

MATEUS NUNES DA SILVA

UM MODELO BASEADO EM DEEP AUTOENCODER PARA CONVERSÃO DE
IMAGENS VISÍVEIS EM INFRAVERMELHO NO RECONHECIMENTO DE
VEIAS DA PALMA COM REDES SIAMESAS

CURITIBA

2024

MATEUS NUNES DA SILVA

UM MODELO BASEADO EM DEEP AUTOENCODER PARA CONVERSÃO
DE IMAGENS VISÍVEIS EM INFRAVERMELHO NO RECONHECIMENTO DE
VEIAS DA PALMA COM REDES SIAMESAS

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de Mestre em Informática.

Área de concentração: Ciência da Computação

Orientador: Dr. Eduardo Kugler Viegas

Coorientador: Dr. Altair Olivo Santin

CURITIBA

2024

Dados da Catalogação na Publicação
Pontifícia Universidade Católica do Paraná
Sistema Integrado de Bibliotecas – SIBI/PUCPR
Biblioteca Central
Luci Eduarda Wielganczuk – CRB 9/1118

S586m
2024
Silva, Mateus Nunes da
Um modelo baseado em *deep autoencoder* para conversão de imagens visíveis em infravermelho no reconhecimento de veias da palma com redes siamesas / Mateus Nunes da Silva ; orientador: Eduardo Kugler Viegas ; coorientador: Altair Olivo Santin. – 2024.
49 f. : il. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2024

Bibliografia: f. 45-49

1. Informática. 2. Biometria. 3. Reconhecimento de padrões, 4. Redes neurais siamesas. I. Viegas, Eduardo Kugler. II. Santin, Altair Olivo. III. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática. IV. Título.

CDD. 20. ed. – 004

Curitiba, 16 de outubro de 2024.

78-2024

DECLARAÇÃO

Declaro para os devidos fins, que **Mateus Nunes da Silva** defendeu a dissertação de Mestrado intitulada “**UM MODELO BASEADO EM DEEP AUTOENCODER PARA CONVERSÃO DE IMAGENS VISÍVEIS EM INFRAVERMELHO NO RECONHECIMENTO DE VEIAS DA PALMA COM REDES SIAMESAS**”, na área de concentração Ciência da Computação no dia 04 de outubro de 2024, no qual foi aprovado.

Declaro ainda, que foram feitas todas as alterações solicitadas pela Banca Examinadora, cumprindo todas as normas de formatação definidas pelo Programa.

Por ser verdade firmo a presente declaração.

Documento assinado digitalmente
gov.br EMERSON CABRERA PARAISO
Data: 17/10/2024 09:48:12-0300
Verifique em <https://validar.it.gov.br>

Prof. Dr. Emerson Cabrera Paraiso
Coordenador do Programa de Pós-Graduação em Informática

Agradecimentos

Concluir este trabalho representa não apenas a superação de um grande desafio acadêmico, mas também uma jornada de crescimento pessoal, marcada por lições de vida e humildade. Este projeto foi uma tarefa desafiadora, que testou meus limites, mas também me proporcionou momentos de profunda reflexão e aprendizado. Atravessar essa jornada não teria sido possível sem o apoio incondicional e a confiança de tantas pessoas que estiveram ao meu lado, tanto nos momentos de conquista quanto nas horas mais difíceis. Primeiramente, agradeço à minha família, que sempre esteve presente me oferecendo amor, suporte e incentivo. Em especial, quero expressar minha gratidão à minha mãe, Sandra Abreu, cujo carinho, compreensão e força sempre me guiaram. Sua confiança em mim foi uma fonte constante de inspiração. Gostaria de dedicar um agradecimento muito especial ao meu tio e professor, Vilmar Abreu, por seu exemplo de integridade, disciplina e determinação. Sua influência foi fundamental ao longo da minha formação, tanto acadêmica quanto pessoal. Sou imensamente grato ao meu orientador e amigo, Eduardo Kugler Viegas, por sua orientação, paciência e incansável dedicação ao longo deste processo. Sua habilidade em enxergar o potencial em mim e sua orientação técnica foram cruciais para que eu alcançasse este momento. Estendo minha gratidão ao professor Altair Olivo Santin, cujas contribuições e insights enriquecedores ajudaram a moldar o rumo deste trabalho. Aos meus amigos e colegas, agradeço profundamente pelo apoio, paciência e pelas palavras de encorajamento que me sustentaram durante toda essa jornada. Vocês foram fundamentais, oferecendo companhia, conselhos e, muitas vezes, os momentos de descontração necessários para que eu pudesse seguir em frente. Este trabalho não é apenas o resultado de muitos meses de esforço, mas sim o reflexo do apoio e do incentivo de cada uma dessas pessoas. A todas elas, meu mais sincero muito obrigado.

Lista de Algoritmos

Algoritmo 1 – Pseudocódigo Autoencoder	31
Algoritmo 2 – Pseudocódigo Rede Siamesa	34

Lista de Ilustrações

Figura 1 - Fluxo tradicional	6
Figura 2 - Fluxograma geral	27
Figura 3 - Pré-processamento e extração da ROI	30
Figura 4 - Visão geral	37
Figura 5 - Gráfico de perda durante o treinamento da rede autoencoder	39
Figura 6 - Imagens NIR geradas de diferentes indivíduos	40
Figura 7 - Gráfico de perda durante o treinamento da rede siamesa	41
Figura 8 - Tempo de comparação em função do número de imagens processadas	43

Lista de Tabelas

Tabela 1 - Trabalhos disponíveis na literatura _____	25
Tabela 2 - Comparação de precisão do modelo de acordo com o tipo de entrada usada ___	41
Tabela 3 - Matriz de Confusão _____	42
Tabela 4 - Comparação do tempo de execução do modelo no Raspberry Pi e no Desktop GPU _____	43

Sumário

Capítulo 1	1
1.1 Motivação e Hipótese	2
1.2 Objetivos	4
1.3 Contribuições	5
1.4 Organização do Documento	5
Capítulo 2	6
2.1 Autenticação biométrica	6
2.2 Autenticação biométrica baseada em veias	8
2.3 Infravermelho próximo (NIR)	9
2.4 Redes Neurais Artificiais	10
2.4.1. CNN – Redes Neurais Convolucionais	11
2.4.1.1 Camada Conv (Convolucional)	12
2.4.1.2 Camada de Pooling	12
2.4.1.3 Função de Ativação – ReLU	13
2.5 Deep learning	13
2.5.1 Deep Autoencoder	14
2.5.2 Rede siamesas	15
2.6 Classificadores	15
2.6.1 Algoritmo KNN (K-Nearest Neighbors)	16
Capítulo 3	18
3.1 Extração de características através de imagens em espectro visível	18
3.2 Biometria utilizando deep learning	19
3.3 Mudança espectral	21
3.4 Redes siamesas	22
3.5 Considerações Finais	23
Capítulo 4	26
4.1 Visão Geral	26
4.2 Estrutura aquisição de dados	28
4.3 Pré-processamento e extração da ROI	28

4.4 Mudança espectral _____	30
4.5 Rede siamesa e Comparação _____	32
4.6 Discussão _____	35
Capítulo 5 _____	36
5.1 Protótipo _____	36
5.2 Construção do modelo _____	37
5.3 Biometria de veias da palma _____	39
Capítulo 6 _____	44
Referências bibliográficas _____	45

Resumo

O reconhecimento de padrões de veias da palma da mão oferece uma característica única de identificação pessoal. Infelizmente, essas técnicas normalmente exigem um sensor de câmera de Infravermelho Próximo (NIR) para extrair o padrão venoso do indivíduo, o que dificulta sua ampla implantação. Este trabalho propõe um novo esquema viável de verificação de veias da palma usando um *Deep Autoencoder* e uma Rede Siamesa, implementado em três etapas. Primeiro, capturamos a palma da mão do indivíduo usando um sensor de câmera tradicional de espectro visível e realizamos tarefas de pré-processamento para corrigir o posicionamento impreciso, reduzindo a necessidade de acessórios de suporte para a palma da mão. Em segundo lugar, eliminamos a necessidade de um sensor NIR ajustando um modelo de *Autoencoder* para converter imagens do espectro visível em suas contrapartes infravermelhas. Em terceiro lugar, as imagens geradas são processadas por uma rede Siamesa leve usando uma função de perda contrastiva para verificação individual. Experimentos realizados em um conjunto de dados publicamente disponível com mais de cem indivíduos confirmaram a viabilidade de nossa proposta. O esquema atinge até 0,97 de taxa de *true negatives*, com apenas 0,01 de decréscimo em comparação com abordagens tradicionais baseadas em NIR. Além disso, a identificação individual pode ser realizada em menos de 6 segundos em um ambiente com recursos limitados graças à implementação do modelo leve.

Palavras-Chave: Verificação de Veias da Palma, Autoencoder Profundo, Redes Siamesa, Dispositivos com Recursos Limitados.

Abstract

Palm vein pattern recognition offers a unique personal identification feature. Unfortunately, these techniques typically require a Near Infrared (NIR) camera sensor to extract the individual's venous pattern, challenging their wide deployment. This paper proposes a new feasible palm vein verification scheme using a *Deep Autoencoder* and a Siamese Network, implemented threefold. First, we capture the individual's palm using a traditional visible spectrum camera sensor and perform preprocessing tasks to correct imprecise positioning, easing palm support accessories requirements. Second, we eliminate NIR sensor requirement by fine-tuning a *Deep Autoencoder* model to convert images from the visible spectrum to their infrared counterparts. Third, generated images are processed by a lightweight Siamese network using a contrastive loss function for individual verification. Experiments conducted on a publicly available dataset with over a hundred individuals confirmed the feasibility of our proposal. Our scheme reaches up to 0.97 of true-negative rate, with only 0.01 decrease compared to traditional NIR-based approaches. In addition, individual identification can be conducted in less than 6 seconds in a resource-constrained environment thanks to our lightweight model's implementation.

Keywords: Palm Vein Verification, Deep Autoencoder, Siamese Networks, Resource-constrained Devices.

Capítulo 1

Introdução

Nos últimos anos, diversas estratégias biométricas têm sido pesquisadas e implementadas globalmente devido à sua praticidade em atestar a identidade de um indivíduo. A autenticação biométrica, que possibilita a verificação simultânea de diversas características físicas, tem se mostrado um método altamente eficaz para a identificação pessoal, conforme descrito por Rui e Yan em seu estudo sobre autenticação biométrica voltada para uma identificação segura e preservação da privacidade (Rui et al. 2019). Dentre as várias abordagens propostas, os processos biométricos que utilizam os padrões de veias na palma da mão são particularmente promissores. Isso ocorre porque os padrões das veias, protegidos na camada subcutânea e únicos para indivíduos vivos, proporcionam uma camada adicional de segurança, tornando-os particularmente difíceis de falsificar, conforme destacado por (Adiraju et al. 2020).

Um método de autenticação fraco pode resultar em sérias consequências, tanto no ambiente digital, como o vazamento de dados sensíveis, quanto no físico, como a entrada de pessoas não autorizadas em áreas restritas. A tecnologia de reconhecimento de veias da palma da mão se destaca no campo da biometria por várias razões: 1) É extremamente difícil de falsificar. 2) Funciona eficazmente em condições de medição sem contato ou com contato mínimo. 3) Apresenta alta resistência a interferências, conforme descrito por (Yang et al. 2018).

Além de estarem protegidos contra lesões e choques na hipoderme, a combinação das artérias radial e ulnar e seus ramos cria um padrão único e complexo quando o indivíduo possui um sistema circulatório funcional. Essa complexidade torna a falsificação, replicação e roubo dessa característica particularmente desafiadores para os atacantes (Guo et al. 2020). Como as veias da palma não são totalmente visíveis, as abordagens atuais geralmente dependem de um sensor de câmera de Infravermelho Próximo (NIR) para extrair informações

relevantes da palma (Rastogi et al. 2020). Nesse processo, a hemoglobina desoxigenada na camada venosa absorve e reflete a luz, ressaltando assim o padrão venoso do indivíduo.

1.1 Motivação e Hipótese

Os padrões venosos presentes na palma da mão oferecem uma característica biométrica única e diferenciada, similar a impressões digitais, retina e face. Diversos estudos demonstram a eficácia das veias da palma para o reconhecimento individual (Zhou e Kumar 2011; Ajmire et al. 2018), consolidando-as como uma alternativa segura e prática para autenticação biométrica. No entanto, a aplicação prática dessa tecnologia enfrenta desafios significativos, especialmente em relação ao desempenho computacional e à necessidade de sensores especializados que muitas vezes são inacessíveis tanto pela disponibilidade quanto pelo custo.

Atualmente, os sensores NIR (Near-Infrared) são amplamente utilizados para a extração de veias da palma. Esses sensores emitem luz infravermelha, que é absorvida pela hemoglobina desoxigenada nas veias, enquanto os tecidos circundantes a transmitem, criando uma imagem detalhada do padrão venoso (Ajmire et al. 2018; Cho et al. 2020). Entretanto, essa abordagem exige configurações de hardware personalizadas, dificultando sua aplicação em larga escala (Rui et al. 2019). A precisão na captura das veias da palma depende do correto posicionamento da mão sob o sensor NIR, o que frequentemente requer a instalação de acessórios que servem como guia para garantir um alinhamento adequado (Barra et al. 2019). Além disso, as câmeras infravermelhas utilizadas para esse fim são complexas e dispendiosas de fabricar, envolvendo processos intrincados como a sobreposição sucessiva de camadas de semicondutores. Essa complexidade não só aumenta o custo dos dispositivos, como também limita sua adoção em produtos eletrônicos de consumo e dispositivos móveis.

A implementação de sistemas de autenticação baseados em veias da palma da mão, apesar de seus benefícios, como a não necessidade de contato físico, enfrenta desafios técnicos como variantes de escala, rotação, e a presença de anéis ou dedos fechados que podem interferir na captura precisa das veias. Além disso, a exposição prolongada à radiação infravermelha, necessária para capturar os padrões venosos, pode ser potencialmente

prejudicial, e a integração de sensores NIR implica em custos adicionais que impactam a viabilidade econômica do sistema.

Devido a essas restrições, as técnicas atuais de autenticação biométrica baseadas em veias da palma utilizam predominantemente Redes Neurais Profundas (DNNs) para a tarefa de verificação (Jia et al. 2021). Essas redes, embora eficazes, requerem um aumento nos parâmetros das arquiteturas DNNs para alcançar maiores precisões, o que eleva os custos computacionais de inferência, tornando essas soluções inadequadas para dispositivos com recursos limitados, como dispositivos móveis (Horng et al. 2022). Em contextos práticos, os sistemas de identificação de veias da palma da mão, como o Fujitsu PalmSecure™ e LG G8 ThinQ, têm sido desenvolvidos e mantidos comercialmente, mas os detalhes técnicos desses sistemas permanecem confidenciais devido a segredos comerciais e ambos emergem como soluções caras.

Nos últimos anos, técnicas de *deep learning* têm ganhado destaque na área de reconhecimento biométrico, especialmente em cenários onde a precisão e a robustez são fundamentais. Modelos baseados em Redes Neurais Convolucionais (CNNs) têm sido amplamente utilizados para extrair características de imagens de veias, graças à sua capacidade de capturar padrões complexos através de camadas profundas. Essas redes são capazes de aprender representações discriminativas das veias da palma, diferenciando indivíduos com alta precisão. Além disso, abordagens baseadas em Redes Siamesas têm mostrado resultados promissores para tarefas de verificação, onde o objetivo é comparar duas imagens e determinar se elas pertencem à mesma pessoa. Essas redes, que compartilham pesos entre os ramos, aprendem uma função de similaridade que pode ser usada para comparar diretamente os padrões venosos extraídos de diferentes imagens.

Outra técnica relevante é o uso de *Autoencoders* Profundos, que podem ser aplicados para redução de dimensionalidade e extração de características latentes das imagens de veias. Ao codificar e decodificar as imagens, esses modelos conseguem captar as características mais importantes das veias, que podem ser usadas posteriormente para classificação ou verificação. A combinação dessas técnicas em um pipeline integrado pode potencialmente eliminar a necessidade de sensores NIR, ao mesmo tempo em que mantém a alta precisão necessária para aplicações de segurança.

Dada a necessidade de soluções que ofereçam alta precisão sem depender de sensores NIR especializados, o desenvolvimento de sistemas biométricos eficientes para dispositivos móveis emerge como uma tarefa crucial. Apesar do poder computacional presente nos dispositivos modernos, a ausência de sensores específicos, como os infravermelhos, representa uma lacuna significativa. Por essa razão, é essencial conceber métodos de autenticação que sejam otimizados em termos de recursos disponíveis e que eliminem a necessidade de hardware especializado.

Como premissa, consideramos que a construção de um método biométrico integralmente composto por redes neurais e métodos baseados em machine learning são capazes de preencher a lacuna mencionada anteriormente. Esta premissa sustenta a hipótese deste trabalho: "A aplicação conjunta de *deep autoencoder* e redes siamesas possibilita a realização da biometria baseada nas veias da palma da mão, eliminando a necessidade de um sensor NIR."

1.2 Objetivos

Diante dos desafios associados aos sistemas atuais de reconhecimento de veias, o principal objetivo deste estudo é propor uma nova estratégia biométrica leve para veias da palma, utilizando um *Deep Autoencoder* como alternativa aos sensores infravermelhos e uma Rede Siamesa responsável por conduzir o processo biométrico. Essa estratégia visa abordar as limitações dos sensores de câmera NIR e dos acessórios de posicionamento da palma, tornando a autenticação biométrica de veias da palma mais viável e econômica, especialmente em dispositivos com recursos limitados. Os objetivos específicos estabelecidos para atingir esse propósito são os seguintes:

- Desenvolver um método de pré-processamento de imagens para aprimorar a qualidade dos dados biométricos.
- Projetar e otimizar uma rede neural autoencoder para realçar características da palma da mão.
- Projetar e treinar um modelo de rede neural siamesa para gerar coeficientes de similaridade entre imagens de veias da palma.

- Definir um limiar de comparação para validar a autenticidade com base no coeficiente gerado pela rede siamesa.
- Avaliar a performance do método proposto em comparação com sistemas tradicionais que utilizam sensores NIR.

Ao alcançar esses objetivos, busca-se promover o desenvolvimento de um protótipo biométrico acessível, com potencial impacto nas práticas de autenticação e segurança da informação em ambientes diversos.

1.3 Contribuições

As principais contribuições deste trabalho são as seguintes:

- Desenvolvimento de um esquema biométrico leve e acessível para veias da palma, implementado através de um *Autoencoder* Profundo e uma Rede Siamesa, que supre a necessidade do sensor de câmera NIR e do acessório de posicionamento da palma para a tarefa de verificação.
- Um protótipo funcional, implementado e validado em um dispositivo com recursos limitados, demonstrando a viabilidade da nossa solução, visando oferecer uma alternativa prática e acessível para a autenticação, com potencial impacto positivo em aplicações reais.

Este trabalho foi publicado na International Conference on Machine Learning and Applications (ICMLA) 2024, Qualis A2, pelos autores Mateus Nunes, Eduardo K. Viegas e Altair O. Santin.

1.4 Organização do Documento

O documento encontra-se organizado da seguinte forma: No Capítulo 1 é realizada a introdução da proposta. O Capítulo 2 aborda a fundamentação teórica. No Capítulo 3 são apresentados os trabalhos relacionados à proposta. O Capítulo 4 detalha a metodologia utilizada para a realização do projeto. No Capítulo 5 é apresentada a análise dos resultados preliminares. Por fim no Capítulo 6 é apresentada a conclusão.

Capítulo 2

Fundamentação

Este capítulo apresenta conceitos relacionados ao processo de autenticação biométrica tradicional e proposto, contemplando técnicas e processos específicos relacionados ao reconhecimento de veias da palma da mão. Compreenderemos as técnicas empregadas nesse processo, os detalhes dos procedimentos envolvidos e as características únicas que tornam a autenticação biométrica baseada em veias palmares uma abordagem única.

2.1 Autenticação biométrica

A biometria refere-se ao reconhecimento de indivíduos com base em suas características físicas ou traços comportamentais. Nesse sentido, a biometria pode ser vista como tão antigo quanto a própria humanidade. A possibilidade de automatizar o processo de reconhecimento e deixar computadores e dispositivos de captura conectados executando esta tarefa levou ao sucesso desenvolvimento e implantação de inúmeras tecnologias biométricas. A biometria vascular surgiu nos últimos anos e é percebida como um recurso atraente, mas ainda inexplorado de muitas perspectivas (Shinzaki 2019).

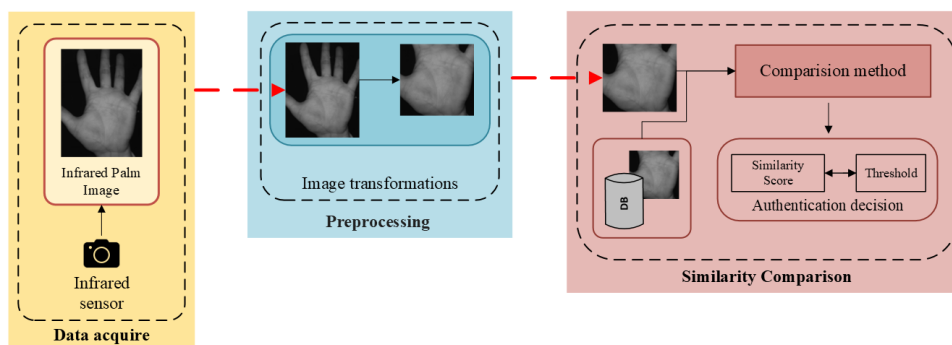


Figura 1 – Fluxo tradicional

A Figura 1 ilustra o fluxo tradicional de autenticação baseado em veias da palma da mão. Primeiro, a imagem NIR é capturada por um sensor infravermelho. Em seguida, a imagem passa por um processo de pré-processamento para corrigir problemas relacionados à posição, rotação e tamanho. Por fim, a imagem é submetida a uma etapa de comparação, na qual é confrontada com uma imagem previamente conhecida de um usuário válido. O resultado é uma taxa de similaridade, que é comparada com um limiar previamente estipulado. Por fim, ocorre a decisão biométrica.

A autenticação biométrica é um método de verificação de identidade que utiliza características físicas exclusivas de um indivíduo para confirmar sua autenticidade. Diferentemente dos métodos tradicionais, como senhas ou cartões, a autenticação biométrica se baseia em atributos únicos que são inerentes à própria pessoa.

Algumas das características biométricas comumente utilizadas incluem impressões digitais, íris, retina, reconhecimento facial, geometria da mão, voz e padrões de veias.

Vantagens da Autenticação Biométrica:

1. Segurança Aprimorada:

As características biométricas são únicas para cada indivíduo, proporcionando um nível elevado de segurança.

Dificulta a falsificação ou roubo de identidade, uma vez que as características biométricas não podem ser facilmente replicadas.

2. Conveniência para Usuários:

Elimina a necessidade de memorizar senhas complexas, reduzindo a carga cognitiva dos usuários.

O processo é geralmente rápido e sem a necessidade de itens adicionais, como cartões ou chaves.

3. Não Repudiabilidade:

A autenticação biométrica reduz a possibilidade de não repudiabilidade, pois a conexão entre a pessoa e sua identidade é mais difícil de negar.

4. Prevenção de Acesso Não Autorizado:

Minimiza o risco de acesso não autorizado, uma vez que as características biométricas são difíceis de replicar ou adulterar.

5. Aprimoramento da Experiência do Usuário:

Oferece uma experiência de usuário mais amigável e eficiente, especialmente em comparação com métodos tradicionais que podem ser esquecidos ou perdidos.

6. Aplicações Diversificadas:

Pode ser aplicada em uma variedade de contextos, desde dispositivos móveis e sistemas de segurança até transações financeiras e controle de acesso físico.

7. Monitoramento Contínuo:

Algumas modalidades biométricas, como o reconhecimento facial, podem permitir o monitoramento contínuo, aumentando a segurança em tempo real.

8. Redução de Fraudes:

Contribui para a redução de fraudes, uma vez que as características biométricas são difíceis de serem replicadas por terceiros.

É importante observar que, embora a autenticação biométrica apresente muitas vantagens, também enfrenta desafios, como questões de privacidade, a necessidade de infraestrutura especializada e a possibilidade de falsificação em certas circunstâncias. Portanto, sua implementação deve ser cuidadosamente considerada e adaptada ao contexto específico de uso.

2.2 Autenticação biométrica baseada em veias

A autenticação biométrica, centrada na análise da camada venosa subcutânea, emerge como uma solução resiliente a fatores externos e fortemente resistente à falsificação. Esta técnica apresenta identificação exclusiva, sendo discernível pela presença de corrente

sanguínea em indivíduos vivos, proporcionando praticidade e a vantagem de não exigir contato direto com objetos, garantindo, assim, condições higiênicas.

Anatomicamente, a singularidade dos padrões vasculares é determinada pelas principais artérias que irrigam a mão, como a artéria ulnar e a artéria radial, que se ramificam para formar uma rede complexa de veias subcutâneas na palma. A artéria ulnar, segue ao longo do lado medial do antebraço e, ao atingir a palma da mão, contribui para a formação do arco palmar superficial, que por sua vez dá origem a ramos que irrigam os dedos e a superfície palmar. A combinação única das ramificações dessas artérias e a subsequente disposição das veias fazem com que cada indivíduo possua um padrão venoso exclusivo, inimitável mesmo entre gêmeos idênticos.

Cada indivíduo possui sistemas vasculares únicos, estabelecendo distinção para além das características das impressões digitais e palmares (Hornig et al. 2022). Segundo a explicação da Fujitsu (Fujitsu 2024), os sistemas de varredura das veias da palma da mão, assim como os das veias dos dedos, utilizam tecnologia baseada em raios infravermelhos e na interação da hemoglobina nas veias com esses raios. A hemoglobina, responsável pelo transporte de oxigênio no sangue, apresenta-se oxigenada nas artérias, enquanto as veias conduzem o sangue desoxigenado de volta ao coração.

Essa abordagem não apenas confere robustez à autenticação biométrica, mas também se destaca pela sua aplicação prática e pela consideração das implicações higiênicas. A análise da camada venosa subcutânea, amparada pela tecnologia de raios infravermelhos, oferece uma camada adicional de segurança, beneficiando-se da singularidade dos padrões vasculares individuais. Essa singularidade, aliada à tecnologia empregada, não apenas reforça a autenticidade do processo, mas também amplia suas possibilidades de aplicação em diversos contextos, desde a segurança de dispositivos até ambientes que demandam elevados padrões de higiene.

2.3 Infravermelho próximo (NIR)

O infravermelho próximo é uma região do espectro eletromagnético situada entre a luz visível e o infravermelho propriamente dito. A região espectral do Infravermelho Próximo

compreende a região do espectro eletromagnético que vai de 750 nm a 2500 nm ($2,65 \times 10^{-19}$ a $7,96 \times 10^{-20}$ J). A absorção de radiação eletromagnética nessa região ocorre quando a energia da radiação tem o mesmo valor da diferença de energia entre dois estados vibracionais; ou seja, o processo está relacionado à ressonância entre a diferença de níveis de energia da molécula e a radiação eletromagnética. Na região do NIR ocorrem combinações e sobretons (sobreposição) das absorções vibracionais de grupamentos químicos, o que significa que os espectros de NIR possuem grande quantidade de informação a respeito da organização da maioria dos compostos orgânicos (Pasquini 2003).

2.4 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) constituem uma classe de modelos computacionais inspirados na estrutura e funcionamento dos neurônios no cérebro humano. Essa abordagem é fundamentada na ideia de que sistemas complexos podem ser simulados e emulados por meio da interconexão de unidades simples, chamadas de neurônios artificiais. Esses modelos foram concebidos para realizar tarefas que envolvem aprendizado, reconhecimento de padrões, e tomada de decisões, em uma variedade de domínios.

As RNAs são compostas por camadas de neurônios interconectados, organizados em três tipos principais: camada de entrada, camadas ocultas e camada de saída. A informação é processada e transmitida através dessas camadas por meio de conexões ponderadas entre os neurônios, conhecidas como sinapses. Cada neurônio recebe entradas, realiza um cálculo ponderado, aplica uma função de ativação e transmite o resultado para os neurônios nas camadas subsequentes.

O aprendizado em RNAs ocorre por meio do ajuste dos pesos das sinapses durante o treinamento do modelo. Este processo é frequentemente supervisionado, utilizando um conjunto de dados de treinamento para otimizar a capacidade da rede em realizar tarefas específicas. Os algoritmos de aprendizado, como o *Backpropagation*, são empregados para minimizar a diferença entre as saídas previstas e os valores reais desejados.

Diferentes arquiteturas de RNAs são desenvolvidas para atender a diversas necessidades. Redes *Feedforward*, onde a informação flui em uma única direção, são comuns em tarefas de classificação. Redes Recorrentes, por outro lado, permitem conexões

retroativas, sendo úteis para lidar com dados sequenciais ou temporais. As redes neurais convolucionais (CNNs) têm alcançado resultados impressionantes em tarefas desafiadoras, como reconhecimento de imagem e detecção de objetos, o que tem ampliado significativamente o interesse dos pesquisadores em tais métodos (Krizhevsky 2009; Krizhevsky et al. 2012; Ciresan et al. 2012).

2.4.1. CNN – Redes Neurais Convolucionais

As Redes Neurais Convolucionais (CNN) representam uma classe especializada de arquiteturas de redes neurais artificiais que se destacam em tarefas relacionadas à visão computacional, processamento de imagens e reconhecimento de padrões espaciais. Essas redes foram projetadas para capturar eficientemente padrões hierárquicos e complexos em dados com estrutura espacial. A estrutura fundamental das CNNs é composta por camadas convolucionais, de pooling (agrupamento) e totalmente conectadas. A camada convolucional é central para a eficácia das CNNs, sendo responsável por aprender filtros ou características locais em diferentes regiões da entrada. A camada de *pooling* reduz a dimensionalidade, preservando características essenciais. As CNNs são amplamente aplicadas em tarefas como reconhecimento de objetos, classificação de imagens, segmentação semântica, detecção de características e até mesmo em aplicações mais avançadas, como reconhecimento de faces e diagnóstico médico por meio de imagens. Uma das principais características da CNN é a sua capacidade de ser escalável (Strigl et al. 2023), sendo ideal para tarefas de visão computacional, pois isso permite processar uma grande quantidade de informações de entrada advindas dos pixels das imagens ou dos frames dos vídeos.

Embora as CNNs tenham obtido notável sucesso, desafios persistem, como a interpretabilidade dos modelos e o manejo de conjuntos de dados desbalanceados. Tendências futuras incluem a exploração de arquiteturas mais eficientes, o desenvolvimento de técnicas de regularização avançadas e a aplicação de CNNs em domínios mais amplos, como processamento de linguagem natural.

2.4.1.1 Camada Conv (Convolutacional)

As Camadas Convolutacionais constituem um componente central nas Redes Neurais Convolutacionais (CNN), uma classe de arquiteturas especializadas para processamento de imagens e tarefas relacionadas à visão computacional. Estas camadas desempenham um papel crucial na extração e aprendizado de padrões espaciais, possibilitando a eficaz identificação de características em dados complexos.

A operação fundamental nas camadas convolutacionais é a convolução, que envolve a aplicação de filtros sobre a entrada da imagem. Cada filtro atua como um detector de características específicas, como bordas, texturas ou formas. Durante o treinamento, os parâmetros desses filtros são ajustados para aprender características relevantes para a tarefa em questão.

As camadas convolutacionais são projetadas para capturar características locais em regiões específicas da entrada, permitindo uma representação mais eficiente dos padrões presentes na imagem. Os filtros convolutacionais percorrem a entrada e geram mapas de ativação, destacando áreas onde características específicas foram identificadas.

2.4.1.2 Camada de Pooling

As Camadas de Pooling constituem um elemento essencial nas Redes Neurais Convolutacionais (CNN), desempenhando um papel significativo na redução da dimensionalidade dos dados e na preservação das características essenciais extraídas pelas camadas convolutacionais. Esta técnica é fundamental para o sucesso das CNNs em tarefas de visão computacional e processamento de imagens.

A operação de pooling é aplicada após as camadas convolutacionais e envolve a redução das dimensões dos mapas de ativação gerados por essas camadas. A forma mais comum é uma camada de agrupamento com filtros de tamanho 2x2 aplicados com um passo de 2 amostras em cada fatia de profundidade (na entrada de 2 ao longo da largura e altura), que encolhe a imagem de entrada para 1/4 do tamanho original, descartando assim 75% das ativações (CS231n 2023).

O principal propósito das camadas de pooling é a redução da dimensionalidade dos dados. Isso é alcançado pela subamostragem dos mapas de ativação, mantendo as características mais importantes e descartando informações redundantes. A redução de dimensionalidade contribui para a eficiência computacional e ajuda a evitar o overfitting, melhorando a capacidade de generalização do modelo.

2.4.1.3 Função de Ativação – ReLU

A Função de Ativação ReLU (Rectified Linear Unit) é crucial nas Redes Neurais Convolucionais (CNNs) para introduzir não linearidades. Definida como $f(x)=\max(0, x)$, a ReLU retorna zero para entradas negativas e mantém as positivas. Essa simplicidade promove eficiência computacional e evita o problema de saturação.

A ReLU acelera o treinamento ao evitar a saturação para entradas positivas, contribuindo para uma propagação eficaz de gradientes. Amplamente aplicada em camadas convolucionais, a ReLU é essencial para aprender características não lineares em tarefas como reconhecimento de objetos.

Apesar das vantagens, a "dying ReLU" é um desafio, onde neurônios podem ficar inativos. Variações, como Leaky ReLU e Parametric ReLU, foram propostas para resolver esse problema, permitindo um pequeno gradiente para entradas negativas.

Em síntese, a Função de Ativação ReLU é uma escolha eficaz e simples em CNNs, desempenhando papel fundamental na capacidade da rede de aprender representações não lineares em tarefas de visão computacional.

2.5 Deep learning

Deep Learning, ou aprendizado profundo, é uma subcategoria de machine learning (aprendizado de máquina) que utiliza redes neurais artificiais profundas para realizar tarefas específicas de forma automática. No contexto da autenticação biométrica, o Deep Learning desempenha um papel significativo ao oferecer abordagens avançadas para a extração e análise de características de grandes quantidades de informações biométricas. De acordo com (LeCun et al.2015), o Deep Learning é definido como métodos que permitem que um modelo

composto de várias camadas aprenda representações dos dados em diversos níveis de abstração. Esses algoritmos mostraram serem capazes de superar os algoritmos tradicionais, chegando muito próximos do desempenho do nível humano e, em alguns casos, até mesmo sendo superiores. O Deep Learning refere-se a um subgrupo das RNAs (Artificial Neural Networks), que são redes compostas por múltiplas camadas (Schmidhuber 2015).

2.5.1 Deep Autoencoder

O *Deep Autoencoder*, refere-se a uma arquitetura avançada de rede neural artificial, especificamente projetada para aprendizado não supervisionado e extração eficiente de características complexas dos dados de entrada. Essa abordagem é "profunda" devido à presença de múltiplas camadas, permitindo a aprendizagem de representações hierárquicas e abstratas dos dados.

O *Deep Autoencoder* para mudança espectrográfica de imagens representa uma abordagem avançada no campo do processamento de imagens, especialmente no que diz respeito à alteração controlada das propriedades espectrais de uma imagem. Esta metodologia se baseia na utilização de redes neurais profundas, conhecidas como autoencoders, com o intuito de aprender representações complexas e hierárquicas dos componentes espectrais das imagens, permitindo transformações significativas nas características visuais.

O termo "*Deep Autoencoder*" refere-se a uma arquitetura neural que possui várias camadas, tornando-a capaz de aprender de maneira eficaz representações abstratas dos dados de entrada. No contexto da mudança espectrográfica de imagens, essa arquitetura é treinada para compreender as correlações e padrões nas informações espectrais, capacitando-a a realizar modificações nas propriedades cromáticas de maneira controlada.

Durante o treinamento, a rede é exposta a pares de imagens, consistindo em versões originais e suas contrapartes com mudanças espectrográficas conhecidas. O objetivo principal é capacitar a rede para mapear automaticamente as características espectrais relevantes e realizar transformações desejadas nas imagens. O processo de retropropagação é empregado para ajustar os pesos da rede, refinando sua capacidade de representação.

A profundidade da rede, refletida no termo "deep", é fundamental para a extração de representações cada vez mais complexas das informações espectrais. Isso permite que a rede aprenda nuances detalhadas nas mudanças de cores e propriedades visuais, proporcionando um controle mais refinado sobre as transformações realizadas.

Essa abordagem possui amplas aplicações em diversos domínios, como processamento de imagens, edição de fotos e correção de cores em vídeos. Sua capacidade de manipular de forma controlada as propriedades espectrais das imagens torna-a valiosa em situações em que a personalização e a otimização das características visuais são críticas.

2.5.2 Rede siamesas

Uma rede siamesa refere-se a uma arquitetura especial de rede neural utilizada para comparar e avaliar a similaridade entre dois inputs, como imagens, sequências de texto ou outras representações. A Rede Neural Siamesa foi proposta no trabalho (Seok et al. 2023), sendo aplicada em um problema de análise de imagens para a autenticação de assinaturas. A rede siamesa é formada por duas sub-redes idênticas. Essas redes, que possuem a mesma estrutura e recebem os mesmos pesos e parâmetros, levam as amostras recebidas em suas entradas para um novo espaço de características. Neste novo espaço, as amostras de mesma classe se encontram mais próximas e as amostras de classes diferentes são distanciadas uma das outras. Em seguida, é feito o cálculo da distância entre as duas amostras no espaço. Esse valor de distância é então processado por outra rede neural (ou um único neurônio), que tem como saída um valor entre 0 e 1. Se as amostras analisadas forem da mesma classe, o valor de similaridade calculado será mais próximo de 1, e se as amostras forem de classes diferentes, então esse valor será próximo de 0. No treinamento, esse valor é comparado com o alvo das entradas, é calculada a perda e os pesos das sub-redes são atualizados via *backpropagation* (Goodfellow et al. 2016).

2.6 Classificadores

Em um contexto de aprendizado supervisionado, os classificadores desempenham um papel fundamental ao automatizar a tarefa de atribuir rótulos a instâncias desconhecidas com base em um conjunto de treinamento rotulado. Esses modelos são essenciais para a

classificação de dados, onde o objetivo é prever a categoria à qual uma nova instância pertence.

Um classificador, conforme definido em (Kotsiantis et al. 2007), é um modelo resultante do treinamento em um conjunto de dados rotulado, capaz de associar classes a instâncias de teste cujas classes são desconhecidas. Esta associação é realizada com base nas informações contidas nos atributos das instâncias.

Os classificadores podem gerar saídas de duas principais naturezas: discreta (categórica) ou contínua, dependendo da natureza da tarefa. Em tarefas de classificação, o modelo aprendido produzirá saídas discretas, representando a classe à qual uma instância pertence. Por outro lado, em tarefas de regressão, as saídas serão valores contínuos.

2.6.1 Algoritmo KNN (K-Nearest Neighbors)

O algoritmo K-Nearest Neighbors (KNN), ou K-Vizinhos Mais Próximos, é uma técnica de aprendizado de máquina que se destaca por sua simplicidade conceitual e eficácia em problemas de classificação e regressão. Fundamentado na ideia de que instâncias semelhantes tendem a pertencer à mesma classe ou compartilhar padrões similares, o KNN é classificado como um método baseado em instância.

A essência do KNN reside na determinação da classe de uma instância desconhecida com base nas classes das K instâncias mais próximas a ela em um espaço multidimensional. A proximidade é frequentemente medida utilizando métricas como a distância euclidiana, onde as instâncias mais próximas têm uma influência maior na decisão de classificação.

O parâmetro "K" denota o número de vizinhos mais próximos considerados para tomar uma decisão de classificação. A escolha adequada de K é crucial, pois valores pequenos podem tornar o modelo sensível a outliers, enquanto valores grandes podem introduzir uma generalização excessiva.

Em tarefas de classificação, a classe da instância desconhecida é determinada pela classe mais frequente entre seus vizinhos mais próximos. Para tarefas de regressão, a saída prevista é frequentemente a média dos valores alvo das instâncias vizinhas.

Uma das vantagens do KNN é sua capacidade de se adaptar a diferentes formas de distribuição de dados e a complexidades não lineares, tornando-o útil em cenários onde a fronteira de decisão é intrincada.

Embora o KNN seja robusto em muitos cenários, sua eficácia pode ser comprometida em conjuntos de dados com muitas dimensões ou quando a distribuição das classes é desigual. Estratégias de normalização de dados e otimização do valor de K podem ser adotadas para mitigar essas limitações.

Capítulo 3

Trabalhos relacionados

A biometria tem sido um tópico de pesquisa amplamente explorado na literatura nos últimos anos (Lien e Vhaduri 2023). As abordagens atuais variam desde o uso do rosto, impressão digital, íris e até mesmo a veia da palma (Minaee et al. 2019). A complexidade encontrada nesse mecanismo permanece um tópico aberto ao debate entre pesquisadores.

Nesse contexto, a análise dos Trabalhos Relacionados apresenta uma visão completa das estratégias adotadas, obstáculos enfrentados e avanços.

3.1 Extração de características através de imagens em espectro visível

De acordo com a revisão da literatura, alguns estudos têm se dedicado à extração de informações relevantes das veias presentes na palma da mão, no entanto, enfrentam desafios decorrentes da falta de informações suplementares provenientes do emprego da luz infravermelha.

Conforme mencionado em (Horng et al. 2022), vários estudos sobre reconhecimento de veias palmares em smartphones utilizam imagens de infravermelho próximo (NIR) capturadas por conjuntos de dados como PolyU, CASIA e Tongji. No entanto, esses estudos dependem de dispositivos personalizados com câmeras especiais e iluminação LED, que são caros e inacessíveis para o usuário comum, além de não considerarem o desempenho em um ambiente de produção real.

(Horng et al. 2022) propõem uma abordagem funcional, mas onerosa, que exige altos recursos computacionais em dispositivos móveis. Após a extração das características, a imagem passa por processos manuais para aprimorar a região de interesse das veias da palma, incluindo correção automática de gama (AGC), equalização adaptativa de histograma (CLAHE), e filtros passa-baixa e passa-alta. Além disso, o algoritmo extrai o canal de

saturação da imagem RGB para destacar as veias. Embora eficaz, essa técnica é desafiadora para implementação em dispositivos móveis devido à alta demanda computacional.

O estudo "A hand-based biometric system in visible light for mobile environments," liderado por Silvio Barra et al., propõe um sistema biométrico para dispositivos móveis usando imagens em luz visível. Embora os autores obtenham um Equal Error Rate (EER) de 0,52%, a abordagem enfrenta limitações significativas devido à facilidade de falsificação e à falta de unicidade do formato da mão.

O trabalho "Palm-Vein Verification Using Images From the Visible Spectrum" propõe a verificação de identidade por veias da palma usando imagens RGB. Embora obtenha uma taxa de erro de 0,5% em 200 imagens e seja computacionalmente eficiente, o método não foi projetado como um protótipo funcional para dispositivos móveis. Isso limita sua aplicabilidade prática em ambientes com recursos restritos e não aborda a viabilidade em contextos reais. A abordagem, embora promissora em termos de resultados, enfrenta desafios significativos para a implementação em dispositivos móveis.

3.2 Biometria utilizando deep learning

As abordagens baseadas em deep learning visam extrair representações de características de alta dimensão para medir similaridades individuais. Por exemplo, (F. Schroff et al. 2015) propõem a arquitetura FaceNet, que aprende uma camada de incorporação para medir a similaridade facial. A abordagem deles oferece alta precisão, mas muitas vezes é inviável para dispositivos com recursos limitados devido às exigências computacionais do cálculo de distância. De maneira semelhante, (H. Benradi et al. 2023) propõem um sistema de reconhecimento facial baseado em uma Rede Neural Convolutiva (CNN) que extrai múltiplas características, incluindo pose, iluminação e contorno. A abordagem deles alcança taxas de detecção significativamente altas, mas não considera a implementação em dispositivos com recursos limitados.

Outra abordagem foi proposta por (J. N. Kolf et al. 2023), que implementa um esquema de reconhecimento facial em um dispositivo com recursos limitados. Os autores propõem a quantização do modelo de DNN e a redução da precisão da representação de cores

das imagens para reduzir os requisitos de memória e processamento. A proposta deles alcança alta precisão, mas não pode ser aplicada ao domínio da veia da palma devido à necessidade de um sensor de câmera NIR.

Infelizmente, as biometrias baseadas em características faciais são vulneráveis a adversários, que podem enganar o sistema para obter um reconhecimento falso, como ao imprimir rostos e usá-los para identificação (Tome et al. 2015). Diante disso, processos biométricos baseados em padrões de veias têm sido o tema de vários trabalhos nos últimos anos. Como exemplo, (H. Qin et al. 2021) propõem um esquema de aumento de dados para melhorar a precisão de DNN na identificação de veias da palma. A abordagem deles melhora significativamente a precisão da detecção, mas desconsidera os custos de processamento e assume a disponibilidade de um sensor de câmera NIR. De maneira semelhante, (F. O. Babalola et al.) extraem características estatísticas binarizadas para fins de identificação de veias da palma. O esquema deles oferece alta precisão por meio de uma identificação baseada em DNN, mas também desconsidera a implementação em dispositivos com recursos limitados.

(S. Barra et al.) recorrem a múltiplas tarefas de pré-processamento para identificação confiável de veias da palma. Os autores extraem características de alto nível através de uma fase de pré-processamento baseada em recorte para alcançar alta precisão. No entanto, eles também assumem a disponibilidade de um sensor de câmera NIR e desconsideram os custos de processamento. (D. Zhong et al.) recorrem a uma abordagem baseada em hash para calcular similaridades a partir de características extraídas de um modelo de DNN. O esquema deles pode reduzir os custos de processamento, mas também assume a disponibilidade de um sensor de câmera NIR.

Mais recentemente, o trabalho (Nayar et al. 2023) aborda a autenticação biométrica utilizando padrões parciais de veias da palma, explorando como imagens reduzidas podem ser efetivas para a identificação individual. No entanto, o estudo ignora completamente a implementação e teste com sensores reais e dispositivos com recursos limitados. A pesquisa se concentra apenas em experimentos laboratoriais com imagens e não considera as dificuldades práticas de aplicar a técnica em cenários do mundo real, como dispositivos móveis ou sistemas de autenticação portáteis que têm restrições de tamanho e capacidade.

Como resultado, ainda há uma lacuna de pesquisa na literatura que precisa ser preenchida para a implantação prática de abordagens de identificação de veias da palma. Na prática, os esquemas propostos devem superar a necessidade de um sensor de câmera NIR, enquanto ainda abordam a limitação de custos de processamento das arquiteturas de DNN amplamente utilizadas. Além disso, os acessórios necessários para posicionar adequadamente a palma do indivíduo a ser verificado representam um desafio significativo. As abordagens atuais assumem um posicionamento preciso da palma, muitas vezes alcançado pela instalação de acessórios de suporte para a palma. Assim, elas frequentemente carecem de etapas de pré-processamento, como recorte de imagem, rotação e identificação, para garantir o alinhamento adequado da palma.

3.3 Mudança espectral

O trabalho (Jameel et al. 2021) apresenta uma abordagem inovadora para a geração de imagens espectrais a partir de diferentes tipos de imagens, utilizando a arquitetura do AutoEncoder. A utilização de deep learning, em particular a arquitetura do *AutoEncoder*, permite que o modelo proposto, GVTI-AE, seja totalmente automatizado, sem a necessidade de pré ou pós-processamento ou entrada do usuário. A mudança espectral de imagens visíveis para infravermelho é um dos aspectos abordados no estudo. Os resultados do estudo demonstram que o modelo GVTI-AE foi capaz de realizar transformações entre diferentes tipos de imagens, como imagens visíveis para infravermelho (V2I), com baixas taxas de erro e altos valores de Structural Similarity Index (SSIM). Os resultados do estudo demonstram que o modelo GVTI-AE supera os métodos do estado-da-arte em termos de precisão e desempenho na transformação de diferentes tipos de imagens, incluindo a mudança espectral de imagens visíveis para infravermelho. Os autores observaram que o modelo foi capaz de produzir resultados perceptualmente realistas em conjuntos de dados amplamente disponíveis, como o Tecnocampus Hand Image Database, Carl dataset e IRIS Thermal/Visible Face Database. Essa abordagem automatizada e perceptualmente realista pode ter aplicações em diversas áreas, incluindo medicina, segurança e detecção de objetos.

3.4 Redes siamesas

O estudo "Finger Knuckleprint Based Personal Authentication Using Siamese Network" (Joshi et al. 2019) apresenta uma abordagem inovadora para autenticação biométrica usando impressões digitais do nó do dedo. O sistema utiliza um modelo de rede neural siamesa para autenticação, baseando-se na similaridade entre imagens das impressões digitais do nó do dedo. Este trabalho está em sintonia com pesquisas similares, como o estudo de Jaswal. (Jaswal et al. 2017), que também emprega imagens de impressões digitais do nó do dedo para autenticação, utilizando o método de DeepMatching.

Outra contribuição significativa vem do trabalho de Kumar (Kumar et al., 2009), que propõe a utilização da superfície do nó do dedo para autenticação, incorporando técnicas de processamento de imagem e aprendizado de máquina. O sistema proposto destaca-se por sua imunidade a problemas enfrentados por sistemas tradicionais baseados em impressões digitais, como desgaste dos padrões devido a atividades manuais em áreas rurais, alcançando uma notável taxa de reconhecimento correto de 99,24% e uma taxa de erro de apenas 0,78%.

No âmbito do reconhecimento facial, a pesquisa "Face Recognition based on Convolution Siamese Networks," de (WU, H. et al. 2017), propõe o uso de redes convolucionais siamesas para melhorar a precisão e robustez na era da inteligência artificial. A técnica demonstrou uma notável precisão de 98,21% em um conjunto de dados específico, superando outras abordagens de reconhecimento facial. A capacidade dessas redes em extrair características profundas e abstratas das imagens faciais destaca-se como uma vantagem significativa, promovendo uma classificação e identificação aprimoradas.

Adicionalmente, recentes pesquisas têm explorado o uso de redes neurais convolucionais profundas para autenticação de veias da palma da mão, (Thapar et al., 2019) propuseram o PVSNet, um framework baseado em redes neurais convolucionais profundas, que utiliza uma rede codificador-decodificador convolucional para extração de características e uma rede Siamesa treinada com a função de perda de triplet e mineração adaptativa de exemplos difíceis. O PVSNet apresentou eficácia na autenticação de veias da palma da mão, com resultados superiores em comparação com outras soluções de aprendizado profundo, indicando sua promissora aplicabilidade em segurança biométrica.

Os trabalhos relacionados enfatizam a eficácia das redes Siamesas em diversas aplicações de autenticação biométrica. Desde o reconhecimento de impressões digitais do nó do dedo até a identificação facial e a autenticação de veias da palma da mão, esses estudos destacam a versatilidade e a precisão dessas redes. A consistência de resultados expressivos, como altas taxas de reconhecimento e baixas taxas de erro, respalda a utilização das redes Siamesas para aprimorar a segurança e a confiabilidade em sistemas biométricos.

3.5 Considerações Finais

A autenticação biométrica por análise das veias da palma da mão representa um desafio complexo, amplamente explorado na literatura científica. Este estudo revisitou trabalhos anteriores para destacar as principais abordagens, lacunas e avanços relevantes na área. Embora muitos métodos utilizem luz infravermelha para capturar detalhes das veias, a dependência de dispositivos caros e inacessíveis é uma limitação significativa. Alternativas como o método de Silvio Barra et al., que utiliza luz visível, mostraram-se promissoras para dispositivos móveis, oferecendo boa eficácia e eficiência, além de maior acessibilidade.

No campo do deep learning, as redes neurais convolucionais (CNN) têm demonstrado eficácia na identificação biométrica. Contudo, essas redes frequentemente exigem alta capacidade computacional, o que limita sua aplicabilidade em dispositivos com recursos restritos. A técnica de mudança espectral, exemplificada pelo modelo "Generating Spectrum Images from Different Types — Visible, Thermal, and Infrared Based on *Autoencoder Architecture* (GVTI-AE)", introduziu uma abordagem automatizada e realista com potencial para aplicações em medicina, segurança e detecção de objetos. As redes siamesas, por sua vez, se destacam pela versatilidade em biometria, desde impressões digitais até a autenticação de veias, apresentando resultados consistentes com altas taxas de reconhecimento e baixas taxas de erro. Essas redes têm o potencial de aprimorar significativamente a segurança e a confiabilidade dos sistemas biométricos. Em suma, as abordagens inovadoras e técnicas discutidas neste estudo confirmam a eficácia do deep learning, especialmente por meio de redes CNN e siamesas, na autenticação biométrica. Essas técnicas não apenas enfrentam os desafios atuais, mas também apontam para avanços significativos no campo do

reconhecimento biométrico, oferecendo novas perspectivas para o aprimoramento da segurança e eficiência dos sistemas biométricos.

A seguir, há uma tabela que apresenta uma comparação dos principais estudos sobre reconhecimento de veias da palma, destacando sensores utilizados, métodos de autenticação, uso de deep learning e resultados de acurácia obtidos.

Trabalho	Sensor NIR	ROI automático	Deep Learning no pré-processamento	Deep Learning na etapa de autenticação	Projetado para Mobile	Método de Autenticação	Acurácia/EER	Base Utilizada
Recognizing Palm Vein in Smartphones Using RGB Images	Não	Sim	Não	Sim	Sim	Veias da Palma da mão	(Equal Error Rate - EER) obtida foi de 0,49%	Tongji, PolyU, Xian Jiaotong University, (XJTU-UP),CASIA
Sistema Biométrico em Luz Visível	Não	Sim	Não	Não	Sim	Análise de Características das Mãos	(Equal Error Rate - EER) obtida foi de 0,52%	Autocoletada
Identificação de Veias com CNN	Sim	Sim	Sim	Sim	Não	Rede Neural Convolutacional	85% Dataset A / 90% Dataset B	CASIA e outro não especificado
Autenticação sem Contato com CNN	Sim	Sim	Sim	Sim	Não	Rede Neural Convolutacional	99,4% de precisão, EER média de 0,0683%	CASIA Multi-Spectral Palmprint
Reconhecimento de Veias com Deep Hashing	Sim	Não	Sim	Sim	Não	DHPN (Deep Hashing Network)	EER notável de 0,0222%	PolyU e XJTU Database:
Autenticação com Siamese Network	Sim	Sim	Sim	Sim	Não	Rede Neural Siamesa	eer 0.66% PolyU / 3,71% CASIA	CASIA e PolyU
Autenticação com Nó do Dedo	Não	Sim	Sim	Sim	Não	Processamento de Imagem e Aprendizado de Máquina	Taxa de reconhecimento correto de 99,24%, Taxa de erro de 0,78%	FKP

Tabela 1 Trabalhos disponíveis na literatura

Capítulo 4

Metodologia

Este capítulo apresenta um novo método de autenticação biométrica projetado para dispositivos móveis, com foco na análise das veias da palma da mão. A proposta integra técnicas de processamento de imagem e aprendizado profundo, visando oferecer uma solução robusta e acessível, especialmente em cenários sem a disponibilidade de sensores ou câmeras NIR (Near-Infrared), além de abordar limitações computacionais. Propomos um esquema de verificação de veias palmares utilizando um *Autoencoder* Profundo e uma Rede Neural Siamesa, que dispensa o uso de câmeras NIR. O fluxo de trabalho do modelo proposto, ilustrado na Figura 2, assume um esquema biométrico tradicional de veias palmares, mas sem a necessidade de sensores NIR. Para alcançar esse objetivo, a solução é composta por duas fases principais: Transformação NIR e Biometria baseada em Redes Neurais Profundas.

4.1 Visão Geral

Propomos um novo esquema viável para verificação de veias palmares utilizando um *Autoencoder* Profundo e uma Rede Siamesa para enfrentar esse desafio. O fluxo de trabalho do modelo proposto é ilustrado na Figura 2. O modelo proposto assume um esquema biométrico tradicional de veias palmares, implementado sem a necessidade de um sensor de câmera NIR. Para atingir esse objetivo, ele opera em duas fases principais: Transformação NIR e Biometria baseada em Redes Neurais Profundas (DNN).

Na primeira fase, o módulo de Transformação NIR lida com os desafios relacionados à aplicação de abordagens biométricas de veias palmares. Especificamente, o esquema proposto foi projetado para dispensar tanto o sensor de câmera NIR quanto a necessidade de acessórios especializados para garantir o posicionamento preciso da palma da mão. Essa abordagem realiza diversas transformações finas na imagem visível capturada para padronizar adequadamente a posição da palma. Em seguida, o espectro da imagem visível

pré-processada é transformado em um espectro próximo ao infravermelho usando um modelo de *Autoencoder* Profundo. Esse objetivo é alcançado ajustando cuidadosamente o *Autoencoder* para gerar imagens em infravermelho a partir do par visível correspondente. A ideia central é aproveitar a capacidade do *Autoencoder* Profundo para converter o espectro da imagem visível, viabilizando a biometria de veias palmares sem a necessidade de um sensor NIR.

Na segunda fase, o módulo DNN Biométrico avalia a similaridade do indivíduo a ser verificado para fins de autenticação. Para isso, o módulo utiliza uma Rede Siamesa que calcula a pontuação de similaridade entre a imagem de entrada e um banco de dados de indivíduos previamente autorizados. Como resultado, o modelo proposto consegue superar as limitações das estratégias biométricas de veias palmares convencionais.

As subseções a seguir descrevem em maior detalhe o funcionamento do modelo proposto, incluindo os módulos que o compõem.

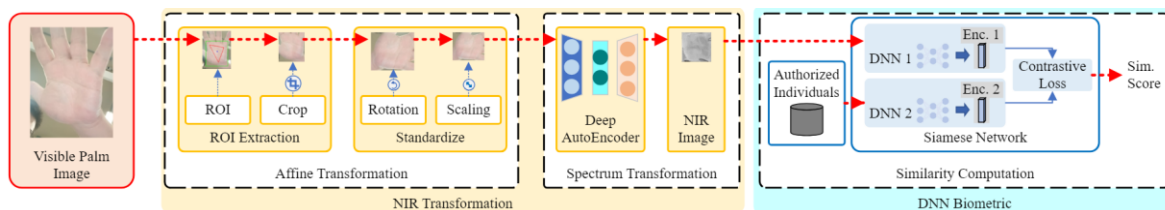


Figura 2 - Fluxograma geral

A Figura 2 representa de forma esquemática o fluxo do método proposto para autenticação biométrica em dispositivos móveis.

No processo de Cadastro do usuário, o primeiro passo consiste na aquisição de dados biométricos, onde imagens das veias da palma da mão são coletadas do usuário. Essas imagens passam por um processo de pré-processamento para melhorar a qualidade e facilitar a extração de características biométricas. Em seguida, é aplicada uma transformação espectrográfica, utilizando um *autoencoder*, para converter as imagens visíveis para o espectro infravermelho. Essa abordagem visa contornar a necessidade de sensores físicos infravermelhos, tornando o método mais prático para dispositivos móveis convencionais. As imagens processadas são então salvas na memória do dispositivo para referência futura.

No processo de Autenticação do usuário, durante a autenticação, novas imagens das veias da palma da mão são adquiridas diretamente do smartphone do usuário, juntamente com a imagem previamente armazenada na memória do dispositivo durante o cadastro. As imagens são submetidas a um processo de pré-processamento semelhante ao do cadastro, seguido pela transformação espectrográfica para convertê-las para o espectro infravermelho. A etapa crucial seguinte envolve o uso de uma rede siamesa, que calcula a dissimilaridade entre as características extraídas das veias da palma da mão. Esse cálculo de dissimilaridade é comparado a um limiar de autenticação para determinar a autenticidade do usuário. Essa abordagem oferece uma representação comparável das duas entradas, contribuindo para uma autenticação prática em dispositivos móveis.

4.2 Estrutura aquisição de dados

As imagens para autenticação podem ser capturadas por uma câmera de smartphone convencional, eliminando a necessidade de hardware especializado. No entanto, é importante observar que os resultados desse processo podem variar conforme a qualidade da câmera e as condições de iluminação, ainda permitindo a autenticação em ambientes cotidianos com iluminação básica.

Questões potenciais, como a rotação da mão, são mitigadas durante as etapas subsequentes de pré-processamento e extração da Região de Interesse (ROI). Este refinamento ocorre de maneira automática, descartando possíveis distorções na imagem que poderiam comprometer a precisão do processo de autenticação. Essa abordagem destaca a robustez e a adaptabilidade do método, contribuindo para sua eficácia em diferentes contextos e cenários do mundo real.

4.3 Pré-processamento e extração da ROI

As abordagens atuais de verificação de veias palmares na literatura geralmente assumem a disponibilidade de um sensor de câmera NIR e o uso de acessórios para garantir o posicionamento adequado da mão do indivíduo. Como resultado, essas abordagens raramente são usadas em ambientes reais devido aos requisitos de implantação pouco práticos.

Para enfrentar esse desafio, o modelo proposto realiza várias tarefas de pré-processamento na imagem coletada. O procedimento geral é ilustrado na Figura 2. Ele começa com a captura de uma imagem visível da palma da mão do indivíduo a ser verificado (Fig. 1, Imagem Visível da Palma). Essa tarefa pode ser realizada usando uma câmera de espectro visível convencional, sem a necessidade de equipamentos para posicionamento da mão, o que facilita a implantação do esquema.

Para garantir que o modelo possa processar adequadamente a imagem coletada, a imagem é usada como entrada no módulo de Transformação Afim (Fig. 1, Transformação Afim). O módulo visa padronizar a imagem de forma adequada para seguir o formato esperado para fins biométricos, implementando duas principais tarefas de pré-processamento: Extração da Região de Interesse (ROI) e Padronização.

Extração da ROI: A imagem da palma coletada do indivíduo nem sempre estará posicionada na mesma região. Pelo contrário, a localização e a rotação da palma podem variar significativamente com base em como o indivíduo posiciona a mão. Para lidar com essa questão, primeiro identificamos a localização da palma. Em seguida, recortamos o centro da palma, gerando uma nova imagem composta apenas pela palma. Esse processo visa garantir que a imagem resultante não contenha nenhuma informação além da palma do indivíduo a ser verificado.

Padronização: A imagem recortada é usada como entrada para o módulo de padronização, que assegura que a palma esteja alinhada quanto à rotação e ao tamanho esperados. Especificamente, primeiro rotacionamos a imagem para padronizar o posicionamento da palma, garantindo que o polegar do indivíduo esteja sempre orientado para a parte superior da imagem. A imagem rotacionada é então escalada para compensar variações na distância entre a palma e a câmera, que podem afetar o tamanho da imagem.

A Figura 3 representa o processo de extração de ROI.

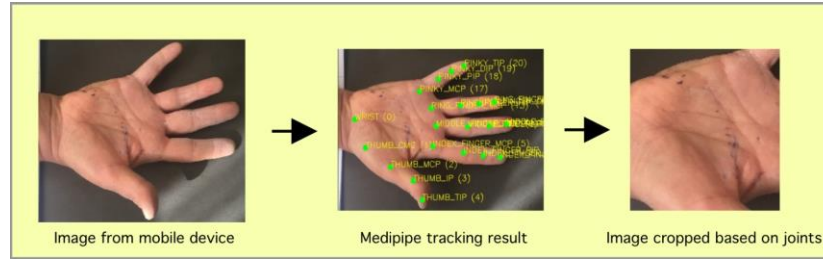


Figura 3 - Pré-processamento e extração da ROI

4.4 Mudança espectral

Como resultado, a etapa anterior produz uma imagem pré-processada e padronizada, levando em consideração variações na fase de aquisição da imagem, incluindo posicionamento inadequado da palma, distância e rotação. Essa característica permite que o modelo proposto opere sem a necessidade de um acessório para posicionamento da mão, uma vez que consideramos essas variações. A fase subsequente lida com a transformação espectral, que converte uma imagem em espectro visível para sua contraparte infravermelha. Para alcançar esse objetivo, o esquema proposto utiliza um *Autoencoder* Profundo.

Dado um *Autoencoder* $f(x^{vis}): x^{vis} \rightarrow x^{\widetilde{nir}}$ que gera a imagem no espectro NIR $x^{\widetilde{nir}}$ com base em uma imagem no espectro visível x^{vis} , sendo $x^{vis} \in R^{n \times n \times 3}$ e $x^{\widetilde{nir}} \in R^{n \times n \times 3}$, onde n corresponde ao tamanho da imagem. O objetivo do *autoencoder* profundo é gerar a imagem no espectro NIR $x^{\widetilde{nir}}$ com base na sua contraparte visível. Para atingir esse objetivo, assumimos a disponibilidade de uma tupla x^{vis} e x^{nir} , coletada sob as mesmas condições ambientais, incluindo distância da câmera, rotação e ângulo. A tupla pode ser utilizada para ajustar o modelo de *Autoencoder* Profundo para gerar imagens NIR com base em suas contrapartes visíveis. Assim, podemos calcular a perda do modelo conforme a seguinte equação

$$Loss(x^{\widetilde{nir}}, x^{nir}) = \sqrt{\frac{1}{n \times n} \sum_{i=1}^n \sum_{j=1}^n (x_{i,j}^{\widetilde{nir}} - x_{i,j}^{nir})^2} \quad (1)$$

Para facilitar a compreensão, o método proposto foi esquematizado no Pseudocódigo 1, denota-se o fluxo sequencial dos módulos. O algoritmo a seguir processa imagens da palma da mão utilizando um *autoencoder* para realçar características relevantes para a biometria. Inicialmente, a imagem da palma é capturada e uma Região de Interesse (ROI) é extraída

para focar na área específica da palma. Essa imagem recortada é então processada pelo *autoencoder*, que gera uma versão em infravermelho destacando detalhes importantes. Por fim, a imagem resultante é apresentada ou salva, completando o fluxo de processamento necessário para o reconhecimento biométrico.

1	<i>algoritmo autoencoder_palma_mao</i>
2	<i># Definindo estruturas de dados</i>
3	
4	<i>var</i>
5	<i>imagem_original,imagem_infravermelho: estrutura_imagem</i>
6	<i>autoencoder: estrutura_autoencoder</i>
7	
8	
9	<i># Função principal para processar a imagem usando o autoencoder</i>
10	
11	
12	<i>função processar_imagem_palma_mao()</i>
13	<i># Inicializar o autoencoder</i>
14	<i>autoencoder ← inicializar_autoencoder()</i>
15	
16	
17	<i># Receber imagem da palma da mão do smartphone</i>
18	<i>imagem_original ← receber_imagem_palma_mao()</i>
19	
20	
21	<i># Aplicar ROI para extrair a região da palma da mão</i>
22	<i>imagem_roi ← aplicar_roi(imagem_original)</i>
23	
24	
25	
26	<i># Aplicar autoencoder à imagem da palma da mão</i>
27	<i>imagem_infravermelho←aplicar_autoencoder(imagem_roi)</i>
28	
29	
30	<i># Apresentar ou salvar a imagem resultante em infravermelho</i>
31	<i>apresentar_imagem(imagem_infravermelho)</i>
32	
33	<i>fim_função</i>
34	<i># Chamar a função principal</i>
35	<i>processar_imagem_palma_mao()</i>
36	
37	<i>fim_algoritmo</i>

Algoritmo 1 Pseudocódigo Autoencoder

4.5 Rede siamesa e Comparação

Finalmente, a imagem NIR gerada pode ser utilizada para fins biométricos. O modelo proposto utiliza uma rede Siamesa para a tarefa de verificação biométrica. Na prática, assumimos a utilização de um banco de dados de indivíduos previamente autorizados (Fig. 1, Indivíduos Autorizados) que devem ser aceitos pelo sistema. A rede Siamesa é então usada para calcular as similaridades entre o indivíduo a ser verificado e os indivíduos previamente autorizados.

Seja \tilde{x}^{nir} a imagem da palma da mão do indivíduo a ser verificado e D um banco de dados de indivíduos previamente autorizados. O objetivo da rede Siamesa é calcular uma pontuação de similaridade entre \tilde{x}^{nir} e cada imagem em D .

Para alcançar esse objetivo, podemos ajustar a rede Siamesa para otimizar os vetores codificados gerados através da função de perda contrastiva conforme a seguinte equação:

$$\sum_{i=1}^N (y_i d_i^2 + (1 - y_i) \max(m - d_i, 0)^2) \quad (2)$$

Onde y_i indica se o par avaliado é similar ou dissimilar, d_i é a distância de cada vetor de codificação na posição i (Fig. 1, Enc. 1 e 2), e m é a margem. Assim, na fase de treinamento, a rede Siamesa é otimizada para minimizar a distância entre pares similares e aumentar a distância entre amostras dissimilares.

Na fase de inferência, utilizamos a rede Siamesa resultante para calcular a pontuação de similaridade em relação a todas as amostras de imagem de D . A pontuação pode ser classificada e usada para encontrar amostras similares de indivíduos através da seguinte equação:

$$\operatorname{argmin}_{x \in \mathcal{D}} k \sqrt{\sum_{i=1}^n (Siamese(\tilde{x}_i^{nir}) - Siamese(x_i))^2} \quad (3)$$

onde a função Siamesa retorna o vetor codificado de tamanho n a partir da imagem de entrada, e k denota o número dos principais indivíduos a serem avaliados. Finalmente, o indivíduo avaliado é autorizado se o k –ésimo indivíduo mais distante estiver abaixo de um limite de aceitação t .

Abaixo, apresenta-se um pseudocódigo simplificado que ilustra o funcionamento da rede siamesa. O algoritmo a seguir é projetado para processar duas imagens utilizando uma rede neural siamesa, com o objetivo de calcular a dissimilaridade entre elas e determinar se essa dissimilaridade ultrapassa um limiar predefinido. Primeiramente, a rede siamesa é inicializada, e as saídas do *autoencoder* e de um banco de dados são recebidas para serem comparadas. O algoritmo então processa essas saídas através da rede siamesa, calcula a dissimilaridade entre elas e a compara com um limiar definido (neste caso, 0,5). Se a dissimilaridade for maior que o limiar, o resultado é considerado verdadeiro; caso contrário, é falso. Por fim, o resultado é apresentado, concluindo o processo de verificação de dissimilaridade.

1	<i>algoritmo rede_siamesa</i>
2	<i>var</i>
3	<i>saida_rede_siamesa_1, saida_rede_siamesa_2, dissimilaridade, limiar: real</i>
4	
5	
6	<i># Função para processar a imagem pela rede siamesa e verificar dissimilaridade</i>
7	
8	
9	<i>função processar_e_verificar_dissimilaridade()</i>
10	<i># Inicializar a rede siamesa</i>
11	<i>inicializar_rede_siamesa()</i>
12	
13	
14	
15	<i># Receber saídas do autoencoder e do banco de dados</i>
16	<i>saida_rede_siamesa_1 ← receber_saida_autoencoder()</i>
17	<i>saida_rede_siamesa_2 ← receber_saida_bd()</i>
18	
19	
20	<i># Implementação do processamento pela rede siamesa</i>
21	<i># Calcular a dissimilaridade entre as saídas da rede siamesa</i>
22	<i>dissimilaridade ← calcular_dissimilaridade(saida_rede_siamesa_1,</i>
23	<i>saida_rede_siamesa_2)</i>
24	
25	
26	
27	<i># Definir um limiar</i>
28	<i>limiar ← 0.5</i>
29	
30	<i># Verificar se a dissimilaridade ultrapassa o limiar</i>
31	<i>resultado_final ← dissimilaridade > limiar</i>
32	
33	
34	<i># Apresentar o resultado final</i>
35	<i>escreva("Resultado Final: " + resultado_final)</i>
36	
37	
38	<i>fim_função</i>
39	
40	
41	<i># Chamar a função principal</i>
42	<i>processar_e_verificar_dissimilaridade()</i>
43	
44	
45	<i>fim_algoritmo</i>

Tabela 2 - Pseudocódigo Rede siamesa

4.6 Discussão

A eficácia e as implicações do novo método proposto para autenticação biométrica de veias palmares são avaliadas à luz da metodologia utilizada. O método, que combina um *Autoencoder* Profundo e uma Rede Neural Siamesa, foi concebido para superar as limitações dos sistemas tradicionais que dependem de sensores NIR e infraestrutura especializada. Durante a avaliação, observou-se que a substituição do sensor NIR por um modelo de transformação espectral, acoplado à funcionalidade das Redes Neurais Profundas, resulta em uma solução prática e aplicável em dispositivos móveis convencionais, sem a necessidade de hardware especializado.

A abordagem adotada, que envolve a pré-processamento das imagens para padronização e a transformação espectral para simular o espectro infravermelho, revelou-se eficaz ao manter altos níveis de precisão na autenticação. Isso demonstra que o modelo proposto não apenas reduz as barreiras de adoção em ambientes reais, mas também oferece uma alternativa viável e inovadora às tecnologias biométricas existentes.

Capítulo 5

Avaliação

Nesta seção, investigamos o desempenho do esquema proposto. Mais especificamente, buscamos responder as seguintes Questões de Pesquisa (RQs):

- RQ1: Qual é o desempenho do módulo de Transformação NIR?
- RQ2: Qual é o desempenho em termos de acurácia do esquema proposto com as imagens NIR transformadas?
- RQ3: Quais são os custos de processamento do protótipo implementado?
- RQ4: Qual a relação entre a quantidade de pessoas cadastradas no protótipo com o desempenho computacional?

As subseções a seguir descrevem detalhadamente o processo de construção do modelo e seu desempenho.

5.1 Protótipo

Implementamos um protótipo proposto considerando um cenário de implantação com recursos limitados, com requisitos mínimos em termos de processamento e capacidades de hardware.

Consideramos uma implantação que utiliza uma webcam de espectro visível tradicional para o processo de coleta de imagens da palma. O protótipo foi implementado em Python, coletando as imagens da webcam através de um módulo Image Collector utilizando a API OpenCV v.4.10. A amostra de imagem coletada é fornecida como entrada para o módulo ROI Extraction, que realiza as fases de Extração de ROI e Recorte usando a API MediaPipe Hands v.0.10.14. A imagem resultante é utilizada como entrada pelo módulo de Padronização, que realiza a rotação e escalonamento da palma usando a API OpenCV v.4.10.

Finalmente, implementamos nossa proposta de Autoencoder Profundo e uma Rede Siamesa utilizando a API Pytorch v.2.16.1. A Figura 4 ilustra a implementação do protótipo proposto.

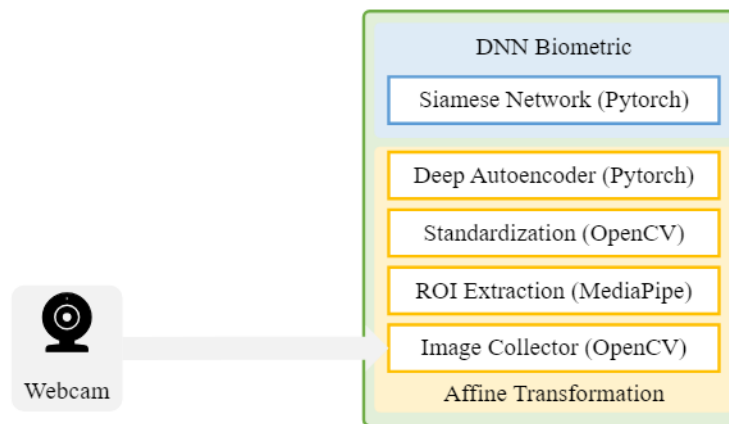


Figura 4- Visão geral

Portanto, nossa implementação de protótipo utiliza hardware tradicional. Essa abordagem assume uma webcam de espectro visível, eliminando a necessidade de um sensor de câmera NIR adicional para extrair as características das veias da palma do indivíduo, facilitando sua implantação.

5.2 Construção do modelo

Para nossa avaliação, utilizamos o conjunto de dados Tecnocampus Hand Image (Faundez-Zanuy et al., 2013). O conjunto de dados fornece imagens das palmas adquiridas simultaneamente nos espectros visível e NIR. Durante a coleta de dados, cada indivíduo coloca sua mão em um equipamento personalizado para garantir uma posição e distância consistentes para a coleta das imagens em ambos os espectros. O conjunto de dados é composto por um total de 1040 imagens para cada espectro, coletadas de 104 indivíduos únicos. Ele foi dividido em conjuntos de treinamento, teste e validação, compostos por 60%, 20% e 20% de indivíduos selecionados aleatoriamente, respectivamente. Essa configuração permite o uso eficaz do conjunto de dados para o treinamento do modelo de *Autoencoder* Profundo, dado a disponibilidade de pares de imagens visíveis e NIR.

O modelo de Autoencoder Profundo proposto (Figura 2) foi implementado e avaliado por meio da arquitetura Attention U-Net (Oktay et al., 2018) sobre o conjunto de dados pré-processado. O *Autoencoder* selecionado é amplamente utilizado para tarefas específicas de

regiões e pixel a pixel, como segmentação ou restauração de objetos, devido ao seu mecanismo de atenção, fornecendo, assim, resultados promissores para a tarefa de transformação espectral. Para reduzir o tempo de inferência de imagem, modificamos a arquitetura padrão do modelo. Reduzimos o decoder para diminuir o tempo de inferência, mantendo apenas uma camada convolucional por nível. Usamos um esquema de atenção multiplicativa em vez do padrão aditivo. Dada nossa arquitetura compacta, não utilizamos camadas de pooling, mas usamos a função de ativação LeakyReLU para evitar o problema de "neurônios mortos". O decoder possui uma configuração de 3 níveis com 32, 64 e 32 filtros, respectivamente. O modelo de *Autoencoder* Profundo foi treinado usando a equação tradicional de Erro Quadrático Médio (RMSE) para calcular a diferença entre a imagem NIR gerada e a imagem NIR de referência (veja Eq. 1).

Da mesma forma, a Rede Siamesa proposta, usada no módulo biométrico DNN (veja Fig. 1), foi implementada como uma rede dupla leve. A rede depende de uma série de camadas convolucionais, de ativação e de pooling máximo para atingir tal objetivo. As camadas convolucionais foram implementadas com 6 e 12 convoluções, com um tamanho de kernel de 3. Ambas as camadas convolucionais utilizam a função de ativação ReLU, seguidas por uma camada dropout de 0,25. Finalmente, a saída da última camada convolucional é achatada em um vetor de 128 neurônios com ativação ReLU e dropout de 0,1, seguido por uma camada de 50 neurônios com ativação ReLU. A rede siamesa gera um vetor codificado de tamanho 50, que é usado para calcular a similaridade entre duas imagens de entrada usando a distância Euclidiana. No geral, o modelo possui aproximadamente 266 mil parâmetros. A rede neural é treinada usando a função de perda contrastiva tradicional (veja Eq. 2).

O *Autoencoder* Profundo e a Rede Siamesa foram treinados usando o conjunto de treinamento e avaliados com o conjunto de teste. Finalmente, a acurácia do modelo resultante é medida no conjunto de validação. O treinamento do modelo foi executado por 1000 épocas com o otimizador Adam, com uma taxa de aprendizado inicial de 0,1 decaída por annealing cosseno. Usamos um peso de momento de 0,9 e um tamanho de lote de 16. O modelo proposto foi construído e avaliado na API TensorFlow, v.2.16.1.

5.3 Biometria de veias da palma

O primeiro experimento visa responder à RQ1 e investigar o desempenho do modelo de *Autoencoder* Profundo para transformar imagens do espectro visível em suas contrapartes NIR. Para alcançar tal objetivo, conduzimos a tarefa de treinamento do modelo de *Autoencoder* Profundo e avaliamos a taxa de erro de reconstrução nas imagens NIR geradas. Lembrando que realizamos a tarefa de treinamento do modelo utilizando as imagens visíveis como entrada do *autoencoder* profundo, e usamos suas contrapartes NIR de referência para avaliar o desempenho na geração de imagens.

A Figura 5 mostra a taxa de convergência do *Autoencoder* Profundo durante a fase de treinamento. O esquema de geração de imagens NIR efetivamente gera imagens NIR. Em termos práticos, o modelo proposto alcançou um valor médio de RMSE de 0,15 no conjunto de teste, com um erro máximo de 0,25 e um mínimo de 0,08. É importante notar que o conjunto de teste consiste em indivíduos que não foram incluídos nos conjuntos de treinamento ou validação. Portanto, os resultados obtidos destacam a eficácia da nossa abordagem proposta na geração de imagens NIR com base em suas contrapartes visíveis.

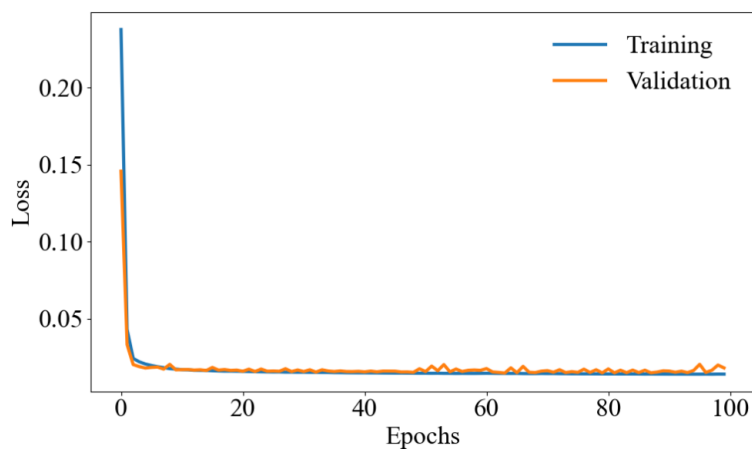


Figura 5 - Gráfico de perda durante o treinamento da rede autoencoder

A Figura 6 mostra exemplos de imagens NIR geradas usando o modelo de *Autoencoder* Profundo proposto. É possível observar que o esquema consegue gerar imagens NIR semelhantes às suas contrapartes NIR de referência. Especificamente, a maioria dos

padrões de veias presentes nas imagens NIR originais também está presente nas imagens geradas.

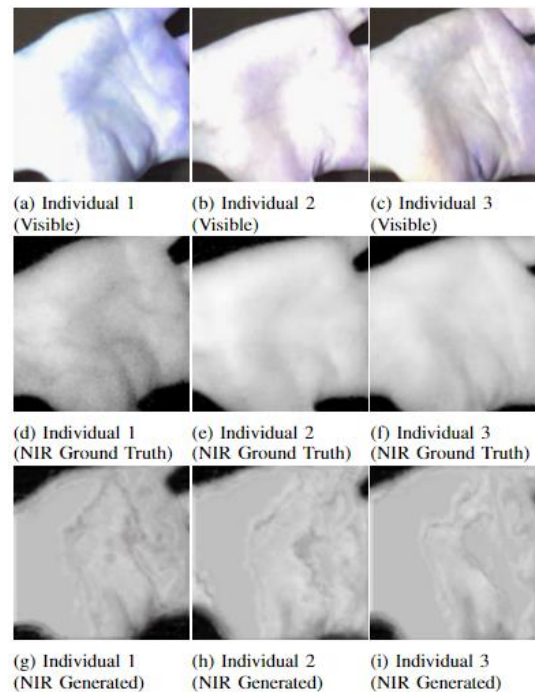


Figura 6 - Imagens NIR geradas de diferentes indivíduos.

O segundo experimento tem como objetivo responder à RQ2 e investiga o desempenho da nossa abordagem biométrica baseada em DNN usando as imagens geradas anteriormente. Para alcançar tal objetivo, realizamos a fase de treinamento da nossa rede siamesa considerando os 50 primeiros indivíduos como autorizados (Fig. 1, Indivíduos Autorizados) e os 51º ao 104º como não autorizados. Assumimos um limite de aceitação t de 0,5 para autorizar indivíduos com base em sua pontuação de similaridade e um valor de k de 5.

A Figura 7 mostra a taxa de convergência da rede siamesa para a verificação de usuários autorizados. Notavelmente, a rede converge usando as imagens geradas pelo modelo de *Autoencoder* Profundo. Isso demonstra a capacidade do modelo de gerar características representativas das imagens que podem ser usadas para autenticar usuários. Na prática, o esquema proposto pode superar a necessidade de um sensor de câmera NIR.

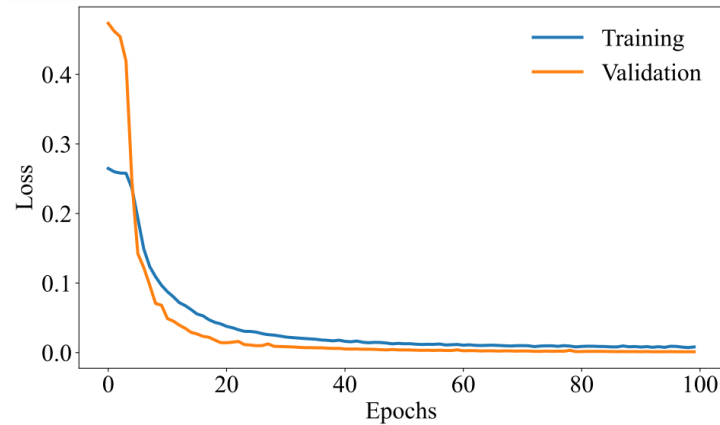


Figura 7 - Gráfico de perda durante o treinamento da rede siamesa

A Tabela 2 mostra o desempenho de verificação da nossa rede siamesa proposta ao usar as imagens geradas pelo *Autoencoder* Profundo em comparação com o uso das imagens NIR capturadas de forma tradicional. Primeiro, é possível notar que o esquema proposto fornece precisões de verificação significativamente altas. O modelo proposto alcança uma taxa de verdadeiros negativos de 0,98 usando as imagens geradas, ou seja, sem a necessidade do sensor de câmera NIR. Em segundo lugar, a acurácia alcançada é semelhante ao desempenho obtido pelas imagens NIR capturadas, com uma redução de apenas 0,01 na taxa de verdadeiros negativos. Esse resultado indica que o esquema proposto pode alcançar altas precisões comparáveis ao padrão que exige a utilização do sensor de câmera NIR.

Input	Acc. (%)	True Positive	True Negative
NIR Originais	99.12	1.00	0.98
NIR Geradas	98.79	1.00	0.97

Tabela 2 Comparação de precisão do modelo de acordo com o tipo de entrada usada.

A Tabela 3 apresenta a matriz de confusão do nosso modelo de verificação de veias da palma da mão, que representa o desempenho do sistema em termos de classificações corretas e incorretas. Esses números resultaram em métricas adicionais de desempenho que são cruciais para a avaliação da eficácia do modelo.

Matriz de Confusão		
Impostor	4873	121
Genuíno	0	5006
	Impostor	Genuíno

Tabela 3 Matriz de Confusão

O F1 Score, que mede a precisão e a sensibilidade do modelo, foi calculado em 0,9881, indicando um equilíbrio favorável entre precisão e recall. O Recall (ou taxa de verdadeiro positivo) atingiu 1,0000 o que significa que não houve erros na identificação de imagens positivas, refletindo uma sensibilidade excepcional do modelo. Por fim, a acurácia geral do sistema foi de 0,9879, evidenciando que aproximadamente 98,79% das classificações foram corretas, solidificando a eficácia da abordagem proposta.

O último experimento visa responder à RQ3 e investiga os requisitos de hardware do modelo proposto. Para alcançar tal objetivo, avaliamos os custos computacionais do protótipo proposto em ambientes Desktop e Raspberry Pi. O ambiente Desktop está equipado com uma CPU Intel i5 de 6 núcleos, 16 GB de memória e uma GPU Nvidia RTX3060 rodando no Ubuntu Linux 24.04. O ambiente Raspberry é um Raspberry Pi 3 Model B, com uma CPU Broadcom de 4 núcleos e 1 GB de memória rodando no Raspberry Pi OS com kernel versão 6.1. O objetivo é investigar mais a viabilidade da implementação das arquiteturas DNN selecionadas em um ambiente com recursos limitados.

A Tabela 4 mostra o tempo médio de processamento de imagens do protótipo em cada ambiente avaliado. O tempo de processamento por evento para o *Autoencoder* Profundo inclui nossas propostas de transformações afins e transformação espectral. Em contraste, o tempo da Rede Siamesa consiste no cálculo de similaridade por usuário autorizado (veja Fig. 1). É possível observar que o esquema proposto demanda, em média, 566 e 36,3 milissegundos por imagem a ser verificada quando executado no Raspberry Pi e na CPU do Desktop, respectivamente. Na prática, o esquema proposto pode autenticar 10 usuários autorizados em menos de 3 segundos considerando um ambiente com recursos computacionais altos e menos de 1 minuto para autenticar 10 usuários considerando um ambiente com recursos limitados. Isso é alcançado sem exigir capacidades de hardware adicionais, como uma GPU dedicada, equipamento para posicionamento da palma e um

sensor de câmera NIR. Como resultado, o modelo proposto abre caminho para a implementação de abordagens biométricas baseadas em veias da palma.

Network	Raspberry Pi (ms)	Desktop GPU (ms)
Rede Autoencoder	180	360
Rede Siamesa	380	1

Tabela 4 Comparação do tempo de execução do modelo no Raspberry Pi e no Desktop GPU.

Durante o processo de autenticação, a comparação da imagem leva, em média, 0,38 segundos com a Rede Siamesa rodando no Raspberry Pi. Para um conjunto de 100 imagens, o tempo total de processamento foi de 1 minuto e 10 segundos.

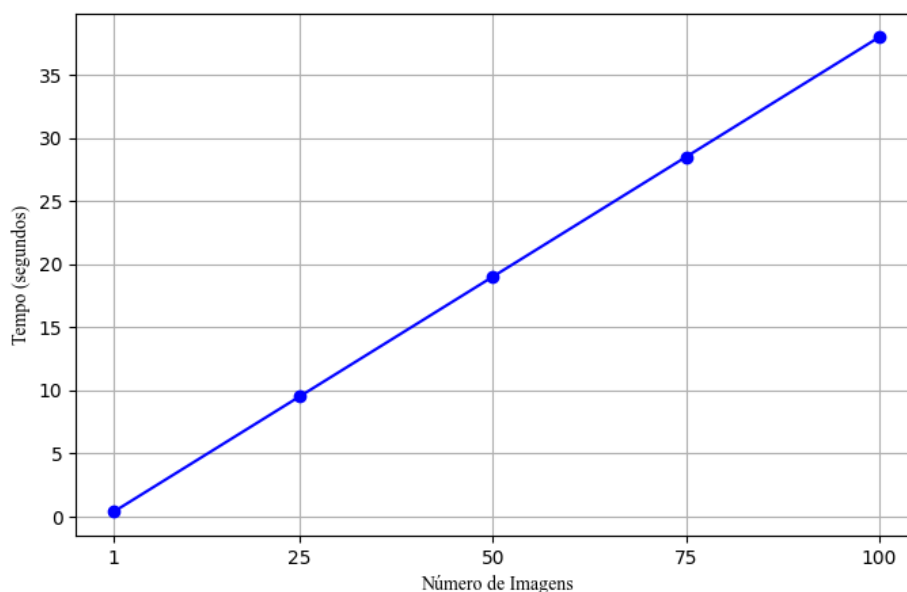


Figura 8: Tempo de comparação em segundos em função do número de imagens processadas.

A figura 8 mostra como o tempo de comparação aumenta proporcionalmente com o número de imagens, com valores de 1, 25, 50, 75 e 100 imagens. Essa relação linear entre o número de imagens e o tempo total de processamento é esperada em sistemas como este, onde cada imagem é processada individualmente em sequência. Não há degradação significativa do tempo de comparação por imagem conforme o número total de imagens aumenta, o que indica que o Raspberry Pi lida bem com a carga de trabalho sem sobrecargas substanciais. Isso sugere que o sistema é escalável, dentro das limitações do tempo de processamento por imagem, para conjuntos maiores de imagens.

Capítulo 6

Conclusão

Neste trabalho, apresentamos um novo esquema de identificação de veias da palma da mão que utiliza um Deep Autoencoder e uma rede Siamesa. O framework proposto tem como objetivo reduzir os requisitos de implantação das técnicas atuais de identificação, realizando a transformação do espectro de imagem por meio de um Deep Autoencoder leve que surge como uma alternativa ao uso de sensores infravermelhos físicos. As imagens RGB das veias da palma são transformadas para suas versões no espectro NIR (infravermelho próximo) através do treinamento de um modelo de Deep Autoencoder. Um desafio encontrado foi a necessidade de pares de imagens visíveis e infravermelhas para treinar o autoencoder, o que limitou a escolha dos datasets. No entanto, esse obstáculo foi superado ao utilizar um conjunto de dados específico e a pós-validação com outros datasets, garantindo a eficiência do modelo proposto. As imagens NIR geradas das veias da palma são então verificadas por uma rede Siamesa para avaliar a similaridade com um banco de dados de usuários previamente autorizados. O desempenho do esquema foi avaliado utilizando um conjunto de dados publicamente disponível, que confirmou a viabilidade da proposta. O esquema pode alcançar taxas de acurácia semelhantes sem a necessidade de um sensor de câmera NIR, permitindo sua implementação também em dispositivos com recursos limitados. Em trabalhos futuros, planejamos investigar possíveis problemas que possam surgir no protótipo funcional em um ambiente de produção real. Além disso, nossa intenção é aprimorar a eficiência e a transparência do modelo, garantindo que ele seja aplicável em cenários de larga escala sem comprometer a qualidade da verificação biométrica. Para isso, iremos integrar técnicas de Explicabilidade em Inteligência Artificial (XAI), tornando o modelo mais interpretável em relação à tomada de decisão e aumentando a transparência e a confiança nos resultados obtidos.

Referências bibliográficas

RUI, Z.; YAN, Z. A Survey on Biometric Authentication: Toward Secure and Privacy-Preserving Identification. *IEEE Access*, v. 7, p. 5994–6009, 2019.

ADIRAJU, R. V. et al. An extensive survey on finger and palm vein recognition system. *Materials Today: Proceedings*, out. 2020.

YANG, L. et al. Finger Vein Recognition With Anatomy Structure Analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 28, n. 8, p. 1892–1905, ago. 2018.

GUO, H. et al. Fooling A Deep-Learning Based Gait Behavioral Biometric System. 1 maio 2020.

RASTOGI, S. et al. NIR Palm Vein Pattern Recognition. 2020 IEEE International Conference for Innovation in Technology (INOCON), 6 nov. 2020.

ZHOU, Y.; KUMAR, A. Human Identification Using Palm-Vein Images. *IEEE Transactions on Information Forensics and Security*, v. 6, n. 4, p. 1259–1274, dez. 2011.

Ajmire, P. E. et al. (2018). *Human Identification using Palm Vein and Image Processing Techniques.* *International Journal of Emerging Technologies in Engineering Research*, 6(3), 123-127, 2018.

CHO, S. et al. Extraction and Cross-Matching of Palm-Vein and Palmprint From the RGB and the NIR Spectrums for Identity Verification. *IEEE Access*, v. 8, p. 4005–4021, 1 jan. 2020.

BARRA, S. et al. A hand-based biometric system in visible light for mobile environments. *Information Sciences*, v. 479, p. 472–485, abr. 2019.

JIA, W. et al. A survey on dorsal hand vein biometrics. *Pattern Recognition*, v. 120, p. 108122, dez. 2021.

HORNG, S.-J. et al. Recognizing Palm Vein in Smartphones Using RGB Images. *IEEE Transactions on Industrial Informatics*, v. 18, n. 9, p. 5992–6002, set. 2022.

SHINZAKI, T. Use Case of Palm Vein Authentication. *Handbook of Vascular Biometrics*, p. 145–158, 14 nov. 2019.

Palm Vein Pattern Authentication Technology Whitepaper. [s.l: s.n.]. Disponível em: <<https://www.fujitsu.com/ca/en/imagesgig5/palmsecure.pdf>>. Acesso em: 8 out. 2024.

PASQUINI, C. Near Infrared Spectroscopy: fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, v. 14, n. 2, p. 198–219, 1 abr. 2003.

KRIZHEVSKY, A. Learning Multiple Layers of Features from Tiny Images. [s.l: s.n.]. Disponível em: <<https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf>>. .

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, v. 60, n. 6, p. 84–90, 24 maio 2012.

CIRESAN, D.; MEIER, U.; JÜRGEN SCHMIDHUBER. Multi-column Deep Neural Networks for Image Classification. 13 fev. 2012.

STRIGL, D.; KOFLER, K.; PODLIPNIG, S. Performance and Scalability of GPU-Based Convolutional Neural Networks. Disponível em:<<https://ieeexplore.ieee.org/document/5452452>>. Acesso em: 23 jan. 2023.

CS231n Convolutional Neural Networks for Visual Recognition. Disponível em:
<<http://cs231n.github.io/>>.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep Learning. *Nature*, v. 521, n. 7553, p. 436–444, maio 2015.

SCHMIDHUBER, J. Deep learning in neural networks: An overview. *Neural Networks*, v. 61, n. 61, p. 85–117, jan. 2015.

SEOK, C. L. et al. Photoplethysmogram Biometric Authentication Using a 1D Siamese Network. *Sensors*, v. 23, n. 10, p. 4634, 1 jan. 2023.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Learning. Disponível em:
<<https://www.deeplearningbook.org/>>.

Kotsiantis, S., Zaharakis, I., Pintelas, P. (2007). *Supervised machine learning: A review of classification techniques.* *Emerging Artificial Intelligence Applications in Computer Engineering,* pp. 3-24.

LIEN, C.-H.; SUDIP VHADURI. Challenges and Opportunities of Biometric User Authentication in the Age of IoT: A Survey. *ACM Computing Surveys*, v. 56, n. 1, p. 1–37, 25 ago. 2023.

MINAEE, S. et al. Biometrics Recognition Using Deep Learning: A Survey. *arxiv.org*, 30 nov. 2019.

CHO, S. et al. Palm-Vein Verification Using Images From the Visible Spectrum. *IEEE Access*, v. 9, p. 86914–86927, 2021.

SCHROFF, F.; KALENICHENKO, D.; PHILBIN, J. FaceNet: A unified embedding for face recognition and clustering. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), jun. 2015.

BENRADI, H.; CHATER, A.; LASFAR, A. A hybrid approach for face recognition using a convolutional neural network combined with feature extraction techniques. IAES International Journal of Artificial Intelligence (IJ-AI), v. 12, n. 2, p. 627, 1 jun. 2023.

JAN NIKLAS KOLF et al. How Colorful Should Faces Be? Harmonizing Color and Model Quantization for Resource-restricted Face Recognition. 25 set. 2023.

TOME, P.; MARCEL, S. On the vulnerability of palm vein recognition to spoofing attacks. 2015 International Conference on Biometrics (ICB), maio 2015.

QIN, H. et al. Multi-Scale and Multi-Direction GAN for CNN-Based Single Palm-Vein Identification. IEEE Transactions on Information Forensics and Security, v. 16, p. 2652–2666, 2021.

BABALOLA, F. O.; BITIRIM, Y.; TOYGAR, Ö. Palm vein recognition through fusion of texture-based and CNN-based methods. Signal, Image and Video Processing, 18 ago. 2020.

ZHONG, D. et al. Palm Vein Recognition with Deep Hashing Network. Lecture notes in computer science, p. 38–49, 1 jan. 2018.

NAYAR, G. R.; THOMAS, T. Partial palm vein based biometric authentication. Journal of Information Security and Applications, v. 72, p. 103390, fev. 2023.

JAMEEL, S. K.; MAJIDPOUR, J. Generating Spectrum Images from Different Types — Visible, Thermal, and Infrared Based on Autoencoder Architecture (GVTI-AE). International Journal of Image and Graphics, p. 2250005, 10 abr. 2021.

JOSHI, J. C. et al. Finger Knuckleprint Based Personal Authentication Using Siamese Network. 1 mar. 2019.

JASWAL, G.; NIGAM, A.; NATH, R. Finger knuckle image based personal authentication using DeepMatching. 1 fev. 2017.

KUMAR, A.; RAVIKANTH, CH. Personal Authentication Using Finger Knuckle Surface. IEEE Transactions on Information Forensics and Security, v. 4, n. 1, p. 98–110, mar. 2009.

WU, H. et al. Face recognition based on convolution siamese networks. 1 out. 2017.

THAPAR, D. et al. PVSNet: Palm Vein Authentication Siamese Network Trained using Triplet Loss and Adaptive Hard Mining by Learning Enforced Domain Specific Features. arXiv (Cornell University), 1 jan. 2019.

FAUNDEZ-ZANUY, M.; MEKYSKA, J.; FONT-ARAGONÈS, X. A New Hand Image Database Simultaneously Acquired in Visible, Near-Infrared and Thermal Spectrums. Cognitive Computation, v. 6, n. 2, p. 230–240, 13 out. 2013.

OKTAY, O. et al. Attention U-Net: Learning Where to Look for the Pancreas. Disponível em: <<https://arxiv.org/abs/1804.03999v3>>.