo que era (possíveis minúcias) ou não era (pontos fora da digital) necessário, além do tratamento da saturação que dificultou a diferenciação das cristas e dos vales.

No segundo experimento feito com as imagens pré-processadas esperou-se que pelo menos com a utilização de uma rede de qualidade filtrando os melhores *frames*, o resultado fosse melhor. No entanto, no registro de imagem as digitais ficaram sobrepostas em planos diferentes o que acabou gerando um resultado que não era pretendido. Dessa forma, foi no-

tado a necessidade de utilizar alguma outra técnica adicional além do pré-processamento, a segmentação.

No terceiro experimento foi então utilizada a ferramenta de segmentação que ajudou definindo melhor, visualmente, as características da digital. Os vales e cristas ficaram bem definidas e demarcadas nas imagens, assim imaginou-se que os pontos-chaves das figuras seriam melhor aproveitados e que o mesmo registro feito anteriormente aplicado agora, geraria melhorias. O resultado obtido foi o mesmo, a sobreposição das digitais em planos diferentes ao invés da junção das imagens. Embora mais um teste tenha apresentado resultados negativos, pontos positivos também foram observados e absorvidos para o teste seguinte.

Os pontos positivos citados anteriormente contribuíram para o ultimo experimento, foi realizada uma soma de imagens sem pré-processamento mas com segmentação. Os resultados foram promissores: ao somar as três figuras, observou-se um aumento na área de contato da impressão digital e o surgimento de pontos que não estavam presentes anteriormente. Ao analisar a tabela, é evidente que o número de minúcias extraídas aumentou com a soma das imagens. Esse aumento é bom e significativo, pois demonstra que a técnica contribuiu para a identificação de características importantes da impressão digital, fundamentais para a correspondência e identificação. Além disso, o emprego dessa técnica pode ser vantajoso na preparação de imagens destinadas ao treinamento de uma rede neural. Com a extração de mais características, ou seja, mais dados, a qualidade do treinamento tende a ser aprimorada.

6 CONCLUSÃO

A identificação biométrica, especialmente através de impressões digitais, oferece alta precisão e segurança, sendo particularmente relevante na área de saúde para recém-nascidos porque garantem a segurança do indivíduo, acesso as informações médicas e precisão na identificação. As impressões digitais possuem características únicas, como minúcias, vales e cristas, em cada indivíduo. A captura de digitais em recém-nascidos é geralmente realizada de forma pousada devido ao tamanho do dedo e ao comportamento da criança, embora a captura rolada ofereça mais detalhes. Com isso, este trabalho visou desenvolver um algoritmo capaz de identificar as melhores partes de múltiplas imagens de digitais pousadas e, por meio do registro de imagem, construir digitais roladas artificiais de alta qualidade.

O registro de imagem é uma técnica que, por meio da correspondência de características, tem como objetivo compor uma imagem a partir de outras com posições diferentes. Suas ferramentas e implementações forneceram um bom suporte para as tarefas, assim como a linguagem Python e suas bibliotecas.

Um dos maiores desafios foi a escassez de trabalhos específicos sobre digitais utilizando a técnica de registro de imagem. Essa técnica é amplamente utilizada na medicina, mas em abordagens que envolvem imagens de tomografia computadorizada. Até o momento, os resultados positivos são limitados, mas considerados promissores. A maior parte dos objetivos foi cumprida, como a extração e escolha de *frames*, o teste das técnicas clássicas de registro de imagem e a utilização da soma como preparação de imagens para entrada nas redes neurais. No entanto, permanecem em aberto os testes nas redes neurais e os testes de melhoria.

Perspectivas futuras englobam a continuidade e aprimoramento do projeto, incluindo a exploração e teste de novas abordagens. Isso envolve a comparação de datasets utilizando TAR/FAR e CMC, a utilização de redes convolucionais para analisar imagens somadas, e a análise de extração para determinar a validade das minúcias. Espera-se também aprimorar as abordagens e técnicas apresentadas ao longo do desenvolvimento do trabalho, visando alcançar resultados ainda mais robustos e aplicáveis na identificação biométrica de recém-nascidos.

REFERÊNCIAS

- ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. *In: IEEE. 2017 international conference on engineering and technology (ICET)*. [S.l.], 2017. p. 1–6.
- ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. *In: 2017 International Conference on Engineering and Technology (ICET)*. [S.l.: s.n.], 2017. p. 1–6.
- ALCANTARILLA, P.; NUEVO, J.; BARTOLI, A. **Fast explicit diffusion for accelerated features in nonlinear scale spaces british machine vision conference (BMVC)**. [S.l.]: Bristol, 2013.
- ALCANTARILLA, P. F.; BARTOLI, A.; DAVISON, A. J. Kaze features. *In: FITZGIBBON, A. et al. (Ed.). Computer Vision – ECCV 2012*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p. 214–227. ISBN 978-3-642-33783-3.
- AMMAR, H.; TAO, Y. Fingerprint registration using genetic algorithms. *In: Proceedings 3rd IEEE Symposium on Application-Specific Systems and Software Engineering Technology*. [S.l.: s.n.], 2000. p. 148–154.
- APPANA, V. *et al.* Similarity score of two images using different measures. *In: 2021 6th International Conference on Inventive Computation Technologies (ICICT)*. [S.l.: s.n.], 2021. p. 741–746.
- BARROS, G. V. P. d. *et al.* Eficiência de redes neurais artificiais na classificação de uso e do solo da bacia hidrográfica do rio japatubá-se. **Revista Brasileira de Meteorologia**, SciELO Brasil, v. 35, p. 823–833, 2020.
- BAY, H.; TUYTELAARS, T.; GOOL, L. V. Surf: Speeded up robust features. *In: LEONARDIS, A.; BISCHOF, H.; PINZ, A. (Ed.). Computer Vision – ECCV 2006*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. p. 404–417. ISBN 978-3-540-33833-8.
- BAZEN, A. M.; GEREZ, S. H. Segmentation of fingerprint images. *In: CITESEER. ProRISC 2001 Workshop on Circuits, Systems and Signal Processing*. [S.l.], 2001. p. 276–280.
- BRADSKI, G.; KAEHLER, A. **Learning OpenCV: Computer vision with the OpenCV library**. [S.l.]: "O'Reilly Media, Inc.", 2008.
- BULSARI, A. Some analytical solutions to the general approximation problem for feedforward neural networks. **Neural Networks**, v. 6, n. 7, p. 991–996, 1993. ISSN 0893-6080. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0893608009800087>.
- DARGAN, S.; KUMAR, M. A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. **Expert Systems with Applications**, v. 143, p. 113114, 2020. ISSN 0957-4174. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0957417419308310>.
- DE LEON F. DE CARVALHO. **Redes Neurais Artificiais**. 2009. Disponível em: <https://sites.icmc.usp.br/andre/research/neural/>. Acesso em: 23 de novembro 2023.
- Emna Kamoun. **Image Registration: From Scale-invariant feature transform to Deep Learning**. 2022. Disponível em: <https://www.sicara.fr/blog-technique/2019-07-16-image-registration-deep-learning>. Acesso em: 2 de novembro 2023.

GONZALES, R. C.; WINTZ, P. **Digital image processing**. [S.l.]: Addison-Wesley Longman Publishing Co., Inc., 1987.

GU, J. *et al.* Recent advances in convolutional neural networks. **Pattern Recognition**, v. 77, p. 354–377, 2018. ISSN 0031-3203. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0031320317304120>.

GULLI, A.; PAL, S. **Deep learning with Keras**. [S.l.]: Packt Publishing Ltd, 2017.

HARRIS, C. R. *et al.* Array programming with numpy. **Nature**, Nature Publishing Group UK London, v. 585, n. 7825, p. 357–362, 2020.

HECKBERT, P. S. Fundamentals of texture mapping and image warping. Citeseer, 1989.

HUNTER, J. **Matplotlib: A 2D graphics environment Computing in Science & Engineering 9 (3): 90-95**. 2007.

InfantID. **INFANTID**. 2024. Disponível em: <https://natosafe.com.br/>. Acesso em: 20 de maio 2024.

JAIN, A.; ROSS, A. Fingerprint mosaicking. *In*: **2002 IEEE International Conference on Acoustics, Speech, and Signal Processing**. [S.l.: s.n.], 2002. v. 4, p. IV–4064–IV–4067.

JAIN, A. K.; CAO, K.; ARORA, S. S. Recognizing infants and toddlers using fingerprints: Increasing the vaccination coverage. *In*: IEEE. **IEEE international joint conference on biometrics**. [S.l.], 2014. p. 1–8.

JAIN, A. K.; KUMAR, A. **Biometric recognition: an overview**. Berlin, Germany: Springer, 2012. 49–79 p.

JAIN, A. K.; ROSS, A. A.; NANDAKUMAR, K. **Introduction to Biometrics**. [S.l.]: Springer Publishing Company, Incorporated, 2011. ISBN 0387773258.

JENKINS, B.; TANGUAY, A. **Handbook of neural computing and neural networks**. [S.l.]: Boston: MIT Press, 1995.

Julião Medeiros. **A importância dos testes de qualidade na captura biométrica**. 2023. Disponível em: <https://www.vsoft.com.br/post/qualidade-impressao-digital>. Acesso em: 12 de novembro 2023.

KO, K. **User's Guide to NIST Biometric Image Software (NBIS)**. [S.l.]: NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD, 2007.

KRAMER, O. Scikit-learn. *In*: _____. **Machine Learning for Evolution Strategies**. Cham: Springer International Publishing, 2016. p. 45–53. ISBN 978-3-319-33383-0. Disponível em: https://doi.org/10.1007/978-3-319-33383-0_5.

KWON, D.; YUN, I. D.; LEE, S. U. Rolled fingerprint construction using mrf-based nonrigid image registration. **IEEE Transactions on Image Processing**, v. 19, n. 12, p. 3255–3270, 2010.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015.

LI, A. *et al.* An improved fast+ surf fast matching algorithm. **Procedia Computer Science**, Elsevier, v. 107, p. 306–312, 2017.

LI, S. Z. **Encyclopedia of Biometrics: I-Z**. Berlin, Germany: Springer Science & Business Media, 2009. v. 1.

LI, S. Z.; JAIN, A. **Encyclopedia of Biometrics**. [S.l.]: Springer Publishing Company, Incorporated, 2015. ISBN 148997489X.

MATOS, R. M. d. Autenticação de usuários através da utilização de sistemas biométricos. 2000.

MOON, Y. *et al.* Template synthesis and image mosaicking for fingerprint registration: an experimental study. *In: 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*. [S.l.: s.n.], 2004. v. 5, p. V-409.

MUNDIAL, B. Databank-world development indicators. **Recuperado de <https://data.worldbank.org/indicator/SP.POP.TOTL>**, 2023.

NASCIMENTOFILHO, J. *et al.* Detecção e correspondência de minúcias em impressões digitais. 2019.

Natosafe. 2022. Disponível em: <https://natosafe.com.br/tipos-de-impressao-digital/#:~:text=As%20min%C3%BAcias%20s%C3%A3o%20classificadas%20por,%3A%20tipo%2C%20posi%C3%A7%C3%A3o%20e%20orienta%C3%A7%C3%A3o>.

Neurotechnology. 2023. Disponível em: <https://www.neurotechnology.com/verifinger.html>.

QIU, Y. *et al.* **IFViT: Interpretable Fixed-Length Representation for Fingerprint Matching via Vision Transformer**. 2024.

RUBLEE, E. *et al.* Orb: An efficient alternative to sift or surf. *In: 2011 International Conference on Computer Vision*. [S.l.: s.n.], 2011. p. 2564-2571.

SEDLAK, A. *et al.* National estimates of missing children: an overview. national incidence studies of missing, abducted, runaway, and throwaway children. **National Center for Missing Exploited Children, Alexandria, VA. www.missingkids.com/en_US/documents/nismart2_overview.pdf**, 2002.

SHAN, B. A novel image correlation matching approach. **J. Multim.**, Citeseer, v. 5, n. 3, p. 268-275, 2010.

SOUTHIER, L. F. P. *et al.* **Systematic Literature Review on Neonatal Fingerprint Recognition**. 2023. Preprint.

TABASSI, E. *et al.* **NIST Fingerprint Image Quality 2**. NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD, 2021. Disponível em: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=920087.

TAREEN, S. A. K.; SALEEM, Z. A comparative analysis of sift, surf, kaze, akaze, orb, and brisk. *In: 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. [S.l.: s.n.], 2018. p. 1-10.

UNICEF United Nations International Children's Emergency Fund. **Ending preventable child deaths: how britain can lead the way**. 2023. Disponível em: https://www.unicef.org.uk/wp-content/uploads/2020/01/Unicef-UK-Ending-Preventable-Child-Deaths_Report-2020.pdf.

UNICEF United Nations International Children's Emergency Fund. **Under-five mortality**. 2023. Disponível em: <https://data.unicef.org/topic/child-survival/under-five-mortality/>.

United Nations. **World Population Prospects 2022**. 2023. Disponível em: <https://population.un.org/wpp/>.

VAGHELA, D.; NAINA, P. *et al.* A review of image mosaicing techniques. **arXiv preprint arXiv:1405.2539**, 2014.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. *In*: SN. **Proceedings of the xxix conference on graphics, patterns and images**. [S.l.], 2016. v. 1, n. 4.

WHO World Health Organization. **Immunization Agenda 2030: A Global Strategy To Leave No One Behind**. 2023. Disponível em: <https://www.who.int/publications/m/item/immunization-agenda-2030-a-global-strategy-to-leave-no-one-behind>.

WHO World Health Organization. **Immunization coverage**. 2023. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/immunization-coverage>.

WU, Y.-c.; FENG, J.-w. Development and application of artificial neural network. **Wireless Personal Communications**, Springer, v. 102, p. 1645–1656, 2018.