

Predição Estruturada de Posições Anatômicas em Conjuntos Decadactilares

Saulo G. de Matos Silva^{1,2} , Glauco V. Pedrosa² 

¹ Instituto Nacional de Identificação, Polícia Federal – Brasília, DF – Brasil

² Programa de Pós-Graduação em Computação Aplicada (PPCA)
Universidade de Brasília (UnB) – Brasília, DF – Brasil

saulo.silva.research@outlook.com, glauco.pedrosa@unb.br

Abstract. *Accurate finger position classification from fingerprint images plays a critical role in civil biometric identification systems. However, missing or inconsistent labels significantly degrade indexing performance. Existing methods classify fingerprints independently, ignoring that each position appears exactly once in a tenprint set. We propose a method integrating Convolutional Neural Networks with linear assignment optimization to compute a globally consistent assignment across ten fingerprints. Evaluated on the SOCOFing dataset using multiple deep learning architectures, independent classification yields poor subject-level consistency (13.3%). The proposed method improves this metric to 60.0%, mitigating the limitations of isolated classification.*

Resumo. *A classificação precisa da posição anatômica a partir de imagens de impressões digitais desempenha um papel crítico em sistemas biométricos de identificação civil. Contudo, rótulos ausentes ou inconsistentes degradam severamente a indexação. Métodos atuais classificam as impressões isoladamente, ignorando a ocorrência única de cada posição no conjunto decadactilar. O método proposto neste trabalho integra Redes Neurais Convolucionais à otimização por atribuição linear para calcular uma atribuição globalmente consistente das dez impressões. Avaliada na base SOCOFing com múltiplas arquiteturas, a classificação independente produz baixa consistência por indivíduo (13,3%). O método proposto eleva essa métrica para 60,0%, mitigando as limitações da classificação isolada.*

1. Introdução

A identidade civil constitui um elemento essencial para a garantia de direitos, o acesso a serviços públicos e o pleno exercício da cidadania, sendo incorporado como uma meta global nos Objetivos de Desenvolvimento Sustentável (Meta 16.9) [United Nations (UN) 1948]. Nesse contexto, as impressões digitais consolidaram-se, ao longo de mais de um século, como o principal mecanismo biométrico para identificação civil, em virtude de seu baixo custo, caráter pouco invasivo e elevado grau de confiabilidade, sustentado pelas propriedades de unicidade e persistência ao longo do tempo [Jain et al. 2025].

Sistemas Automatizados de Identificação Biométrica (*Automated Biometric Identification Systems* – ABIS) desempenham papel central na gestão de bases nacionais de identificação civil em larga escala [Win et al. 2020]. Esses sistemas realizam

comparações entre milhões de impressões digitais e dependem de estratégias de indexação e filtragem para reduzir o espaço de busca. Entre esses filtros, a posição anatômica da impressão digital destaca-se como um dos mais relevantes, ao categorizar os dedos em dez rótulos distintos. Quando corretamente utilizada, essa informação reduz o espaço de busca em um fator de dez. Na ausência ou incorreção desse metadado, o ABIS precisa executar buscas sobre todas as posições possíveis, o que eleva significativamente o custo computacional, o tempo de resposta e, em cenários críticos, compromete a própria recuperabilidade das biometrias [Maltoni et al. 2022].

Tradicionalmente, a coleta decadactilar é realizada por meio da aquisição combinada de impressões digitais roladas e planas. As impressões roladas são coletadas individualmente com o objetivo de maximizar a área útil da imagem e, consequentemente, a quantidade de minúcias disponíveis para a comparação biométrica. As impressões planas, por sua vez, são adquiridas simultaneamente no padrão 4-4-2, garantindo a ordem dos dedos e fornecendo a referência de posição anatômica. A integração desses dois tipos de captura permite associar cada impressão rolada à sua respectiva posição.

Contudo, esse procedimento apresenta limitações operacionais. O uso de sensores de dedo único inviabiliza a coleta simultânea no padrão 4-4-2; processos manuais estão sujeitos a erros, como inversão das mãos ou captura fora de sequência; e digitalizações em larga escala enfrentam grande variabilidade de formulários, o que dificulta a extração automática da posição anatômica. Além disso, migrações de bases legadas e compartilhamentos massivos de dados frequentemente resultam em grandes volumes de imagens sem metadados de posição. Nessas condições, a ausência ou inconsistência dessa informação compromete a eficiência dos sistemas de identificação e pode inviabilizar a recuperação correta de conjuntos decadactilares, mesmo quando as biometrias estão disponíveis na base.

Para lidar com esse problema, diversos trabalhos propõem a classificação automática da posição anatômica diretamente a partir da imagem da impressão digital [Al-Wajih et al. 2022, Shehu et al. 2018a, Hsiao et al. 2022]. Redes Neurais Convolucionais (CNNs) têm sido amplamente empregadas nesse contexto, assim como em tarefas relacionadas à classificação de tipos fundamentais (ex., arcos, presilhas e verticilos), gênero e outras características biométricas. Entretanto, a maior parte da literatura formula a classificação da posição anatômica como uma tarefa de predição independente, processando cada impressão isoladamente.

No contexto da identificação civil, cada uma das dez posições deve ocorrer exatamente uma vez. Assumindo a independência entre as predições com acurácia p , a probabilidade de o conjunto estar correto é $P_{\text{conjunto}} = p^{10}$. Consequentemente, mesmo para $p = 0,8$, a acurácia do conjunto degrada para cerca de 10%, evidenciando a fragilidade da classificação isolada. Este trabalho modela explicitamente essa restrição global como um problema de otimização por atribuição, no qual a matriz de probabilidades produzida pelo classificador é utilizada como matriz de custo para determinar um mapeamento um-para-um entre impressões digitais e posições anatômicas.

Este artigo propõe uma abordagem híbrida, que integra Redes Neurais Convolucionais à Otimização por Atribuição Linear para a classificação da posição anatômica em nível decadactilar. Experimentos conduzidos na base pública SOCOFing demonstram

que a imposição explícita da restrição de unicidade por meio de otimização global mitiga sistematicamente as limitações da classificação isolada e produz ganhos expressivos de desempenho em nível de indivíduo, tornando a abordagem particularmente relevante para sistemas de identificação civil em larga escala.

2. Trabalhos Relacionados

A classificação automática de impressões digitais é considerada um problema desafiador de reconhecimento de padrões devido à pequena variação entre classes distintas e à grande variabilidade intraclasse [Maltoni et al. 2022]. Esta seção revisa o estado da arte na aplicação de CNNs, abrangendo desde a classificação de tipos fundamentais e atributos biométricos até a determinação específica da posição anatômica.

2.1. Classificação de Impressões Digitais com CNNs

Com o avanço recente das Redes Neurais Convolucionais (CNNs), observou-se um crescimento significativo de estudos aplicados à biometria por impressões digitais, abrangendo desde a classificação de tipos fundamentais até a inferência de atributos biométricos e características físicas.

O trabalho de [Michelsanti et al. 2017] utilizou as arquiteturas VGG-F e VGG-S com *transfer learning* e *data augmentation* para classificar a base NIST SD4 em 4 classes, alcançando acurácias de 94,4% e 95,05%, respectivamente. Embora essas redes possuam dezenas de milhões de parâmetros, [Jian et al. 2020] propuseram uma CNN leve com apenas 816 mil parâmetros, combinada a pré-processamentos para extração de regiões de interesse, atingindo 93% de acurácia em 5 classes na mesma base.

A metodologia detalhada em [Militello et al. 2021] avaliou as arquiteturas AlexNet, GoogLeNet e ResNet com *transfer learning* nas bases NIST SD4 e PolyU, considerando 4, 5 e 8 classes. Os resultados demonstraram forte dependência da qualidade do conjunto de dados, com acurácias de até 96,85% na NIST SD4 e superiores a 99,5% na PolyU.

Além dos tipos fundamentais, diversas pesquisas exploram a inferência de atributos biométricos a partir das impressões digitais. Na classificação por gênero, estudos anteriores [Shehu et al. 2018a, Hsiao et al. 2022] utilizaram ResNet34 e VGG16, respectivamente, obtendo acurácias em torno de 75% a 79%. O estudo em [Iloanusi and Ejiogu 2020] avançou ao demonstrar que a fusão de decisões de múltiplos dedos (polegar, médio e anelar) supera significativamente os classificadores unimodais, elevando a acurácia global para 91,3%.

Lidando com rótulos de posição, [Kim et al. 2020] abordaram a classificação de lateralidade (mão esquerda ou direita), obtendo 96,80% de acurácia com ResNet50. Estudos semelhantes foram conduzidos por [Shehu et al. 2018a] e [Hsiao et al. 2022], com resultados variando entre 93,5% e 94,4%. De forma mais simplificada, [Al-Wajih et al. 2022] propuseram uma CNN binária para distinguir polegar de não-polegar, atingindo cerca de 90% de acurácia nas bases NIST SD4 e SOCOFing.

2.2. Classificação da Posição Anatômica

A classificação automática das dez posições anatômicas constitui uma tarefa mais complexa, pois requer a distinção simultânea entre lateralidade e tipo de dedo. Em

[Shehu et al. 2018a], os autores utilizaram a base SOCOFing, composta por 6.000 imagens balanceadas (600 por posição), e aplicaram modelos ResNet com *transfer learning*, alcançando 76,72% de acurácia média em uma tarefa de 10 classes. Os autores observaram desempenho superior para os polegares (acima de 93%), enquanto os dedos centrais (indicador, médio e anelar) apresentaram maiores taxas de confusão, com acurácias próximas a 66%.

Outra pesquisa [Hsiao et al. 2022] utilizou uma base privada do Departamento de Polícia de Nova Taipé, composta por 10.000 imagens extraídas de 1.000 cartões decadaactilares. A arquitetura VGG16 obteve 84,8% de acurácia, superando significativamente o desempenho de especialistas humanos, que atingiram apenas 40%. Embora polegares e mínimos apresentem padrões morfológicos mais distintivos, as CNNs demonstraram capacidade superior de extração de características latentes dos dedos centrais.

Comparações diretas com parte da literatura são limitadas por divergências metodológicas. Alguns estudos restringem a classificação a cinco classes, desconsiderando a lateralidade exigida em sistemas de identificação [Giudice et al. 2020, Rim et al. 2021]. Outros trabalhos não descrevem adequadamente as estratégias de particionamento dos dados [Singh et al. 2021, Maiti et al. 2024] ou reportam métricas superestimadas obtidas no próprio conjunto de treinamento [Oleiwi et al. 2022, Pillai et al. 2024], comprometendo a validade dos resultados apresentados.

Apesar dos avanços na classificação da posição anatômica, a literatura existente formula a tarefa predominantemente como um problema de predição independente, no qual cada impressão é classificada isoladamente, sem a imposição explícita de restrições globais sobre o conjunto decadaactilar.

Recentemente, estratégias de “previsão seguida de otimização” têm sido empregadas para integrar modelos de *machine learning* a problemas de decisão com restrições combinatórias. Como um exemplo fora da biometria, modelos preditivos [Pessach et al. 2020] utilizaram probabilidades produzidas por classificadores para construir matrizes de custo destinadas à alocação ótima de candidatos a vagas de emprego. No contexto da biometria por impressões digitais, contudo, ainda são escassos os trabalhos que explorem formalmente a classificação da posição anatômica como um problema de predição estruturada. Essa lacuna motiva a abordagem proposta neste trabalho.

3. Método Proposto

Este trabalho propõe um método para a classificação da posição anatômica que trata o conjunto completo de dez dedos de um indivíduo como uma única unidade de decisão. Ao invés de classificar cada impressão de forma isolada, o método formula explicitamente o problema como uma tarefa de predição estruturada em nível de indivíduo.

O método integra um classificador probabilístico baseado em Redes Neurais Convolucionais (CNNs) a um modelo de otimização combinatória, impondo a restrição de unicidade inerente ao domínio biométrico, segundo a qual cada posição anatômica deve ocorrer exatamente uma vez em um conjunto válido de dez dedos.

A Figura 1 ilustra o fluxo tradicional de predição por imagem, no qual cada imagem é processada isoladamente e a decisão final resulta da escolha local da classe mais provável. Esse fluxo assume independência entre as predições e pode gerar atribuições

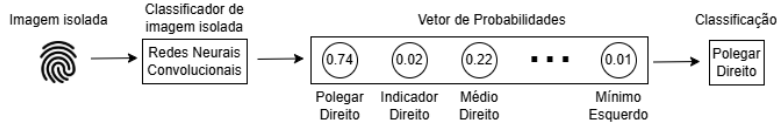


Figura 1. Fluxo de classificação independente, onde cada impressão digital é predita isoladamente.

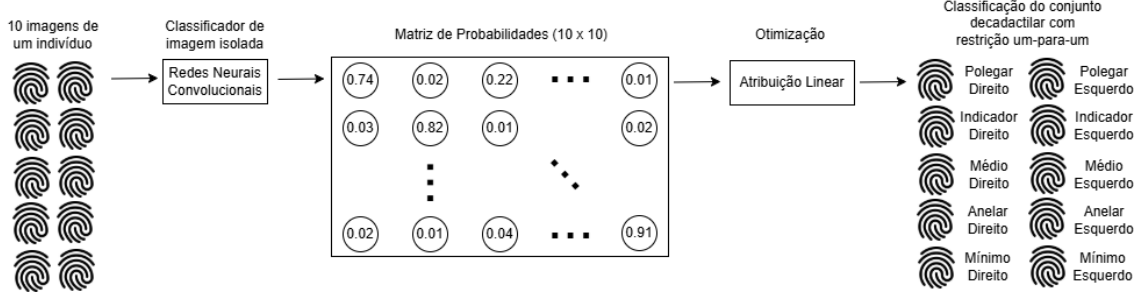


Figura 2. Fluxo do método proposto, integrando probabilidades de CNN com otimização global por atribuição.

inconsistentes quando aplicado ao conjunto completo de um indivíduo.

Em contraste, a Figura 2 apresenta o fluxo do método proposto, no qual as predições individuais são integradas por meio de uma etapa explícita de otimização, responsável por impor restrições globais de consistência.

3.1. Formulação do Problema

Seja $X = \{x_1, x_2, \dots, x_{10}\}$ o conjunto de imagens de impressões digitais pertencentes a um mesmo indivíduo, e seja $Y = \{1, 2, \dots, 10\}$ o conjunto das posições anatômicas possíveis. O objetivo consiste em determinar uma função de atribuição

$$f : X \rightarrow Y, \quad (1)$$

tal que f seja uma bijeção, garantindo que cada imagem seja atribuída a exatamente uma posição anatômica e que nenhuma posição seja repetida, assumindo que cada conjunto decadactilar completo contém exatamente uma ocorrência de cada posição anatômica.

Inicialmente, o método processa cada imagem $x_i \in X$ de forma independente por meio de uma CNN treinada para estimar a distribuição condicional

$$P(y \mid x_i), \quad y \in Y. \quad (2)$$

Essa etapa produz, para cada imagem, um vetor de probabilidades sobre as dez classes possíveis. A agregação das saídas do classificador para as dez imagens resulta em uma matriz de probabilidades

$$\mathbf{P} = [p_{ij}] \in \mathbb{R}^{10 \times 10}, \quad (3)$$

na qual $p_{ij} = P(y = j \mid x_i)$.

A Figura 2 apresenta um exemplo de matriz de probabilidades, onde a codificação por cores evidencia o grau de confiança do classificador. Embora essa matriz contenha evidências discriminativas relevantes, ela não assegura, por si só, a consistência global do conjunto, uma vez que múltiplas imagens podem apresentar alta probabilidade para a mesma classe de posição anatômica.

3.2. Otimização por Atribuição Linear

Para impor a restrição de unicidade, o método proposto define uma matriz binária de decisão $\mathbf{A} = [A_{ij}] \in \{0, 1\}^{10 \times 10}$, na qual $A_{ij} = 1$ indica a atribuição da imagem x_i à posição j . O problema de otimização consiste em maximizar a função objetivo:

$$\max_{\mathbf{A}} \sum_{i=1}^{10} \sum_{j=1}^{10} A_{ij} \cdot C_{ij} \quad (4)$$

sujeita às restrições

$$\sum_{j=1}^{10} A_{ij} = 1, \quad \forall i \in \{1, \dots, 10\}, \quad (5)$$

$$\sum_{i=1}^{10} A_{ij} = 1, \quad \forall j \in \{1, \dots, 10\}. \quad (6)$$

Essas restrições garantem uma correspondência um-para-um entre imagens e posições anatômicas, assegurando a validade decadactilar da solução.

A matriz de custos $\mathbf{C} \in \mathbb{R}^{10 \times 10}$ é derivada diretamente da matriz de probabilidades $\mathbf{P}$, tal que $C_{ij} = p_{ij}$. Conforme ilustrado na Figura 2, a solução seleciona exatamente uma célula por linha e por coluna, redistribuindo rótulos de forma coerente. Ao integrar classificação probabilística e otimização combinatória, o método supera as limitações da predição por imagem (Figura 1), atenuando a propagação de erros e garantindo consistência em nível de indivíduo.

4. Configuração Experimental

Esta seção detalha a configuração experimental. Primeiro, descrevemos os conjuntos de dados SOCOFing e o protocolo de particionamento. Em seguida, apresentamos a estratégia de treinamento das CNNs, a implementação da otimização e os recursos computacionais utilizados.

4.1. Conjuntos de Dados e Particionamento

O conjunto de dados SOCOFing consolidou-se como uma referência em pesquisas sobre classificação de impressões digitais por disponibilizar, de forma aberta, rótulos associados a impressões alteradas, gênero, mão, tipo de dedo e posição anatômica. Utilizamos o subconjunto denominado “Reais”, que contém 6.000 imagens de 600 indivíduos africanos [Shehu et al. 2018b], conforme ilustrado na Figura 3.

Os responsáveis pelo conjunto coletaram as imagens com sensores ópticos Hamster Plus (HSDU03P) e SecuGen SDU03P, originalmente configurados para resolução de 500 ppi e tamanho de 260×300 pixels. Entretanto, as imagens disponibilizadas no SOCOFing encontram-se no formato BMP, apresentam dimensões reduzidas de 96×103 pixels e resolução aproximada de 200 ppi, indicando a aplicação de processos de recorte e redução de resolução não documentados.

O conjunto de dados foi particionado utilizando um protocolo disjuncto por indivíduo nas proporções de 70% (treinamento: 420 indivíduos, 4.200 imagens), 15%

(validação: 90 indivíduos, 900 imagens) e 15% (teste: 90 indivíduos, 900 imagens). Essa estratégia evita qualquer sobreposição entre treinamento, validação e teste. O subconjunto de treinamento foi utilizado para o aprendizado do modelo, a validação para definir hiperparâmetros e arquiteturas, e o teste reservado exclusivamente para avaliação final, garantindo a análise da capacidade de generalização do método.



Figura 3. Amostra de um conjunto decadactilar completo de um único indivíduo da base SOCOFing.

4.2. Estratégia de Treinamento

Os classificadores foram treinados utilizando *transfer learning* [Yosinski et al. 2014] com base em modelos pré-treinados no conjunto ImageNet. Quatro arquiteturas foram selecionadas: VGG16 [Simonyan and Zisserman 2014], ResNet50 [He et al. 2016], DenseNet121 [Huang et al. 2017] e EfficientNetB0 [Tan and Le 2019]. A adaptação nos modelos restringiu-se à substituição da camada densa final, originalmente projetada para 1000 classes (ImageNet), por uma nova camada de 10 classes, destinada às posições anatômicas.

A configuração de treinamento adotada consistiu em uma estratégia em duas fases: treinamento apenas da última camada (20 épocas, *learning rate* de $1e-3$) e *fine-tuning* completo (30 épocas, *learning rate* de $1e-4$). Utilizamos *batch size* de 64, otimizador AdamW [Loshchilov and Hutter 2017] com *weight decay* de $1e-4$, *Early Stopping* [Goodfellow et al. 2016], função de perda *Cross-Entropy* e *data augmentation* com corte, rotação, translação, cisalhamento e supressão aleatórios [Zhong et al. 2020].

4.3. Implementação da Otimização

Após a classificação pelas CNNs, foi implementado um pós-processamento para garantir que, no conjunto de 10 imagens de um indivíduo, cada imagem receba uma classificação sem repetição. Esse método trata cada conjunto de imagens como um sistema único, diferentemente da classificação padrão que avalia cada imagem isoladamente. O problema foi modelado como uma Atribuição Linear (*Linear Assignment Problem*), onde o objetivo é encontrar a correspondência um-para-um ótima entre as dez imagens de entrada e as dez classes possíveis. A entrada do otimizador por atribuição é uma matriz de probabilidades gerada pelo classificador, conforme a Figura 2.

4.4. Recursos Computacionais

O método foi implementado em Python e os experimentos foram conduzidos no Google Colab. Para a etapa de treinamento, a biblioteca PyTorch e *Automatic Mixed Precision* (AMP) foram utilizadas para acelerar cálculos em GPU, em uma instância com NVIDIA A100-SXM4-40GB. A avaliação e otimização foram realizadas em uma instância com NVIDIA T4. A implementação da otimização utilizou o solver GLOP da biblioteca Google OR-Tools.

5. Resultados Experimentais

Esta seção avalia o desempenho da classificação e o impacto da otimização em dois níveis: por imagem (dedo isolado) e por indivíduo (acerto simultâneo dos 10 dedos). Inicialmente, analisamos o conjunto de validação para selecionar os melhores modelos.

5.1. Classificação Independente por Imagem

A Tabela 1 apresenta o desempenho da classificação independente. As quatro arquiteturas mostraram variação reduzida (inferior a 3%) no nível de imagem, indicando que o desempenho máximo alcançado pode ser resultado das características limitantes dos dados ou da convergência das redes. A Tabela 1 destaca a fragilidade da métrica em nível de indivíduo: apenas 14,4% dos indivíduos tiveram todos os dez dedos classificados corretamente (EfficientNetB0), confirmando que pequenos erros locais se acumulam e comprometem a consistência global.

Tabela 1. Desempenho no Conjunto de Validação: Acurácia da Classificação Independente (Imagem vs. Indivíduo)

Modelo	Imagem	Indivíduo
VGG16	74,7% (672)	11,1% (10)
ResNet50	75,2% (677)	8,9% (8)
DenseNet121	76,3% (687)	7,8% (7)
EfficientNetB0	76,9% (692)	14,4% (13)

5.2. Otimização por Atribuição

Para mitigar o acúmulo de erros, foi aplicada a otimização por atribuição utilizando a matriz de probabilidades como matriz de custo. A Tabela 2 demonstra o desempenho dessa abordagem. A otimização gerou ganhos significativos em consistência global, onde a acurácia por indivíduo saltou de 14,4% para 58,9% (EfficientNetB0), um aumento expressivo que confirma a eficácia da restrição imposta. Com base nestes resultados, selecionamos o modelo EfficientNetB0 para a avaliação final.

5.3. Resultados no Conjunto de Teste

O desempenho obtido no conjunto de teste esteve alinhado com as tendências identificadas na fase de validação, conforme a Tabela 3. Na avaliação por imagem, o SOCOFing apresentou acurácia de 77,2% com EfficientNetB0, ampliada para 87,8% após a otimização.

O efeito mais significativo da abordagem proposta é observado na avaliação por indivíduo. No SOCOFing, a acurácia aumentou de 13,3% para 60,0% com a incorporação

Tabela 2. Impacto da Matriz de Custo na Acurácia por Imagem e por Indivíduo (Validação)

Modelo	Imagem	Indivíduo
VGG16	86,3% (777)	53,3% (48)
ResNet50	88,2% (794)	56,7% (51)
DenseNet121	86,1% (775)	54,4% (49)
EfficientNetB0	88,9% (800)	58,9% (53)

Tabela 3. Avaliação Final no Teste: Classificação e Otimização

Método	Imagem	Indivíduo
Classificação	77,2% (695)	13,3% (12)
Otimização	87,8% (790)	60,0% (54)

da otimização, um acréscimo de aproximadamente 46 pontos percentuais. Esse resultado indica que a otimização contribui para mitigar as limitações da classificação independente, tornando a identificação adequada a um cenário operacional.

A Tabela 4 detalha o desempenho da classificação independente por classe. Os resultados evidenciam que os polegares apresentam os maiores F1-Scores (90,8% e 95,4%), enquanto os dedos anelares registram os piores desempenhos (F1-Scores de 67,4% e 62,3%). A matriz de confusão (Tabela 5) mostra que as falhas se concentram majoritariamente no mesmo lado: dos 205 erros totais, 146 ocorreram entre dedos da mesma mão e apenas 59 envolveram a troca de mãos. Além disso, as predições incorretas ocorreram sobretudo entre dedos adjacentes. Destaca-se que os quatro erros mais frequentes do modelo envolveram os dedos anelares, evidenciando a semelhança com os dedos vizinhos. Por exemplo, o classificador apontou o anelar esquerdo como mínimo esquerdo 19 vezes, e o médio direito como anelar direito 13 vezes.

Tabela 4. Métricas de desempenho por classe para a classificação independente por imagem no conjunto de teste.

Classe	Precisão	Revocação	F1-Score
Polegar Direito (PD)	94,0%	87,8%	90,8%
Indicador Direito (ID)	88,0%	81,1%	84,4%
Médio Direito (MD)	83,1%	71,1%	76,6%
Anelar Direito (AD)	63,1%	72,2%	67,4%
Mínimo Direito (mD)	67,9%	84,4%	75,2%
Polegar Esquerdo (PE)	98,8%	92,2%	95,4%
Indicador Esquerdo (IE)	81,6%	78,9%	80,2%
Médio Esquerdo (ME)	80,0%	66,7%	72,7%
Anelar Esquerdo (AE)	61,3%	63,3%	62,3%
Mínimo Esquerdo (mE)	65,7%	74,4%	69,8%

5.4. Resultados Comparativos

A Tabela 6 apresenta uma comparação com a literatura. A acurácia independente do modelo proposto (77,2%) mostrou-se comparável aos 76,7% reportados por

Tabela 5. Matriz de confusão da classificação independente no conjunto de teste.
As linhas representam as classes reais e as colunas as classes preditas.

	PD	ID	MD	AD	mD	PE	IE	ME	AE	mE
PD	79	2	4	2	2	0	1	0	0	0
ID	3	73	3	2	5	0	0	0	1	3
MD	1	3	64	13	2	0	0	1	5	1
AD	0	0	6	65	12	0	2	1	4	0
mD	0	0	0	8	76	0	1	0	1	4
PE	0	1	0	1	2	83	2	0	1	0
IE	0	2	0	1	4	1	71	6	2	3
ME	1	0	0	5	2	0	3	60	14	5
AE	0	1	0	4	1	0	3	5	57	19
mE	0	1	0	2	6	0	4	2	8	67

[Shehu et al. 2018a], que emprega uma arquitetura ResNet18 no mesmo conjunto de dados. Essa proximidade indica alinhamento com o estado da arte e sugere um possível limite prático para a classificação isolada por imagem neste conjunto. Consequentemente, o modelo fornece uma base consistente para a otimização em nível decadactilar. Na análise por classe, os resultados de [Hsiao et al. 2022] e [Shehu et al. 2018a] também registraram maiores taxas de acerto nos polegares e piores desempenhos nos dedos anelares.

Tabela 6. Comparação com trabalhos relacionados na tarefa de classificação da posição anatômica por imagem em dez classes.

Publicação	Dados	Modelo	Acurácia
Método Proposto	SOCOFing	EfficientNetB0	77,2%
[Shehu et al. 2018a]	SOCOFing	ResNet18	76,7%
[Hsiao et al. 2022]	Privado	VGG16	84,8%

Ademais, diferentemente de trabalhos correlatos, este estudo estende a análise para o nível de indivíduo, considerando o conjunto completo de dez impressões. Essa formulação avalia a consistência global e evidencia limitações da classificação isolada não capturadas por métricas exclusivamente baseadas em imagem, contextualizando a relevância da abordagem para cenários operacionais realistas.

6. Conclusão

Este trabalho apresentou uma abordagem que integra quatro arquiteturas de redes neurais convolucionais (VGG16, ResNet50, DenseNet121 e EfficientNetB0) a um modelo de otimização por atribuição linear para a classificação da posição anatômica. Ao empregar a matriz de probabilidades como matriz de custo, o método trata o conjunto decadactilar como uma entidade única, reduzindo as limitações da classificação independente.

Os resultados no conjunto de dados SOCOFing confirmam a eficácia da estratégia. A classificação isolada apresentou melhor desempenho nos polegares e pior nos anelares, com erros concentrados majoritariamente entre dedos adjacentes da mesma mão. Observou-se que a classificação isolada resultou em baixas taxas de acurácia por indivíduo, alcançando apenas 13,3% no conjunto de dados avaliado. Em contraste, a

incorporação da restrição de unicidade por otimização mitigou os erros e elevou significativamente a acurácia para 60,0%.

Como trabalhos futuros, pretende-se investigar cenários mais desafiadores, envolvendo conjuntos de dados de dedos ausentes ou repetidos, nos quais a restrição de atribuição deverá ser flexibilizada. Adicionalmente, o valor da função objetivo da otimização será explorado como um indicador de confiança. Por fim, uma direção promissora consiste no desenvolvimento de arquiteturas de *Deep Learning* capazes de incorporar essas restrições de forma *end-to-end*, unificando aprendizado e otimização em um único modelo.

Referências

- Al-Wajih, Y. A., Hamanah, W. M., Abido, M. A., Al-Sunni, F., and Alwajih, F. (2022). Finger type classification with deep convolution neural networks. In *ICINCO*, pages 247–254.
- Giudice, O., Litrico, M., and Battiato, S. (2020). Single architecture and multiple task deep neural network for altered fingerprint analysis. In *2020 IEEE International Conference on Image Processing (ICIP)*, pages 813–817.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- Hsiao, C. T., Lin, C. Y., Wang, P. S., and Wu, Y. T. (2022). Application of convolutional neural network for fingerprint-based prediction of gender, finger position, and height. *Entropy*, 24.
- Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2261–2269.
- Iloanusi, O. N. and Ejiogu, U. C. (2020). Gender classification from fused multi-fingerprint types. *Information Security Journal: A Global Perspective*, 29(5):209–219.
- Jain, A. K., Ross, A. A., Nandakumar, K., and Swearingen, T. (2025). *Introduction to Biometrics*. Springer International Publishing, Cham.
- Jian, W., Zhou, Y., and Liu, H. (2020). Lightweight convolutional neural network based on singularity roi for fingerprint classification. *IEEE Access*, 8:54554–54563.
- Kim, J., Rim, B., Sung, N.-J., and Hong, M. (2020). Left or right hand classification from fingerprint images using a deep neural network. *Computers, Materials & Continua*, 63(1).
- Loshchilov, I. and Hutter, F. (2017). Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.
- Maiti, D., Basak, M., and Das, D. (2024). Deep learning method for multi-attribute analysis of fingerprint images. *Computer Science Journal of Moldova*, 32(2).
- Maltoni, D., Maio, D., Jain, A. K., and Feng, J. (2022). *Handbook of Fingerprint Recognition*. Springer Cham, 3rd edition.

- Michelsanti, D., Ene, A. D., Guichi, Y., Stef, R., Nasrollahi, K., and Moeslund, T. B. (2017). Fast fingerprint classification with deep neural networks. In *VISIGRAPP 2017 - Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, volume 5, pages 202–209. SciTePress.
- Militello, C., Rundo, L., Vitabile, S., and Conti, V. (2021). Fingerprint classification based on deep learning approaches: Experimental findings and comparisons. *Symmetry*, 13.
- Oleiwi, B. K., Abood, L. H., and Farhan, A. K. (2022). Integrated different fingerprint identification and classification systems based deep learning. In *2022 International Conference on Computer Science and Software Engineering (CSASE)*, pages 188–193.
- Pessach, D., Singer, G., Avrahami, D., Chalutz Ben-Gal, H., Shmueli, E., and Ben-Gal, I. (2020). Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming. *Decision Support Systems*, 134:113290.
- Pillai, R., Sharma, N., Upadhyay, D., Dangi, S., and Gupta, R. (2024). Precision in biometric authentication: Cnn-driven fingerprint classification. In *2024 3rd International Conference for Innovation in Technology (INOCON)*, pages 1–5.
- Rim, B., Kim, J., and Hong, M. (2021). Fingerprint classification using deep learning approach. *Multimedia Tools and Applications*, 80:35809–35825.
- Shehu, Y. I., Ruiz-Garcia, A., Palade, V., and James, A. (2018a). Detailed identification of fingerprints using convolutional neural networks. In *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, pages 1161–1165. Institute of Electrical and Electronics Engineers Inc.
- Shehu, Y. I., Ruiz-Garcia, A., Palade, V., and James, A. (2018b). Sokoto coventry fingerprint dataset. *arXiv preprint arXiv:1807.10609*.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Singh, T., Bhisikar, S., Satakshi, and Kumar, M. (2021). Fingerprint identification using modified capsule network. In *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–6.
- Tan, M. and Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR.
- United Nations (UN) (1948). Universal Declaration of Human Rights. UN General Assembly Resolution 217 A (III). Adopted on December 10, 1948.
- Win, K. N., Li, K., Chen, J., Viger, P. F., and Li, K. (2020). Fingerprint classification and identification algorithms for criminal investigation: A survey. *Future Generation Computer Systems*, 110:758–771.
- Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? *Advances in neural information processing systems*, 27.
- Zhong, Z., Zheng, L., Kang, G., Li, S., and Yang, Y. (2020). Random erasing data augmentation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 13001–13008.