

UNIVERSIDADE ESTADUAL PAULISTA  
“JÚLIO DE MESQUITA FILHO”  
Instituto de Geociências e Ciências Exatas – IGCE  
Curso de Bacharelado em Ciências da Computação

GABRIEL MATTHIESEN LOPES

# **Reconhecimento de Impressões Digitais com Aprendizado Não Supervisionado**

Orientador: Daniel Carlos Guimarães Pedronette

Rio Claro - SP

2017

# Reconhecimento de Impressões Digitais com Aprendizado Não Supervisionado

Relatório de Trabalho de Conclusão de Curso na modalidade Trabalho de Graduação realizado no período Integral, apresentado no 2º semestre de 2017 à disciplina ES/TG do Curso de Bacharelado em Ciências da Computação do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Rio Claro, segundo a Portaria IGCE/DTA no 043/2012.

Aluno: Gabriel Matthiesen Lopes

Orientador: Daniel Carlos Guimarães Pedronette

Rio Claro - SP

2017

006.6    Lopes, Gabriel Matthiesen  
L864r      Reconhecimento de impressões digitais com aprendizado  
             não supervisionado / Gabriel Matthiesen Lopes. - Rio Claro,  
             2017

             36 f. : il., figs., tabs.

             Trabalho de conclusão de curso (bacharelado - Ciências  
             da Computação) - Universidade Estadual Paulista, Instituto de  
             Geociências e Ciências Exatas

             Orientador: Daniel Carlos Guimarães Pedronette

             1. Computação gráfica. 2. Recuperação de imagens  
             baseada em conteúdo. 3. Biometria de impressão digital. 4.  
             Aprendizado não supervisionado. I. Título.

# Resumo

Devido ao fato de que impressões digitais são únicas, elas são uma maneira eficaz de identificar as pessoas. Atualmente, existem sistemas que utilizam essa forma de identificação, sendo que suas aplicações mais comuns são para controle de acesso e reconhecimento de criminosos. No entanto, esses sistemas, assim como quaisquer outros, estão suscetíveis a erros. Logo, considerando sua notável atuação em questões de segurança, é importante reduzir suas ocorrências. Assim, é grande o interesse em maneiras de produzir resultados mais corretos, o que motiva o desenvolvimento de algoritmos mais eficientes e outros métodos que permitam a filtragem de erros. Nesse trabalho foi utilizada uma abordagem de recuperação de imagens baseada em seus conteúdos em imagens de impressões digitais, que foram disponibilizadas pelas edições da *Fingerprint Verification Competition (FVC)* dos anos de 2002 e 2004. Essa abordagem retorna listas ordenadas nas quais as imagens nas primeiras posições são aquelas que mais se assemelham à imagem buscada. Em seguida, foi aplicado sobre esta lista, um método de aprendizado não supervisionado, chamado RL-Sim\*. Os resultados, com e sem o uso desse método, foram comparados, para avaliar os efeitos desse pós-processamento nos resultados finais. Realizada a comparação, evidenciou-se que a aplicação do método RL-Sim\* provocou, de modo geral, uma melhora nos resultados finais, demonstrando ser uma opção interessante à identificação biométrica por meio de impressões digitais.

**Palavras-chave:** recuperação de imagens baseada em conteúdo. biometria de impressão digital. aprendizado não supervisionado.

# Abstract

Due to the fact that fingerprints are unique, they are an effective way of identifying people. Currently, there are systems that use this way of identification, and its most common applications are in access control and criminal recognition. However, such systems, like any other, are prone to errors, and considering their significant influence in matters related to security, it is important to reduce its occurrences. Therefore, there is a great interest in ways to produce more correct results, which motivates development of more effective algorithms as well as other methods that allow errors to be filtered. In this paper, a *CBIR*(Content-Based Image Retrieval) approach is used in fingerprint images provided by the *Fingerprint Verification Competition* (FVC) editions of the years 2002 and 2004. This approach provides ordered lists, in which the images most similar to the query image appear in the first positions. Then, on this list, an unsupervised learning method, called RL-Sim\*, is applied. The results, with and without the use of this method, were compared to evaluate the effects of this post-processing phase on the final results. This comparison evidenced that applying RL-Sim\* caused a general improvement on the results, demonstrating that it is an interesting option to fingerprint identification.

**Keywords:** content-based image retrieval. fingerprint biometrics. unsupervised learning.

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>6</b>
<b>2</b>	<b>TRABALHOS RELACIONADOS</b>	<b>8</b>
<b>2.1</b>	<b>Biometria</b>	<b>8</b>
2.1.1	Biometria da face, da íris e da retina	8
2.1.2	Biometria de impressão digital	9
<b>2.2</b>	<b>Aprimoramento da qualidade da imagem</b>	<b>11</b>
2.2.1	Aprimoramento espacial	11
2.2.2	Aprimoramento de histograma	11
<b>2.3</b>	<b>Recuperação de imagens baseada no conteúdo</b>	<b>12</b>
<b>2.4</b>	<b>Aprendizado de máquina</b>	<b>13</b>
2.4.1	Aprendizado supervisionado	13
2.4.2	Aprendizado não supervisionado	13
2.4.3	Aprendizado semi-supervisionado	14
2.4.4	Aprendizado por reforço	14
<b>3</b>	<b>RECONHECIMENTO DE IMPRESSÕES DIGITAIS BASEADO EM APRENDIZADO NÃO SUPERVISIONADO</b>	<b>15</b>
<b>3.1</b>	<b>Binarização e Inversão</b>	<b>15</b>
<b>3.2</b>	<b>Esqueletização</b>	<b>16</b>
<b>3.3</b>	<b>Localização de minúcias</b>	<b>19</b>
<b>3.4</b>	<b>Descritores de minúcias</b>	<b>20</b>
<b>3.5</b>	<b>Comparação</b>	<b>22</b>
<b>3.6</b>	<b>Aprendizado Não Supervisionado</b>	<b>22</b>
3.6.1	RL-Sim	22
3.6.2	RL-Sim*	24
<b>4</b>	<b>AVALIAÇÃO EXPERIMENTAL</b>	<b>25</b>
<b>4.1</b>	<b>Coleções de imagens</b>	<b>25</b>
4.1.1	FVC 2002	25
4.1.2	FVC 2004	25
4.1.3	Coleções Compostas	26
<b>4.2</b>	<b>Protocolo Experimental</b>	<b>26</b>
4.2.1	Precisão	27
4.2.2	<i>Mean Average Precision</i>	27
<b>4.3</b>	<b>Resultados Quantitativos</b>	<b>27</b>

4.4	Resultados Visuais . . . . .	29
5	CONCLUSÃO . . . . .	32
	REFERÊNCIAS . . . . .	33

# 1 Introdução

Com o avanço da tecnologia e da ciência, a vida humana tornou-se mais confortável. A grande quantidade de imagens digitais utilizadas no dia-a-dia é apenas um dos produtos desse avanço observado em várias áreas, comerciais e/ou acadêmicas, facilitando diversas atividades como a locomoção, a aquisição de conhecimento, a realização de transações, a manutenção da segurança, o diagnóstico e tratamento de condições médicas, entre outras.

No entanto, deve-se atentar que a crescente dependência nessa fonte de informação, por mais benéfica e promissora que seja, também implica na potencialização dos problemas que podem ser causados por ela, desde uma simples inconveniência até o comprometimento da vida de pessoas. Esse é mais um motivo responsável pela importância de imagens no cotidiano, logo, é natural o constante interesse em aprimorar processos que as utilizam.

A identificação biométrica é um exemplo de processo que tem como base o uso de imagens. Ela consiste no uso de características físicas dos seres humanos com o objetivo de identificá-los. Alguns exemplos dessas características são a face, a íris, a retina e a impressão digital. Esta última é única entre seres humanos, portanto é uma escolha atraente para a identificação de pessoas. Suas principais aplicações são no controle de acesso e identificação criminal (NUNES et al., 2017; FERREIRA et al., 2017), assim exercendo notável influência em questões de segurança. Embora seja possível realizar a análise de suas características manualmente, o processo é custoso em relação ao tempo, esforço humano e dinheiro. Logo, técnicas para analisá-las com uso de computadores (ALI et al., 2016) se fazem extremamente convenientes. Além disso, é necessário que tais técnicas possuam alto grau de precisão para evitar erros que podem comprometer a segurança de instituições e pessoas. Esses erros podem tender a um dos opostos, sendo eles: a confirmação de pessoas erradas ou negação de pessoas certas.

A impressão digital é uma característica interessante para a identificação biométrica, pois, não sofre alterações naturais significativas no decorrer da vida de uma pessoa, além de ser distinta até mesmo entre gêmeos idênticos. Suas formas de aquisição, por meio de sensores ou datilograma, também são pouco invasivas. No entanto, não é possível realizar a análise a partir de uma imagem de câmera de vigilância por exemplo, tornando o método inadequado para identificação de pessoas em movimento e em grandes quantidades. A qualidade da imagem coletada desempenha um papel de extrema importância (RAJPUT; SURALKAR, 2013), pois, ela afeta a eficiência final do sistema, embora, na prática, cada coleta gerará resultados diferentes devido, por exemplo, a pressão exercida pelos dedos nos aparelhos. Para fins desse trabalho, as imagens que correspondem a um cadastro de usuários ou suspeitos em um sistema real, foram obtidas de bases disponíveis on-line (MAIO et al., 2002; MAIO et al., 2004).



Sistemas computadorizados que realizam identificação biométrica podem utilizar uma modelagem definida por Sistemas de Busca de Imagens Baseados no Conteúdo (*Content-Based Image Retrieval – CBIR*) (GUDIVADA; RAGHAVAN, 1995). A aplicação mais comum nos sistemas CBIR consiste em retornar imagens de uma coleção que mais se assemelham visualmente à imagem utilizada para consulta. A biblioteca *OpenCV* fornece a implementação de alguns desse métodos, sendo que ela foi utilizada nesse trabalho (HOWSE et al., 2015).

Em geral, essa lista de resultados é criada com base em uma métrica de distância, sendo que imagens com menor distância são mais similares e, dessa forma, exibidas nas primeiras posições da lista ordenada de resultados. No entanto, mesmo atualmente não é uma tarefa simples definir métricas adequadas para comparação de imagens, e o desenvolvimento de outros algoritmos é um processo difícil.

Assim, métodos de aprendizado não-supervisionado (GHAHRAMANI, 2004) têm sido desenvolvidos com o objetivo de aumentar a eficácia de tarefas de busca realizada por Sistemas CBIR. Tais métodos utilizam os resultados de algoritmos específicos de cada área para definir métricas mais eficazes que, então, são utilizadas nos resultados anteriormente obtidos.

Este trabalho apresenta uma abordagem que utiliza a biblioteca *OpenCV* e baseia-se em conceitos de *CBIR* e métodos de aprendizado não supervisionado. Utilizando um algoritmo que permite organizar imagens de impressões digitais mais similares baseando-se em suas características (HOWSE et al., 2015), chamadas de minúcias, será obtida uma lista ordenada. Em seguida, sobre essa lista será utilizado um método de aprendizado não supervisionado (OKADA; PEDRONETTE; TORRES, 2015) para gerar uma nova lista com uma métrica mais eficaz e ambas serão comparadas, para então ser feita uma análise sobre a efetividade desse método nesse cenário de aplicação.

As seções desse trabalho estão organizadas da seguinte forma : a Seção 2 apresenta tópicos e estudos relacionados com áreas envolvidas, seguida pela Seção 3, que detalha os algoritmos utilizados nesse trabalho. Na Seção 4 são apresentadas as coleções de imagens utilizadas e os resultados dos testes, sendo que a Seção 5 contém as observações finais, baseadas no experimento realizado.

## 2 Trabalhos Relacionados

Nesta seção são apresentados tópicos relevantes para o contexto do trabalho. Primeiramente, é introduzido o tema da biometria na Seção 2.1, com exemplos de duas formas de biometria na Seção 2.1.1, seguido da biometria por impressão digital, utilizada nesse trabalho, mais detalhadamente na Seção 2.1.2. Discorre-se sobre as imagens e processos utilizados a fim de aumentar sua qualidade na Seção 2.2. Em seguida, a forma que as consultas são comumente realizadas é apresentada na Seção 2.3. Por fim, trata-se dos métodos de aprendizado de máquina na Seção 2.4, que podem ser utilizados para refinar os resultados do reconhecimento.

### 2.1 Biometria

A Biometria refere-se ao uso de certas características do seres humanos para fins de identificá-los, sendo algumas delas, a face, a íris, a retina, a voz, e a impressão digital. Cada uma dessas formas possui diferentes aspectos. Nesse trabalho, foi utilizada a impressão digital como foco. Apesar de ser custoso para uma pessoa analisá-la manualmente, ela é de alta confiabilidade, pois, é única, até mesmo entre gêmeos idênticos e não costuma sofrer alterações naturais com o passar do tempo, além de não ser facilmente fraudada (CUSTÓDIO, 2007). Assim, utilizar computadores para analisá-la é um meio de se evitar o problema do esforço humano. Além disso, essa forma de biometria possui outros fatores favoráveis, como a baixa invasividade e custo dos sensores (ALI et al., 2016). Descreve-se essa biometria mais detalhadamente na Seção 2.1.2, enquanto que na Seção 2.1.1 são mencionadas as biometrias da face, da íris e da retina.

#### 2.1.1 Biometria da face, da íris e da retina

O reconhecimento de faces, por exemplo, é facilmente realizado pelo cérebro humano, apesar do processo exato ser desconhecido (CARNEIRO, 2012), e um dos notáveis benefícios é de que a coleta de imagens pode ser feita não só em ambiente controlado, como também a partir de câmeras que capturam uma grande quantidade de pessoas simultaneamente. No entanto, existem os reveses de uma mesma face poder apresentar variações de expressões e pêlos faciais, de ângulos, bem como da existência de cirurgias plásticas. Também pode ser utilizado o termograma da face, isto é, o padrão formado pelo calor emitido pelos vasos sanguíneos, o qual pode ser detectado por sensores infra-vermelho, e que possui o benefício de não ser afetado por cirurgias plásticas além de acreditar-se que é único para cada pessoa (JAIN; HONG; PANKANTI, 2000).

A íris e a retina, ambas pertencentes ao olho humano, são formas altamente únicas (DAUGMAN, 2004; BORGEM; BOURS; WOLTHUSEN, 2008). A íris, apesar de ser relativamente protegida, também é de fácil visualização, além de não ser facilmente alterada artificialmente. O reconhecimento da retina se baseia nos padrões dos vasos sanguíneos que a constituem e possui aplicações médicas como na detecção de problemas relacionados à hipertensão ou à diabetes (SÁNCHEZ et al., 2008). As desvantagens dessa forma estão mais relacionadas com a coleta do que com o reconhecimento em si (BORGEM; BOURS; WOLTHUSEN, 2008), pois é mais invasiva e os sensores são mais caros.

### 2.1.2 Biometria de impressão digital

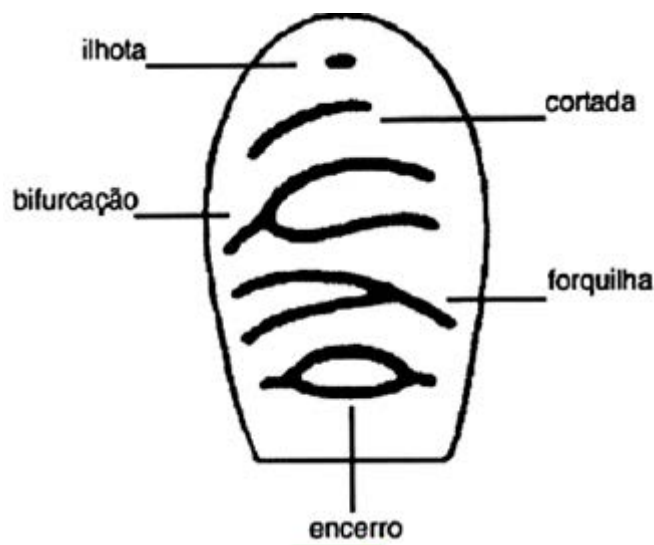
Atualmente, é comum observarmos as impressões digitais sendo utilizadas como uma forma de identificação (biometria de impressão digital) pelas razões mencionadas anteriormente, estando entre as suas aplicações mais comuns, o controle de acesso a locais e recursos, como terminais de autoatendimento (CUSTÓDIO, 2007) e a área forense, no reconhecimento de criminosos ou pessoas desaparecidas (NUNES et al., 2017). Viável e legal, inclusive em situações de flagrante, esta última visa agilizar o trabalho da polícia judiciária, considerando-se que a biometria de impressões digitais, que já é utilizada nas urnas eletrônicas para permitir o voto, tem sido utilizada para fins de cadastro da população carcerária na maioria dos estados brasileiros (FERREIRA et al., 2017).

Deve-se ressaltar que erros ou fraudes em quaisquer das aplicações mencionadas, implicam em um risco para a segurança da população e/ou instituições, o que contribui para que estudos sejam desenvolvidos em torno de seu aprimoramento. Assim, é grande o interesse em maneiras de produzir resultados mais corretos, o que motiva o desenvolvimento de algoritmos mais eficientes e outros métodos que permitam a filtragem de erros, a exemplo desse trabalho de conclusão de curso que utiliza uma abordagem de recuperação de imagens baseada em seus conteúdos (GUDIVADA; RAGHAVAN, 1995) em imagens de impressões digitais (ALI et al., 2016), que retorna listas ordenadas, nas quais as imagens nas primeiras posições são aquelas que mais se assemelham à imagem consultada.

Para Rajput e Suralkar (2013, p. 11): "o desempenho de qualquer reconhecedor de impressões digitais depende muito da qualidade da imagem de impressão digital [...] mas imagens de impressões digitais raramente estão em perfeita qualidade". Isso se deve ao fato de existirem fatores que podem interferir na qualidade da imagem, tais como: a condição da pele, a qualidade do aparelho utilizado para coleta e a resistência do indivíduo em cooperar com o processo.

Para as impressões digitais, as características de interesse para discriminá-las são padrões resultantes das ramificações dos vasos nos dedos, representados na **Figura 1**, a seguir.

Figura 1 – Representação dos padrões de uma impressão digital.



Fonte: Nunes et al. (2017, p. 20)

As imagens inicialmente coletadas não são tão nítidas, a exemplo da **Figura 2**, que se refere à uma das imagens disponibilizadas para a edição da FVC de 2002 (MAIO et al., 2002), utilizada nesse trabalho, antes de ser aprimorada.

Figura 2 – Coleta inicial de uma impressão digital



Fonte : Maio et al. (2002)

De particular interesse sobre essas regiões são suas terminações e bifurcações, chamadas de minúcias, que dão nome a uma das possíveis abordagens para a biometria de impressões digitais. No entanto, para que as minúcias sejam detectadas com mais precisão, não é interessante utilizar uma imagem da mesma forma que ela foi coletada, como exemplificado na **Figura 2**.

Logo, independente da abordagem, convém submeter as imagens a uma série de processos que façam com que as minúcias estejam visíveis o mais nitidamente possível.

Tais processos são detalhados na Seção 2.2, a seguir.

## 2.2 Aprimoramento da qualidade da imagem

Como já mencionado, existem diversos fatores que podem afetar a qualidade da imagem coletada. Além da qualidade dos aparelhos afetarem as imagens em si, deve-se ressaltar que esse problema é agravado quando pretende-se comparar imagens coletadas por sensores diferentes, mesmo que o procedimento seja replicado com exatidão (BORRA; REDDY; REDDY, 2016).

Assim, além de melhorar as imagens individualmente, o aprimoramento da qualidade das imagens também traz o benefício de padronizá-las. Os processos são divididos em dois principais domínios, o espacial e o de histograma (RAJPUT; SURALKAR, 2013), enquanto que Borra, Reddy e Reddy (2016) mencionam um terceiro, no qual a imagem é decomposta em valores de senos e cossenos, através da transformada de Fourier e melhorada em blocos independentes.

### 2.2.1 Aprimoramento espacial

Processos desse tipo focam na vizinhança de cada pixel para aplicar alguma função sobre ele, alterando seus valores RGB. Tais funções são também chamadas de filtros, sendo os dois mais comuns os filtros "passa-alta (frequência)" e os filtros "passa-baixa (frequência)", embora outros também existam (BORRA; REDDY; REDDY, 2016).

Os filtros "passa-alta" tornam um pixel mais escuro se sua vizinhança é mais clara e vice-versa, aumentando o contraste da imagem ao realçar as suas bordas, assim, tornando as minúcias mais evidentes. Já um filtro "passa-baixa", torna um pixel mais escuro se sua vizinhança é mais escura ou, mais claro, se sua vizinhança é mais clara, diminuindo o contraste da imagem ao suavizar as bordas. O ruído de "sal e pimenta", observável quando pixels são incorretamente preenchidos com a cor branca e/ou preta, é tradicionalmente combatido com o filtro de mediana, por exemplo. Ambos os tipos de filtros mencionados também possuem uma "versão" utilizada no aprimoramento de histograma.

### 2.2.2 Aprimoramento de histograma

Um histograma, no contexto de imagens computadorizadas, é a representação da frequência de ocorrência de cada valor RGB presente em uma imagem. Um histograma com alta variação representa uma imagem com alto grau de contraste, enquanto que um histograma mais homogêneo representa uma imagem com tons mais balanceados (BORRA; REDDY; REDDY, 2016).

Imagens em duas dimensões de impressões digitais são geralmente processadas em tons de cinza, ou seja, os valores de RGB de um pixel são iguais, assim como foram nesse trabalho. Por esse motivo, os histogramas aqui evidenciam apenas variação na clareza/escurecimento da imagem.

Filtros "passa-alta" e "passa-baixa" possuem a mesma função que no aprimoramento espacial. A diferença é que em vez de analisarem cada pixel e sua vizinhança, eles utilizam o histograma de uma imagem.

## 2.3 Recuperação de imagens baseada no conteúdo

A abordagem tradicional de busca e consulta de imagens, baseada na análise de texto, apresenta limitações, como a necessidade de haver um ser humano para transformar uma imagem em metadados, além das limitações intrínsecas a esse sistema como, por exemplo, classificações com termos similares ou ofuscação de informação. No contexto de impressões digitais, tais limitações são reforçadas pelo fato de que as imagens não são intuitivamente descritas detalhadamente por seres humanos. Deve-se ressaltar, no entanto, que em outros contextos, como no reconhecimento de objetos (TORRALBA *et al.*, 2003), essa forma pode ser mais razoável, e existem mecanismos de busca híbridos (ENSER, 2000), que a utilizam juntamente com a abordagem baseada em conteúdo, explicada a seguir.

A abordagem baseada em conteúdo consiste em utilizar a informação contida na imagem em si, identificando áreas de interesse e baseando-se nelas para determinar imagens similares (GUDIVADA; RAGHAVAN, 1995). Essa abordagem é mais popular no que se refere a imagens, não somente no contexto de biometria, como também na área médica (MÜLLER *et al.*, 2004) por exemplo, devido à alta quantidade de informações que elas podem fornecer (ENSER, 2000). No caso desse trabalho, foi utilizada essa abordagem e, como mencionado na Seção 2.1.2, as áreas de interesse são as minúcias. Entretanto, existem outros métodos que podem ser utilizados, como o de análise de textura, que visa evitar os problemas relacionados à coleta (THAIYALNAYAKI *et al.*, 2010), já mencionados na Seção 2.1.2 e o de funções wavelets (onduletas), que produz resultados eficientes mesmo sem qualquer tipo de processamento das imagens coletadas (TICO *et al.*, 2001).

Essa abordagem retorna as imagens, de um banco de dados, em uma lista, sendo que as que estão nas primeiras posições, possuem conteúdo mais similar à imagem inserida para consulta, com base em alguma métrica de distância. O algoritmo utilizado nesse trabalho (HOWSE *et al.*, 2015) faz uso de técnicas de aprimoramento das imagens, tema abordado na Seção 2.2. Após a execução do algoritmo, utilizou-se uma técnica de aprendizado de máquina, explanado na Seção 2.4, a seguir.

## 2.4 Aprendizado de máquina

Entende-se por "Aprendizado de Máquina" qualquer algoritmo que possua como objetivo aprimorar resultados de uma operação, sendo um campo interdisciplinar que faz uso de conhecimento pertinente à estatística, à informática, à engenharia, à ciência cognitiva, entre outras áreas da ciência e da matemática.

Supondo que uma máquina receba uma sequência de entrada (*input set*)  $E = (e_1, e_2, \dots)$  com o objetivo de produzir uma sequência de saída (*output set*)  $S = (s_1, s_2, \dots)$ , o aprendizado de máquina é usado para aprimorar  $S$  no contexto analisado.

Ghahramani (2004) distingue quatro formas de se realizar o aprendizado de máquina, explanados nas respectivas seções: "aprendizado supervisionado" (Seção 2.4.1), "aprendizado não supervisionado" (Seção 2.4.2) e "aprendizado por reforço" (que possui duas variações) (Seção 2.4.4). Também aborda-se o aprendizado semi-supervisionado na Seção 2.4.3.

### 2.4.1 Aprendizado supervisionado

No aprendizado supervisionado, um conjunto de entradas de treinamento  $T$  (*training set*), com propriedades já conhecidas, é utilizado antes de  $E$ , com o objetivo de "afinar" o processo baseando-se na extração de características importantes (LEE, 2005) de  $T$ , que acredita-se que também serão observadas em  $E$ , para formular um modelo. Entradas com propriedades conhecidas também são chamadas de "rotuladas" (*labelled*). Áreas comuns de sua aplicação são a regressão estatística (quantitativa) e a classificação (qualitativa) (GHAHRAMANI, 2004; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Dois fatores principais de um algoritmo que afetam os resultados finais são a seleção da característica e o viés (*bias*) de uma sequência de treinamento. O primeiro depende de uma escolha correta das características a serem utilizadas, por parte de indivíduos versados no assunto, pois uma escolha infeliz, de característica irrelevante, por exemplo, não fornecerá boa base para treinamento e piorará os resultados (LEE, 2005). O segundo pode ser explicado como a variação nos resultados quando utilizadas diferentes  $T$  que possuem qualidade semelhante.

### 2.4.2 Aprendizado não supervisionado

No aprendizado não supervisionado, a máquina não recebe qualquer sequência de treinamento e utiliza como referência exclusivamente a entrada  $E$ , com atributos não rotulados (*unlabelled*), de modo que os algoritmos de aprendizado não supervisionado focam em encontrar relações entre os elementos recebidos. Resultados insatisfatórios podem representar problemas na obtenção de  $E$  ou acusar a ineficácia da relação encontrada, fornecendo bases para avaliar e aprimorar processos e tomadas de decisões (GHAHRAMANI, 2004).



Tal abordagem foi empregada, por exemplo, em pesquisas relacionadas a redes neurais (SIMÕES; COSTA, 2007) e domínios *fuzzy* (YONAMINE et al., 2002), ambas para fins de classificação de elementos de um conjunto. O algoritmo de aprendizado não supervisionado utilizado nesse trabalho foi o RL-Sim\* (OKADA; PEDRONETTE; TORRES, 2015), detalhado na Seção 3.6.2.

### 2.4.3 Aprendizado semi-supervisionado

O aprendizado semi-supervisionado é um intermediário entre os aprendizados supervisionado e não supervisionado. Ele utiliza em parte uma sequência de treinamento para um ajuste inicial e, então, utiliza a sequência de entrada, com propriedades desconhecidas, a fim de melhorar esse ajuste.

Devido a pequena quantidade de imagens rotuladas, o efeito que elas exercem na consulta é amplificado, logo, imagens com ruídos podem causar a degradação da qualidade dos resultados (ZHU, 2005). Esse fato serve como motivação especial para o estudo de formas de limitar a propagação de erros ocasionados por imagens rotuladas ruidosas, como a baseada na "cooperação" de imagens com rótulos comuns e na "competição" de imagens com rótulos diferentes, a fim de identificar ruídos (BREVE; ZHAO; QUILES, 2015).

Outros exemplos de algoritmos desse tipo são os algoritmos de Baum-Welch, usado para encontrar parâmetros de cadeias ocultas de Markov (BAGGENSTOSS, 2001), e que serviu de base para algoritmos de reconhecimento de fala (ZHANG; WANG; ZHAO, 2007), e algoritmos por exemplo.

### 2.4.4 Aprendizado por reforço

No aprendizado por reforço,  $S$  corresponde a uma sequência de ações que a máquina (no contexto de inteligência artificial, comumente chamada de agente) realiza sobre o ambiente que, por sua vez, possui consequências  $C = (c_1, c_2, \dots)$  positivas ou negativas. O ambiente pode ser estático, fornecendo a mesma  $E$  durante toda a operação, como também, pode ser dinâmico, havendo outras máquinas que realizam ações sobre ele e sofrem consequências.

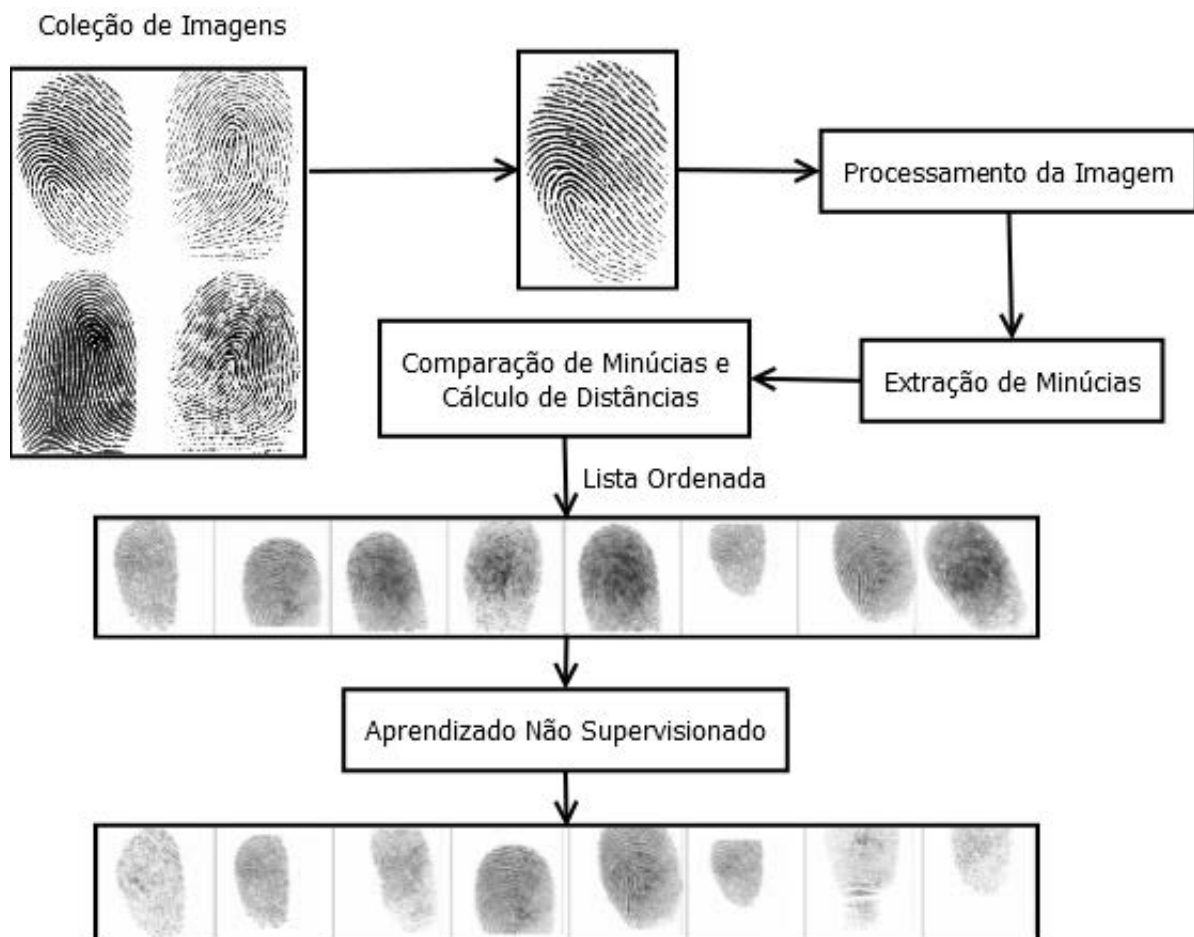
Em ambos os casos, a máquina tem como objetivo agir de modo a maximizar a ocorrência de consequências positivas e minimizar a ocorrência de consequências negativas durante seu período de execução, tendo que levar em consideração as ações de outras máquinas na segunda variação. Essa forma de aprendizado possui grande relação com a teoria de decisão (GHAHRAMANI, 2004; SUTTON et al., 2000).



## 3 Reconhecimento de Impressões Digitais

Nesta seção, são detalhados os algoritmos de reconhecimento de impressões digitais utilizado nesse trabalho (HOWSE et al., 2015), englobando: Os processos de aprimoramento da imagem inicial, nas Seções 3.1 e 3.2, os processos de localização e descrição das minúcias, nas Seções 3.3 e 3.4, o processo de comparação de duas imagens de impressões digitais na Seção 3.5, finalizando com os algoritmos de aprendizado não supervisionado, na Seção 3.6. A implementação de todos os passos dessa seção é possível com o uso da biblioteca *OpenCV*. Esse processo é ilustrado de forma simplificada na **Figura 3**.

Figura 3 – Processo de reconhecimento de impressões digitais



### 3.1 Binarização e Inversão

O primeiro passo para realçar as minúcias é a distinção mais acentuada do que é, na imagem, a impressão digital de fato ou o fundo (*background*), bem como, a filtragem de ruídos. Tendo em vista que, nas imagens em tons de cinza, as impressões digitais tendem à cor preta, sendo que o fundo é branco, convém separá-las em apenas preto e branco,

transformando-as em uma imagem binarizada. Surge, então, a necessidade de definir um específico tom de cinza, tal que todos os tons mais claros que ele sejam transformados em branco e que todos os tons mais escuros sejam transformados em preto, na imagem binarizada.

Esse processo é denominado *thresholding*, havendo diversas formas de fazê-lo. A mais simples, consiste em escolher um tom (limiar) global arbitrário para realizar essa separação. No entanto, essa forma é considerada inadequada para esse trabalho, pois, além de serem processadas várias imagens, elas, possivelmente, possuem luminosidades significativamente diferentes, mesmo quando obtidas por meio de um mesmo sensor.

Nesse caso, optou-se por utilizar o *Otsu thresholding* (OTSU, 1979), que analisa o histograma de cada imagem para determinar o melhor limiar para cada uma. Os histogramas das imagens de impressões digitais são geralmente bimodais, isto é, possuem dois picos expressivos, sendo que o Otsu thresholding seleciona um valor intermediário a esses picos, de tal forma que a variância é minimizada. Além disso, para aumentar a eficiência dos próximos passos (HOWSE et al., 2015), também inverteu-se as cores da imagem binarizada, como mostra a **Figura 4**.

Figura 4 – Comparação da imagem inicial com sua forma binarizada e invertida



## 3.2 Esqueletização

Apesar de ser possível extrair os pontos de interesse após feita a binarização, foi utilizado o algoritmo de Zhang-Suen (ZHANG; SUEN, 1984; CHEN et al., 2012) para reduzir a impressão digital ao seu "esqueleto", isto é, à menor região central possível, a fim de tornar as minúcias ainda mais explícitas.

Seja  $P1(i, j)$  o pixel sendo analisado a cada iteração, e  $P2, P3, \dots, P9$  os pixels ao seu redor, conforme ilustra a Tabela 1 na página seguinte:

Tabela 1 – Pixel P1 sendo analisado, e sua vizinhança

P9(i-1,j-1)	P2(i-1,j)	P3(i-1,j+1)
P8(i,j-1)	P1(i,j)	P4(i,j+1)
P7(i+1,j-1)	P6(i+1,j-1)	P5(i+1,j+1)

O algoritmo separa cada iteração em duas sub-iterações (ZHANG; SUEN, 1984), aqui chamadas de  $\alpha_1$  e  $\alpha_2$  para fins de simplificação. Em  $\alpha_1$ , P1 é removido, ou seja, torna-se da cor de fundo ( $P1 = 0$ ), se todas as seguintes condições forem verdadeiras (ZHANG; SUEN, 1984):

$$2 \leq B(P1) \leq 6 \tag{3.1}$$

$$A(P1) = 1 \tag{3.2}$$

$$P2 \&\& P4 \&\& P6 = 0 \tag{3.3}$$

$$P4 \&\& P6 \&\& P8 = 0 \tag{3.4}$$

$A(P1)$  corresponde ao número de vezes que a sequência 01 é lida ao se percorrer a vizinhança de P1 no sentido horário, ou seja:

$$\begin{aligned} A(P1) = & (P2 = 0 \&\& P3 = 1) + (P3 = 0 \&\& P4 = 1) \\ & + (P4 = 0 \&\& P5 = 1) + (P5 = 0 \&\& P6 = 1) \\ & + (P6 = 0 \&\& P7 = 1) + (P7 = 0 \&\& P8 = 1) \\ & + (P8 = 0 \&\& P9 = 1) + (P9 = 0 \&\& P2 = 1) \end{aligned}$$

$B(P1)$  corresponde ao número de pixels adjacentes que não são da cor de fundo, ou seja:

$$B(P1) = P2 + P3 + P4 + P5 + P6 + P7 + P8 + P9$$

Ao fim de  $\alpha_1$ , são removidos os pixels localizados nas bordas abaixo e à esquerda, bem como, nas pontas acima e à direita, devido a 3.3 e 3.4.

Para a execução de  $\alpha_2$ , as condições 3.1 e 3.2 são mantidas, enquanto que 3.3 e 3.4 são substituídas por 3.5 e 3.6, onde:

$$P2 \ \&\& \ P4 \ \&\& \ P8 == 0 \tag{3.5}$$

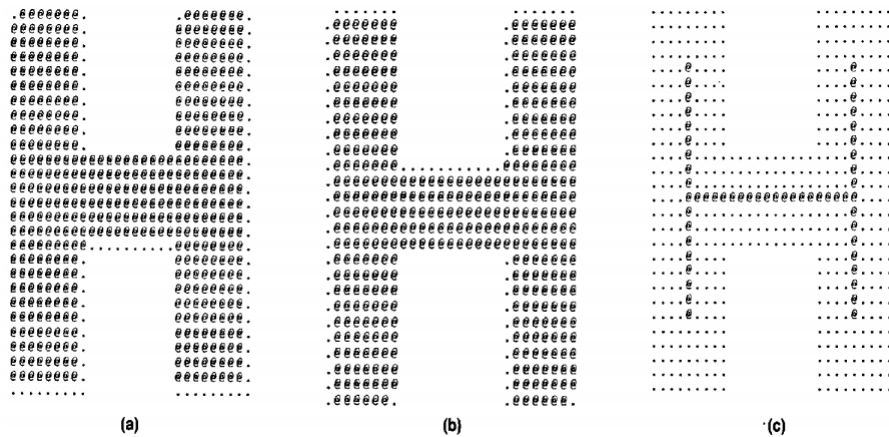
$$P2 \ \&\& \ P6 \ \&\& \ P8 == 0 \tag{3.6}$$

Ao fim de  $\alpha 2$ , são removidos os pixels localizados nas bordas acima e à direita, bem como, nas pontas abaixo e à esquerda, devido a 3.5 e 3.6 respectivamente.

A condição 3.1 garante que as terminações do esqueleto da imagem não sejam removidas, enquanto que a condição 3.2 previne a remoção de pontos pertencentes ao esqueleto que estão localizados entre as terminações (ZHANG; SUEN, 1984).

As iterações se repetem para todos os pixels preenchidos e param quando nenhum outro pixel pode ser removido. As duas imagens a seguir mostram os resultados desse algoritmo. A Figura 5 mostra, da esquerda para a direita, as fases em que houve a execução de apenas  $\alpha 1$ , apenas  $\alpha 2$ , e do algoritmo por completo, respectivamente, sobre uma representação do caractere "H".

Figura 5 – Passos do algoritmo de Zhang-Suen, "@" denota pixels preenchidos, "." denota pixels que foram removidos e espaços vazios denotam pixels que já pertenciam ao background da imagem antes da execução.



Fonte : Zhang e Suen (1984)

Já a Figura 6, apresenta o resultado final do algoritmo sobre a imagem binarizada que consta na Figura 4.

Figura 6 – Comparação da imagem binarizada com sua esqueletização



### 3.3 Localização de minúcias

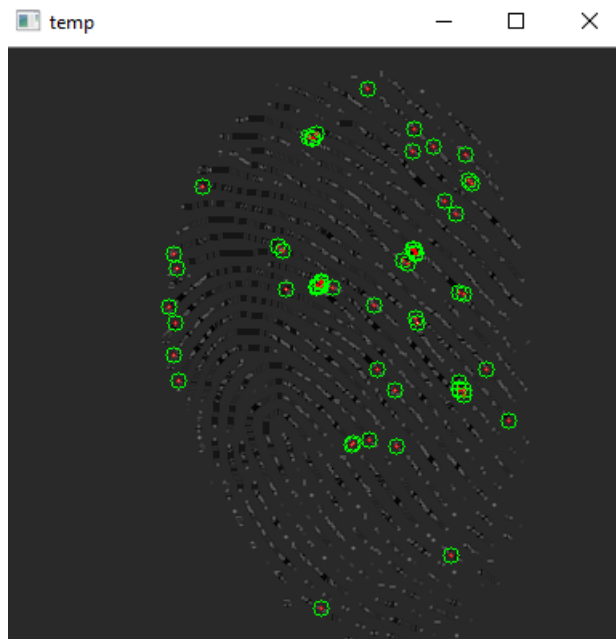
Em posse da imagem esqueletizada, podemos identificar a posição das minúcias, ou seja, terminações e bifurcações. Cabe notar que em ambas as situações, a região em que estão presentes possui um alto grau de contraste, logo uma forma de realizar essa identificação é localizar tais regiões.

Para isso, foi utilizado o algoritmo *Harris Corner and Edge Detector* (HARRIS; STEPHENS, 1988), que analisa a imagem em blocos, nesse caso de dimensões  $2 \times 2$ , em torno de cada pixel  $P(x, y)$ , calculando as respectivas covariâncias e armazenando-as na matriz  $M(i, j)$ . A Equação 3.7 a seguir mostra como é calculado o próximo passo.

$$dst(x, y) = detM(x, y) - k.(trM(x, y))^2 \quad (3.7)$$

O parâmetro  $k$  é denominado parâmetro livre e pode ser determinado empiricamente, mas costuma assumir valores entre 0.04 e 0.15. Em outras palavras, o algoritmo acusa altas variações de contraste em posições próximas, que representam os pontos de interesse para a análise de minúcias.

Para selecionar os melhores pontos, deve-se aplicar um *threshold* para manter apenas os valores mais altos. No caso do trabalho, primeiramente a matriz  $dst(x, y)$  foi normalizada entre 0 e 255, e o valor escolhido de *threshold* foi 125. Essa escolha foi, de certa forma, arbitrária, por se tratar de um valor intermediário, mas, outros valores podem melhorar ou piorar os resultados (HOWSE et al., 2015). Assim, as posições dos pontos acima desse *threshold* foram passados para a próxima etapa. A Figura 7 mostra os pontos selecionados.

Figura 7 – Bordas mais expressivas, ou seja, acima do valor de *threshold*.

### 3.4 Descritores de minúcias

Com as posições de interesse selecionadas, o próximo passo é utilizar um método que permita distinguir o conteúdo de cada uma delas, para que seja possível distingui-las e, futuramente, utilizar essa informação para encontrar as mesmas minúcias em imagens que, provavelmente, possuirão variações de posicionamento, orientação, *zoom*, luminosidade, entre outras possibilidades.

Para essa finalidade, utiliza-se descritores que, como o nome sugere, descrevem padrões locais nas imagens. Existe uma variedade de descritores, sendo que entre os mais comuns estão o SIFT (*scale invariant feature transform*) e o SURF (*speeded up robust features*). No entanto, nesse trabalho foi utilizado parcialmente o ORB (*Oriented FAST and rotated BRIEF*) que, ao contrário dos anteriores, é de uso livre (*BSD licensed*), além de utilizar menos recursos computacionais (RUBLEE et al., 2011).

Por padrão, esse descritor utiliza o algoritmo FAST (*Features from accelerated segment test*) (ROSTEN; DRUMMOND, 2006) para encontrar as minúcias e, então, as ordena pelo fator de Harris, mostrado na Equação 3.7, enquanto que nesse trabalho foi utilizado o algoritmo *Harris Corner and Edge Detector* (HARRIS; STEPHENS, 1988) também para encontrá-las. As equações 3.8, 3.9 e 3.10 a seguir mostram os passos preliminares utilizados pelo algoritmo ORB para obter uma descrição dos pontos invariante a rotações.

$$m_{pq} = \sum_{x,y} x^p y^q I(x, y) \quad (3.8)$$

$$C = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (3.9)$$

$$\theta = \text{atan2}(m_{01}, m_{10}) \quad (3.10)$$

Nas equações acima,  $I(x, y)$  é um dos pontos encontrados pelo algoritmo *Harris Corner and Edge Detector* (HARRIS; STEPHENS, 1988),  $C$  é um ponto chamado de centroide e  $\theta$  é o ângulo do vetor  $\vec{IC}$ .

Em seguida, o algoritmo utiliza uma variação do descritor BRIEF (*Binary Robust Independent Elementary Features*) (CALONDER et al., 2010). A princípio, esse descritor sofre uma significativa perda de precisão quando há variações de rotação (RUBLEE et al., 2011), logo, o valor  $\theta$  calculado acima foi utilizado para aprimorá-lo.

Um descritor BRIEF consiste, inicialmente, em uma *string* de 0s e 1s, formada com base na Equação 3.11 (RUBLEE et al., 2011) a seguir.

$$\tau(p; x; y) = 1 : p(x) < p(y), 0 : p(x) \geq p(y) \quad (3.11)$$

$$f_n = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x_i; y_i) \quad (3.12)$$

Nas Equações 3.11 e 3.12,  $x$  e  $y$  representam pixels dentro de uma região da imagem, em que  $p(x)$  indica a intensidade do pixel, e a Equação 3.12 mostra como é construído um vetor desses valores, sendo que o valor de  $n$  recomendado é 256.

Em seguida, é construída uma matriz  $2 \times n$ ,  $S$ , sobre a qual é aplicada a matriz de rotação pelo ângulo  $\theta$  calculado anteriormente, produzindo uma versão direcionada (*steered*) do descritor anterior, que permite comparação de pontos com variações de rotação. Os passos explanados constam nas Equações 3.13, 3.14 e 3.15 a seguir, que resultam nos novos descritores *steered-BRIEF* (RUBLEE et al., 2011):

$$S = \begin{bmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \end{bmatrix} \quad (3.13)$$

$$S_\theta = R_\theta S \quad (3.14)$$

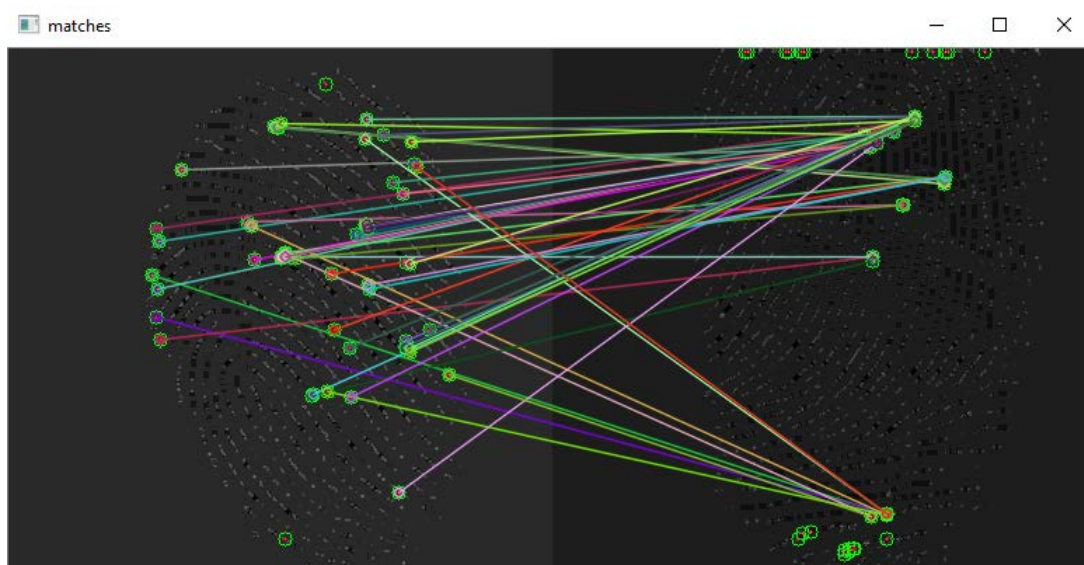
$$g_n(p, \theta) = f_n(p) | (x_i, y_i) \in S_\theta \quad (3.15)$$



## 3.5 Comparação

Identificados e descritos todos os pontos de interesse de várias impressões digitais, resta compará-los. Para isso, foi calculada a distância de Hamming entre os descritores. Ela foi multiplicada pela distância euclidiana entre as posições das minúcias, sendo que os melhores, isto é, aqueles que possuíam menores distâncias entre cada par, foram selecionados. Um dos exemplos é mostrado na **Figura 8** a seguir.

Figura 8 – Comparação de duas imagens, com suas bordas em destaque



## 3.6 Métodos de Aprendizado Não Supervisionado

Nessa seção, será explicado o método de aprendizado não supervisionado utilizado nesse trabalho, chamado RL-Sim\* (OKADA; PEDRONETTE; TORRES, 2015), na Seção 3.6.2, para aprimorar os resultados do processo *CBIR* detalhado nas seções anteriores, bem como uma versão anterior desse algoritmo, chamada RL-Sim (PEDRONETTE; TORRES, 2013), na Seção 3.6.1.

### 3.6.1 RL-Sim

O algoritmo RL-Sim (*Ranked List Similarities*) (PEDRONETTE; TORRES, 2013) utiliza os resultados de algum processo *CBIR*, que retorna uma lista de resultados quando é realizada algum tipo de consulta. A ordem dessa lista é um fator decisivo, pois, é esperado que os resultados nas primeiras posições, isto é, aqueles que são menos distantes da imagem de consulta, tenham maiores probabilidades de estarem corretos.



Assim, o RL-Sim não leva mais em consideração as distâncias das imagens entre si para compará-las, mas, sim, passa a focar na ordem das listas geradas no processo anterior, pois elas possuem importante informação contextual sobre conjuntos de imagens, em vez de apenas sobre uma imagem e outra. Isso é feito com base na suposição de que duas imagens similares, provavelmente, possuirão listas similares. Logo, essa informação é explorada com o objetivo de causar alterações nas ordens previamente estabelecidas, a fim de aproximar, ainda mais, as imagens corretas. Pode-se dizer, então, que as distâncias agora passam a ser entre listas, ao invés de entre imagens.

Considerando-se a "vizinhança" de uma imagem como sendo as  $k$  primeiras imagens ( $Top-k$ ) retornadas em sua lista, visa-se definir um descritor mais preciso, utilizando-se esse valor  $k$  como parâmetro para medida de correlação, em relação ao descritor específico das minúcias. Supondo que esse aumento de precisão ocorrerá, convém realizar esse processo mais de uma vez, definindo-se o valor  $t$  como sendo o número de iterações realizadas, para sucessivamente redefinir a distância entre imagens e aprimorar os resultados. Ao final das iterações, esse novo descritor define as novas distâncias entre as imagens, as quais formam uma lista definitiva.

Seja  $N(i, k)$  a representação do conjunto definido pelas  $k$  primeiras imagens na vizinhança de uma imagem  $img_i$ , a qual possui uma lista ordenada  $\tau_i$ . Antes da execução do RL-Sim,  $\tau_i$  foi definida pelo descritor visual (ORB, no caso desse trabalho). Considerando uma outra imagem,  $img_j$ , com sua respectiva lista  $\tau_j$ , que por sua vez possui vizinhança  $N(j, k)$ , define-se uma métrica de correlação entre as duas listas por  $d(\tau_i, \tau_j, k)$ . Com base em  $d(\tau_i, \tau_j, k)$ , define-se uma nova métrica de distância  $\rho_c(img_i, img_j) = d(\tau_i, \tau_j, k)$ . Uma vez que a entrada e saída do processo podem ser modeladas como listas ordenadas, uma métrica iterativa pode ser definida como :  $\rho_c^{(t+1)}(img_i, img_j) = d(\tau_i^{(t)}, \tau_j^{(t)}, k)$ .

Como as imagens não relevantes são sucessivamente removidas das primeiras posições, espera-se que novas imagens, potencialmente mais similares, tomem seus lugares, assim é razoável aumentar também o valor de  $k$  com o passar das iterações, redefinindo a métrica como  $\rho_c^{(t+1)}(img_i, img_j) = d(\tau_i^{(t)}, \tau_j^{(t)}, k + t)$ .

Realizadas  $t$  iterações, chega-se à representação final da distância entre duas imagens  $\hat{\rho}(img_i, img_j) = \rho_c^{(T)}(img_i, img_j)$ , que será o valor na posição  $(i, j)$  da nova matriz de distâncias, a qual será usada para formar as listas ordenadas finais.

A métrica de distância utilizada foi a de interseção, que sucessivamente incrementa valores de  $k$ , como consta na Equação 3.16.

$$\psi(\tau_i, \tau_j, k) = \frac{\sum_{k_c=1}^k |\mathcal{N}(i, k_c) \cap \mathcal{N}(j, k_c)|}{k} \quad (3.16)$$

### 3.6.2 RL-Sim\*

O algoritmo RL-Sim\* (OKADA; PEDRONETTE; TORRES, 2015) é um aprimoramento sobre o algoritmo RL-Sim (PEDRONETTE; TORRES, 2013), devido ao fato de que as medidas de correlação produzem uma perda de eficácia quando não há interseções nas vizinhanças de duas imagens.

Esse algoritmo divide as listas em três segmentos (OKADA; PEDRONETTE; TORRES, 2015). O primeiro ocorre quando as vizinhanças de duas imagens possuem imagens em comum. Nesse caso, a nova distância entre elas será definida pela medida de correlação, assim como no RL-Sim definida entre 0 e 1. O segundo segmento ocorre quando uma imagem pertence à vizinhança da outra, mas, o restante de suas vizinhanças não possuem interseção. Nessa situação, não se utiliza a medida de correlação, apenas se incrementa o valor anterior por 1. O terceiro segmento comporta as imagens que não pertencem à vizinhança, ou seja, estão posicionadas após as  $k$  primeiras posições da lista da imagem analisada, caso em que a distância anterior é incrementada por 2, garantindo que essas imagens permaneçam ao final das listas, por não serem similares.

## 4 Avaliação Experimental

Nessa seção se apresenta o material usado como base de dados de impressão digital na Seção 4.1, o protocolo seguido na utilização das imagens e na realização dos testes na Seção 4.2, bem como, os resultados quantitativos e visuais gerados por eles nas Seções 4.3 e 4.4, respectivamente.

### 4.1 Coleções de imagens

As coleções utilizadas nesse trabalho foram obtidas dos *sites* de duas edições da *Fingerprint Verification Competition*, que tinha como objetivo avaliar algoritmos comerciais, acadêmicos e/ou independentes, de reconhecimento de impressões digitais. A seleção dessa base de imagens foi feita pelo fato de já ter sido utilizada para comparar uma grande variedade de algoritmos e propiciar a comparação de imagens coletadas ou geradas a partir de diferentes tecnologias e práticas, além de ser de fácil acesso.

#### 4.1.1 FVC 2002

Para a *Fingerprint Verification Competition* do ano de 2002 (FVC2002)([MAIO et al., 2002](#)), foram disponibilizadas 4 coleções, denominadas DB1, DB2, DB3 e DB4, todas com 880 imagens, correspondendo a 110 dedos e 8 coletas para cada dedo. A diferença entre as coleções está relacionada às tecnologias utilizadas para obtenção de suas imagens, sendo elas: o sensor óptico "TouchView II" da Identix, o sensor óptico "FX2000" da Biometrika, o sensor capacitivo "100 SC" da Precise Biometrics, e o "SFinGe v2.51" que é um gerador sintético de impressões digitais ([CAPPELLI, 2004](#)), respectivamente. A nomenclatura das imagens segue o padrão N\_M, onde N (1 a 110) refere-se ao dedo e M (1 a 8) refere-se à coleta.

As coleções utilizadas para testes nesse trabalho são compostas pelas 80 últimas imagens das coleções originais citadas acima, disponíveis para *download* no *site* da FVC2002. Essas parcelas das imagens foram fornecidas aos participantes da competição para fins de refinar os parâmetros dos algoritmos antes que eles fossem submetidos para a avaliação, sendo elas denominadas DB1\_B, DB2\_B, DB3\_B e DB4\_B, respectivamente. Assim, estas 80 imagens representam 10 dedos diferentes com 8 coletas para cada dedo.

#### 4.1.2 FVC 2004

Para a *Fingerprint Verification Competition* do ano de 2004 (FVC2004)([MAIO et al., 2004](#)), foram disponibilizadas 4 coleções, denominadas DB1, DB2, DB3 e DB4, todas

com 880 imagens, correspondendo a 110 dedos e 8 coletas para cada dedo. Inicialmente foram coletadas imagens de 120 dedos e 12 coletas para cada dedo, mas, algumas não foram incluídas posteriormente. Para fins de diferenciação em relação às coleções da FVC2002, nesse documento, elas serão referidas como DB5, DB6, DB7 e DB8, respectivamente. A diferença entre as coleções está relacionada às tecnologias utilizadas para obtenção de suas imagens, sendo elas: o sensor óptico "V300" da CrossMatch, o sensor óptico "U.are.U 4000" da Digital Persona, o sensor térmico "FingerChip FCD4B14CB" da Atmel, e o SFinGe v3.0 que é um gerador sintético de impressões digitais (CAPPELLI, 2004), respectivamente. A nomenclatura das imagens segue o padrão N\_M, onde N (1 a 110) refere-se ao dedo e M (1 a 8) refere-se à coleta.

As coleções utilizadas para testes nesse trabalho são compostas pelas 80 últimas imagens das coleções originais citadas acima, disponíveis para *download* no *site* da FVC2004. Essas parcelas das imagens foram fornecidas aos participantes da competição para fins de refinar os parâmetros dos algoritmos antes que eles fossem submetidos para a avaliação. Assim, estas 80 imagens representam 10 dedos diferentes com 8 coletas para cada dedo. Nesse documento, elas serão referidas como DB5\_B, DB6\_B, DB7\_B e DB8\_B, respectivamente.

Na edição do ano de 2004, foram introduzidas propositalmente distorções nas imagens com a finalidade de diminuir a precisão dos algoritmos (MAIO *et al.*, 2004). Tais distorções incluem um posicionamento ligeiramente mais vertical do dedo do que o padrão, diferenças na pressão exercida durante a coleta, rotações exageradas e alteração na umidade dos dedos.

### 4.1.3 Coleções Compostas

Com a intenção de aumentar a amostragem para testes, as coleções acima foram combinadas para criar outras 2 novas coleções; DB320, que é a união de DB1\_B, DB2\_B, DB3\_B e DB4\_B; e DB640, que é a união de DB1\_B, DB2\_B, DB3\_B, DB4\_B, DB5\_B, DB6\_B, DB7\_B e DB8\_B. A nomenclatura segue o mesmo padrão das coleções anteriores, com a diferença que a numeração foi alterada para não haver conflito (por exemplo, a imagem 101\_1 da DB2\_B passou a ser 111\_1 em DB320 e DB640).

## 4.2 Protocolo Experimental

Dentro de cada coleção, todas as imagens foram utilizadas para consulta. Em posse das matrizes de distâncias correspondentes à cada coleção, sem e com a aplicação do RL-Sim\*, foram calculadas duas métricas para avaliar a precisão dos resultados, a Precisão e a MAP (*Mean Average Precision*) brevemente explicadas nas Seções 4.2.1 e 4.2.2, respectivamente. Para as seções seguintes, considere  $L = (l_1, l_2, l_3, l_4, l_5, \dots, l_{10})$  uma lista

ordenada resultante de algum processo *CBIR* na qual os elementos pares são resultados incorretos (irrelevantes) e os elementos ímpares são resultados corretos (relevantes) da consulta.

### 4.2.1 Precisão

A Precisão é uma métrica calculada a partir da interseção do conjunto de imagens recuperadas com o conjunto de imagens relevantes, como mostra a Equação 4.1, a seguir:

$$\text{Precisão} = \frac{|\{\textit{imagens recuperadas}\} \cap \{\textit{imagens relevantes}\}|}{|\{\textit{imagens recuperadas}\}|} \quad (4.1)$$

Essa métrica também é referida como  $P@n$ , onde  $n$  é o número de imagens recuperadas considerado. Se apenas as 5 primeiras imagens foram consideradas, diz-se que foi calculada a métrica  $P@5$ , e assim por diante. Considerando-se  $L$ , no conjunto  $(l_1, l_2, l_3, l_4, l_5)$  apenas  $l_1$ ,  $l_3$  e  $l_5$  são relevantes, logo, tem-se que  $P@5 = 3/5 = 60\%$ .

### 4.2.2 Mean Average Precision

A *Mean Average Precision* (MAP) considera a média das precisões de cada imagem relevante. Em outras palavras, cada vez que uma imagem relevante é encontrada, calcula-se sua precisão da mesma forma que foi calculada na Equação 4.1. Em seguida, divide-se a soma desses valores pelo número de imagens relevantes. Considerando-se  $L$ , tem-se :  $MAP = (1/1 + 2/3 + 3/5 + 4/7 + 5/9)/5 \cong 68\%$

## 4.3 Resultados Quantitativos

As tabelas a seguir mostram as medidas de eficácia do algoritmo aplicado para todas as coleções apresentadas na Seção 4.1. A **Tabela 2** mostra os resultados da avaliação quando não foi utilizado o RL-Sim\*. Em seguida, serão mostrados os resultados com o RL-Sim\* como pós-processamento, para as mesmas coleções, variando-se os valores de  $K$ , isto é, da vizinhança considerada dentro de cada lista, para 10, 15 e 20, respectivamente nas **Tabelas 3, 4 e 5**. Para todas elas, 3 iterações ( $t = 3$ ) foram realizadas. Os maiores valores para MAP para cada coleção foram destacados em negrito. Na última coluna das tabelas é apresentado o tempo de execução (exec) da fase posterior ao cálculo das distâncias das imagens; isto é, a soma dos tempos da leitura da matriz de distâncias, da execução do RL-Sim\* (quando utilizado), e da avaliação dos resultados.

Tabela 2 – Resultados sem o RL-Sim\*

	MAP	P@10	P@15	P@20	exec
DB1_B	44.70%	32.13%	25.08%	21.31%	59 ms
DB2_B	42.88%	32.12%	24.50%	20.69%	47 ms
DB3_B	33.43%	22.87%	18.17%	14.87%	38 ms
DB4_B	31.32%	19.62%	16.83%	15.44%	33 ms
DB5_B	39.20%	29.00%	22.42%	18.50%	41 ms
DB6_B	32.75%	22.50%	19.67%	16.81%	32 ms
DB7_B	48.33%	37.12%	28.83%	23.50%	39 ms
DB8_B	34.67%	22.75%	18.50%	17.44%	32 ms
DB320	26.85%	20.25%	15.19%	12.38%	656 ms
DB640	23.13%	18.33%	13.55%	10.81%	2087 ms

Tabela 3 – Resultados com o RL-Sim\* para K = 10

	MAP	P@10	P@15	P@20	exec
DB1_B	55.16%	42.13%	31.50%	25.56%	62 ms
DB2_B	<b>52.29%</b>	40.38%	31.00%	24.38%	58 ms
DB3_B	38.73%	27.37%	22.58%	19.88%	60 ms
DB4_B	30.79%	21.75%	17.33%	15.19%	62 ms
DB5_B	41.67%	29.75%	24.67%	21.75%	62 ms
DB6_B	33.27%	24.50%	18.42%	15.75%	67 ms
DB7_B	74.37%	55.63%	40.67%	32.31%	58 ms
DB8_B	40.81%	30.75%	24.50%	21.06%	71 ms
DB320	33.51%	25.00%	19.35%	16.06%	1185 ms
DB640	<b>29.33%</b>	22.66%	16.61%	13.61%	5383 ms

Tabela 4 – Resultados com o RL-Sim\* para K = 15

	MAP	P@10	P@15	P@20	exec
DB1_B	53.34%	38.88%	30.17%	25.62%	85 ms
DB2_B	48.94%	37.50%	28.83%	23.31%	88 ms
DB3_B	39.42%	29.00%	24.83%	22.19%	77 ms
DB4_B	30.55%	21.00%	17.83%	15.75%	83 ms
DB5_B	41.23%	29.12%	24.42%	22.06%	82 ms
DB6_B	34.22%	25.75%	20.42%	17.56%	74 ms
DB7_B	<b>75.04%</b>	55.13%	39.42%	32.56%	85 ms
DB8_B	<b>42.41%</b>	30.37%	25.25%	22.31%	84 ms
DB320	<b>33.61%</b>	24.97%	20.31%	17.06%	1502 ms
DB640	28.83%	22.39%	16.71%	13.73%	5189 ms

Tabela 5 – Resultados com o RL-Sim\* para  $K = 20$ 

	MAP	P@10	P@15	P@20	exec
DB1_B	<b>55.76%</b>	41.38%	33.25%	27.94%	106 ms
DB2_B	45.32%	33.00%	27.25%	23.00%	94 ms
DB3_B	<b>40.48%</b>	28.62%	24.92%	22.81%	99 ms
DB4_B	<b>34.58%</b>	24.37%	20.25%	18.38%	109 ms
DB5_B	<b>42.30%</b>	30.38%	24.83%	22.37%	97 ms
DB6_B	<b>35.88%</b>	26.25%	22.08%	18.31%	103 ms
DB7_B	74.02%	55.25%	41.58%	32.50%	103 ms
DB8_B	40.85%	30.88%	25.50%	22.88%	97 ms
DB320	32.91%	24.91%	19.85%	16.75%	1707 ms
DB640	28.83%	22.56%	16.64%	13.52%	6208 ms

Observa-se que os valores das métricas foram superiores para quase todas as coleções quando foi utilizado o RL-Sim\* como pós-processamento, uma vez que nenhum dos maiores valores de MAP está na **Tabela 2**. Em particular, observa-se que para a coleção DB4\_B, os valores de  $k=10$  e  $k=15$  deterioraram os resultados, sendo piores do que quando não foi aplicado o RL-Sim\*, mas com  $k=20$ , atingiu seu maior valor de MAP. A melhora mais notável ocorreu para a coleção DB7\_B, que tinha 48.33% na MAP sem o RL-Sim\*, mas que atingiu 75.04% nessa mesma métrica quando  $k = 15$ . Além disso, as coleções compostas naturalmente apresentariam valores menores de modo geral, pois o número de imagens foi multiplicado por 4 e 8, nas coleções DB320 e DB640 respectivamente, mas o número de imagens para cada dedo (8) permaneceu constante, ou seja, a proporção de imagens incorretas aumentou expressivamente. Também é visível que, dentro de pequenas variações, quanto maior a vizinhança e/ou coleção considerada(s), naturalmente mais tempo foi necessário para computar o processo completamente.









































































## 4.4 Resultados Visuais

A seguir serão mostradas tabelas que mostram os primeiros 8 resultados de consultas pelas 8 primeiras impressões digitas da coleção DB1\_B. Os nomes das impressões digitais na cor verde representam impressões digitais pertencentes à mesma pessoa (acertos), enquanto que nomes na cor vermelha representam impressões de uma outra pessoa (erros). A porcentagem de acertos é exibida na última coluna das tabelas, enquanto que as legendas fornecem a média simples dessas porcentagens. A vizinhança considerada foi de 15 imagens ( $k = 15$ ) e 3 iterações ( $t = 3$ ) foram realizadas em ambas.

Na **Tabela 6 na próxima página** constam os resultados para consultas sem a aplicação do RL-Sim\*.



Tabela 6 – Consultas sem uso do RL-Sim\*









































































 101_1	 101_3	 101_6	 106_6	 108_7	 106_5	 101_2	 101_7	 106_3	50.00%
 101_2	 101_3	 107_8	 101_1	 103_6	 105_8	 106_6	 106_5	 104_6	25.00%
 101_3	 101_6	 103_6	 101_1	 107_6	 106_3	 108_7	 106_2	 101_2	37.50%
 101_4	 101_6	 106_3	 103_6	 104_7	 104_6	 108_7	 101_3	 106_4	25.00%
 101_5	 101_3	 106_5	 104_6	 101_1	 106_7	 107_7	 103_6	 104_7	25.00%
 101_6	 101_3	 106_3	 101_7	 101_1	 107_6	 108_7	 103_6	 106_2	37.50%
 101_7	 101_6	 107_6	 108_7	 106_3	 106_6	 106_5	 104_5	 101_1	25.00%
 101_8	 101_3	 106_2	 101_1	 101_6	 101_2	 101_7	 108_7	 106_5	62.50%

Porcentagem média : 35,93%

Na **Tabela 7** constam os resultados para buscas pelas mesmas 8 imagens do DB1\_B, agora com a aplicação do RL-Sim\*.



Tabela 7 – Consultas com uso do RL-Sim\*

 101_1	 101_8	 101_3	 106_8	 101_6	 101_7	 101_2	 110_6	 101_5	75.00%
 101_2	 101_1	 101_8	 106_8	 101_3	 101_6	 101_7	 101_5	 110_6	75.00%
 101_3	 101_6	 101_8	 101_1	 101_7	 106_8	 110_6	 101_2	 101_4	75.00%
 101_4	 101_7	 110_6	 104_8	 109_3	 110_8	 101_6	 105_5	 101_3	37.50%
 101_5	 102_6	 103_6	 107_2	 103_4	 102_1	 107_3	 103_1	 103_3	0.00%
 101_6	 101_7	 101_3	 101_8	 101_1	 110_6	 106_8	 101_2	 101_4	75.00%
 101_7	 101_6	 101_3	 101_8	 101_1	 110_6	 106_8	 101_2	 101_4	75.00%
 101_8	 101_1	 101_3	 106_8	 101_6	 101_7	 101_2	 110_6	 101_4	75.00%

Porcentagem média : 60,93

É visível o aumento no número de imagens corretas na segunda tabela, mais especificamente nas primeiras posições, evidenciado também pela média das porcentagens. No entanto, para a imagem 101\_5, na célula destacada foi observada uma diminuição no número de acertos, ao ponto que as 8 primeiras imagens retornadas são incorretas. Uma provável causa é a baixa qualidade da imagem em questão, na qual parte da impressão digital não está visível, que já sugeria que ela apresentaria resultados imprecisos a princípio.

## 5 Conclusão

Nesse trabalho foi apresentado o tema da biometria de impressão digital, bem como, uma série de processos utilizados para a sua identificação, incluindo-se uma abordagem que tem como objetivo aprimorar os resultados desses processos.

Observou-se que a **Tabela 7** possui muito mais acertos que a **Tabela 6**, especialmente nas primeiras posições. Adicionalmente, as métricas calculadas, analisando-se todas as imagens como busca, também evidenciam que os resultados foram mais precisos quando o RL-Sim\* foi utilizado. Assim, verificou-se que o RL-Sim\* é uma alternativa interessante à identificação biométrica por meio de impressões digitais e, em relação à formulação de novos algoritmos, é menos trabalhosa.

Apesar disso, alguns resultados específicos podem ser piorados com a aplicação do RL-Sim\*, como a consulta pela imagem 101\_5 mostrada na **Tabela 7**, na qual as 8 primeiras posições da lista foram preenchidas com imagens incorretas, ao contrário dos 2 acertos na **Tabela 6**, bem como da coleção DB4\_B nas **Tabelas 3 e 4**, mostrando que tanto a qualidade da imagem quanto a vizinhança considerada no RL-Sim\* também podem piorar os resultados.

Por não estar restrito à modelagem da área em questão, o RL-Sim\* pode ser usado em outras aplicações que forneçam listas de resultados ordenadas pelas suas distâncias, seguindo a abordagem de recuperação de imagens baseada em seus conteúdos.

## Referências

- ALI, M. M. H. et al. Fingerprint recognition for person identification and verification based on minutiae matching. In: *2016 IEEE 6th International Conference on Advanced Computing (IACC)*. [S.l.: s.n.], 2016. p. 332–339. Citado 3 vezes nas páginas 6, 8 e 9.
- BAGGENSTOSS, P. M. A modified baum-welch algorithm for hidden markov models with multiple observation spaces. *IEEE Transactions on Speech and Audio Processing*, v. 9, n. 4, p. 411–416, May 2001. ISSN 1063-6676. Citado na página 14.
- BORGEN, H.; BOURS, P.; WOLTHUSEN, S. D. Visible-spectrum biometric retina recognition. In: IEEE. *Intelligent Information Hiding and Multimedia Signal Processing, 2008. IHHMSP'08 International Conference on*. [S.l.], 2008. p. 1056–1062. Citado na página 9.
- BORRA, S. R.; REDDY, G. J.; REDDY, E. S. A broad survey on fingerprint recognition systems. In: *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. [S.l.: s.n.], 2016. p. 1428–1434. Citado na página 11.
- BREVE, F. A.; ZHAO, L.; QUILES, M. G. Particle competition and cooperation for semi-supervised learning with label noise. *Neurocomputing*, v. 160, n. Supplement C, p. 63 – 72, 2015. ISSN 0925-2312. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0925231215001277>>. Citado na página 14.
- CALONDER, M. et al. Brief: Binary robust independent elementary features. *Computer Vision–ECCV 2010*, Springer, p. 778–792, 2010. Citado na página 21.
- CAPPELLI, R. Sfinge: an approach to synthetic fingerprint generation. 01 2004. Citado 2 vezes nas páginas 25 e 26.
- CARNEIRO, L. N. das V. Reconhecimento de face invariante a iluminação baseado em uma abordagem supervisionada. *Universidade Federal de Ouro Preto*, 2012. Citado na página 8.
- CHEN, W. et al. Improved zhang-suen thinning algorithm in binary line drawing applications. In: IEEE. *Systems and Informatics (ICSAI), 2012 International Conference on*. [S.l.], 2012. p. 1947–1950. Citado na página 16.
- CUSTÓDIO, K. V. A. Estudo do uso de biometria para autenticação em terminais de auto-atendimento. *UNICEUB - Centro Universitário de Brasília*, 2007. Disponível em: <<http://repositorio.uniceub.br/bitstream/123456789/3127/2/20168377.pdf>> Citado 2 vezes nas páginas 8 e 9.
- DAUGMAN, J. How iris recognition works. *IEEE Transactions on circuits and systems for video technology*, IEEE, v. 14, n. 1, p. 21–30, 2004. Citado na página 9.
- ENSER, P. Visual image retrieval: seeking the alliance of concept-based and content-based paradigms. *Journal of Information Science*, Sage Publications Sage CA: Thousand Oaks, CA, v. 26, n. 4, p. 199–210, 2000. Citado na página 12.

- FERREIRA, F. H. B. et al. Aplicação de identificação biométrica por impressão digital na lavratura do auto de prisão em flagrante. Araranguá, SC, 2017. Citado 2 vezes nas páginas 6 e 9.
- GHAHRAMANI, Z. Unsupervised learning. In: *Advanced lectures on machine learning*. [S.l.]: Springer, 2004. p. 72–112. Citado 3 vezes nas páginas 7, 13 e 14.
- GUDIVADA, V.; RAGHAVAN, V. V. Content-based image retrieval systems. In: *Computer*. [S.l.: s.n.], 1995. v. 28, p. 18 – 22. Citado 3 vezes nas páginas 7, 9 e 12.
- HARRIS, C.; STEPHENS, M. A combined corner and edge detector. In: MANCHESTER, UK. *Alvey vision conference*. [S.l.], 1988. v. 15, n. 50, p. 10–5244. Citado 3 vezes nas páginas 19, 20 e 21.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. Overview of supervised learning. In: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer New York, 2009. p. 9–41. ISBN 978-0-387-84858-7. Disponível em: <[https://doi.org/10.1007/978-0-387-84858-7\\_2](https://doi.org/10.1007/978-0-387-84858-7_2)>. Citado na página 13.
- HOWSE, J. et al. Unsupervised learning. In: *OpenCV 3 Blueprints*. [S.l.: s.n.], 2015. Citado 5 vezes nas páginas 7, 12, 15, 16 e 19.
- JAIN, A.; HONG, L.; PANKANTI, S. Biometric identification. *Communications of the ACM*, ACM, v. 43, n. 2, p. 90–98, 2000. Citado na página 8.
- LEE, H. D. *Seleção de atributos importantes para a extração de conhecimento de bases de dados*. Tese (Doutorado) — Universidade de São Paulo, 2005. Citado na página 13.
- MAIO, D. et al. Fvc2002: Second fingerprint verification competition. In: IEEE. *Pattern recognition, 2002. Proceedings. 16th international conference on*. [S.l.], 2002. v. 3, p. 811–814. Citado 3 vezes nas páginas 6, 10 e 25.
- MAIO, D. et al. Fvc2004: Third fingerprint verification competition. *Biometric Authentication*, Springer, p. 31–35, 2004. Citado 3 vezes nas páginas 6, 25 e 26.
- MÜLLER, H. et al. A review of content-based image retrieval systems in medical applications—clinical benefits and future directions. *International journal of medical informatics*, Elsevier, v. 73, n. 1, p. 1–23, 2004. Citado na página 12.
- NUNES, F. T. et al. Técnicas de biometria baseadas em padrões faciais e sua utilização na segurança pública. Araranguá, SC, 2017. Citado 3 vezes nas páginas 6, 9 e 10.
- OKADA, C. Y.; PEDRONETTE, D. C. G.; TORRES, R. da S. Unsupervised distance learning by rank correlation measures for image retrieval. In: ACM. *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*. [S.l.], 2015. p. 331–338. Citado 4 vezes nas páginas 7, 14, 22 e 24.
- OTSU, N. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, IEEE, v. 9, n. 1, p. 62–66, 1979. Citado na página 16.
- PEDRONETTE, D. C. G.; TORRES, R. d. S. Image re-ranking and rank aggregation based on similarity of ranked lists. *Pattern Recognition*, Elsevier, v. 46, n. 8, p. 2350–2360, 2013. Citado 2 vezes nas páginas 22 e 24.

- RAJPUT, S.; SURALKAR, S. Comparative study of image enhancement techniques. *International Journal of Computer Science and Mobile Computing-A Monthly Journal of Computer Science and Information Technology*, v. 2, n. 1, p. 11–21, 2013. Citado 3 vezes nas páginas 6, 9 e 11.
- ROSTEN, E.; DRUMMOND, T. Machine learning for high-speed corner detection. *Computer Vision–ECCV 2006*, Springer, p. 430–443, 2006. Citado na página 20.
- RUBLEE, E. et al. Orb: An efficient alternative to sift or surf. In: IEEE. *Computer Vision (ICCV), 2011 IEEE international conference on*. [S.l.], 2011. p. 2564–2571. Citado 2 vezes nas páginas 20 e 21.
- SÁNCHEZ, C. I. et al. A novel automatic image processing algorithm for detection of hard exudates based on retinal image analysis. *Medical Engineering & Physics*, Elsevier, v. 30, n. 3, p. 350–357, 2008. Citado na página 9.
- SIMÕES, A. d. S.; COSTA, A. H. R. Aprendizado não-supervisionado em redes neurais pulsadas de base radial: um estudo da capacidade de agrupamento para a classificação de pixels. *Sba: Controle & Automação Sociedade Brasileira de Automatica*, SciELO Brasil, v. 18, n. 2, p. 251–264, 2007. Citado na página 14.
- SUTTON, R. S. et al. Policy gradient methods for reinforcement learning with function approximation. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2000. p. 1057–1063. Citado na página 14.
- THAIYALNAYAKI, K. et al. Finger print recognition using discrete wavelet transform. *International Journal of Computer Applications*, International Journal of Computer Applications, 244 5 th Avenue, # 1526, New York, NY 10001, USA India, v. 1, n. 24, p. 96–100, 2010. Citado na página 12.
- TICO, M. et al. Fingerprint recognition using wavelet features. In: *ISCAS 2001. The 2001 IEEE International Symposium on Circuits and Systems (Cat. No.01CH37196)*. [S.l.: s.n.], 2001. v. 2, p. 21–24 vol. 2. Citado na página 12.
- TORRALBA, A. et al. Context-based vision system for place and object recognition. In: IEEE. *null*. [S.l.], 2003. p. 273. Citado na página 12.
- YONAMINE, F. S. et al. Aprendizado não supervisionado em domínios fuzzy–algoritmo fuzzy c-means. *São Carlos: UFSCAR*, 2002. Citado na página 14.
- ZHANG, T.; SUEN, C. Y. A fast parallel algorithm for thinning digital patterns. *Communications of the ACM*, ACM, v. 27, n. 3, p. 236–239, 1984. Citado 3 vezes nas páginas 16, 17 e 18.
- ZHANG, X.; WANG, Y.; ZHAO, Z. A hybrid speech recognition training method for hmm based on genetic algorithm and baum welch algorithm. In: *Second International Conference on Innovative Computing, Informatio and Control (ICICIC 2007)*. [S.l.: s.n.], 2007. p. 572–572. Citado na página 14.
- ZHU, X. *Semi-Supervised Learning Literature Survey*, *Computer Sciences, University of Wisconsin-Madison*. [S.l.]n. 1530, 2005. Citado na página 14.