



Hand Vein Biometric Recognition Approaches Based on Wavelet, SVM, Artificial Neural Network and Image Registration

Daniel Felix de Brito^[1,A] and Lee Luan Ling^[1,B]

^[1]Faculty of Electrical and Computer Engineering (FEEC), State University of Campinas (UNICAMP), Campinas-SP, Brazil.

^[A]fbrito.daniel@gmail.com

^[B]lee@decom.fee.unicamp.br

Abstract This paper describes in detail different hand vein recognition methods based on Wavelet-SVM, Wavelet-ANN and Image Registration. A new image segmentation and processing algorithm is proposed to efficiently locate vein regions and suitable for feature extraction (wavelet coefficients and normalized vein images) and classification (SVM, ANN and Image Registration). For real time recognition and high recognition rate, we proposed an integrated system which combines three above mentioned classification methods. The simulation results reveal that the proposed integrated system achieves 1% *false rejection rate* (FRR) and 0.02% *false acceptance rate* (FAR).

Resumo Este artigo descreve em detalhes diferentes métodos para o reconhecimento biométrico com base nas veias das mãos. Os métodos têm como base Wavelets com SVM, Wavelets com Rede Neural Artificial e Registro de Imagens. Um novo método de segmentação de imagens é proposto para localizar de modo adequado e eficiente a região das veias, permitindo a extração de características (com Wavelets ou a partir de uma imagem normalizada) e a classificação (com SVM, Rede Neural e Registro de Imagens). Para uma alta taxa de reconhecimento realizada em tempo real nós propomos um sistema híbrido que combina os três métodos de classificação mencionados. Os resultados de testes revelam que o sistema híbrido fornece uma *taxa de falsa rejeição* (FRR) de 1% para uma *taxa de falsa aceitação* (FAR) de 0.02%.

Keywords: Hand vein biometric, Support Vector Machine, Artificial Neural Network, Image Registration, Wavelet.

Palavras-chave: Biometria de veias das mãos, Máquina de vetores de suporte, Rede Neural Artificial, Registro de Imagens.

1 Introdução

Biometria pode ser definida como uma técnica de reconhecimento de padrões que identifica um indivíduo através de suas características fisiológicas ou comportamentais [1]. Entre as características biométricas mais utilizadas estão: impressão digital, íris, assinaturas, reconhecimento facial, veias das mãos, sendo que cada tipo biométrico apresenta certas vantagens e desvantagens quando aplicadas na prática [2].

Atualmente, o uso da biometria para identificação pessoal é uma técnica amplamente aplicada em diversas áreas envolvendo o acesso à segurança de dados e a confirmação de identidade pessoal no cotidiano. Podemos citar alguns exemplos de seu uso: caixas eletrônicos; smartphones, tablets e notebooks; controle

eletrônico de acesso aos locais de trabalho; identificação de usuários de planos de saúde. Em alguns casos, para aumentar ainda mais as taxas de correta identificação individual e o grau de confiabilidade no sistema que identificará o indivíduo, são utilizados sistemas biométricos multimodais, envolvendo mais de uma característica biométrica na identificação, como por exemplo, um sistema capturar a imagem da íris e da impressão digital e utilizar ambas para a identificação [3][4][5][6].

Neste projeto, utilizamos as veias das mãos como característica biométrica. Um sistema biométrico que utiliza a imagem das veias de uma das mãos considera o fato da hemoglobina, presente no sangue que flui nas veias, ser sensível à luz infravermelha. Sendo assim, ao iluminar a mão com uma luz infravermelha faz com que a hemoglobina do sangue absorva parte da luz e destaque o formato e a localização das veias, o que pode ser observado em uma imagem capturada por uma câmera infravermelha [7]. Logo, a variação dos formatos e das localizações das veias na região iluminada com a luz infravermelha permite a diferenciação entre os indivíduos, formando padrões que permitem determinar exclusivamente um indivíduo com base nessa característica biométrica. Na literatura, diversos estudos têm sido realizados com base nas veias das mãos para a autenticação pessoal [8][9][10][11].

O intenso interesse por essa área de pesquisa é justificado por algumas vantagens do uso da biometria de veias em relação aos demais tipos biométricos (íris, impressão digital, face, entre outros), incluindo: alta variabilidade inter-classes e baixa variabilidade intra-classes; baixo custo para a montagem de um protótipo para aquisição de imagens das veias; sem contato físico entre o equipamento de aquisição e a característica biométrica (veias), o que poderia interferir na correta captura, como é o caso da impressão digital; alto desempenho em termos de taxas de acerto no processo de reconhecimento.

Um sistema biométrico com base nas veias das mãos normalmente é composto de quatro unidades funcionais em cascata, ilustradas na Figura 1.



Figura 1: Fluxograma das etapas de um sistema biométrico de veias.

- **Aquisição da imagem:** através de um dispositivo de captura, posicionar a mão e iluminar a região do dorso ou da palma com luz infravermelha e capturar uma imagem digital da região iluminada;
- **Localização das veias na imagem:** utilizando técnicas de processamento de imagens digitais, dividir a imagem capturada em regiões e encontrar a região que representa as veias;
- **Normalização:** mapear a região das veias, que pode ter diferentes dimensões de acordo com cada captura, sempre para uma imagem de dimensões $N \times N$;
- **Extração das características:** utilizar uma técnica de reconhecimento de padrões que possibilite a diferenciação entre as imagens das veias entre indivíduos distintos;
- **Armazenamento de características de referências:** armazenar em uma base de dados as características extraídas para posterior comparação;
- **Identificação ou autenticação:** comparar o padrão atual com o padrão capturado previamente e armazenado em uma base de dados.

O processo realizado em cada unidade é extremamente relevante para o sistema como um todo, influenciando o resultado das próximas unidades e o desempenho global do sistema. Métodos e algoritmos de processamento mais simples requerem a colaboração do usuário no correto posicionamento da mão para uma correta captura da imagem. Por outro lado, as abordagens mais sofisticadas consideram algum tipo de posicionamento incorreto pelo usuário e efetuam os ajustes necessários para identificar corretamente.

As seções seguintes apresentam em detalhes os procedimentos para cada etapa. Na seção 2 são apresentadas características gerais de um dispositivo de captura e uma ilustração de sua estrutura interna. Na seção 3 são descritos os filtros utilizados para melhorar a qualidade da imagem digital adquirida, os

procedimentos para dividir a imagem em regiões e assim localizar o contorno da mão seguida da localização das veias. Na seção 3.6 são descritos os procedimentos para transformar as regiões de veias de imagens capturadas em momentos distintos, seja da mesma pessoa ou de pessoas diversas, sempre em uma região de mesma dimensão. A seção 4 apresenta as seguintes técnicas para extrair padrões de características das regiões das veias: Transformada Wavelet com SVM, Transformada Wavelet com Redes Neurais e Registro de Imagens. Por fim, as seções 5 e 6 apresentam os procedimentos adotados na implementação de todos os métodos apresentados na seções anteriores, bem como discussões sobre os resultados obtidos e as conclusões finais.

2 Aquisição da Imagem

A unidade de aquisição da imagem consiste na captura de imagens das veias da parte superior (dorso) ou inferior (palma) da mão utilizando um dispositivo de captura. A estrutura geral de um dispositivo de captura é ilustrada na Figura 2, onde leds infravermelhos (normalmente na faixa de 850 nm) são usados para iluminar a região de interesse das veias. Nessa faixa de frequência, a hemoglobina absorve parte da energia da luz fazendo com que as veias sejam mais nitidamente realçadas na imagem capturada por uma câmera CCD comum ou uma câmera infravermelha. Uma câmera CCD comum tem menor custo, porém menor qualidade de imagem, em comparação com o custo e a qualidade de uma câmera infravermelha.

A Figura 3 ilustra imagens capturadas por uma câmera CCD¹. Para uma câmera CCD, é necessário remover o filtro infravermelho que está presente nas câmeras comerciais e utilizar um filtro de luz visível. A distância entre a mão e a câmera costuma variar entre 8 a 20 cm.

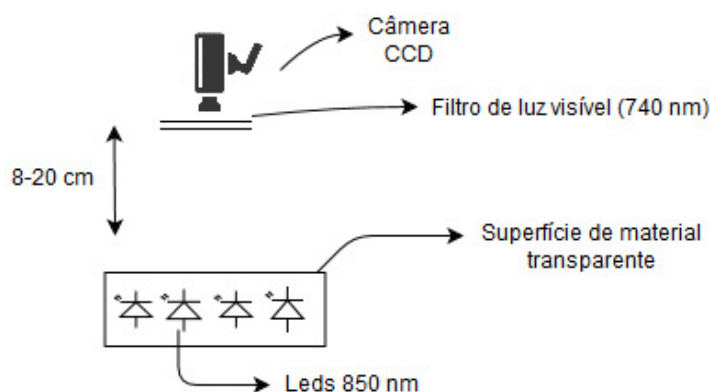


Figura 2: Visão geral de um protótipo da estrutura de um dispositivo de captura de imagens de veias da mão. Nesse caso, a palma da mão deve ser posicionada sobre a superfície de material transparente, com a palma para baixo.

Detalhes de alguns protótipos de dispositivos de aquisição podem ser encontrados em [7][9][10][11]. Um estudo mais completo sobre a aquisição de imagens das veias e o uso de feixes de luz próximos ao infravermelho é feito por Wang *et al.*[13].

Alternativamente ao uso de um dispositivo de captura, pode-se obter a imagem a partir de um banco de imagens já capturadas. Este segundo caso normalmente ocorre nos estudos na literatura quando o foco do estudo não é a aquisição da imagem em si, mas uma ou mais das outras etapas do processo de reconhecimento, como, por exemplo, apresentar uma nova técnica para a extração de características.

Neste estudo, utilizamos as imagens do banco de dados compilado por Badawi [12], que possui imagens capturadas da região do dorso das mãos. Três motivos principais levaram a escolha desse banco de imagens dentre os disponíveis:

1. O banco de imagens é utilizado por outros pesquisadores na literatura, facilitando a comparação dos resultados de testes com artigos do Estado da Arte;

¹Imagens do banco de dados [12]

2. As imagens foram capturadas por uma imagem CCD, o que apresenta um maior desafio na identificação uma vez que as imagens são de menor qualidade em relação a imagens de outros banco de dados disponíveis;
3. A quantidade de imagens disponíveis por pessoa permite separá-las em dois conjuntos: um conjunto de treino e um conjunto de teste;
4. Alta variabilidade dos voluntários: homens e mulheres com idades entre 18 e 65 anos.

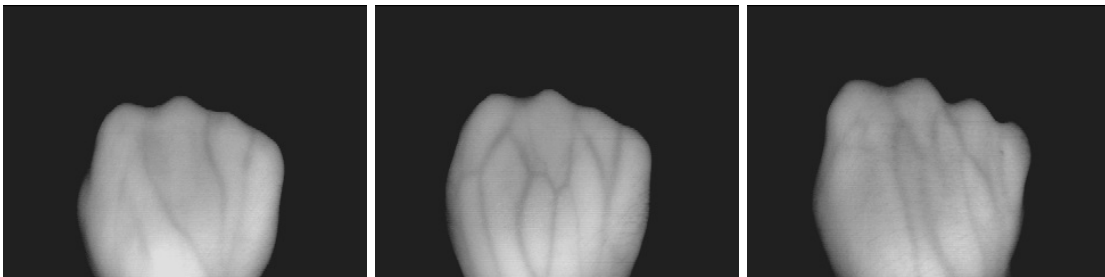


Figura 3: Exemplos de imagens de veias do dorso da mão.

3 Pré-Processamento: Segmentação e Normalização

A segunda etapa do sistema biométrico desenvolvido trata-se da segmentação da imagem capturada com o objetivo de localizar e delimitar a região de veias na imagem capturada. Para atingir este objetivo, a técnica desenvolvida, em linhas gerais, determina o contorno da mão na imagem e posteriormente, com base neste contorno localizado, identifica os pixels da imagem que contêm a região das veias. Conforme ilustradas na Figura 4, o procedimento completo implementado neste trabalho subdivide-se em quatro fases: Filtragem para melhoria da imagem (Filtro Mediana, filtragem no domínio da frequência e ajuste de contraste); Geração de uma imagem binária; Determinação da região de interesse; e Eliminação de ruídos.

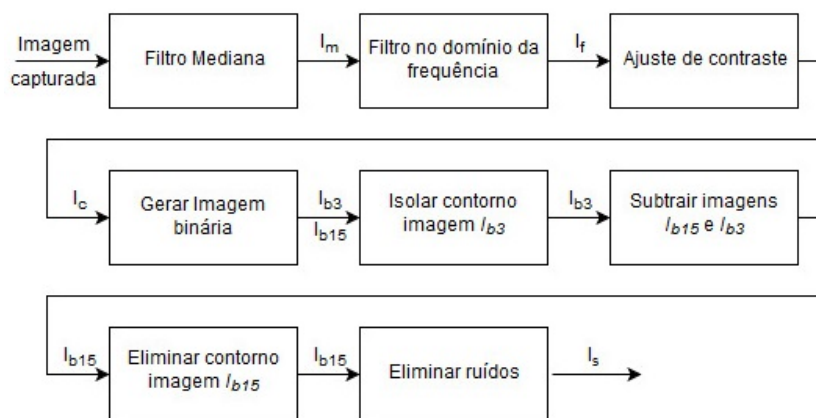


Figura 4: Visão geral das etapas do processo de segmentação da imagem.

3.1 Filtragem

Esta etapa de filtragem tem como objetivos melhorar a qualidade geral da imagem através da redução de ruídos gerados na fase de aquisição e realçar as bordas na imagem, em especial as bordas que formam o contorno da mão e as veias. Este procedimento de realce é fundamental para o próximo passo de processamento da imagem que é a determinação da região da mão utilizando bordas realçadas presentes na imagem, bem como localizar as veias. O processo de melhoria da qualidade da imagem consiste na aplicação de dois filtros e no ajuste de contraste.

3.1.1 Filtro Mediana

O primeiro filtro aplicado é um Filtro Mediana. Este filtro tem como objetivo eliminar ruídos gerados no processo de aquisição da imagem, ocasionados principalmente por: variações da luz ambiente e da incidência de luz solar; variações de temperatura; acúmulo de poeira na lente; defeitos de fábrica em transdutores fotossensíveis que formam o sensor de captura. Em geral, o ruído gera valores de brilho que tendem a ser muito diferentes dos valores de brilho de seus vizinhos adjacentes.

Basicamente, o procedimento de filtragem é feito da seguinte forma: para cada pixel p da imagem, ordena-se, em ordem crescente, uma lista com os valores de brilho dos $n \times n$ pixels adjacentes a p . Considera-se n um valor ímpar. O valor de brilho do pixel p é substituído pelo valor de brilho da posição mediana da lista que foi ordenada.

A Figura 5 ilustra este procedimento, onde o pixel central da região em análise, com valor de brilho igual a 7 (Figura 5a) receberá o valor 8 (Figura 5b), após a ordenação dos valores de brilho de sua região vizinha (Figura 5c).

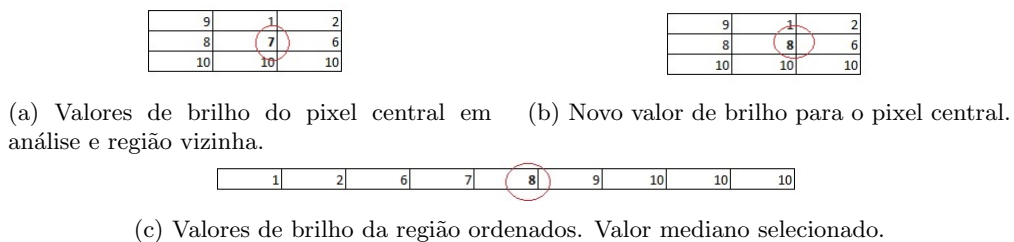


Figura 5: Exemplo da escolha de um novo valor de brilho para um Filtro Mediana de dimensão 3x3.

Foram testados diversos valores de n para a dimensão do filtro da mediana e feita a análise visual das imagens resultantes. Os melhores resultados foram obtidos com o valor de n igual a 3, ou seja, um filtro de dimensão 3x3, o qual reduziu os ruídos sem distorcer as bordas presentes na imagem, o que ocorreu com a aplicação de filtros de dimensões maiores, como 15x15, 21x21.

Ao final desta etapa, temos uma imagem com os ruídos reduzidos, com regiões mais uniformes, ou seja, os valores de brilhos dos pixels de cada região estão muito próximos ao valor de brilho médio da região. A imagem resultante é denominada I_m , conforme o fluxograma apresentado na Figura 4. As Figuras 6a e 6b reproduzem uma imagem original e a imagem após a aplicação do Filtro da Mediana (imagem I_m), respectivamente.



Figura 6: Exemplos de imagens com o resultado da aplicação de filtros Mediana.

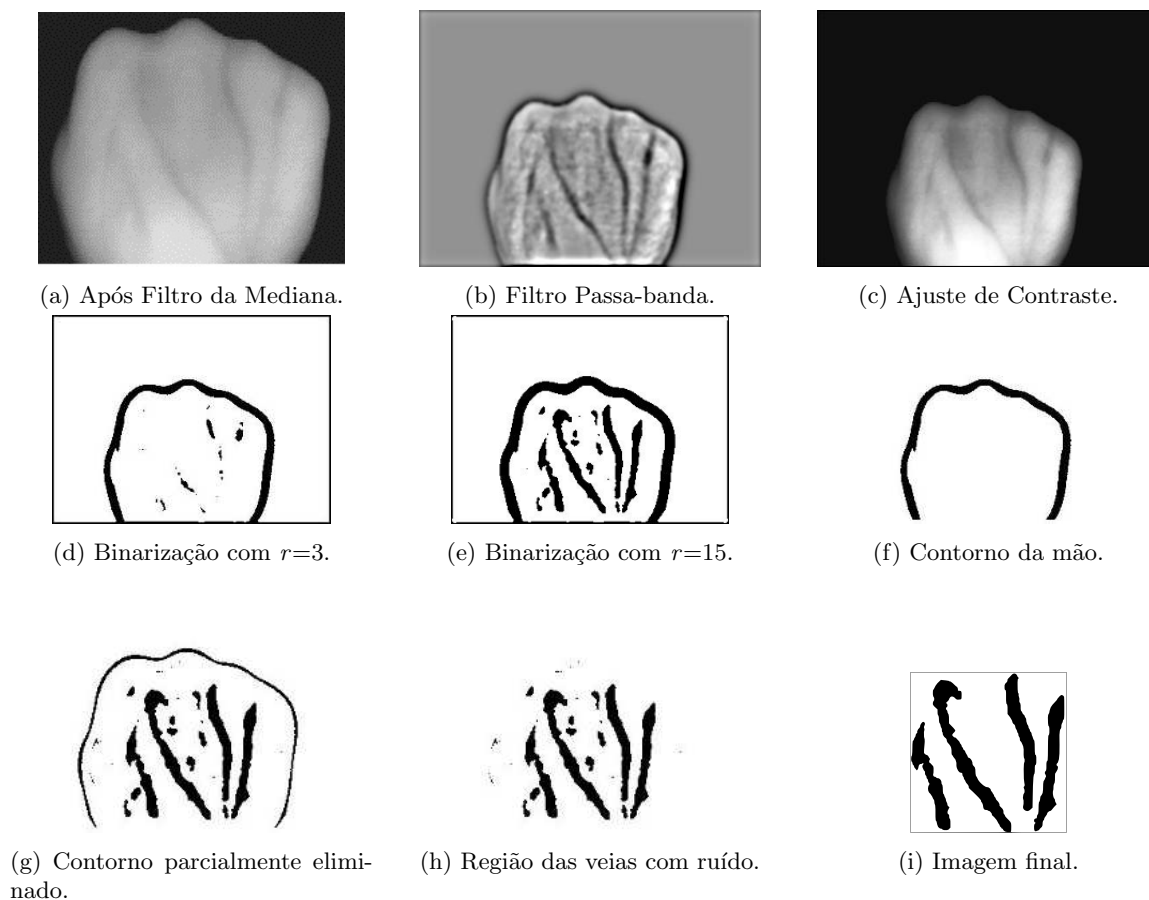


Figura 7: Imagens resultantes de cada etapa do processo de segmentação.

3.1.2 Filtragem no domínio da frequência

O segundo filtro aplicado tem como objetivo realçar as bordas da região que formam as veias. Trata-se de um filtro passa-banda ou passa-faixa (*FFT bandpass filter*), sendo aplicado no domínio da frequência. Este filtro permite a passagem das frequências de uma certa faixa de valores, enquanto rejeita o restante. Nesse projeto utilizamos a implementação *open-source* disponibilizada no projeto ImageJ [14]. Para utilizar esta implementação, é necessário definir alguns parâmetros. Os dois principais parâmetros são os valores que determinam a faixa de frequências permitidas pelo filtro, denominados no ImageJ como *Filter Large Structures Down* e *Filter Large Structures Up*. Após testes com diversos valores, adotou-se para os dois parâmetros os valores 5 e 20, respectivamente. Para os demais parâmetros, os valores utilizados não foram alterados, ou seja, foram utilizados valores *default* definidos pela própria implementação. O filtro é aplicado sobre a imagem I_m (resultante da etapa anterior), produzindo como resultado a imagem I_f , conforme ilustrado no fluxograma da Figura 4. A Figura 7b ilustra o resultado desta etapa. Comparando-se com a Figura 7a, nota-se que as veias foram realçadas.

3.1.3 Ajuste de contraste

Esta etapa utiliza dois conceitos de processamento de imagens: saturação e contraste. A saturação indica a pureza ou a intensidade de uma cor em particular. O contraste é a diferença entre a região clara e escura da imagem. Deste modo, esta etapa tem como objetivo reduzir o efeito de regiões muito claras e muito escuras, em relação as demais. Este efeito é causado, principalmente, pelo excesso de luz em algumas regiões e pela falta de luz em outras, além de reflexos, durante a captura da imagem. Em outras palavras, esta etapa busca equalizar a imagem, ou seja, que a diferença entre os valores de brilhos das

regiões mais claras e mais escuras não seja tão grande, mas mais uniforme.

O ajuste de contraste tem como base as informações do histograma dos pixels da imagem. Utilizando a implementação *open-source* ImageJ [14] para o ajuste de contraste, é possível definir uma porcentagem do número de pixels da imagem que serão saturados, bem como se ocorrerá uma equalização do histograma e uma normalização dos valores de brilho.

Foram testados diferentes valores de porcentagem de saturação e foi feita uma análise visual dos resultados obtidos. O valor de porcentagem de saturação de 0,3% apresentou o melhor resultado em termos de realce da região das veias na imagem. Foram realizadas a equalização do histograma e a normalização dos valores de brilho. Estes procedimentos foram aplicados na imagem I_f , resultante da etapa anterior. Como resultado, temos a imagem I_c , ilustrada na Figura 7c, com o realce das veias do dorso da mão.

3.2 Geração da imagem binária

Esta etapa é responsável pela produção de duas imagens binárias, denominadas I_{b_3} e $I_{b_{15}}$. Ambas as imagens são obtidas de modo independente a partir da imagem I_c , conforme ilustrado no fluxograma da Figura 4. Uma imagem binária consiste em uma imagem com apenas dois valores possíveis para cada pixel. A partir de um valor referência T , denominado limiar de corte, os pixels com valores acima de T formarão o objeto e receberão valor v_1 . Caso contrário, serão classificados como fundo e receberão valor v_2 .

O valor T pode ser um único valor, denominado limiar global, ou um valor calculado de modo independente para cada posição ou região da imagem, denominado limiar local. Nesse projeto, utilizou-se um limiar local, calculado a partir da Fórmula 1, proposta por Phansalkar *et al.* [15]. O valor de T na Fórmula 1 é calculado para cada pixel da imagem, considerando seu valor de brilho e os valores de brilho de seus pixels vizinhos. Apesar de Phansalkar *et al.* terem sugerido uma área retangular para delimitar a vizinhança em torno da posição corrente, nesse projeto delimitou-se a vizinhança a uma área circular de raio 0,5.

Os parâmetros da Fórmula 1 são: m é a média e d o desvio padrão dos valores dos brilhos dos pixels dentro da área circular delimitada; p e q são constantes cujos valores recomendados por Phansalkar *et al.* são 2 e 10, respectivamente, e foram os valores utilizados; k é um parâmetro cujo valor recomendado por Phansalkar *et al.* é 0,25 e foi o valor utilizado; $T(x,y)$ é o valor de limiar local para a posição em análise. O parâmetro r possui valor 3 para gerar a imagem I_{b_3} e 15 para gerar a imagem $I_{b_{15}}$, ilustradas nas Figuras 7d e 7e, respectivamente.

$$T = m * (1 + p * \exp(-q * m) + k * ((d/r) - 1)) \quad (1)$$

3.3 Determinação da região de interesse

Esta etapa tem como objetivo isolar a região que compõe as veias. Para isto, são utilizadas as imagens binárias I_{b_3} e $I_{b_{15}}$. Este procedimento é composto de três passos, descritos a seguir. A seção 3.5 descreve detalhadamente porque são utilizadas duas imagens binárias e não apenas uma.

3.3.1 Isolamento do contorno das mãos na imagem I_{b_3}

O primeiro passo utiliza a imagem binária I_{b_3} . Conforme é possível perceber ao analisar a Figura 7d, ao aplicar a Fórmula 1 para valor de r baixo, restaram os pixels que compõem o contorno da mão e pixels em outras áreas. Sendo assim, é necessário eliminar esses pixels de outras áreas e selecionar apenas o contorno da mão.

O método utilizado para atingir este objetivo consiste em localizar cada região contínua da imagem binária I_{b_3} . Denomina-se região contínua uma região em que qualquer pixel dessa região cumpre dois critérios: o primeiro, possuir um valor de brilho igual a v_1 , ou seja, pertencer a um objeto e não ao fundo; e ao menos um de seus vizinhos imediatos também tenha valor de brilho v_1 .

Para cada região, são calculados o ponto médio (X_m, Y_m) e o desvio-padrão, considerando as coordenadas x e y de cada pixel da região. A região escolhida como sendo a que representa o contorno da mão é àquela que possui o maior valor de desvio-padrão. A Figura 7f é o resultado desta operação na

imagem I_{b_3} (Figura 7d). Como pode ser notado ao analisar visualmente as Figuras 7f e 7d a região do contorno da mão é a região que possui os pixels menos agrupados em determinada área, motivo pelo qual adotou-se o critério de escolha a região com o maior valor de desvio-padrão.

3.3.2 Subtração das imagens $I_{b_{15}}$ e I_{b_3}

O segundo passo tem como objetivo afinar a região que representa o contorno da mão na imagem $I_{b_{15}}$, ilustrada na Figura 7e, utilizando a informação do contorno da mão isolado na imagem I_{b_3} , ilustrado na Figura 7f. Esse processo de afinar uma região é conhecido em Morfologia como Erosão.

Para atingir este objetivo, a imagem $I_{b_{15}}$ é subtraída da imagem I_{b_3} : caso um pixel pertença ao contorno da mão na imagem I_{b_3} e o pixel na posição correspondente na imagem $I_{b_{15}}$ seja um pixel de algum objeto, o pixel na imagem $I_{b_{15}}$ deixará de pertencer ao objeto e passará a ser considerado como fundo. Em termos de implementação, o valor de brilho do pixel será alterado de v_1 para v_2 .

O resultado deste passo é ilustrado na Figura 7g. Ao analisar a Figura 7g, comparando-se com a Figura 7e, nota-se que a imagem $I_{b_{15}}$ foi afinada parcialmente na região do contorno da mão.

3.3.3 Eliminação do contorno da mão na imagem $I_{b_{15}}$

Este terceiro passo tem como objetivo eliminar a região do contorno da mão na imagem $I_{b_{15}}$. O procedimento adotado é similar ao passo 3.3.1. No entanto, ao invés de manter somente a região com o maior valor de desvio-padrão, elimina-se justamente essa região que representa a região do contorno da mão. O resultado é ilustrado na Figura 7h. Conforme pode ser observado na Figura 7h o resultado final consiste em uma imagem com as regiões das veias, sem o contorno da mão, embora ainda persistam alguns pixels que não pertencem às veias. Esses pixels são considerados ruídos e precisam ser eliminados.

3.4 Eliminação de ruídos e imagem final

Esta é a última etapa do processo de segmentação com o objetivo de eliminar ruídos ainda presentes após a localização dos pixels que compõem as veias. Para atingir este objetivo, localizamos cada uma das regiões contínuas da imagem da Figura 7h. Remove-se da imagem as regiões que tiverem um total de pixels menor do que 10% da região com o maior número de pixels. O resultado desta operação é uma imagem em que estarão presentes apenas os pixels das veias, ilustrada na Figura 7i.

É importante ressaltar, para fins de implementação, que a imagem final deve possuir dimensões equivalente apenas a região final das veias, conforme ilustrada pela moldura ao redor da região das veias na Figura 7i. Para isto, é suficiente verificar a localização (coordenadas) dos pixels mais extremos em cada um dos lados (direito, esquerdo, superior e inferior). Este detalhe é importante para a etapa de normalização em que poderão ser realizadas alterações na escala da imagem final. Destaca-se que a imagem final contém apenas a região das veias. Sendo assim, além de eliminar os ruídos, estes procedimentos encerram a etapa de Segmentação da Imagem e a imagem resultante é utilizada na etapa seguinte de Normalização.

3.5 Subtração das Imagens $I_{b_{15}}$ e I_{b_3}

Antes de continuar o processo de reconhecimento biométrico, convém explicar porque no processo de segmentação foram utilizadas duas imagens binárias, I_{b_3} e $I_{b_{15}}$, bem como a etapa de subtração entre elas.

Quando utiliza-se valores maiores para r na Fórmula 1, a região das veias na imagem binária são melhor preservadas. No entanto, em algumas imagens, as regiões que formam as veias conectam-se à região que forma o contorno da mão, formando uma única região, conforme ilustrado na Figura 8. Deste modo, ao se determinar a região com o maior valor de desvio-padrão, eliminava-se o contorno da mão e também os pixels que compõem as veias, pois estavam todos unidos.

A solução encontrada foi aplicar a técnica de limiar local para r igual a 3 e gerar uma imagem binária em que o contorno da mão é preservado, enquanto a região das veias é praticamente eliminada. Além disso, o uso de r igual a 3 faz com que o contorno da mão seja uma borda mais fina, quando comparada ao contorno produzido na imagem gerada para r igual a 15, o que é possível verificar ao

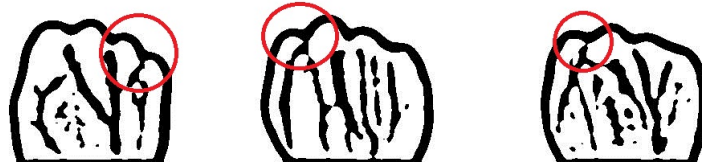


Figura 8: Exemplos de imagens em que o contorno da mão e as veias formam uma única região.

analisar visualmente as Figuras 7d e 7e). Deste modo, ao efetuar a subtração dos contornos, elimina-se justamente essa conexão entre o contorno da mão e as veias na imagem que preserva melhor as veias, $I_{b_{15}}$. Por este motivo foram utilizadas duas imagens binárias com valores de r distintos.

3.6 Normalização

A imagem final produzida no processo de segmentação, I_{final} , que contém apenas a região das veias, conforme ilustrado na Figura 7i, pode possuir dimensões diferentes para imagens tanto do mesmo indivíduo quanto de pessoas distintas. Isso ocorre tanto pela possibilidade de alterações de escala durante o processo de aquisição da imagem quanto pelo processo de segmentação em si, que pode resultar em maior ou menor quantidade de pixels ao seu final. Sendo assim, convém que seja feita uma normalização da imagem I_{final} , para que o processo de extração de características, etapa seguinte, atue sobre imagens padronizadas quanto às suas dimensões.

Assim, a etapa de normalização tem como objetivo produzir uma imagem de dimensões $n \times n$ a partir da imagem I_{final} . Para atingir este objetivo, a técnica de Transformação Afim ou *affine transformation* foi utilizada. Esta técnica consiste basicamente em efetuar as operações geométricas básicas através de um mapeamento linear aplicando uma matriz de transformação sobre os valores dos pixels da imagem. Foi utilizada uma implementação disponível em Java na classe *AffineTransform* que utilizava a matriz ilustrada na Figura 9. Os parâmetros s_x e s_y indicam, respectivamente, os fatores de alteração de escala nas direções x e y . O valor desses fatores dependem das dimensões da imagem inicial e final.

$$\begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Figura 9: Matriz de Transformação para normalização da imagem I_{final} .

Para os estudos experimentais, os valores de dimensões $n \times n$ utilizados para a imagem normalizada foram de 100×100 e 25×25 . As imagens normalizadas em 100×100 são utilizadas no Registro de Imagens (seção 4.4), enquanto as imagens 25×25 são utilizadas na SVM e na Rede Neural (seções 4.2 e 4.3).

4 Extração de Características e Classificação

Esta seção descreve os métodos de classificação usados para identificação pessoal com base nas características das veias das mãos. Dois conjuntos de características são formados, um a partir da Transformada Wavelet aplicada na imagem das veias e o outro a partir da própria imagem da veias. Três métodos de classificação são avaliados: SVM com Transformada Wavelet, Redes Neurais com Transformada Wavelet e Registro de Imagem usando imagens das veias.

A seção está organizada da seguinte forma: a seção 4.1 faz uma revisão sucinta sobre Transformada Wavelet, SVM, Rede Neural e os conceitos de *FAR* e *FRR*. As seções 4.2 e 4.3 apresentam os métodos de classificação SVM e Rede Neural com a Transformada Wavelet. A seção 4.4 descreve o procedimento de classificação de Registro de Imagem.

4.1 Resumo teórico

4.1.1 Transformada Wavelet

Wavelets são um conjunto de funções obtidas a partir de operações de dilatações e translações de uma função principal, conhecida como Wavelet mãe [16]. Wavelets possuem formatos variantes contínua e discreta. Exemplos de famílias de funções Wavelets muito utilizadas na literatura são: Haar, Daubechies, Symlet e Coiflet [17][18][19].

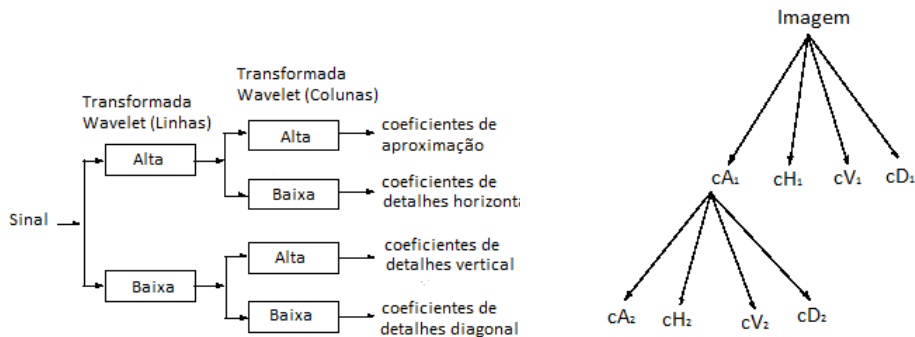
A Transformada Wavelet é uma ferramenta de análise de funções temporais que faz a decomposição de um sinal em termos de wavelets. Uma vantagem da Transformada de Wavelet é que as funções wavelets adaptam-se a cada ponto do sinal, permitindo a análise em diferentes resoluções temporais. Ao aplicar a Transformada Wavelet em um sinal, gera-se um novo sinal que contém dois componentes: um de baixa frequência e um de alta frequência. O componente de baixa frequência contém os denominados *coeficientes de aproximação*, que representam uma versão suavizada do sinal original. Por outro lado, o componente de alta frequência é composto pelos *coeficientes de detalhes*, que representam os detalhes do sinal original.

Na área de Processamento Digital de Imagens, a Transformada Wavelet pode ser utilizada para a decomposição de uma imagem em sub-imagens, também chamadas de componentes ou sub-bandas, sendo que cada componente possui informações e níveis distintos de detalhes. Normalmente, utiliza-se a Transformada Wavelet discreta em duas dimensões, em um processo conhecido como Decomposição de Wavelet em duas dimensões, ou *2D DWT* (*2-D discrete wavelet decomposition*).

O processo *2D DWT* consiste basicamente na aplicação duas vezes da Transformada Wavelet em um sinal uni-dimensional. Assim, considerando uma imagem representada por uma matriz bi-dimensional, em que cada elemento da matriz indica o valor de brilho local da imagem, a primeira aplicação da transformada utiliza os valores de brilho das linhas da matriz. Sobre o resultado da primeira operação de Transformada Wavelet, aplica-se novamente a transformada, mas desta vez considerando os valores de brilho das colunas. Como consequência da aplicação da *2D DWT* são gerados quatro componentes: os coeficientes de aproximação, coeficientes de detalhes horizontal, coeficientes de detalhes vertical e coeficientes de detalhes diagonal. Detalhes sobre os coeficientes gerados podem ser consultados em [20].

A Figura 10a ilustra o processo de aplicação da Transformada Wavelet, em que primeiramente aplica-se a transformada considerando os valores de brilho das linhas, e para cada uma das duas imagens resultantes, aplica-se a transformada considerando os valores de brilho das colunas. Na Figura 10a, *Alta* indica o componente de alta frequência e *Baixa* o componente de baixa frequência.

Diferentes níveis de decomposição podem ser realizados, sendo o componente que representa os *coeficientes de aproximação* de um nível decomposto no nível seguinte, formando uma decomposição multi-nível. A Figura 10b ilustra a decomposição multi-nível, em que cA_i , cH_i , cV_i , cD_i representam, respectivamente, os coeficientes de aproximação, de detalhes horizontal, de detalhes vertical e de detalhes diagonal no nível i .



(a) Componentes resultantes da aplicação da *2D DWT*.

(b) Representação de uma decomposição multi-nível da Transformada Wavelet.

Figura 10: Organogramas das estruturas gerais da aplicação da Transformada de Wavelets.

4.1.2 SVM - Support Vector Machine

SVM é uma técnica de aprendizado de máquina supervisionado que pode ser utilizada para separar dados em classes. O SVM tem como base a teoria de aprendizado estatístico de Vapnik and Chervonenkis [21]. O funcionamento básico de uma SVM consiste em encontrar o hiperplano que separa um conjunto de dados de treinamento em duas classes. Neste processo, o SVM busca o hiperplano que maximiza a distância entre as classes, ou seja, o SVM não busca apenas um hiperplano, mas um hiperplano ótimo. A Figura 11a ilustra a representação de duas classes e o hiperplano ótimo para um conjunto de dados.

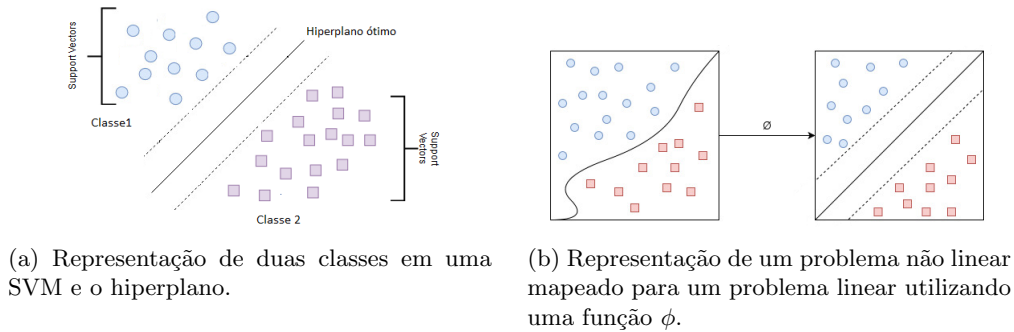


Figura 11: Representação gráfica de conceitos relativos à SVM.

O SVM pode ser usado para solucionar um problema com dados não linearmente separáveis. Para isto, o conjunto de dados é mapeado de seu espaço original para um novo espaço de maior dimensão, denominado *espaço de características* e a separação de dados é feita no espaço expandido. Esse novo conjunto de dados é separado por uma SVM linear. Na literatura o termo *Kernel* é utilizado para designar uma função que faz este mapeamento entre espaços. A Figura 11b ilustra a transformação ou expansão de espaço, transformando um problema não linearmente separável em linearmente separável, através da execução de uma função ϕ . Detalhes sobre o mapeamento entre espaços, bem como de diversas funções de mapeamento podem ser obtidos em [22].

4.1.3 Redes Neurais

Rede Neural Artificial é um modelo computacional composto por diversos elementos de processamento chamados de neurônios. Este modelo é amplamente utilizado no aprendizado de máquina e reconhecimento de padrões. Em sua forma mais tradicional, os neurônios são estruturados em camadas, dividindo-se em: uma camada de entrada, onde são recebidos os dados; uma ou mais camadas intermediárias e uma camada de saída. Uma rede neural com camadas intermediárias é denominada de *Multilayer Perceptron* (MLP).

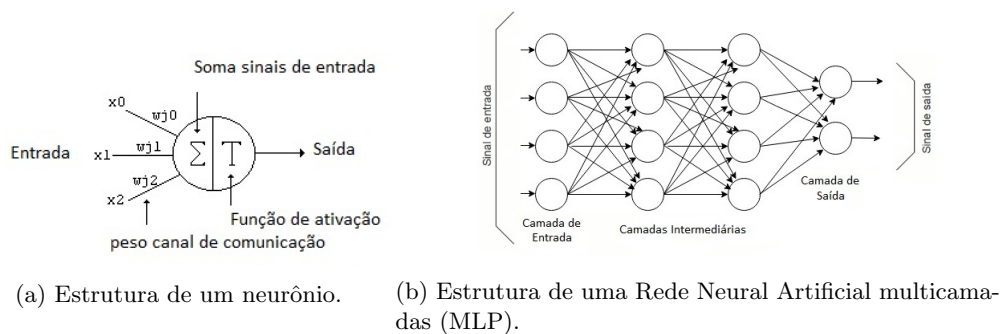


Figura 12: Ilustrações de elementos de uma Rede Neural Artificial.

A saída de um neurônio de uma camada é uma das entradas para cada um dos neurônios da camada seguinte. A ligação entre dois neurônios é chamada de canal de comunicação e possui um peso associado.

Assim, o sinal de saída de um neurônio será multiplicado pelo respectivo peso do canal antes de servir de entrada para o neurônio seguinte correspondente. Pesos distintos podem ser atribuídos a cada canal de comunicação. A entrada de um neurônio será a soma ponderada de todos os sinais de entrada, conforme ilustrado na Figura 12a. O neurônio possui internamente uma função associada, denominada função de ativação. A função de ativação produzirá um sinal de saída caso o somatório dos sinais de entrada ultrapasse um determinado valor limiar.

A Figura 12b ilustra uma estrutura geral de Rede Neural Artificial multicamadas (MLP). Um algoritmo para treinar uma rede MLP é o *Backpropagation* [23]. Nesse algoritmo, os dados são apresentados na camada de entrada e os resultados são analisados na camada de saída. O sinal de saída é analisado de acordo com o padrão desejado, e caso diferente, contabiliza-se o erro. Assim, durante a fase de treinamento da rede neural, a partir da camada de saída, os valores de erros são propagados pelas camadas intermediárias até a camada de entrada, fazendo com que os pesos dos canais de comunicações sejam alterados conforme o nível de erro propagado. Esse processo de ajuste de pesos nos canais de comunicações é repetido até obter-se um resultado de saída desejado.

4.1.4 Conceitos de FAR e FRR

No processo de autenticação, um erro de Falsa Aceitação ocorre quando o sistema aceita a amostra testada como legítima, embora a amostra não pertença a um membro da classe genuína. Nesse caso, o sistema aceita um impostor como usuário legítimo. O erro de Falsa Rejeição ocorre quando o sistema rejeita uma amostra da classe genuína, negando acesso a um usuário legítimo. Neste trabalho o desempenho do sistema de autenticação pessoal é avaliado de acordo com as taxas de Falsa Aceitação (FAR - *False Accept Rate*) e Falsa Rejeição (FRR - *False Reject Rate*).

4.2 Extração de características com Transformada Wavelet e classificação com SVM

Nesta seção apresentamos como utilizamos os métodos a Transformada Wavelet e SVM para extrair as características discriminantes da imagem e efetuar a classificação individual. Neste trabalho, utilizamos a família de funções wavelets Daubechie e Coiflet, aplicando a Transformada Wavelet multi-nível. Os melhores resultados obtidos foram com 4 níveis de decomposição. Assim, todos os *coeficientes de aproximação* gerados formam um único *vetor de características* que identifica exclusivamente uma pessoa, sendo este vetor usado como dados de entrada para o classificador SVM.

4.2.1 Banco de dados

Para implementar e avaliar o classificador SVM, desenvolvemos um sistema utilizando a versão multi-classes da SVM disponível em Python², particularmente a versão multi-classes *one-versus-one*. Para o treinamento e os testes utilizamos as imagens da base de dados fornecidas por [12]. Essa base de dados possui um total de 500 imagens de 50 pessoas distintas, sendo 10 imagens por pessoa divididas em 5 imagens da mão direita e 5 imagens da mão esquerda. As imagens foram coletadas de forma diversificada entre homens e mulheres de 16 a 65 anos. Foram utilizadas 3 imagens para treinamento e 2 para testes, das 50 pessoas disponíveis, totalizando assim 150 imagens para treinamento e 100 amostras para testes, sendo todas as imagens utilizadas da mão esquerda.

4.2.2 Treinamento da SVM

Em termos de implementação da SVM, para o treinamento, o *vetor de características* é fornecido como parâmetro para a função *SVC*, juntamente com qual classe ele pertence, ou seja, de qual pessoa é a amostra. O principal parâmetro dessa função é qual função *kernel* utilizar. A função kernel que apresentou os melhores resultados foi a *rbf*. Os valores dos parâmetros *gamma* e *C* (parâmetro de penalidade de erro) dependem da função Wavelet utilizada e podem ser consultados na Tabela 3 na seção 5. Os demais parâmetros foram mantidos os valores *default*. Detalhes sobre a teoria envolvida e a implementação

²<http://scikit-learn.org/> Acessado em 01/06/2018.

da função *SVC* podem ser consultados em [24]. Foram utilizadas 3 amostras para cada pessoa para o treinamento, totalizando 150 amostras.

4.2.3 Testes da SVM

Após o treinamento do modelo SVM, foram efetuados dois testes separadamente, sendo o primeiro para avaliar a taxa FRR o segundo para avaliar a taxa FAR. Para os testes, o sistema define dois parâmetros de entrada: a amostra de teste e a pessoa a quem alega-se ser a amostra. Como resultado, o sistema indica se a pessoa é aceita ou rejeitada.

Para avaliar a taxa FRR, o seguinte procedimento foi efetuado:

1. Dentre as 100 imagens disponíveis para testes, uma amostra foi aleatoriamente sorteada, sendo denominada a_m . Conhecemos a priori a pessoa p a quem pertence a amostra a_m . Assim, a_m e p são entradas do sistema;
2. A amostra a_m é apresentada ao classificador SVM que irá indicar pertencer à pessoa p_i ;
3. Caso p seja igual a p_i , houve um acerto; caso contrário, um usuário legítimo será negado pelo sistema e, portanto, um erro de FRR.

Para a avaliação da FRR, são efetuados 100 testes: são apresentadas como entrada para o sistema cada uma das 100 amostras, sempre indicando p como entrada a pessoa a quem de fato pertence a amostra.

Para avaliar a taxa FAR, o seguinte procedimento foi efetuado:

1. Dentre as 100 imagens disponíveis para testes, uma amostra foi aleatoriamente sorteada, sendo denominada a_m . Conhecemos a priori a pessoa p a quem pertence a amostra a_m . Assim, a_m e p são entradas do sistema;
2. A amostra a_m é apresentada ao classificador SVM que indica pertencer à pessoa p_i ;
3. Caso p seja diferente de p_i , houve um acerto, já que um impostor foi negado pelo sistema. Caso contrário, um usuário impostor foi aceito pelo sistema e, portanto, um erro de FAR.

Para a avaliação da FAR, para cada amostra sorteada são testadas 49 possibilidades: são fornecidos como entrada para o sistema a amostra sorteada e o valor de p como sendo cada uma das outras 49 pessoas que não sejam a pessoa a qual de fato a amostra pertence. Sendo 100 amostras disponíveis para teste, há um total de 4900 testes.

4.3 Extração de características com Transformada Wavelet e classificação com Rede Neural

Os procedimentos adotados para gerar o *vetor de características*, efetuar o treinamento e os testes no caso da Rede Neural foram os mesmos utilizados na SVM, descritos na seção 4.2. No caso da Rede Neural, a função Python utilizada foi a *MLPClassifier*, cujos detalhes teóricos e de implementação podem ser consultados em [25]. Os melhores valores encontrados para os parâmetros *alpha* e *hidden_layer_sizes* (total de camadas intermediárias para a Rede Neural multicamadas) da função *MLPClassifier* dependem da função Wavelet utilizada e podem ser consultados na Tabela 3 na seção 5. Os demais parâmetros foram mantidos os valores *default*.

4.4 Classificação por Registro de Imagens com base em imagem de veia

A técnica de Registro de Imagens é um processo de alinhamento de duas imagens no mesmo sistema de coordenadas espaciais através da aplicação das operações geométricas básicas de alterações de escala, rotação e translação, cujo objetivo é determinar se duas imagens pertencem a mesma classe. O processo de alinhamento possui duas etapas: na primeira etapa efetuam-se as operações geométricas enquanto a segunda etapa verifica-se o resultado do alinhamento.

4.4.1 Primeira etapa: operações geométricas para o alinhamento de imagens

Particularmente nesse projeto, para esta etapa, em termos de implementação, utilizamos o seguinte procedimento: sejam duas imagens I_1 e I_2 com as mesmas dimensões, que contêm pixels de veias e de fundo; seja uma imagem I_3 que contém apenas pixels de fundo. Efetua-se o mapeamento de I_1 e I_2 para I_3 . A imagem I_1 é mapeada sem alterações, enquanto sobre a imagem I_2 efetua-se operações geométricas de alteração de escala e rotação antes do mapeamento e de translação durante o mapeamento. A dimensão da imagem I_3 deve ser ao menos 4 vezes maior do que a dimensão de I_2 para que não sejam perdidos pixels devido aos procedimentos realizados sobre I_2 .

O processo de mapeamento consiste em dois passos: primeiramente calcula-se o centroide da região de veias para cada imagem I_1 e I_2 ; posteriormente, de acordo com a distância de cada pixel de veia em relação ao centroide, calcula-se a coordenada correspondente em relação ao centro da imagem I_3 . O cálculo do centroide (X_c, Y_c) é feito de acordo com as Equações 2 e 3, em que x e y são as coordenadas horizontal e vertical, respectivamente, de um pixel de veia da imagem e N é o número total de pixels que são pixels de veia na imagem.

$$X_c = \frac{\sum_{i=1}^N x_i}{N} \quad (2)$$

$$Y_c = \frac{\sum_{i=1}^N y_i}{N} \quad (3)$$

As novas coordenadas de cada pixel (x, y) são (x', y') e calculadas de acordo com as Equações 4 e 5, em que: (X_c, Y_c) são as coordenadas do centroide da imagem a ser mapeada (X_{c_1} e X_{c_2}); (X_{c_3}, Y_{c_3}) são as coordenadas do ponto central da imagem I_3 ; Δ_x e Δ_y indicam valores de deslocamento no intervalo de $[-10, 10]$.

$$x' = |x - X_c| + X_{c_3} + \Delta_x \quad (4)$$

$$y' = |y - Y_c| + Y_{c_3} + \Delta_y \quad (5)$$

A escala da imagem I_2 é alterada nas seguintes proporções da dimensão original: 0,8; 0,9; 1,1 e 1,2. A rotação é alterada, em graus, de $\{-8, -6, -4, -2, 2, 4, 6, 8\}$. A alteração de escala e a rotação são implementadas em Java utilizando as funções *getScaleInstance* e *rotate*, respectivamente, da classe *AffineTransform*, cujos detalhes podem ser consultados em [26].

4.4.2 Segunda etapa: verificação do resultado do processo de alinhamento

Durante o alinhamento ocorre uma sobreposição quando um pixel de veia de I_1 e um pixel de veia de I_2 são mapeados para a mesma posição em I_3 . A verificação do resultado do processo de alinhamento ocorre de acordo com o valor de três taxas de sobreposição, determinadas pelas Equações 6, 7 e 8, onde N_1 e N_2 correspondem ao número total de pixels de veias das imagens I_1 e I_2 , respectivamente. Assim T_3 indica o total de pixels sobrepostos em relação a soma dos pixels de veias de ambas as imagens, enquanto T_1 e T_2 indicam o total de pixels sobrepostos em relação aos pixels de veias das imagens I_1 e I_2 , respectivamente.

$$T_1 = \frac{(\text{Número total de pixels sobrepostos})}{N_1} \quad (6)$$

$$T_2 = \frac{(\text{Número total de pixels sobrepostos})}{N_2} \quad (7)$$

$$T_3 = \frac{(\text{Número total de pixels sobrepostos})}{N_1 + N_2} \quad (8)$$

Após cada cálculo das taxas T_1 , T_2 e T_3 , é verificado se os percentuais estão de acordo com limiares pré-definidos, e caso estejam, o processo de Registro de Imagens é encerrado e o sistema aceita como legítimo o usuário. Se após todas as tentativas não forem atingidos os limiares, o sistema rejeita o

usuário. Após diversos testes, foi definido o limiar de 56% para cada uma das taxas T_1 , T_2 e T_3 . A Figura 13 exemplifica o conceito de Registro de Imagem através do alinhamento entre duas imagens de veias de pessoas distintas, assumindo que quanto maior o número de pixels sobrepostos, maior o indicativo das duas amostras pertencerem ao mesmo indivíduo. A decisão da classificação depende de limiares de sobreposição conforme discutido.

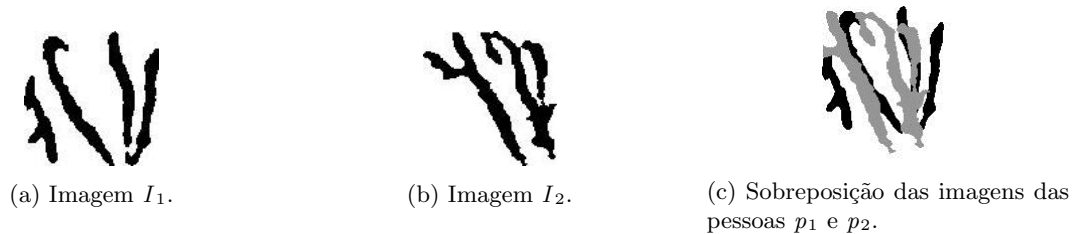


Figura 13: Exemplo de amostras para o Registro de Imagens.

5 Estudo Experimental

Para o estudo experimental, um sistema completo com todas as etapas (Segmentação, Normalização, Extração de Características e Classificação) foi desenvolvido. Os métodos de Segmentação, Normalização e Registro de Imagens foram implementados em Java, enquanto os métodos de Transformada Wavelet, SVM e Rede Neural foram desenvolvidos em Python utilizando as ferramentas Scikit-learn³ e PyWavelets⁴. Para os testes, foram utilizadas as imagens fornecidas pela base de dados [12], descrito na seção 4.2.1.

Primeiramente, todas as imagens do conjunto de dados foram segmentadas e normalizadas. Nenhuma imagem foi descartada. Posteriormente, foram feitas quatro avaliações distintas sobre a etapa de Extração de Características e Classificação das veias: SVM, Rede Neural e Registro de Imagens individualmente e uma avaliação híbrida envolvendo em conjunto as 3 técnicas mencionadas anteriormente. Para as técnicas de SVM e Rede Neural, conforme mencionado na seção 4.2, os dados consistem em um vetor de características formado pela concatenação dos *coeficientes de aproximação* resultantes da aplicação da Transformada Wavelet sobre a imagem normalizada em 4 níveis. Já a técnica de Registro de Imagens utiliza a própria imagem normalizada. As seções a seguir descrevem os resultados obtidos com cada técnica testada.

5.1 Resultados experimentais da 2D DWT com SVM e Rede Neural

As Tabelas 1 e 2 apresentam os resultados obtidos considerando 2 tipos de funções wavelets. Foram testados de 1 a 6 níveis de decomposição para a 2D DWT, sendo que com 4 níveis os melhores resultados foram obtidos. Para escolher os melhores modelos de classificadores, foram testados diversos valores para os parâmetros exigidos para as funções de classificação *SVC*, no caso do SVM, e *MLPClassifier*, para a Rede Neural. A Tabela 3 apresenta os valores testados, bem como aqueles que permitiram obter os melhores resultados.

Tabela 1: Taxa de acertos utilizando a função wavelet Coiflet 7 na Transformada Wavelet.

	Coiflet 7		
	FRR (%)	FAR (%)	Acertos (%)
SVM	8	0,16	91,84
Rede Neural	9	0,18	90,82

³<http://scikit-learn.org/> Acessado em 22/02/2019.

⁴<https://pypi.org/project/PyWavelets/> Acessado em 22/02/2019

Tabela 2: Taxa de acertos utilizando a função wavelet Daubechies 12 na Transformada Wavelet.

	Daubechies 12		
	FRR (%)	FAR (%)	Acertos (%)
SVM	7	0,14	92,86
Rede Neural	8	0,16	91,84

Tabela 3: Valores de parâmetros testados para os classificadores SVM e Rede Neural. Foram testados de 1 a 6 níveis de decomposição para a 2D DWT para cada um dos valores de parâmetros.

Classificador		Melhores Parâmetros (com 4 níveis de decomposição da 2D DWT)	
		Coiflet7	Daubechies 12
SVM	$10^{-5} \leq \gamma \leq 10^5$, $\forall \gamma$ múltiplo de 10; Para cada valor de γ , $10^{-5} \leq c \leq 10^5$, $\forall c$ múltiplo de 10.;	$\gamma = 0.01$; $c = 0.0001$.	$\gamma = 0.0001$; $c = 10$.
Rede Neural	$\alpha = 0.1$ e $\alpha = 0.01$; $10 \leq \text{hidden_layer_sizes} \leq 120$, $\forall \text{hidden_layer_sizes}$ múltiplo de 10.	$\alpha = 0.1$; $\text{hidden_layer_sizes} = 60$.	$\alpha = 0.1$; $\text{hidden_layer_sizes} = 80$.

5.2 Resultados experimentais da classificação com Registro de Imagens

No caso do Registro de Imagens, 3 imagens do mesmo indivíduo são usadas como amostras de referência e 2 imagens para testes. Assim, no processo de avaliação, cada imagem de teste deve ser comparada com 3 imagens de referência do suposto indivíduo, com a decisão de aceitar ou rejeitar baseada no critério majoritário. Isto é, a autenticação positiva é declarada quando a imagem de teste é validada, com as taxas de sobreposições maiores do que o limiar estabelecido, com 2 ou mais imagens de referência.

Diversos valores de limiar de sobreposição foram avaliados. O valor de limiar de 56% de sobreposição oferece o melhor desempenho com a taxa de Falsa Rejeição (FRR) igual a 1% com uma taxa de Falsa Aceitação (FAR) nula (zero), sendo que estas taxas foram obtidas através de 100 e 4900 comparações, respectivamente.

5.3 Resultados experimentais da classificação com um Sistema Híbrido

No Sistema Híbrido, as três técnicas são consideradas conjuntamente. No processo de autenticação híbrido, o teste da amostra é feito simultaneamente pelos classificadores SVM e Rede Neural. Caso ambos os classificadores indiquem o mesmo resultado, esta será a resposta final do sistema híbrido (positivo ou negativo). Quando as respostas divergirem, o sistema híbrido executa a técnica de Registro de Imagens, que fornece a resposta definitiva do teste.

A aplicação do sistema híbrido tem como objetivo otimizar a performance do sistema de autenticação em termos de rapidez computacional fornecida pelas técnicas SVM e Rede Neural combinada com a alta taxa de acerto fornecida pelo Registro de Imagem. A melhor taxa de acertos obtida para o sistema híbrido foi utilizando a SVM com Wavelet Daubechies 12 e a Rede Neural com Coiflet 7. A Tabela 4 compara os desempenhos das técnicas de autenticação estudadas.

5.4 Resultados experimentais do desempenho computacional

Os testes foram realizados com uma configuração computacional simples (Windows 7, processador Intel i5 com 8Gb de memória RAM). O tempo médio da etapa de segmentação foi de aproximadamente 0,4 segundos. O tempo médio no caso do Registro de Imagens foi de 0,84 segundos para autenticar positivamente um usuário legítimo e 5,27 segundos para negar um usuário impostor. Note que o tempo médio

Tabela 4: Comparativo de desempenho entre os diversos métodos propostos.

	FRR (%)	FAR (%)	Acertos (%)
SVM com Daubechies 12	7	0,14	92,86
Rede Neural com Coiflet 7	9	0,18	90,82
Registro de Imagens	1	0	99,00
Sistema híbrido	1	0,02	98,98

para detectar um impostor é consideravelmente maior devido a necessidade de se testar e executar todas as possíveis alterações de escalas, rotações e translações antes da aplicação rejeitar a amostra, negando acesso ao indivíduo. Por outro lado, no caso de um usuário genuíno, normalmente, apenas pequenos ajustes e poucas operações são exigidos, o que indica que um usuário legítimo pode ser rapidamente validado.

Tabela 5: Desempenho em termos de processamento computacional.

	Coiflet 7	
	Tempo de treinamento (s)	Tempo de teste (s)
SVM	11,28	2,51
Rede Neural	64,73	2,32

Tabela 6: Desempenho em termos de processamento computacional.

	Daubechies 12	
	Tempo de treinamento (s)	Tempo de teste (s)
SVM	4,55	1,07
Rede Neural	23,32	1,01

Os tempos médios de treinamento e testes para os classificadores SVM e Rede Neural são apresentados nas Tabelas 5 e 6, respectivamente. Ao analisar estas tabelas, é possível verificar que os tempos de treinamento da SVM são bem inferiores aos exigidos para a Rede Neural Artificial, apesar de ambos possuírem os tempos de testes muito próximos e relativamente curtos. Nota-se que para autenticar um usuário legítimo, os tempos dos três métodos utilizados são muito próximos. No entanto, para negar um usuário impostor, o tempo no caso do Registro de Imagens é praticamente 5 vezes superior ao tempo necessário para a SVM e para a Rede Neural. Este fato motivou a proposta do sistema híbrido, integrando as técnicas estudadas para otimizar o ganho de desempenho geral no sistema de autenticação, uma vez que o Registro de Imagem será utilizado somente quando os classificadores SVM e Rede Neural tiverem resultados divergentes.

5.5 Comparativo entre os resultados obtidos e de métodos da literatura

As Tabelas 7 e 8 apresentam um comparativo de resultados entre os métodos propostos e os métodos frequentemente usados na literatura. O melhor desempenho apresentado por [27] é de 3% de FRR, e uma FAR igual a 0,0202% utilizando a mesma base de dados usada neste projeto. Isso significa que a nossa bordagem utilizando apenas o Registro de Imagens permite uma redução da taxa de Falsa Rejeição (FRR) de 3% para 1% e ao mesmo tempo um decremento da taxa de Falsa Aceitação (FAR) de 0,0202% para zero. Quando utilizado o sistema híbrido, também reduzimos a taxa de falsa rejeição para 1% mantendo praticamente a mesma taxa de falsa aceitação.

Em relação a outros métodos de reconhecimento do Estado da Arte que utilizam diferentes bases de dados, o método descrito possui um desempenho, em termos de taxas de erro, compatível com os melhores métodos da literatura. O trabalho de Wang *et al.* [28] também utiliza decomposição wavelet. No entanto,

os autores escolheram a função wavelet Haar e apresentam taxas de erros iguais (EER, quando FAR é igual a FRR) de 2,067%, sendo que para uma taxa FAR nula (zero), a taxa FRR é superior a 6%, enquanto para uma taxa FAR próxima de 0,02%, a taxa FRR é superior a 4%.

Tabela 7: Comparativo do desempenho entre o sistema híbrido proposto e métodos da literatura para o mesmo banco de dados de imagens.

	FRR (%)	FAR (%)
Mohamed Shahin <i>et al.</i> [27]	3	0,02
Registro de Imagens	1	0
Sistema híbrido	1	0,02

Tabela 8: Comparativo do desempenho entre o sistema híbrido proposto e métodos da literatura para banco de dados diferente.

	FRR (%)	FAR (%)
Wang <i>et al.</i> [28] - Wavelet Haar	> 6	0
Registro de Imagens	1	0
Sistema híbrido	1	0,02

As melhores performances obtidas nesse trabalho podem ser justificadas, principalmente, por dois fatores. O primeiro fator é a melhoria na etapa de localização das veias na imagem, através de uma melhor combinação e sequência de aplicação de filtros considerando as especificidades de uma imagem do dorso de uma das mãos. Além disso, houve uma melhor escolha de parâmetros dos filtros utilizados. Estas melhorias permitiram que a região adicional não pertencente às veias fosse eliminada e as regiões efetivamente relevantes preservadas.

O segundo fator deve-se ao método de extração de características. O método de Registro de Imagens apresentado engloba a maioria das possíveis variações de escala, de rotação e de translação da imagem das veias. Já o sistema híbrido, quando comparado com o Registro de Imagens, mantém baixas taxas de erro e simultaneamente reduz o tempo de processamento computacional necessário para negar um usuário impostor para 3,58 segundos, embora aumente o tempo para aceitar um usuário legítimo para 3,14 segundos. Ao compararmos o sistema híbrido com o uso dos classificadores SVM ou Rede Neural individualmente, nota-se uma expressiva queda nas taxas de erro, mantendo-se um tempo computacional reduzido.

6 Conclusões

Ao analisar os dados obtidos pode-se concluir que funções wavelets, das famílias Daubechies e Coiflet, integradas com SVM e Rede Neural Artificial multicamadas podem ser utilizadas no reconhecimento biométrico das veias das mãos com relativa eficiência em termos de taxas de acertos e alta eficiência em questões de tempo computacional. No entanto, ainda são necessários ajustes na etapa de Segmentação da imagem para que uma melhor localização e extração das regiões das veias não atrapalhe os métodos de aprendizado de máquina e deste modo, possam ser atingidas taxas de erros muito baixas obtidas com o método de Registro de Imagens e de métodos apresentados na literatura. Observa-se também que o método de Registro de Imagens é menos susceptível a não perfeita localização das veias na etapa de Segmentação, quando comparado aos métodos de aprendizado de máquina.

É possível concluir também que pode haver a utilização de um sistema híbrido para aproveitar a ótima eficiência em termos de taxas de reconhecimento do método de Registro de Imagens com o ótimo desempenho em termos de tempo de reconhecimento apresentado pelos métodos de aprendizado de máquina.

Referências

- [1] S. Srivastava, S. Bhardwaj, S. Bhargava, *et al.*, “Fusion of palm-phalanges print with palmprint and dorsal hand vein,” *Applied Soft Computing*, vol. 47, pp. 12–20, 2016.
- [2] A. K. Jain, K. Nandakumar, and A. Ross, “50 years of biometric research: Accomplishments, challenges, and opportunities,” *Pattern Recognition Letters*, vol. 79, pp. 80–105, 2016.
- [3] R. Parkavi, K. C. Babu, and J. A. Kumar, “Multimodal biometrics for user authentication,” in *Intelligent Systems and Control (ISCO), 2017 11th International Conference on*, pp. 501–505, IEEE, 2017.
- [4] G. Arora, P. L. Pavani, R. Kohli, and V. Bibhu, “Multimodal biometrics for improvised security,” in *Innovation and Challenges in Cyber Security (ICICCS-INBUSH), 2016 International Conference on*, pp. 1–5, IEEE, 2016.
- [5] V. Mhaske and A. Patankar, “Multimodal biometrics by integrating fingerprint and palmprint for security,” in *Computational Intelligence and Computing Research (ICCIC), 2013 IEEE International Conference on*, pp. 1–5, IEEE, 2013.
- [6] M. K. Shahin, A. M. Badawi, and M. E. Rasmy, “Multimodal biometric system based on near-infrared dorsal hand geometry and fingerprints for single and whole hands,” *World Academy of Science, Engineering and Technology*, vol. 56, pp. 1107–1122, 2011.
- [7] P.-O. Ladoux, C. Rosenberger, and B. Dorizzi, “Palm vein verification system based on sift matching,” in *International Conference on Biometrics*, pp. 1290–1298, Springer, 2009.
- [8] R. B. Trabelsi, A. D. Masmoudi, and D. S. Masmoudi, “Hand vein recognition system with circular difference and statistical directional patterns based on an artificial neural network,” *Multimedia Tools and Applications*, vol. 75, no. 2, pp. 687–707, 2016.
- [9] Y. Wang, W. Xie, X. Yu, and L.-K. Shark, “An automatic physical access control system based on hand vein biometric identification,” *IEEE Transactions on Consumer Electronics*, vol. 61, no. 3, pp. 320–327, 2015.
- [10] R. B. Trabelsi, A. D. Masmoudi, and D. S. Masmoudi, “A novel biometric system based hand vein recognition,” *Journal of Testing and Evaluation*, vol. 42, no. 4, pp. 809–818, 2013.
- [11] C.-B. Hsu, S.-S. Hao, and J.-C. Lee, “Personal authentication through dorsal hand vein patterns,” *Optical Engineering*, vol. 50, no. 8, p. 087201, 2011.
- [12] A. M. Badawi, “Hand vein biometric verification prototype: A testing performance and patterns similarity,” *IPCV*, vol. 14, pp. 3–9, 2006.
- [13] L. Wang, G. Leedham, and S.-Y. Cho, “Infrared imaging of hand vein patterns for biometric purposes,” *IET computer vision*, vol. 1, no. 3, pp. 113–122, 2007.
- [14] <https://imagej.nih.gov/ij/docs/guide/146-29.html>. Acessado em 10/05/2018.
- [15] N. Phansalkar, S. More, A. Sabale, and M. Joshi, “Adaptive local thresholding for detection of nuclei in diversity stained cytology images,” in *Communications and Signal Processing (ICCSP), 2011 International Conference on*, pp. 218–220, IEEE, 2011.
- [16] T. Dócusse, A. Pereira, N. Marranghello, R. Guido, J. Furlani, P. Maturana, and R. Romano, “Aplicação da transformada wavelet para melhoria de visualização de microcalcificações em mamografias digitais.”
- [17] I. Daubechies, *Ten lectures on wavelets*, vol. 61. Siam, 1992.

- [18] R. Singh, R. E. Vasquez, and R. Singh, "Comparison of daubechies, coiflet, and symlet for edge detection," in *Visual Information Processing VI*, vol. 3074, pp. 151–160, International Society for Optics and Photonics, 1997.
- [19] <https://www.mathworks.com/help/wavelet/g/introduction-to-the-wavelet-families.html>. Acessado em 10/05/2018.
- [20] G. O. F. da Silva, G. A. Boggione, and L. M. G. Fonseca, "Fusao de imagens de sensoriamento remoto utilizando a transformada wavelet haar," *Simpósio Brasileiro de Sensoriamento Remoto*, vol. 12, pp. 6175–6182, 2007.
- [21] V. N. Vapnik and A. Y. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities," in *Measures of complexity*, pp. 11–30, Springer, 2015.
- [22] S. Khellat-kihel, N. Cardoso, J. Monteiro, M. Benyettou, *et al.*, "Finger vein recognition using gabor filter and support vector machine," in *Image Processing, Applications and Systems Conference (IPAS), 2014 First International*, pp. 1–6, IEEE, 2014.
- [23] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural Network Design*. USA: Martin Hagan, 2nd ed., 2014.
- [24] scikit-learn.org/stable/modules/svm.html. Acessado em 10/05/2018.
- [25] http://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html. Acessado em 10/05/2018.
- [26] <https://docs.oracle.com/javase/7/docs/api/java/awt/geom/AffineTransform.html>. Acessado em 10/05/2018.
- [27] M. Shahin, A. Badawi, and M. Kamel, "Biometric authentication using fast correlation of near infrared hand vein patterns," *International Journal of Biological and Medical Sciences*, vol. 2, no. 3, pp. 141–148, 2007.
- [28] Y.-D. Wang, Q.-Y. Yan, and K.-F. Li, "Hand vein recognition based on multi-scale lbp and wavelet," in *Wavelet Analysis and Pattern Recognition (ICWAPR), 2011 International Conference on*, pp. 214–218, IEEE, 2011.