Universidade Federal da Bahia
Instituto de Matemática

Programa de Pós-Graduação em Ciência da Computação

AUTENTICAÇÃO CONTÍNUA DE INDIVÍDUOS BASEADA EM ALGORITMOS DE DETECÇÃO DE ANOMALIAS

Matheus Magalhães Batista dos Santos

DISSERTAÇÃO DE MESTRADO

Salvador
22 de Janeiro de 2020

MATHEUS MAGALHÃES BATISTA DOS SANTOS

AUTENTICAÇÃO CONTÍNUA DE INDIVÍDUOS BASEADA EM ALGORITMOS DE DETECÇÃO DE ANOMALIAS

Esta Dissertação de Mestrado foi apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal da Bahia, como requisito parcial para obtenção do grau de Mestre em Ciência da Computação.

Orientador: Maurício Pamplona Segundo

Salvador
22 de Janeiro de 2020

Sistema de Bibliotecas - UFBA

Santos, Matheus.

Autenticação Contínua de Indivíduos baseada em Algoritmos de Detecção de anomalias / Matheus Magalhães Batista dos Santos – Salvador, 2014.

77p.: il.

Orientador: Prof. Dr. Maurício Pamplona Segundo.

Dissertação (Mestrado) – Universidade Federal da Bahia, Instituto de Matemática, 2014.

1. Autenticação Contínua. 2. Biometria. 3. Detecção de Anomalias. I. Pamplona, Maurício. II. Universidade Federal da Bahia. Instituto de Matemática. III Título.

CDD – XXX.XX

CDU – XXX.XX.XXX

AGRADECIMENTOS

Gostaria de agradecer a toda minha família e amigos pelo apoio. Em especial para os que ajudaram na revisão deste documento: Gabriel Dahia, Leone de Jesus e Fernando Medeiros.

RESUMO

Métodos de autenticação como senhas e cartões de acesso se tornaram comuns no dia-a-dia da sociedade. Devido a uma preocupação cada vez maior com a segurança, a biometria passou a ser uma forma de controle de acesso comum. Porém, assim como outros métodos de controle de acesso, eles só realizam a verificação da identidade do usuário apenas uma vez. Nenhuma verificação adicional é realizada posteriormente e com isso, não há garantia que o usuário permitido é o mesmo a utilizar um sistema ou recurso durante toda a sua utilização. Para resolver esse problema, a autenticação contínua realiza a verificação constantemente. Inúmeros esforços foram feitos para melhorar o desempenho da verificação na autenticação contínua, como o uso de biometrias cada vez mais seguras, mas não há muitos trabalhos que visam melhorar o método de autenticação contínua em si. Um bom sistema de autenticação contínua deve evitar considerar um usuário genuíno como um invasor e deve detectar invasores o mais rápido possível. Sistemas de detecções de anomalias visam detectar comportamentos que destoam do padrão. Para tais sistemas é desejável que um comportamento padrão não seja considerado uma anomalia e que anomalias sejam detectadas o mais rápido possível. Devido a similaridade de alguns dos objetivos da autenticação contínua e da detecção de anomalias, este trabalho se propõe a investigar técnicas de detecção de anomalias que podem ser utilizadas na autenticação contínua para torná-la segura independentemente do tipo de biometria utilizada. Após selecionar e implementar a melhor técnica para o contexto de autenticação contínua, comparamos o método proposto com o estado-da-arte e os resultados mostram que o novo método é melhor e não precisa de treinamento, adaptando-se facilmente a qualquer biometria.

Palavras-chave: Autenticação Contínua; Biometria; Detecção de Anomalias

ABSTRACT

Authentication methods such as passwords and access cards have become common in the everyday life of society. Due to increasing security concerns, biometrics has become a common form of access control. However, like other access control methods, they only perform identity verification only once. No additional checks are performed, so there is no guarantee that the allowed user is the same during the entire time of system utilization. To address this problem, continuous authentication constantly checks the identity. Numerous efforts have been made to improving the performance of the verification in continuous authentication, such as the use of increasingly secure biometrics, but there are not many works that aim improving the continuous authentication method itself. A good continuous authentication system should avoid consider a genuine user like an impostor and should detect impostors as soon as possible. Anomaly detection systems aim to detect behaviors that differ from a default behavior. For those systems, it is desired that a default behavior does not be considered an anomaly and that anomalies be detected as soon as possible. Due to the similarities between some of the objectives of the anomaly detection and continuous authentication, this work proposes to investigate anomaly detection techniques that can be used in the continuous authentication to make it safer regardless of the type of biometrics used. After select and implement the best technique for the continuous authentication context, we compare the proposed method with the state-of-art and the results show that the new method is better and does not need training, adapting easily to any biometric.

Keywords: Continuous Authentication; Biometrics; Anomaly Detection

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SIGLAS

CASIA CASIA NIR-VIS 2.0 41
CDF Cumulative Distribution Function 41
CNN Convolutional Neural Network 66
COF Connectivity-based Outlier Factor 25
ECG Eletrocardiograma 66
EER Equal Error Rate 73
FAR False Acceptance Rate 32
FCM Fuzzy C-Means 14
FRGC Face Recognition Grand Challenge 41
FRR False Rejection Rate 32
ICA Independent Component Analysis 72
LDA Linear Discriminant Analysis 30
LFW Labeled Faces in the Wild 41
LOF Local Outlier Factor 22
LTST Long-Term ST Database 69
MDEF Multi-granularity Deviation Factor 25
MITDB MIT-BIH Arrhythmia Database 68
MLII Modified Limb Lead I 70
MLP Multi Layer Perception 72
MLV1 Modified Limb V1 48
MTCNN Multi-Task Convolutional Neural Networks 48
NIR Near Infrared 41
NSRDB MIT-BIH Normal Sinus Rhythm Database 69
ODIN Outlier Detection using In-degree Number 25
PCA Principal Component Analysis 72
PCM Possibilistic C-Means 54
PTB PTB Diagnostic ECG Database 69
ReLU Rectified Linear Units 67

Capítulo 1

INTRODUÇÃO

Com o passar dos anos, métodos tradicionais de autenticação como senhas ou cartões de acesso se tornaram arriscados em ambientes que demandam um controle de segurança mais rígido. A biometria foi uma solução adotada em muitos sistemas para suprir essa demanda (DUGELAY et al., 2002; KONG et al., 2004). Diversas características biométricas podem ser utilizadas, dentre elas: face (JANAKIRAMAN et al., 2005; NIINUMA; PARK; JAIN, 2010), impressão digital (SIM et al., 2007), voz (DAMOUSIS; TZOVARAS; BEKIARIS., 2008), íris (MOCK et al., 2012), eletroencefalograma (NAKANISHI; BABA; MIYAMOTO, 2009) e eletrocardiograma (AGRAFIOTI; HATZINAKOS, 2009).

Os sistemas tradicionais de reconhecimento realizam a verificação da identidade do usuário apenas uma vez, o que não garante que um ataque de um usuário não autorizado, posterior a uma autenticação válida, possa ser feito. Para solucionar esse problema, a autenticação contínua (ou verificação contínua) realiza a verificação da identidade do usuário constantemente, garantindo assim que o usuário autorizado seja o mesmo durante toda a utilização do sistema. A autenticação contínua é bastante importante em ambientes de alto risco, onde o custo de um uso do sistema por alguém não autorizado é alto, como no controle de aviões, computadores de bancos, departamentos de defesa e outras aplicações que lidem com grandes quantidades de dinheiro ou que afetem a segurança de vidas humanas. Nesses casos, é desejável que o sistema se torne inoperante quando um usuário autorizado não possa ser autenticado (JANAKIRAMAN et al., 2005).

A autenticação contínua utilizando biometria foi abordada em diversos trabalhos (ALTINOK; TURK, 2003; SILVA; SEGUNDO, 2015; FLIOR; KOWALSKI, 2010a; JANAKIRAMAN et al., 2005; LEGGETT et al., 1991; SANTOS; SEGUNDO, 2015; MONACO et al., 2012; NIINUMA; PARK; JAIN, 2010; PAMPLONA et al., 2013; TSAI et al., 2014). Recentemente existe um foco maior em aplicações de autenticação contínua para *smartphones* (GASCON et al., 2014; CROUSE et al., 2015; ROY; HALEVI; MEMON, 2014; SITOVÁ et al., 2016). Cada biometria possui vantagens e desvantagens. Por exemplo, impressões digitais são bastante confiáveis, mas são incômodas ao usuário, exigindo que o dedo esteja constantemente em contato com o sensor. Os métodos de autenticação

contínua na literatura diferem mais quanto ao tipo de biometria utilizada, cada um visando extrair o máximo da biometria escolhida para sua aplicação específica, além de ressaltar suas vantagens em relação às outras biometrias.

Por mais seguro que um sistema seja, independente do método utilizado para autenticar usuários, o mesmo não é inviolável. Quando um invasor consegue burlar o sistema de segurança (biométrico ou tradicional), é desejável que a invasão seja detectada, e para isso, sistemas de detecção de intrusos foram desenvolvidos para detectar quando um usuário mal intencionado obtém acesso indevido ao sistema ou transmite dados sigilosos. Uma forma de fazer essa detecção é considerar a entrada de um invasor como uma anomalia. Uma anomalia é um comportamento que destoa do comportamento padrão e no contexto de detecção de intrusos pode significar uma possível violação do sistema. Uma vez detectada a anomalia um conjunto de decisões podem ser tomadas, como considerá-la uma invasão ao sistema e bloquear o acesso. Atualmente existem diversos sistemas de detecção de anomalias, principalmente para prevenção de ataques cibernéticos. Tais sistemas monitoram a todo momento diferentes dados como tráfego de rede, pacotes enviados em uma rede, comandos executados em terminais e etc, e foram desenvolvidos para detectar quando um comportamento foge do padrão observado ao longo do tempo (PATCHA; PARK, 2007; CHANDOLA; BANERJEE; KUMAR, 2009).

É crucial que uma anomalia seja detectada o mais rápido possível, pois uma simples anomalia no tráfego de rede de um computador pode significar que um hacker está enviando informações sigilosas para outro computador não autorizado (KUMAR, 2005). Além disso, é importante que um sistema de detecção de anomalias não considere comportamentos normais como anomalias (*e.g.* uma compra usual feita no cartão de crédito seja considerada uma fraude), pois isso gera uma sobrecarga desnecessária e recursos podem ser negados a usuários autorizados (CHANDOLA; BANERJEE; KUMAR, 2009).

É possível observar que sistemas de autenticação contínua e de detecção de anomalias possuem objetivos em comum: detectar invasores o mais rápido possível enquanto garantem o acesso aos usuários autorizados. Também é possível observar que é necessário uma nova abordagem para autenticação contínua que não se limite ao tipo de biometria utilizada, pois desde 2007 (SIM et al., 2007) não foi observado um esforço para melhorar o método de autenticação contínua em si, independente da biometria utilizada.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

O principal o objetivo deste trabalho é: *verificar a hipótese de que técnicas de detecção de anomalia podem melhorar o desempenho da autenticação contínua e torná-la mais segura e robusta ao tipo de biometria utilizada.*

1.1.2 Objetivos Específicos

Para verificar a hipótese este trabalho se propõe a:

1. Investigar técnicas de detecção de anomalias que podem ser adaptadas ao contexto de autenticação contínua, avaliando suas vantagens e desvantagens.

2. Propor uma nova abordagem de autenticação contínua utilizando uma ou mais técnicas promissoras de detecção de anomalias.

3. Analisar o desempenho da nova abordagem com diferentes biometrias.

4. Comparar o desempenho da nova abordagem com outras técnicas estado-da-arte.

1.2 ORGANIZAÇÃO DO DOCUMENTO

O restante deste documento está organizado da seguinte maneira:

- **Capítulo 2** introduz o funcionamento básico de um sistema de autenticação contínua assim como o estado-da-arte.
- **Capítulo 3** apresenta as técnicas de detecção de anomalia investigadas e suas conformidades no contexto de autenticação contínua.
- **O Capítulo 4** introduz as biometrias utilizadas neste trabalho.
- **O Capítulo 5** propõe um sistema de autenticação contínua utilizando uma técnica de detecção de anomalia para biometrias diferentes.
- **O Capítulo 6** descreve a metodologia utilizada nos experimentos, os resultados obtidos e a discussão.
- **Capítulo 7** conclui este trabalho apresentando sugestões de trabalhos futuros.

Capítulo 2

AUTENTICAÇÃO CONTÍNUA

A Figura 2.1 ilustra as etapas do funcionamento de um sistema de autenticação contínua genérico utilizando biometria. A primeira etapa é a aquisição de uma amostra biométrica através de um sensor (*e.g.* uma câmera obtendo uma imagem e detectando faces presentes nela para um sistema de reconhecimento facial). Na segunda etapa, a amostra poderá ser pré-processada a fim de melhorar o desempenho do algoritmo que irá descrevê-la (*e.g.* padronização facial para que as faces sigam um mesmo padrão ou remoção de cílios em uma imagem dos olhos para reconhecimento de íris). Uma vez pré-processada a amostra, ela será submetida a algum algoritmo de descrição na terceira etapa, resultando em uma estrutura de dados que possibilite comparações (*e.g.* passar a imagem de uma face por uma rede neural de convolução (Convolutional Neural Network (CNN)) e obter um vetor de 256 valores que descrevem tal face (WU et al., 2015)).

Caso o usuário não esteja logado, é feito o login do mesmo, onde a amostra de login é armazenada. Por fim, o algoritmo de autenticação contínua compara os descritores com a identidade que o usuário diz ter, podendo ocorrer uma fusão das similaridades ao longo do tempo para a tomada de decisão. Por exemplo, caso a distância entre a amostra atual e a amostra do usuário reivindicado seja maior que um limiar, o sistema bloqueia o acesso.

2.1 TRABALHOS RELACIONADOS

A forma de digitar foi a característica biométrica pioneira utilizada na autenticação contínua. Um exemplo de sua utilização é na realização de exames a distância. Estudantes poderiam ceder o acesso a outras pessoas para realização do exame, uma vez que os tradicionais métodos de usuário/senha não conseguem lidar com isso. Através de características como velocidade de digitação, letras por determinada unidade de tempo, e tempo entre o apertar e o soltar de teclas, esta característica é capaz de validar constantemente o usuário do sistema (FLIOR; KOWALSKI, 2010b; GUNETTI; PICARDI, 2005; LEGGETT et al., 1991). Tal biometria é vista na literatura até nos dias atuais (MONACO; TAPPERT, 2018; ACAR et al., 2018; SINGH et al., 2018; FENU; MARRAS;

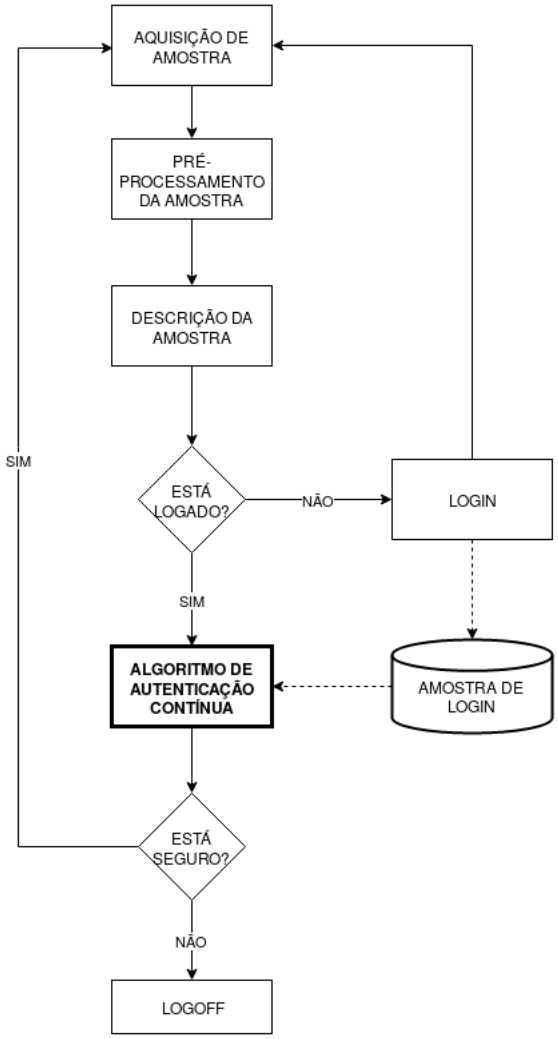


Figura 2.1: Diagrama de estágios em um sistema de autenticação contínua.

BORATTO, 2018). Apesar dos benefícios, é necessário muito tempo para se detectar um impostor, e comandos nocivos ao sistema podem ser digitados rapidamente antes que o sistema invalide o acesso.

Combinando facilidade de aquisição de amostras, frequência de captura e capacidade de distinguir indivíduos, a biometria facial foi alvo de diversas pesquisas para autenticação contínua (PAMPLONA et al., 2013; SIM et al., 2007; SILVA; SEGUNDO, 2015; SANTOS; SEGUNDO, 2015; JANAKIRAMAN et al., 2005; NIINUMA; PARK; JAIN, 2010).

Atualmente, existe um aumento no interesse pela autenticação contínua para *smartphones* (BARBELLO, 2016). Logo, diferente biometrias podem ser utilizadas, como a voz (KUNZ et al., 2011; FENG; FAWAZ; SHIN, 2017) e a forma de tocar na tela (FIERREZ et al., 2018; MONDAL; BOURS, 2018). Alguns trabalhos também combinam mais de uma biometria para melhorar o desempenho do processo de autenticação contínua (KU-

MAR et al., 2005; FENU; MARRAS; BORATTO, 2018; MURMURIA et al., 2015; SIM et al., 2007).

Apesar dos diferentes métodos na literatura, em geral, cada um deles está fortemente atrelado à biometria escolhida e a sua aplicação, o que dificulta muito a comparação entre trabalhos. Por exemplo, é difícil comparar sistemas que usam a forma de digitar ou de tocar na tela, com um sistema de autenticação contínua facial: os primeiros tem uma frequência de captura muito menor (*e.g.* um usuário pode não utilizar o teclado nem tocar na tela do *smartphone* caso esteja assistindo a um vídeo) quando comparado à captura de faces (*e.g.* é muito provável que haja uma face visível enquanto se utiliza qualquer sistema). Além disso, o contexto influencia muito: em um exame online onde o usuário escreve textos, a forma de digitar pode ser muito mais adequada do que um sistema que utiliza faces.

Apesar da incerteza de como comparar trabalhos de autenticação contínua, Sim et al. (2007) definiram novas métricas e critérios de desempenho para autenticação contínua multibiométrica, mas que também são aplicáveis a sistemas monobiométricos. Foram definidos três critérios. O primeiro deles é que diferentes modalidades biométricas possuem diferentes níveis de confiabilidade (*e.g.* impressão digital é considerada mais confiável que faces). Essa diferença deve ser levada em consideração no método de fusão das biometrias. O segundo critério é que observações antigas devem ser esquecidas para ressaltar a incerteza da presença do usuário permitido. O último critério é que um sistema de autenticação contínua deve ser capaz de determinar a segurança do sistema a qualquer instante do tempo, mesmo nos momentos em que não houve observação biométrica. Tais critérios se tornaram referência na literatura de autenticação contínua. Os últimos dois serão melhor detalhados na subseção a seguir.

2.2 ESTADO-DA-ARTE

Para atingir os 3 critérios de desempenho descritos anteriormente, um sistema de autenticação contínua deve integrar as amostras biométricas ao longo do tempo. Sim et al. (2007) também propuseram uma nova maneira de realizar autenticação contínua multimodal que integra as observações biométricas ao longo do tempo utilizando Modelos Ocultos de Markov (HMM, *Hidden Markov Models*) (RABINER, 1989). O HMM é uma sequência de estados que emite observações ao longo do tempo. Aplicando-o para o contexto de segurança, cada estado pode assumir o valor *Seguro* ou *Invadido* e existe uma probabilidade associada a cada transição, como ilustrada na Figura 2.2. O estado *Seguro* significa que o usuário genuíno continua utilizando o sistema, enquanto o estado *Invadido* significa que um invasor está utilizando o sistema. As observações ao longo do tempo podem ser quaisquer amostras biométricas. Diferente dos modelos comuns que buscam a sequência mais provável de estados dado um histórico de observações, a ideia é inferir a probabilidade do sistema estar no estado *Seguro* dado um histórico de observações.

Sim et al. (2007) propuseram um método holístico para calcular essas probabilidades ao longo do tempo e comparou com outros métodos de integração temporal. Dentre eles, o método *Temporal-First* foi modificado por Pamplona et al. (2013), e se tornou o estado-da-arte em integração temporal. O método de Pamplona et al. (2013) consiste em

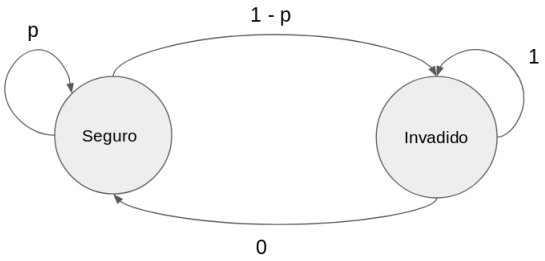


Figura 2.2: Diagrama de transição dos estados.

calcular para cada amostra ao longo de um histórico de observações $\mathcal{Z}_t$ a probabilidade do sistema estar seguro, chamada de P_{seguro}, no momento t desse histórico.

Com isso, podemos modelar o problema de autenticação contínua como sendo de calcular a todo instante a probabilidade do sistema estar seguro (P_{seguro}). Um bom sistema de autenticação contínua deve manter o valor de P_{seguro} alto enquanto o usuário permitido utiliza o sistema, assim como deve manter o valor de P_{seguro} baixo caso um invasor esteja utilizando o sistema, como ilustrado na Figura 2.3. Isso facilita a comparação entre trabalhos.

Pamplona et al. (2013) definiu o valor de P_{seguro} segundo a Equação 2.1:

$$P_{seguro} = \frac{2^{-\frac{\Delta t}{K}} \times P(x = seguro|\mathcal{Z}_t)}{P(x = seguro|\mathcal{Z}_t) + P(x = invadido|\mathcal{Z}_t)} \tag{2.1}$$

onde Δt é o tempo que se passou desde a última amostra capturada e K é uma taxa de decaimento que reduz a probabilidade do sistema estar seguro se o intervalo é muito grande, indicando a incerteza sobre a segurança se não houver novas amostras biométricas. Por exemplo, para $K = 10$, após 10 segundos sem amostras, o valor de P_{seguro} cai pela metade (PAMPLONA et al., 2013), como ilustrado na Figura 2.4. O valor de K depende da aplicação: em um cenário de alta segurança, K deve ser o mais baixo possível, como por exemplo em um caixa eletrônico de banco; já em um desktop doméstico, o valor de K pode ser maior. Com isso é possível determinar a segurança do sistema a qualquer instante do tempo, mesmo nos momentos em que não houve observação biométrica, atingindo um dos critérios de desempenho propostos por Sim et al. (2007).

As Equações 2.2 e 2.3 correspondem aos termos expandidos da Equação 2.1:

$$P(x = seguro|\mathcal{Z}_t) \propto P(z_t|x = seguro) + 2^{\frac{u-t}{K}} \times P(x = seguro|\mathcal{Z}_u) \tag{2.2}$$

$$P(x = invadido|\mathcal{Z}_t) \propto P(z_t|x = invadido) + 2^{\frac{u-t}{K}} \times P(x = invadido|\mathcal{Z}_u) \tag{2.3}$$

onde $z_t \in \mathcal{Z}_t$ corresponde a última amostra capturada, u o tempo da última amostra antes de t, z_u. A Figura 2.5 ilustra tal histórico de observações caso o sistema utilize faces como

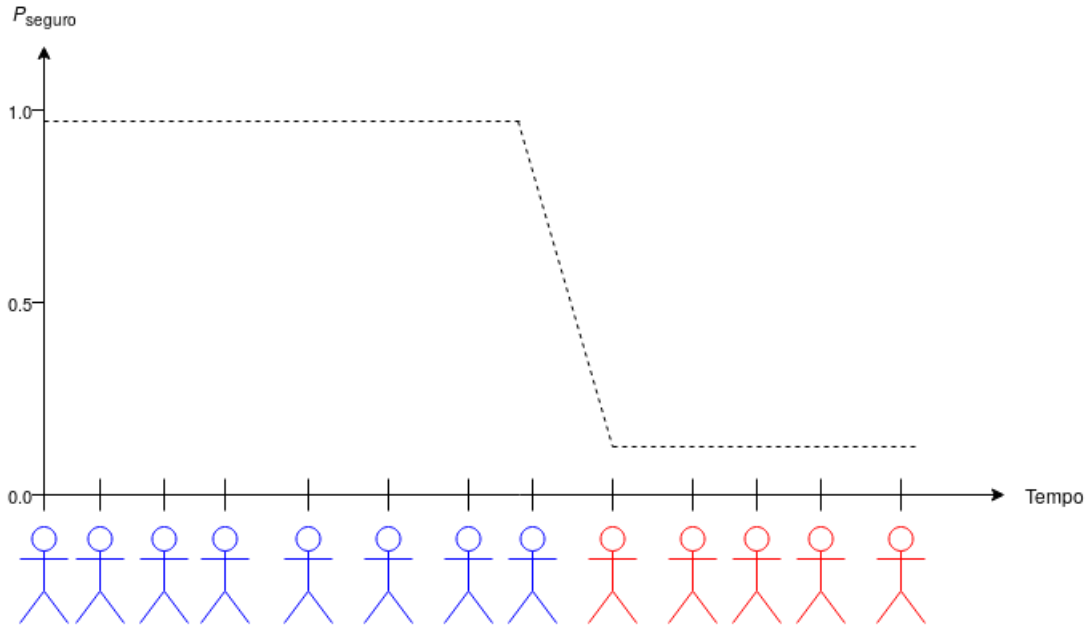


Figura 2.3: Ilustração do comportamento de um bom sistema de autenticação contínua. O usuário permitido (em azul) tem suas amostras biométricas capturas ao longo do tempo, até que em um determinado momento, um invasor assume o sistema (em vermelho). Enquanto o usuário permitido utiliza o sistema, o valor de P_{seguro} se mantém alto. Porém, no momento em que um invasor começa a utilizar o sistema, o valor de P_{seguro} cai e se mantém baixo.

amostras. Para garantir que o usuário atual continua sendo o usuário genuíno, antigas observações são esquecidas sem a necessidade de manter um histórico de observações, apenas as probabilidades da última observação e da atual são utilizadas. Desta forma é atingido o critério de desempenho de que observações antigas devem ser esquecidas para ressaltar a incerteza da presença do usuário permitido. O sistema deve assumir que exista um método de login confiável, e portanto, está seguro no momento do login. Com base nisso temos que $P(x = seguro|\mathcal{Z}_0) = 1$ e $P(x = invadido|\mathcal{Z}_0) = 0$.

No momento do login é armazenada a amostra de login. Toda vez que uma nova amostra z_t é adquirida, é calculada a distância entre essa amostra e a de login. Essa distância pode ser classificada como genuína (*i.e.* é uma distância entre amostras da mesma pessoa) ou impostora (*i.e.* é uma distância entre pessoas diferentes). O valor de $P(z_i|x = seguro)$ é a probabilidade da distância entre a amostra z_i e a amostra de login ser equivalente a uma distância entre duas amostras de uma mesma pessoa. Já $P(z_i|x = invadido)$ é a probabilidade da distância entre a amostra z_i e a amostra de login ser equivalente a uma distância entre pessoas diferentes. Sim et al. (2007) calculou estas probabilidades por meio de histogramas de distâncias entre amostras da mesma pessoa e de pessoas diferentes. Já Pamplona et al. (2013) simplificou este processo usando funções

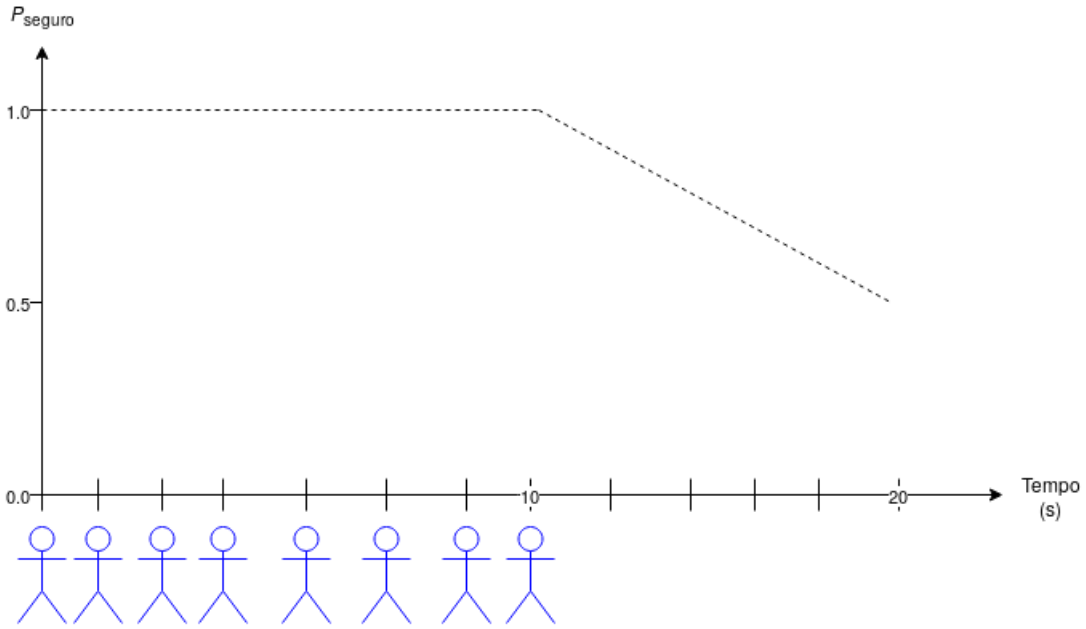


Figura 2.4: Ilustração do decaimento do valor de P_{seguro} na ausência de observações para $K = 10s$. Amostras do usuário permitido estavam sendo capturadas constantemente até o momento $t = 10s$. Após esse momento, não houveram mais capturas de novas amostras. Após 10 segundos sem novas amostras o valor de P_{seguro} caiu pela metade, ressaltando a incerteza sobre a segurança do sistema.

de distribuição acumulada (Cumulative Distribution Function (CDF)), como ilustrado na Figura 2.6:

Utilizando as CDFs é possível calcular $P(z_i|x = seguro)$ e $P(z_i|x = invadido)$, através das Equações 2.4 e 2.5:

$$P(z_i|x = seguro) \propto 1 - \frac{1}{2}\left[1 + erf\left(\frac{distância - \mu_{seguro}}{\sigma_{seguro} \times \sqrt{2}}\right)\right] \tag{2.4}$$

$$P(z_i|x = invadido) \propto \frac{1}{2}\left[1 + erf\left(\frac{distância - \mu_{invadido}}{\sigma_{invadido} \times \sqrt{2}}\right)\right] \tag{2.5}$$

sendo que os parâmetros μ e σ são respectivamente a média e desvio padrão das distâncias entre amostras genuínas e impostoras, e *distância* é a distância entre a amostra atual z_t e a amostra de login. Os parâmetros μ e σ podem ser calculados em uma base de dados fazendo a combinação entre amostras de uma mesma pessoa e de pessoas diferentes.

Ambas as técnicas são bastante similares e atingem os novos critérios de desempenho propostos por Sim et al. (2007). As principais vantagens dessas integrações temporais

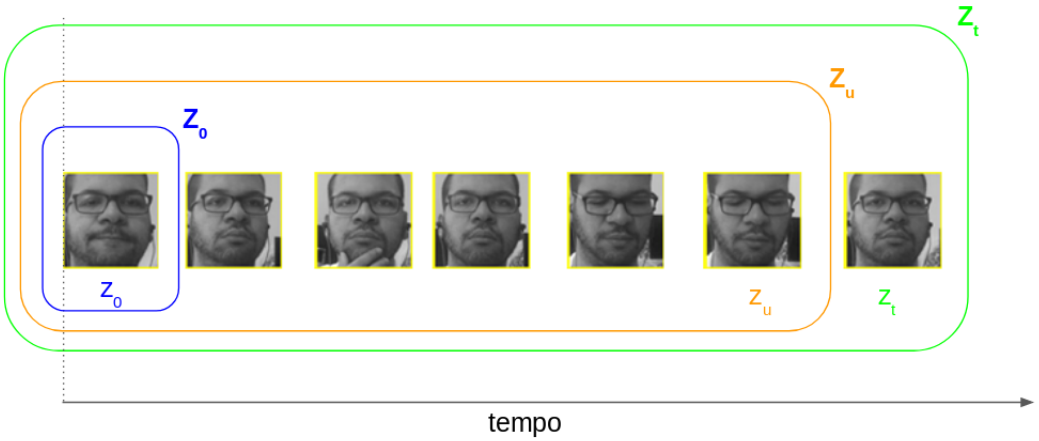


Figura 2.5: Ilustração do histórico de observações $\mathcal{Z}_t$ ao longo do tempo.

são o baixo custo computacional e os bons resultados reportados pelos autores. Porém, o fator crucial dos métodos é a forma como são calculados os valores de $P(z_i|x = seguro)$ e $P(z_i|x = invadido)$, onde ambos partem do pressuposto de que as distâncias entre amostras biométricas seguem a distribuição de uma base de treino, o que pode não ser aplicável em um cenário real ou quando uma biometria ou descritor diferente forem utilizados. Além disso, este método é bastante limitado a quão bem o descritor consegue separar os genuínos dos invasores. Caso haja uma grande intersecção entre as distribuições das distâncias dos genuínos e dos impostores, o desempenho do sistema irá cair. Por fim, os parâmetros μ e σ dependem do tamanho da base onde foram calculados. Bases muito pequenas tendem a ter uma grande diferença entre os genuínos e impostores por serem mais fáceis enquanto bases muito grandes tendem a aproximar essas duas distribuições. Sim et al. (2007) e Pamplona et al. (2013) não realizaram um estudo sobre o número ótimo de indivíduos para o cálculo dos parâmetros e nem reportaram o impacto causado pela escolha da base em seus resultados.

Uma possível solução é utilizar técnicas de detecção de anomalia para o cálculo de $P(z_i|x = seguro)$ e $P(z_i|x = invadido)$ (2.4, 2.5) e manter as equações 2.1, 2.2 e 2.3 que foram utilizadas por Pamplona et al. (2013). É importante ressaltar que $P(z_i|x = seguro)$ e $P(z_i|x = invadido)$ não precisam ser necessariamente valores de probabilidade. Basta que esses valores determinem o quão seguro um sistema pode estar, dada a última observação z_t.

No capítulo seguinte é estudada a tarefa de detecção de anomalia, assim como a descrição de algumas técnicas e como elas podem ser adaptadas para o contexto de autenticação contínua.

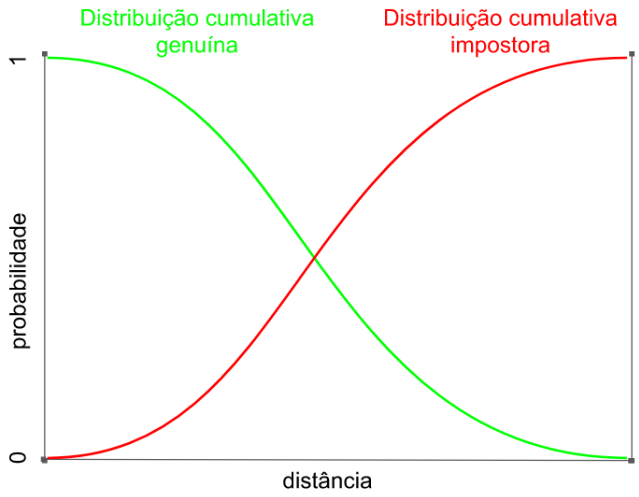


Figura 2.6: Ilustração de uma CDF.

Capítulo 3

DETECÇÃO DE ANOMALIAS

Uma anomalia (ou *outlier*) pode ser descrita como algo que foge de um comportamento padrão esperado. A tarefa de detectar uma anomalia é muito importante em diversas áreas, pois elas podem representar um grande perigo ou dano em seu respectivo contexto. Por exemplo, uma anomalia em um padrão de compras com cartão de crédito pode significar que o cartão foi roubado (ALESKEROV; FREISLEBEN; RAO, 1997). Uma anomalia em uma imagem médica pode indicar um tumor (SPENCE; PARRA; SAJDA, 2001). Na computação, uma anomalia no tráfego de rede pode representar que um computador foi invadido e está enviando informações confidenciais (KUMAR, 2005; CHANDOLA; BANERJEE; KUMAR, 2009).

Com o avanço da Internet e da utilização cada vez maior de computadores, sistemas de detecção de intrusos se tornaram um complemento importante à segurança cibernética. Quando um invasor consegue passar pela segurança do sistema (*e.g.* firewalls, protocolos, sistemas de login), ele ainda pode ser detectado como um invasor e as devidas medidas podem ser tomadas. Em sistemas críticos é desejável que essa detecção seja feita em tempo real, para que os administradores possam tomar as medidas de segurança o mais rápido possível (HODGE; AUSTIN, 2004). Tradicionalmente, sistemas de detecção de intrusos são divididos em sistemas de detecção de assinaturas, sistemas de detecção de anomalias e sistemas híbridos. Um sistema de detecção de assinatura identifica padrões previamente conhecidos como sendo maliciosos (*e.g.* comandos para enviar dados para computadores fora de uma rede interna podem compor uma assinatura de um invasor tentando vazar dados). A principal vantagem desse tipo de sistema é a baixa taxa de alarmes falsos, pois ele sabe exatamente o que é o comportamento de um intruso. Em contrapartida, é necessário que o sistema conheça a maior quantidade possível de assinaturas invasoras, o que na prática pode não ser possível. Sistemas de detecção de anomalias detectam padrões que destoem do comportamento considerado normal (*e.g.* um grupo de comandos nunca executados previamente em um terminal pode ser uma anomalia). Como ela não se baseia em padrões previamente conhecidos como sendo invasores, esse tipo de sistema pode detectar ataques nunca vistos anteriormente. Uma outra vantagem é que os padrões de comportamento considerados normais dependem de vários

fatores, o que dificulta a tarefa invasora de descobrir os fatores necessários para simular um comportamento normal. Entretanto, tais sistemas possuem uma taxa alta de alarmes falsos (PATCHA; PARK, 2007). Além disso, definir o que é um comportamento normal pode ser uma tarefa muito difícil (CHANDOLA; BANERJEE; KUMAR, 2009). Sistemas híbridos incorporam técnicas de detecção de assinatura e anomalias. No contexto de autenticação contínua, existe muita incerteza sobre a presença do usuário permitido e não é possível definir uma assinatura para um invasor, portanto, apenas os sistemas de detecção de anomalias foram investigados para este trabalho.

Sistemas de autenticação contínua e detecção de anomalias possuem objetivos em comum. Ambos os sistemas devem detectar uma invasão o mais rápido possível enquanto garantem que os usuários permitidos continuem a utilizar o sistema. Por esse motivo, este trabalho propõe uma análise de técnicas de detecção de anomalias para o contexto de autenticação contínua. Nas subseções a seguir são detalhadas tais técnicas e como elas podem ser adaptadas para a autenticação contínua.

3.1 AGRUPAMENTO

Como uma anomalia é um objeto que destoa de um comportamento padrão, é natural que algoritmos de agrupamento possam ser utilizados para encontrar anomalias (TAN; STEINBACH; KUMAR, 2005; PORTNOY; ESKIN; STOLFO, 2001; RAMASWAMY; RASTOGI; SHIM, 2000). Algoritmos de agrupamento buscam descobrir grupos de objetos que compartilhem as mesmas características. Grupos são compostos de objetos fortemente relacionados uns com os outros. Apesar de agrupamento e detecção de outliers estarem correlacionados (*e.g.* um outlier pode ser definido como um objeto que não pertence a nenhum grupo após o agrupamento), eles foram analisados separadamente.

É possível modelar a autenticação contínua como um problema de dois grupos: do usuário genuíno e do invasor. Utilizando um histórico de amostras biométricas como objetos e particionando esse histórico nos dois grupos, as pertinências de uma ou mais amostras podem ser utilizadas como valores de $P(z_i|x = seguro)$ e $P(z_i|x = invadido)$. Com base nessa abordagem, podemos analisar os tipos de agrupamento e sua conformidade com o contexto de autenticação contínua.

Agrupamento hierárquico encontra grupos formando uma árvore, onde cada nó representa um grupo formado pela união de nós filhos. Esta árvore pode ser construída de duas formas: aglomerativa ou divisiva. Na abordagem aglomerativa a árvore é construída a partir de cada elemento representando um grupo e a cada interação esses grupos são combinados até formar a raiz da árvore contendo todos os elementos. Na divisiva todos os elementos formam um único grupo inicialmente (a raiz) e a cada iteração são gerados subgrupos menores até atingir um critério de parada. Esses algoritmos tem custo computacional elevado e diferentes parâmetros a serem escolhidos (critério de parada, critério de fusão ou divisão de subgrupos, etc) e são mais utilizados quando se deseja obter uma taxonomia do conjunto de dados. Para autenticação contínua isso é desnecessário, visto que a modelamos como um problema de apenas dois grupos.

Diferentemente do agrupamento hierárquico, o agrupamento particional obtém um único particionamento dos dados. Esse particionamento pode ser de dois tipos: *hard*

(também chamado de *crisp*) ou flexível. No agrupamento *hard* cada objeto do conjunto de dados só pode pertencer a um único grupo. Tal abordagem é muito exclusiva. Durante a autenticação contínua existe muita incerteza sobre a pertinência de uma amostra ao grupo do usuário genuíno ou do invasor (*e.g.* em um sistema que utiliza faces, variações de pose e expressão facial podem fazer com que uma face seja incorretamente classificada como invasora). Tal incerteza não é levada em conta em grupos disjuntos, e, por isso, não é muito adequada para autenticação contínua.

O agrupamento flexível cria grupos com sobreposição para ilustrar cenários onde um objeto pode pertencer a mais de um grupo ao mesmo tempo. Este cenário é mais próximo do contexto de autenticação contínua, onde muitas vezes existe uma incerteza quanto a pertinência de uma amostra aos dois grupos. Como exemplo de agrupamento particional flexível tem-se o agrupamento fuzzy, onde cada objeto possui um grau de pertinência que varia de 0 a 1 para cada grupo (TAN; STEINBACH; KUMAR, 2005; HARTIGAN; WONG, 1979). Outro exemplo de agrupamento flexível é o agrupamento probabilístico, que calcula a probabilidade de cada objeto pertencer a cada grupo. Porém tais algoritmos se utilizam de distribuições estatísticas, o que gera dois problemas: distribuições dependem muito do tamanho da base onde foram calculadas, podendo não representar bem um cenário real e muitas vezes são muito complicadas de serem calculadas.

Dos tipos de agrupamento, o fuzzy é o que mais se aproxima do contexto de autenticação contínua. Ele consegue capturar a incerteza sobre as amostras atribuindo um grau de pertinência a ambos os grupos. O próprio conceito de segurança é fuzzy (BRIDGES; VAUGHN et al., 2000). Deste modo, sua utilização ajuda a suavizar a fronteira que separa usuários permitidos de usuários invasores. Por esses motivos, o agrupamento fuzzy foi estudado com mais aprofundamento e é detalhado a seguir.

3.1.1 Fuzzy C-Means

O algoritmo Fuzzy C-Means (FCM) (BEZDEK; EHRLICH; FULL, 1984) é uma versão fuzzy do algoritmo K-means (HARTIGAN; WONG, 1979). O algoritmo consiste em escolher centroides aleatórios para cada grupo fuzzy e atribuir valores de pertinência para cada objeto. Após essa inicialização, o algoritmo atualiza os centroides para cada grupo e recalcula as pertinências repetidamente até que os centroides não mudem mais ou alguma condição de parada seja atingida (*e.g.* se a diferença entre o erro da última iteração e a atual estiver abaixo de algum limiar).

Dado um conjunto de objetos $P = \{x_1, ..., x_n\}$, onde cada objeto x_i é um objeto d-dimensional (*i.g.* vetor com d posições), uma coleção de grupos fuzzy $G_1, ..., G_k$, é o subconjunto de todos os possíveis subgrupos fuzzy de P. A pertinência de um objeto x_i para um grupo G_j é dada por w_{ij}, um valor entre 0 e 1. Para que os grupos possam formar uma partição pseudo-fuzzy duas condições devem ser cumpridas (TAN; STEINBACH; KUMAR, 2005):

1. A soma dos graus de pertinências de cada objeto deve ser igual a 1 ($\sum_{j=1}^{k} w_{ij} = 1$).

2. Todos os objetos pertencem a todos os grupos e nenhum grupo G_j pode conter todos os objetos com pertinência igual a 1 ($0 < \sum_{i=1}^{n} w_{ij} < n$).

Assim como o K-means, o FCM tenta minimizar a Soma dos Erros Quadrados (SEQ). Os detalhes do FCM são descritos abaixo.

1. **Cálculo do SEQ**: A Soma dos Erros Quadrados (SEQ) modificada para o FCM é dada pela equação:

$$SEQ(G_1, ..., G_k) = \sum_{j=1}^{k} \sum_{i=1}^{n} (w_{ij})^m dist(x_i, c_j)^2 \tag{3.1}$$

   onde $dist(x_i, c_j)$ é a distância entre o centroide c_j do grupo G_j e um objeto x_i (*e.g.* distância euclidiana entre dois objetos), e m é a constante de fuzzificação que controla a influência das pertinências, sendo $1 < m < \infty$.

2. **Cálculo dos centroides**: Assim como no K-means, devemos achar um centroide que minimize a SEQ para otimização do problema. A definição de um centroide fuzzy é bastante similar as definições tradicionais, porém o centroide fuzzy leva em consideração todos os objetos, cada objeto com sua contribuição dada pelo grau de pertinência. Para um grupo G_j, o seu centroide c_j é definido pela equação:

$$c_j = \frac{\sum_{i=1}^{n} (w_{ij})^m x_i}{\sum_{i=1}^{n} (w_{ij})^m} \tag{3.2}$$

3. **Atualização das pertinências**: Intuitivamente, para minimizar a SEQ, a pertinência de um dado objeto x_i a um grupo G_j deve ser alta se a distância entre esse objeto e o centroide do grupo c_j for pequena, e a pertinência deverá ser baixa se a distância for grande. Como a soma das pertinências de um objeto deve ser igual a 1, esse cálculo deve ser normalizado:

$$w_{ij} = \frac{1}{\sum_{q=1}^{k} \left(\frac{dist(x_i, c_j)}{dist(x_i, c_q)}\right)^{\frac{2}{m-1}}} \tag{3.3}$$

Um exemplo de funcionamento do Fuzzy C-Means é ilustrado no Algoritmo 1:

Algoritmo 1 Algoritmo básico do Fuzzy C-Means

Selecione aleatoriamente objetos em P como centroides para as partições e atribua valores a cada pertinência w_{ij} usando a Equação 3.3;
repita
    Compute os centroides de cada partição usando a Equação 3.2;
    Atualize cada pertinência w_{ij} usando a Equação 3.3;
até *Centroides não mudem ou alguma condição de parada seja alcançada*;

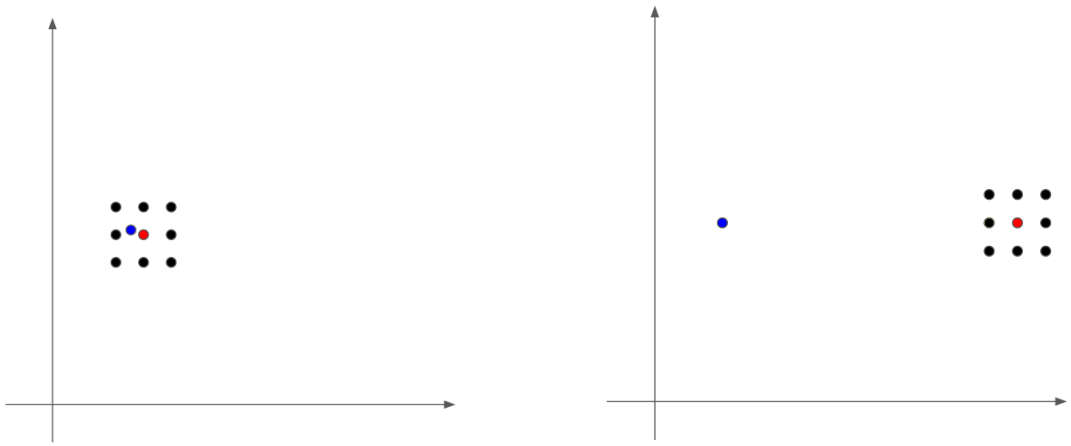
Para ser utilizado no contexto de autenticação contínua, podemos definir que o FCM formará dois grupos, um do usuário permitido $G_{genuíno}$ e um do invasor $G_{invasor}$. No momento inicial do login, a amostra de login é a única amostra que o sistema tem certeza que é segura. Com isso, é preferível que tal amostra seja utilizada como protótipo do grupo genuíno, garantindo assim que qualquer outra amostra subsequente só terá pertinência alta ao grupo genuíno se estiver próxima da amostra mais genuína. Após armazenar a amostra de login, o sistema deverá guardar um conjunto X, tal que $X \subset \mathcal{Z}_t$, onde X é composto pelas últimas n amostras biométricas (exceto a amostra de login). A cada nova amostra z_t, deve-se atualizar o conjunto X e rodar o algoritmo FCM sobre o novo X. O grau de pertinência de z_t a $G_{genuíno}$, chamado $w_{t,genuíno}$, pode ser utilizado no lugar de $P(z_t \mid x = seguro)$ e a pertinência a $G_{invasor}$, chamado $w_{t,invasor}$, para $P(z_t \mid x = invadido)$ nas Equações 2.2 e 2.3. O funcionamento completo do sistema é exemplificado pelo Algoritmo 2:

Algoritmo 2 Algoritmo de autenticação contínua utilizando Fuzzy C-Means

armazena a amostra de login;
X = últimas n amostras em $\mathcal{Z}_t$;
repita
  se z_t = *nova amostra válida* **então**
    $X = \{X - \{z_v\}\} \cup \{z_t\}$, sendo z_v a observação mais antiga em X;
    $c_{genuíno}$ = amostra de login;
    $c_{invasor} = argmax_{x_i \in X}(dist(x_i, c_{genuíno}))$;
    Atualize cada pertinência $w_{i,j}$ de $x_i \in X$ utilizando a Equação 3.3;
    repita
      Calcule $c_{invasor}$ de X utilizando a Equação 3.2;
      Atualize cada pertinência $w_{i,j}$ de X utilizando a Equação 3.3;
    até $c_{invasor}$ *não mude*;
    $P(z_t \mid x = seguro) = w_{t,genuíno}$;
    $P(z_t \mid x = invadido) = w_{t,invasor}$;
  fim
  Use as Equações 2.1, 2.2 e 2.3 para determinar P_{seguro};
até $P_{seguro} < limiar$;

No decorrer da autenticação contínua, amostras do usuário genuíno se parecem entre si e com a amostra inicial de login, enquanto amostras do usuário invasor se parecem entre si, mas são diferentes da de login. Isso significa que amostras do usuário genuíno e do invasor formam grupos diferentes. Intuitivamente, enquanto o usuário genuíno estiver utilizando o sistema, mesmo o algoritmo formando um grupo do invasor, as amostras terão alto grau de pertinência ao grupo genuíno, pois tais amostras estão bastante próximas do centroide do grupo genuíno (os dois grupos estarão muito próximos). Quando um invasor assume o sistema, o grupo invasor se afasta do grupo genuíno e as amostras invasoras terão um baixo grau de pertinência ao grupo genuíno e alto grau ao invasor, pois elas estão mais próximas do centroide invasor do que do centroide genuíno. Ambas as situações são ilustradas em um cenário em duas dimensões na Figura 3.1. Escolhendo um tamanho do

histórico pequeno possibilita-se que o sistema seja rápido, pois serão calculadas poucas distâncias.



(a) (b)

Figura 3.1: Ilustrações 2D de um dado momento no tempo após aplicado o FCM , onde o usuário genuíno está utilizando o sistema (a) e quando um invasor está utilizando o sistema (b). O ponto azul corresponde à amostra de login e o vermelho ao centroide invasor.

Apesar de promissor, o FCM possui um grande problema: se uma amostra do usuário genuíno está próxima do centroide invasor (composto por amostras do usuário genuíno) e não tão próxima da amostra de login por conta das variações intraclasse, o FCM atribuirá uma alta pertinência ao grupo invasor e baixa ao grupo genuíno. Essa situação é ilustrada na Figura 3.2. Na prática, o algoritmo não seria capaz de distinguir um usuário genuíno de um invasor, ambos seriam considerados sempre invasores. Uma possível solução seria aumentar a constante de fuzzificação, fazendo os grupos ficarem mais difusos, porém isso melhoraria o caso onde o usuário genuíno está presente, mas pioraria quando um invasor está presente. Esse problema do FCM é causado pela restrição de que a soma das pertinências de um objeto deve somar um. Isso é modificado no algoritmo Possibilistic C-Means (PCM) (KRISHNAPURAM; KELLER, 1993).

3.1.2 Possibilistic C-Means

O PCM é uma variação do FCM , onde a restrição de que a soma das pertinências de um objeto tem que somar um é removida. Consequentemente, a função objetivo é modificada para evitar a solução trivial, segundo a equação:

$$SEQ(G_1, ..., G_k) = \sum_{j=1}^{k}\sum_{i=1}^{n}(w_{ij})^m dist(x_i, c_j)^2 + \sum_{j=1}^{k}\eta_j \sum_{i=1}^{n}(1 - w_{ij})^m \qquad (3.4)$$

e consequentemente os graus de pertinência são dados pela equação:

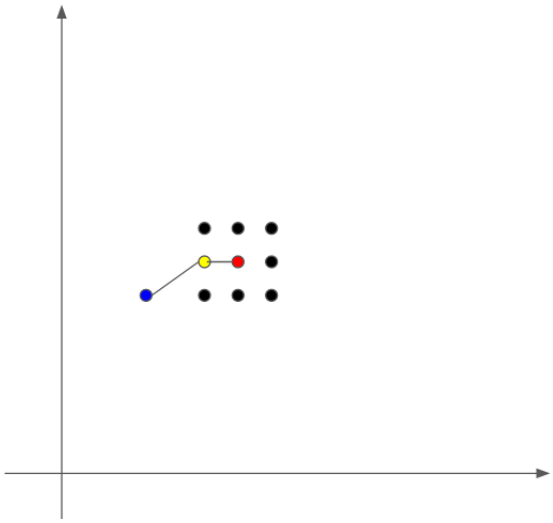


Figura 3.2: Problema do FCM enquanto o usuário permitido utiliza o sistema. O ponto amarelo representa a última amostra no histórico. Amostras do mesmo usuário permitido se parecem entre si, portanto o centróide invasor tende a estar mais próximo da última amostra do que a amostra de login.

$$w_{ij} = \frac{1}{1 + (\frac{dist(x_i, c_j)}{\eta_j})^{\frac{1}{m-1}}} \tag{3.5}$$

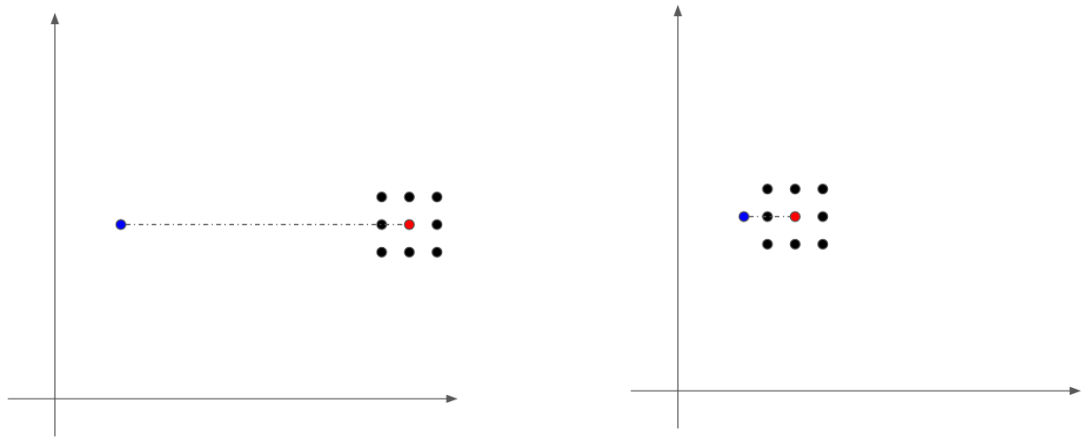
A constante η_j define qual deve ser a distância entre uma amostra e um centroide de um grupo j para que a pertinência da amostra seja 0.5. Em outras palavras, se a distância entre uma amostra e um centroide de um grupo j for bem menor que η_j, sua pertinência será alta (*e.g.* próxima de 1). Se a distância é exatamente η_j, a pertinência será 0.5. Se a distância for bem maior que η_j, a pertinência será baixa (*e.g.* próxima de 0). Existem diversas abordagens para a definição da constante η_j. Ela pode ser definida a priori e não ser alterada, pode ser calculada dinamicamente ao longo das iterações do algoritmo ou apenas uma vez (KRISHNAPURAM; KELLER, 1993). O método mais comum de definir seu valor é dado pela equação abaixo:

$$\eta_j = \Phi \frac{\sum_{i=1}^{n} (w_{ij})^m dist(x_i, c_j)^2}{\sum_{i=1}^{n} (w_{ij})^m} \tag{3.6}$$

Onde em geral $\Phi = 1$. Em cenários práticos de autenticação contínua, o cálculo de η_j utilizando essa equação faria com que as pertinências para o grupo genuíno e invasor ficassem muito parecidas. Por exemplo, considerando o exemplo da Figura 3.1b, como todas as distâncias entre as amostras e a amostra de login são altas e suas pertinências iniciais são baixas, o valor de $\eta_{genuíno}$ será alto também (*e.g.* próximo a média das

distâncias), e consequentemente um invasor pode ter uma pertinência ao grupo genuíno próxima ou maior que 0.5. Portanto, é necessária uma maneira mais intuitiva e eficiente de definir os valores de η_j.

Supondo uma métrica de distância que varia de 0 a 100, a Figura 3.3a ilustra um cenário hipotético onde o centroide invasor está quase na distância máxima da amostra de login. Intuitivamente, isso significa que muito provavelmente um invasor está utilizando o sistema e que há menos certeza sobre a segurança do sistema. Consequentemente, o valor de $\eta_{genuíno}$ deveria ser baixo e $\eta_{invasor}$ alto, fazendo com que seja mais difícil que um invasor se passe pelo usuário genuíno.



(a) (b)

Figura 3.3: Cenário hipotético após aplicado o PCM , onde o centroide invasor está próximo da distância máxima possível da amostra de login (a). Cenário hipotético onde o centroide invasor está muito próximo da amostra de login (b).

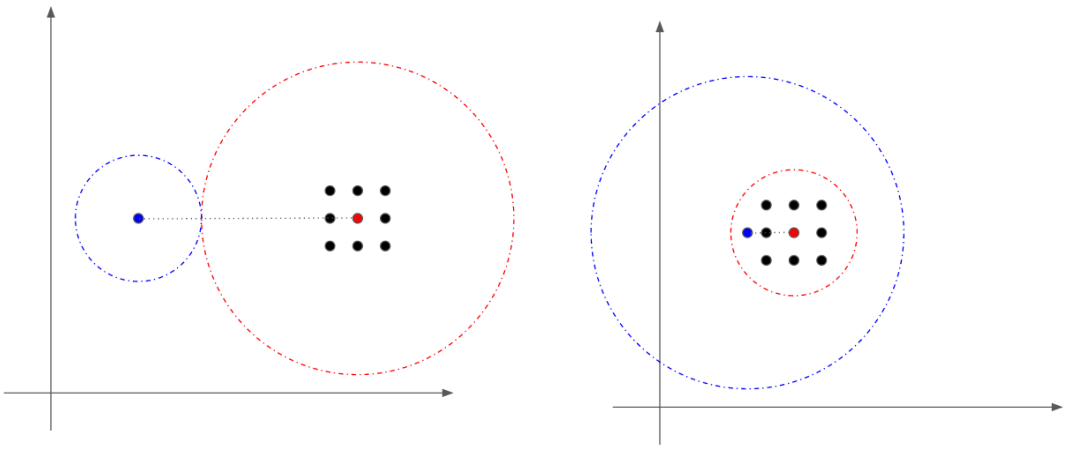
De maneira análoga, a Figura 3.3b ilustra o caso contrário, quando o centroide invasor está muito próximo da amostra do login. Intuitivamente, isso significa que muito provavelmente o usuário genuíno está utilizando o sistema e há mais certeza sobre a segurança do sistema. O valor de $\eta_{genuíno}$ deveria ser alto e $\eta_{invasor}$ baixo, reduzindo o impacto de variações intra-classe. A Figura 3.4 ilustra um cenário dessa intuição. Utilizando essa intuição, podemos definir o valor de $\eta_{invasor}$ pela Equação 3.7:

$$\eta_{invasor} = dist(c_{genuíno}, c_{invasor}) \tag{3.7}$$

Para os valores de $\eta_{genuíno}$, algumas abordagens são possíveis. A primeira delas é uma abordagem otimista, definida pela Equação 3.8:

$$\eta_{genuíno} = DIST\_MAX - dist(c_{genuíno}, c_{invasor}) \tag{3.8}$$

onde $DIST\_MAX$ é a distância máxima permitida pela métrica (*e.g.* utilizando a distância cosseno (1 - similaridade cosseno), temos que $DIST\_MAX = 2$). Nesta abor-



(a) (b)

Figura 3.4: Ilustração do valor de η_j quando um invasor está utilizando o sistema (a). A linha tracejada azul corresponde ao valor de $\eta_{genuíno}$ e a linha tracejada vermelha ao valor de $\eta_{invasor}$. Ilustração do valor de η_j quando o usuário genuíno está utilizando o sistema (b).

dagem estamos supondo que amostras de indivíduos diferentes estarão bem próximos da distância máxima, como ilustrado na Figura 3.5a.

Essa abordagem pode trazer o seguinte problema: na prática, descritores não conseguem separar tão bem amostras de indivíduos diferentes. A Figura 3.5 ilustra o que acontece com os valores de $\eta_{genuíno}$ para diferentes descritores. Na Figura 3.5c o descritor usa apenas metade da distância máxima. Com isso, uma amostra do usuário invasor terá pertinência ao grupo genuíno de um pouco menor que 0.5. Uma forma de contornar essa situação é reduzir o valor da distância máxima, visto que nem todos os descritores conseguirão utilizar todo seu alcance. Podemos definir um novo valor para $\eta_{genuíno}$ utilizando a Equação 3.9:

$$\eta_{genuíno} = DIST\_MAX * 0.5 - dist(c_{genuíno}, c_{invasor}) \tag{3.9}$$

A Figura 3.6 ilustra a utilização da Equação 3.9 usando o mesmo descritor que na Figura 3.5c. Também é possível reduzir mais ainda, utilizando a Equação 3.10:

$$\eta_{genuíno} = DIST\_MAX * 0.333 - dist(c_{genuíno}, c_{invasor}) \tag{3.10}$$

As abordagens anteriores levam em consideração apenas a distância entre os centroides, porém a distância entre a amostra de login e a última amostra podem dar informação sobre a segurança do sistema. Podemos então incluir essa distância na subtração, como mostrado na Equação 3.11. Desta forma, para amostras do usuário permitido, $dist(c_{genuíno}, z_t)$ é pequena, não reduzindo tanto o valor de $\eta_{genuíno}$. Por outro lado, para amostras do usuário invasor, $dist(c_{genuíno}, z_t)$ irá reduzir mais ainda o valor de $\eta_{genuíno}$.

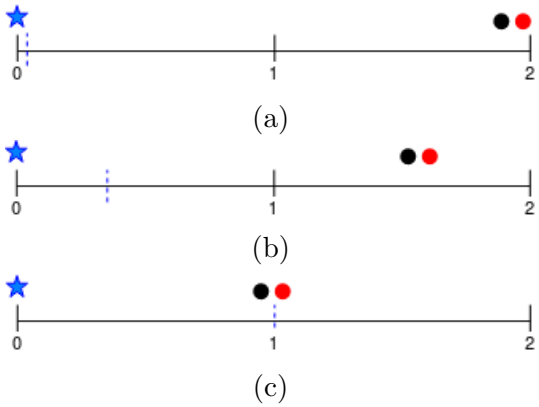


Figura 3.5: Ilustração de diferentes cenários considerando a distância cosseno quando um invasor está utilizando o sistema. A estrela em azul representa a amostra de login, o ponto vermelho representa o centroide do grupo invasor, e o ponto preto representa a última amostra. A linha azul traçejada representa o valor de $\eta_{genuíno}$. Utilizando a abordagem otimista, temos 3 descritores diferentes em ordem decrescente de capacidade de separar indivíduos: (a), (b) e (c).

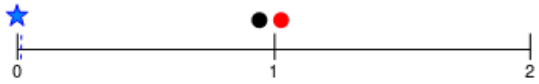


Figura 3.6: Ilustração da utilização do mesmo descritor que a Figura 3.5c, porém com a Equação 3.9.

$$\eta_{genuíno} = DIST\_MAX - dist(c_{genuíno}, c_{invasor}) - dist(c_{genuíno}, z_t) \tag{3.11}$$

Uma limitação da técnica é que não é possível utilizar uma métrica que não possua uma distância máxima (*e.g.* distância euclidiana). O funcionamento completo do algoritmo utlizando o PCM de maneira otimista é descrito no Algoritmo 3:

Algoritmo 3 Algoritmo de autenticação contínua utilizando Possibilistic C-Means

X = últimas n amostras em $\mathcal{Z}_t$;
repita
  se z_t = *nova amostra válida* **então**
    $X = \{X - z_v\} \cup z_t$, sendo z_v a observação mais antiga em X;
    $c_{genuíno}$ = amostra de login;
    $c_{invasor} = max_{x_i \in X}(dist(x_i, c_{genuino}))$;
    Calcule $\eta_{genuíno}$ e $\eta_{invasor}$ usando as Equações 3.8 e 3.7;
    Atualize cada pertinência $w_{i,j}$ de $x_i \in X$ utilizando a Equação 3.5;
    repita
      Calcule $c_{invasor}$ de X utilizando a Equação 3.2;
      Calcule $\eta_{genuíno}$ e $\eta_{invasor}$ usando as Equações 3.8 e 3.7;
      Atualize cada pertinência $w_{i,j}$ de X utilizando a Equação 3.5;
    até $c_{invasor}$ *não mude*;
    $P(z_t \mid x = seguro) = w_{t,genuíno}$;
    $P(z_t \mid x = invadido) = w_{t,invasor}$;
  fim
  Use as Equações 2.1, 2.2 e 2.3 para determinar P_{seguro};
até $P_{seguro} < limiar$;

3.2 DETECÇÃO DE OUTLIERS

A detecção de outliers é muito utilizada na mineração de dados para diversas aplicações. Os tipos mais comuns de detecção de outliers são os baseados em proximidade e os baseados em densidade. Os algoritmos baseados em proximidade partem da premissa que outliers são objetos distantes da maioria dos outros objetos do conjunto. Apesar de sua complexidade quadrática, para conjuntos com poucos objetos como no contexto de autenticação contínua, esses algoritmos são rápidos. A maior desvantagem dessa técnica é que a proximidade que define um outlier é variável e depende de cada conjunto de dados. Algoritmos baseados em densidade partem da premissa que outliers pertencem a regiões de baixa densidade. Ambas as técnicas são parecidas, compartilham vantagens e desvantagens e foram utilizadas para sistemas de detecções de anomalias (BYERS; RAFTERY, 1998; GUTTORMSSON et al., 1999; LIAO; VEMURI, 2002). As técnicas mais promissoras são mostradas nas subseções seguintes.

3.2.1 Fator Local de Outlier

O Fator Local de Outlier (Local Outlier Factor (LOF)) (BREUNIG et al., 2000) é uma medida do grau de anomalia de um objeto dentro de um conjunto de dados. Para compreender como o LOF é calculado, Breunig et al. (2000) introduziu as seguintes definições:

Definicão 1. *(k-distância de um objeto p):*

Dado um conjunto de dados D, para qualquer inteiro positivo k, a *k-distância* do objeto p (denotada por *k-dist(p)*), é definida como a distância *d(p,x)* entre p e um objeto

$x \in D$, tal que: pelo menos k objetos $x' \in D - \{x\}$ possuem *d(p,x')* $\leq$ *d(p,x)* e pelo menos $k-1$ objetos $x' \in D - \{x\}$ possuem *d(p,x')* $<$ *d(p,x)*. Em outras palavras, a *k-dist(p)* é a distância entre o objeto p e o seu k-ésimo vizinho mais próximo.

Definição 2. *(vizinhança de k-distância de um objeto p):*

A vizinhança de k-distância de um objeto p (denotada por $N_k(p)$) contém todos os objetos cujas distâncias não são maiores que a *k-distância* (i.e. $N_k(p) = \{q \in D - \{p\} \mid d(p,q) \leq$ *k-dist(p)*$\}$. Em outras palavras, os objetos q são os k vizinhos mais próximos de p.

Definição 3. *(distância de alcance de um objeto p em relação ao objeto x):*

A distância de alcance de um objeto p em relação ao objeto x (denotada por $dist - alc_k(p,x)$) é definida pela Equação 3.12.

$$\textit{dist-alc}_k(p,x) = \max\{\textit{k-dist(x)},\ \textit{dist(p,x)}\} \tag{3.12}$$

A Figura 3.7 ilustra as notações descritas anteriormente. É possível notar que mesmo para um objeto próximo (*i.e.* p_1 na Figura 3.7), a distância atual é substituída pela *k-distância* de x quando estamos calculando a distância de alcance de p_1 em relação x. Desta forma, as variações de *d(p,x)* são significantemente reduzidas.

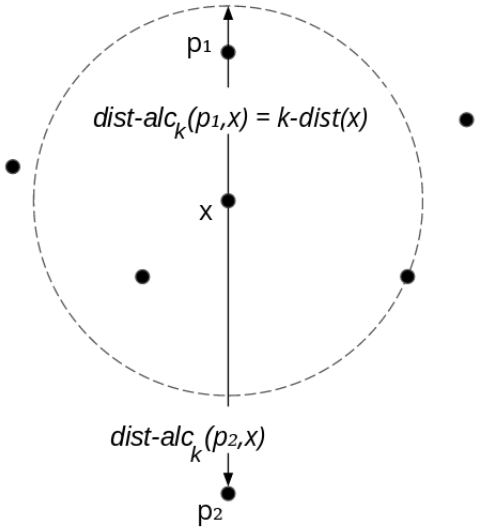


Figura 3.7: dist-alc(p_1,x) e dist-alc(p_2,x) para $k = 4$.

Definição 4. *(densidade local de alcance de um objeto p):*

A densidade local de alcance de um objeto p (denotada por $dla_k(p)$) é o inverso da média das distâncias de alcance dos seus k vizinhos mais próximos (Equação 3.13). Em

algoritmos de agrupamento baseados em densidade, k é chamado de *minPts* e define quantos objetos são necessários para se definir um grupo. No LOF este parâmetro é utilizado como uma medida de volume para determinar a densidade na vizinhança de um objeto p. A ideia é que um objeto muito distante dos seus *minPts* vizinhos mais próximos terá densidade local de alcance pequena, enquanto um objeto muito próximo de seus vizinhos mais próximos terá densidade grande.

$$dla_{minPts}(p) = 1/\left(\frac{\sum_{x \epsilon N_{minPts}(p)} dist\text{-}alc_{minPts}(p,x)}{|N_{minPts}(p)|}\right) \tag{3.13}$$

Definicão 5. *(local outlier factor (LOF) de um objeto p):*

Por fim, o valor LOF para um objeto p é a média da taxa das densidades locais de alcance de p e seus *minPts* vizinhos mais próximos (Equação 3.14). Desta forma, quanto menor a densidade local de alcance de um objeto p e maior a densidade local de alcance dos seus *minPts* vizinhos mais próximos, maior será seu valor de LOF. A Figura 3.8 ilustra o caso onde um objeto possui um valor LOF alto.

$$LOF_{minPts}(p) = \frac{\sum_{x \epsilon N_{minPts}(p)} \frac{dla_{minPts}(x)}{dla_{minPts}(p)}}{|N_{minPts}(p)|} \tag{3.14}$$

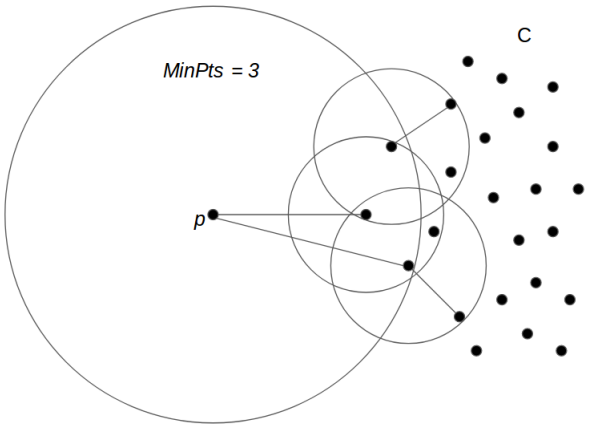


Figura 3.8: Neste exemplo o ponto p possui uma densidade local de alcance muito maior do que os seus 3 vizinhos mais próximos e consequentemente um valor de LOF alto.

Para utilizar o LOF em autenticação contínua, é possível utilizar n amostras biométricas no login, e para cada amostra subsequente é calculado seu LOF. Armazenando apenas as amostras de login, muito provavelmente as amostras subsequentes do usuário genuíno estarão próximas das amostras de login, e consequentemente terão LOF baixo enquanto que as amostras do usuário invasor estarão mais distante das amostras de login, e consequentemente terão LOF mais alto.

Devido ao LOF ser uma medida do grau de anomalia de um objeto, ele pode variar muito a depender do conjunto de dados (*e.g.* um valor de LOF de um dado objeto em conjunto de dados pode representar uma anomalia, porém o mesmo valor em um conjunto diferente de dados pode não representar uma anomalia). Portanto, é necessário determinar uma maneira de mensurar a segurança do sistema de autenticação contínua com base nos valores de LOF. Com isso, seria necessário realizar um treinamento em uma base de dados, o que vai de encontro com a proposta deste trabalho.

3.2.2 Variantes do LOF

Existem diversas extensões do LOF. Tang et al. (2002) propôs um variação do LOF chamada Fator de Outlier baseado em Conectividade (Connectivity-based Outlier Factor (COF)). A única diferença entre o LOF e o COF é a maneira como é definido o conjunto dos k vizinhos mais próximos. No COF esse conjunto é calculado de maneira incremental. Inicialmente, o objeto mais próximo é adicionado ao conjunto dos vizinhos mais próximos. O próximo objeto a ser adicionado é aquele que possuir a menor distância entre ele e algum objeto pertencente ao conjunto de vizinhos mais próximos. Novos objetos são adicionados até que se alcance a quantidade k desejada. Após calculado o grupo dos k vizinhos mais próximos, o valor de COF é calculado da mesma forma que o LOF. A Figura 3.9 ilustra a principal diferença entre as duas abordagens. O ponto n_1 é um outlier, mas por possuir uma densidade alta, dificilmente seria identificado como outlier usando o LOF.

Uma versão mais simples do LOF é o Outlier Detection using In-degree Number (ODIN). O ODIN de um objeto p é a quantidade de k-vizinhos mais próximos que também possuem p como um dos k-vizinhos mais próximos. O inverso do ODIN é o grau de anomalia de um objeto (HAUTAMAKI; KARKKAINEN; FRANTI, 2004), similar ao valor de LOF. Outra versão simples do LOF é o Multi-granularity Deviation Factor (MDEF). O MDEF de um objeto é o desvio padrão das densidades locais de seus vizinhos mais próximos (incluindo o próprio objeto). O inverso do MDEF é o grau de anomalia do objeto (IDÉ; PAPADIMITRIOU; VLACHOS, 2007). Essas variações podem ser usadas para autenticação contínua substituindo o valor de LOF, mas possuem a mesma necessidade de treino.

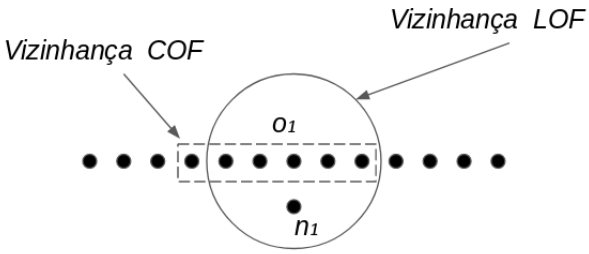


Figura 3.9: Diferença entre as vizinhanças do LOF e do COF.

3.2.3 Outilers baseados em Distância

Ramaswamy, Rastogi e Shim (2000) propuseram um algoritmo para particionar o conjunto de dados para reduzir a complexidade de encontrar outliers. Apesar de ser utilizado para conjuntos grandes de dados, a ideia de usar a distância do k-ésimo vizinho mais próximo pode ser utilizada separadamente para o contexto de autenticação contínua.

Dado um inteiro positivo k e um objeto p, denotamos $k - dist(p)$ como a distância do k-ésimo vizinho mais próximo de p. É possível notar que objetos com alto valor de $k-dist(p)$ possuem vizinhança mais esparsa, o que intuitivamente indica que tais objetos são outliers em potencial, pois objetos que pertencem a um grupo denso tem $k - dist(p)$ menores. Dessa forma, basta ordenar os objetos pelas suas distâncias até seu k-ésimo vizinho mais próximo, e selecionar os n primeiros objetos como os n outliers do conjunto de dados (RAMASWAMY; RASTOGI; SHIM, 2000).

Para o contexto de autenticação contínua, basta armazenar os valores de $k - dist(p)$ das n primeiras amostras no login e comparar com o valor de $k - dist(p)$ da amostra atual z_t. Um valor de $k - dist(z_t)$ muito diferente das primeiras amostras coletadas no login pode significar que um invasor está utilizando o sistema. De maneira análoga ao sistema utilizando LOF é necessário treinamento para descobrir o quanto tais valores podem variar para serem considerados genuínos ou invasores, e por isso, esta técnica possui a mesma desvantagem de utilizar LOF.

Tais treinamentos mencionados previamente, muitas vezes resultam em calcular a probabilidade de um determinado valor pertencer a uma distribuição específica. Portanto, métodos estatísticos estão muito presentes em detecção de anomalias e são descritos na seção seguinte.

3.3 MÉTODOS ESTATÍSTICOS

Métodos estatísticos criam um modelo para um conjunto de dados considerado normal, através de alguma distribuição estatística. Eles partem da premissa que uma anomalia tem pequena probabilidade de pertencer ao modelo gerado. Por exemplo, se os dados considerados normais pertencem a uma distribuição gaussiana, é possível calcular a média e o desvio padrão desse conjunto de dados e utilizar isso para calcular a probabilidade de uma amostra pertencer a essa distribuição. Apesar de ser rápido e bastante utilizado na detecção de anomalias, a principal desvantagem desses métodos é que nem todos os cenários irão pertencer a uma distribuição e muitas vezes é muito difícil encontrar uma distribuição que defina bem o modelo. No contexto de autenticação contínua isso é mais difícil ainda, visto que, não existem bases de dados específicas para tal tarefa. As subseções a seguir mostram algumas técnicas que poderiam ser utilizadas para autenticação contínua.

3.3.1 Teste de Resíduo Normalizado Máximo

O teste de Grubbs (também conhecido como *maximum normed residual test*) é um dos métodos mais antigos de detecção de outlier em bases de dados univariados (GRUBBS, 1969, 1950; STEFANSKY, 1972; ANSCOMBE; GUTTMAN, 1960).

Assumindo que os objetos em um conjunto de dados se aproximam de uma distribuição

Gaussiana, para cada objeto de teste z, podemos calcular seu score G_z segundo a Equação 3.15:

$$G_z = \frac{|z - \mu|}{\sigma} \tag{3.15}$$

onde μ é a média e σ o desvio padrão do conjunto de dados. O valor de G_z é comparado com um valor de referência baseado na distribuição de G com nível de significância α, como mostrado na Tabela 3.1. Um objeto z é denominado um outlier se seu score G_z for maior que o valor de referência G.

Tabela 3.1: Tabela do teste de Grubbs.

	Nível de Significância α			
tamanho do conjunto	0,1	0,05	0,025	0,01
3	1,148	1,153	1,154	1,155
4	1,425	1,462	1,481	1,492
5	1,602	1,671	1,715	1,749
6	1,729	1,822	1,887	1,944
7	1,828	1,938	2,02	2,097
...	...	...	...	...

Uma maneira simples de usar o teste de Grubbs na autenticação contínua é criar um histórico de distâncias (inicialmente composto apenas pela distância entre a amostra de login e as três primeiras amostra após ela) e para cada amostra subsequente z_t, adicionar sua distância da amostra de login no histórico, calcular seu G_{z_t} e comparar com o valor de referência. O sistema bloqueia o acesso ao usuário caso $G_{z_t} > G$. Como é utilizada a distância entre a amostra atual e a de login, quanto menor a distância entre as duas, mais genuína tal amostra é. Entretanto, como o teste de Grubbs utiliza o módulo de $z - u$, caso uma amostra tenha um valor bem menor que a média, o teste pode considerá-la como outlier, quando na verdade tal amostra é mais genuína que as demais. Por isso, é necessário modificar o método padrão do teste de Grubbs, removendo o módulo na Equação 3.15 e calculando uma nova tabela de Grubbs. Além disso, essa abordagem tem duas grandes desvantagens. A primeira é que uma grande variação intraclasse pode fazer com que a amostra atual seja considerada um outlier e o sistema bloqueie o acesso ao usuário genuíno. A segunda é que o histórico aumentaria ao longo do tempo tornado o cálculo de μ e σ cada vez mais custoso. Uma possível solução para o tamanho do histórico é definir que ele será composto apenas pelas distâncias entre as x primeiras amostras e a amostra de login. A cada nova amostra que chega z_t, ela é incluída no histórico apenas para o cálculo de G_{z_t}, sendo descartada do histórico logo em seguida. Outra desvantagem dessa abordagem é que por se basear em média e desvio padrão previamente calculados, exige treinamento em uma base de dados.

3.3.2 Teste sequencial de razão de probabilidades

A técnica de hipótese binária chamada de Sequential Probability Ratio Test (SPRT) é utilizada para determinar a probabilidade de uma amostra ter sido gerada por uma distribuição normal ou anômala. Dada duas hipóteses H_0 a hipótese nula e H_1 a hipótese alternativa:

$$H_0 : p = p_0 \qquad H_1 : p = p_1 \tag{3.16}$$

A um dado momento t é calculado a soma cumulativa do *log* da razão entre as probabilidades p_0 e p_1 toda vez que uma nova amostra chega:

$$S_i = S_{i-1} + \log(\frac{p_0}{p_1}) \tag{3.17}$$

Para i = 1,2,... e $S_0 = 0$. Aplicando uma limiarização, existem três possibilidades. A primeira é que se $a < S_i < b$, então o sistema deve continuar monitorando. Se $S_i \geq b$, então deverá aceitar H_1, e por fim, se $S_i \leq a$ deverá aceitar H_0 (WALD, 1973). Os parâmetros a e b devem ser escolhidos de maneira a maximizar o objetivo.

Para utilizar a SPRT na autenticação contínua, basta assumir que as duas hipóteses são do sistema estar seguro ou invadido:

$$H_0 : x = seguro \qquad H_1 : x = invadido \tag{3.18}$$

e a partir daí utilizar a Equação 3.19:

$$S_i = S_{i-1} + \log(\frac{P(z_i|x = seguro)}{P(z_i|x = invadido)}) \tag{3.19}$$

Contudo, é necessário estimar os valores de a e b, necessitando de treinamento. Essa técnica também não oferece maneiras de calcular $P(z_i|x = seguro)$ e $P(z_i|x = invadido)$.

3.4 OUTRAS TÉCNICAS

Nesta subseção são citados outros grupos de técnicas de detecção de anomalias e o motivo pelo qual não se enquadram no contexto de autenticação contínua.

Sistemas que exigem grande conjunto de dados para um treino como algoritmos genéticos, redes neurais, geração de regras indutivas, Máquinas de Vetores de Suporte (Support Vector Machine (SVM)), não são facilmente adaptáveis à autenticação contínua, pois não existem bases de dados suficientemente grandes para o treinamento de tais técnicas nesse contexto. Devido a natureza contínua e específica da autenticação contínua, é muito difícil construir uma base de dados grande (*e.g.* para autenticação contínua utilizando reconhecimento facial, seriam necessários milhares de vídeos com longa duração de pessoas diante de uma câmera).

Diante da análise realizada sobre as técnicas de detecção de anomalia, a técnica do PCM foi a escolhida por ser mais promissora, e foi implementada como proposta para este trabalho. No capítulo seguinte são introduzidas as características biométricas utilizadas nesse trabalho.

Capítulo 4

CARACTERÍSTICAS BIOMÉTRICAS UTILIZADAS

Diferentes características biométricas podem ser empregadas no contexto da autenticação contínua. Existem diferentes critérios para se avaliá-las, dentre os quais (JAIN et al., 2004):

1. Acurácia: quão bem as características conseguem diferenciar indivíduos
2. Aceitabilidade: como as pessoas aceitam o uso da característica
3. Permanência: o quanto as características variam ao longo do tempo
4. Coletabilidade: quão fácil é coletar as amostras biométricas
5. Universalidade: quantas pessoas possuem as características

Para validar quão robusto e seguro é o novo sistema proposto, foram escolhidas duas características biométricas distintas: a face, coletada em imagens coloridas, de profundidade e infravermelhas, e o batimento cardíaco, coletado em Eletrocardiograma (ECG). Desta forma estaremos comparando o novo método utilizando biometrias com diferentes características, o que permite avaliar o quanto o método é robusto ao tipo de biometria utilizada.

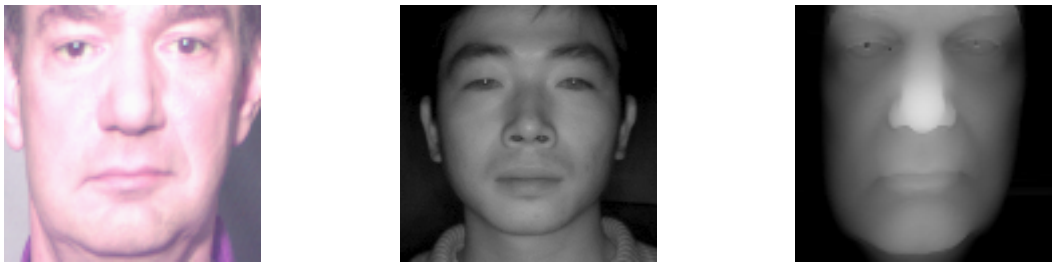
Como o foco do trabalho não está nas características biométricas, apenas uma análise superficial será feita sobre cada uma delas. Cada uma das etapas de processamento de amostras biométricas mostradas na Figura 2.1 foram implementadas e são detalhadas nas seções a seguir.

4.1 RECONHECIMENTO FACIAL

Reconhecimento facial é uma área de pesquisa que cresceu bastante nos últimos anos. Trata-se de uma maneira não incômoda e possivelmente mais natural de identificação, tanto que é a mais utilizada pelos seres humanos (KONG et al., 2005). A análise facial

também permite interpretar expressões faciais, emoções humanas, intenções e comportamentos, peças chaves para sistemas de segurança cada vez mais inteligentes.

O reconhecimento facial é uma característica biométrica de custo relativamente baixo e resultados satisfatórios, e pode ser utilizada na autenticação contínua sem a colaboração do usuário. Existem diferentes propriedades que podem ser utilizadas no reconhecimento facial, como textura (SILVA; SEGUNDO, 2015), refletividade infravermelha (SANTOS; SEGUNDO, 2015) e geometria (PAMPLONA et al., 2013). A textura é uma propriedade obtida em fotos comuns, através da captura da luz visível da cena. Refletividade infravermelha é obtida através de luz invisível aos olhos humanos, muito utilizada em câmeras de segurança, pois não dependem de iluminação ambiente. A geometria captura a profundidade dos objetos e da cena em relação ao sensor. A Figura 4.1 ilustra as três propriedades:



(a) (b) (c)

Figura 4.1: Exemplos de faces capturadas em cor (a), infravermelho (b) e geometria (c).

Uma imagem facial possui uma quantidade muito grande de informações, portanto um dos objetivos de um sistema de reconhecimento facial é selecionar e extrair as características que mais discriminam as faces, reduzindo assim o custo computacional (LI; LIAO, 2003; WEI; ZHIHUA, 2011; LI et al., 2007; ZHENG, 2012; ZHIHUA; GUODONG, 2013). Existem três categorias principais de métodos que podem ser empregados para este fim: métodos baseados em características, métodos holísticos e métodos híbridos. Os baseados em características dependem de características da face, como olhos, nariz e boca, assim como as relações geométricas entre elas. Métodos holísticos levam em consideração a face como um todo. As técnicas pioneiras desta categoria foram a Análise de Componentes Principais (Principal Component Analysis (PCA)), popularmente conhecida como *eigenfaces* (MOGHADDAM; PENTLAND, 1998), e a Análise de Discriminantes Lineares (Linear Discriminant Analysis (LDA)), popularmente conhecida como *fisherfaces* (BELHUMEUR; HESPANHA; KRIEGMAN, 1996). Ambas são baseadas em autovetores e autovalores e criam representações concisas da aparência global de imagens faciais, permitindo comparações entre as faces (FERNANDES; BALA, 2013; LU; PLATANIOTIS; VENETSANOPOULOS, 2003; H; PJ, 2001). Os métodos híbridos combinam métodos baseados em características e holísticos para obter informações sobre as características específicas da face e sobre a mesma como um todo (ZHAO; GRIGAT, 2005). Muitas das técnicas pioneiras se baseiam no conhecimento humano sobre as faces, como por exemplo, dividindo as faces em regiões que são sabidamente mais discriminantes, como ilustrado na Figura 4.2.

(a) (b)

Figura 4.2: (a) Um exemplo de imagem dividida em janelas onde são utilizados descritores de textura e (b) atribuindo pesos para cada região. (AHONEN; HADID; PIETIKÄINEN, 2004).

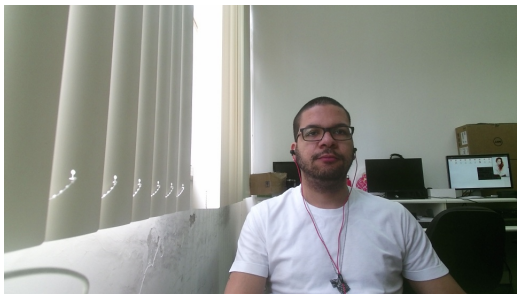
Atualmente, com a avanço nas técnicas de *deep learning*, milhares ou até mesmo milhões de faces são utilizadas no treinamento de redes neurais profundas que aprendem a discriminar faces como um todo, obtendo resultados muito superiores aos métodos anteriores (WU et al., 2015; BALTRUŠAITIS; ROBINSON; MORENCY, 2016; CAO et al., 2018; SCHROFF; KALENICHENKO; PHILBIN, 2015; SIMONYAN; ZISSERMAN, 2014).

4.1.1 Aquisição de Amostra

Os sensores Microsoft Kinect One[1] e Intel RealSense[2] são capazes de capturar imagens 2D, 3D e Near Infrared (NIR) simultaneamente. A Figura 4.3 ilustra uma aquisição de imagens usando o Kinect. Cada uma dessas imagens representa diferentes propriedades faciais que constituem diferentes características biométricas. Assim, ambos os sensores permitiriam a avaliação de diferentes biometrias em nosso sistema de autenticação contínua. O Kinect foi escolhido porque suas imagens têm melhor qualidade em comparação com o Realsense.

Uma vez obtida uma imagem, é necessário detectar a face. Para detectar as faces em 2D e NIR foi utilizada uma rede neural convolucional multi-tarefa (Multi-Task Convolutional Neural Networks (MTCNN)) (ZHANG et al., 2016b). Para detectar as faces em 3D foi utilizado o método de Segundo et al. (2014). Ambos os métodos tratam-se do estado-da-arte da detecção facial em suas respectivas modalidades (2D e 3D). Apesar da MTCNN não ter sido treinada para detectar faces em NIR ela possui bons resultados com as mesmas.

[1] https://www.xbox.com/en-US/xbox-one/accessories/kinect
[2] https://software.intel.com/pt-br/realsense



(a) 2D

(b) NIR

(c) 3D

Figura 4.3: Exemplos de imagens capturadas simultaneamente utilizando o Kinect One.

4.1.2 Pré-processamento e Descrição da Amostra

Para descrever as faces nas três modalidades foi utilizada a Convolutional Neural Network (CNN) proposta por Wu et al. (2015), com o modelo treinado disponibilizado pelos autores, pois a mesma possui bons resultados para as três modalidades (DAHIA; SANTOS; SEGUNDO, 2017). Originalmente treinada para imagens 2D, a CNN de Wu *et al.* obteve um bom desempenho em imagens NIR e 3D na literatura (DAHIA; SANTOS; SEGUNDO, 2017), embora obtenha maior precisão para a modalidade de seu treinamento. Para que uma face possa ser descrita, ela precisa seguir as seguintes especificações (WU et al., 2015): estar em preto e branco; possuir tamanho de 128×128 pixels; olhos alinhados e a 40 pixels da borda superior da imagem; distância entre o centro dos olhos e o centro da boca de 48 pixels. Tal pré-processamento é chamado de normalização facial e visa reduzir variações de pose. A MTCNN detecta pontos faciais (*e.g.* pontos ao redor dos olhos, nariz e boca) que são utilizados para realizar translações, rotações e redimensionamentos necessários para que a face fique no formato aceito pela CNN descritora. Para a normalização em geometria a face 3D é alinhada a uma face modelo, como proposto por Pamplona et al. (2013), e redimensionada para 128×128 pixels. Exemplos de faces normalizadas são exibidos na Figura 4.4:

Após a face normalizada ser utilizada como entrada para a CNN descritora, é obtido um vetor de 256 valores. Para avaliar o desempenho de um sistema de reconhecimento biométrico, as três métricas mais importantes são: a taxa de aceitação falsa (False Acceptance Rate (FAR)), que representa a taxa em que um invasor foi reconhecido como o usuário genuíno; a taxa de falsa rejeição (False Rejection Rate (FRR)), que representa a

(a) 2D (b) NIR (c) 3D

Figura 4.4: Exemplos de faces normalizadas para cada modalidade.

taxa em que um genuíno foi considerado um invasor; e a Taxa de Erro Igual (Equal Error Rate (EER)) é quando se tem taxas de FRR e FAR iguais. Quanto menor o EER de um sistema ou descritor, melhor é seu desempenho.

Os EERs do descritor de Wu *et al.* para imagens 2D, NIR e 3D em conjuntos de dados controlados foram aproximadamente 0,5%, 2% e 7%, respectivamente, reportados no trabalho de Dahia, Santos e Segundo (2017). A Figura 4.5 resume o processo utilizado para faces.

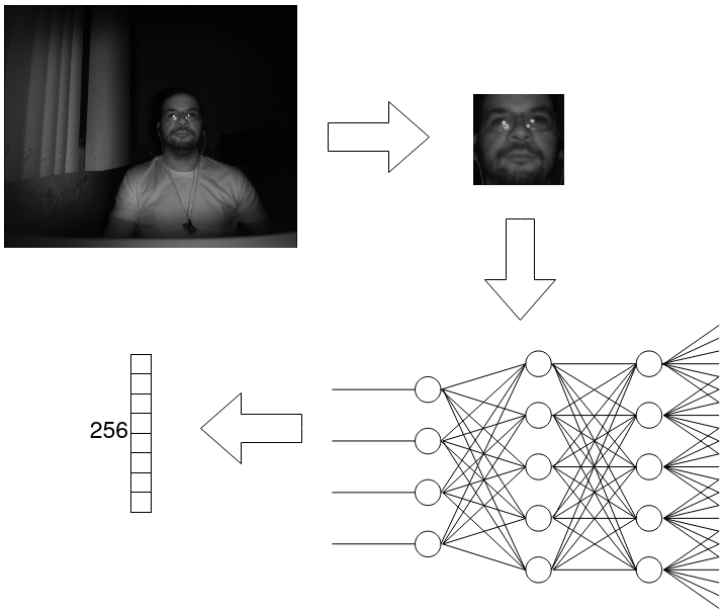


Figura 4.5: Etapas para biometria facial.

4.2 RECONHECIMENTO PELO BATIMENTO CARDÍACO

Nos últimos anos o reconhecimento pelo batimento cardíaco tem sido alvo de pesquisas por diversos motivos, dentre eles, por ser uma das características mais difíceis de se forjar, ser universal, ser discriminativa e por requerer vivacidade, ou seja, apenas enquanto o indivíduo estiver vivo é que se pode coletar novas amostras biométricas (PINTO et al., 2017).

Para capturar os sinais elétricos produzidos pelo coração, é necessário colocar eletrodos no corpo da pessoa. Existem diferentes formas de posicionar tais eletrodos e cada configuração define o que é conhecido como um canal de ECG. O exame mais utilizado captura 12 canais (ou leads), analisando diferentes perspectivas sobre a atividade elétrica do coração. A Figura 4.6 ilustra o posicionamento de parte dos eletrodos. Já na Figura 4.7 podemos ver como diferentes canais geram diferentes formatos de ondas.

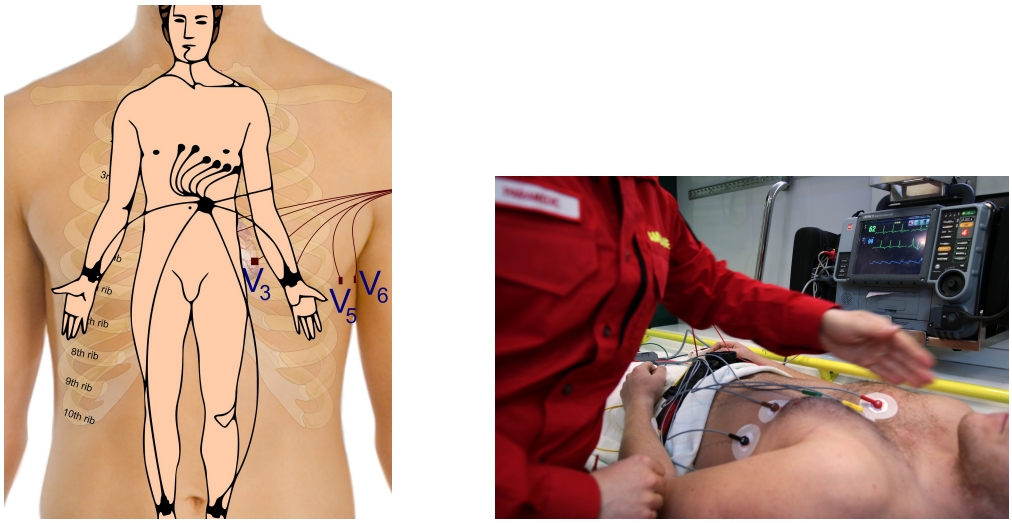


Figura 4.6: Exemplos de posicionamento de eletrodos para ECG.

Um batimento cardíaco consiste basicamente em três partes, como mostrado na Figura 4.8: uma onda P (contração dos átrios) seguida por um complexo QRS (contração dos ventrículos), seguido por uma onda T (relaxamento dos ventrículos). Essas partes variam em termos de localização, amplitude e intervalo das ondas, o que torna o batimento cardíaco diferente de pessoa pra pessoa.

Métodos de reconhecimento de pessoas utilizando ECG podem ser divididos em fiduciais, não-fiduciais e híbridos. Métodos fiduciais detectam e extraem características locais (*e.g.* amplitudes do complexo PQRST) e utilizam essas características como entrada de métodos tradicionais de classificação, como por exemplo, Support Vector Machine (SVM). A Figura 4.9 ilustra tais características.

Já os métodos não-fiduciais utilizam um ou mais batimentos cardíacos inteiros, como por exemplo, calculando os coeficientes de transformadas de ondas (CHAN et al., 2008) ou auto-correlação (PLATANIOTIS; HATZINAKOS; LEE, 2006). A principal vantagem dos métodos não-fiduciais é que eles são independentes de detectores de características, ao contrário dos métodos fiduciais, cujo desempenho é afetado pela acurácia desses detectores. Por outro lado, tais métodos podem ser mais custosos. Métodos híbridos combinam ambas abordagens visando compensar as desvantagens de cada uma.

Após uma busca na literatura sobre métodos de reconhecimento de pessoas utilizando ECG, foram identificados alguns problemas recorrentes em diversos trabalhos: o uso dos mesmos indivíduos no treinamento dos descritores e nos testes; o uso de bases de dados privadas ou atualmente indisponíveis; e ausência de informações importantes como quais

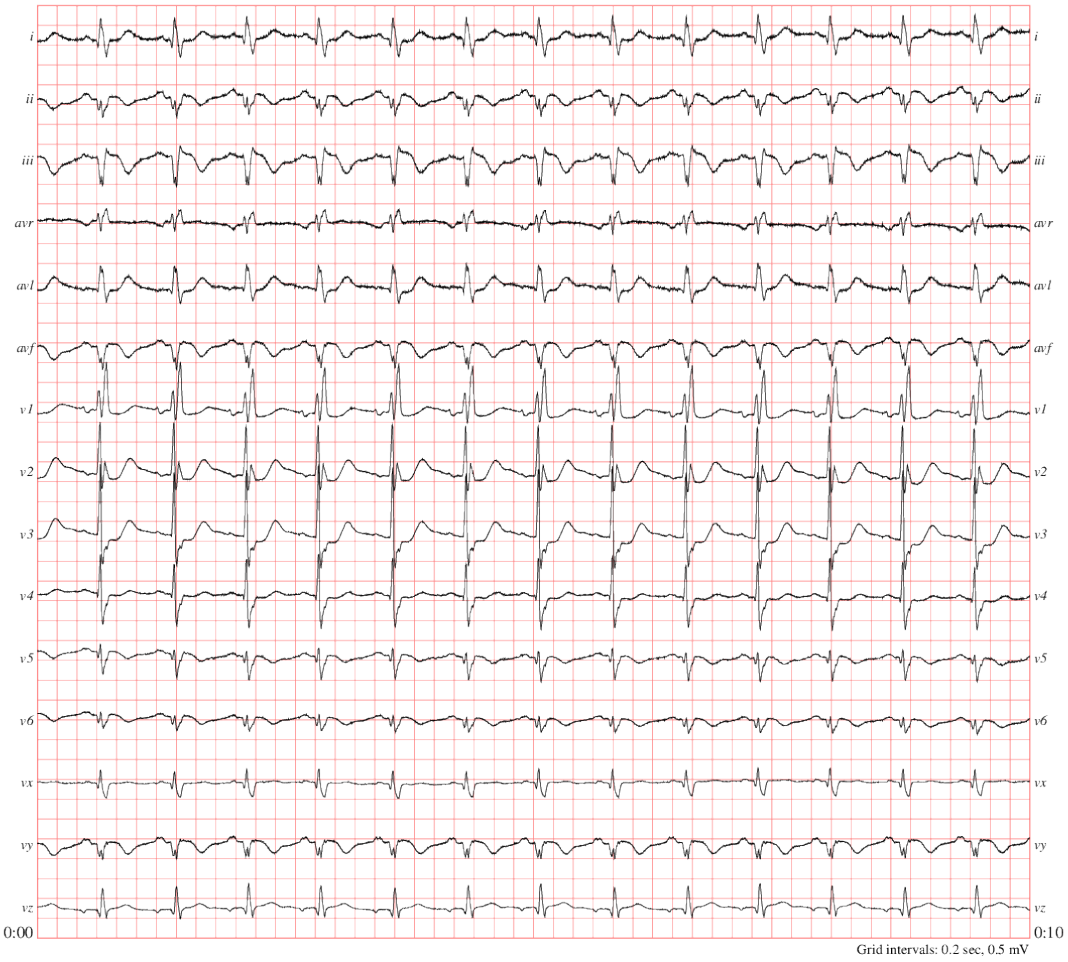


Figura 4.7: Diferentes canais de um ECG para uma mesma pessoa.

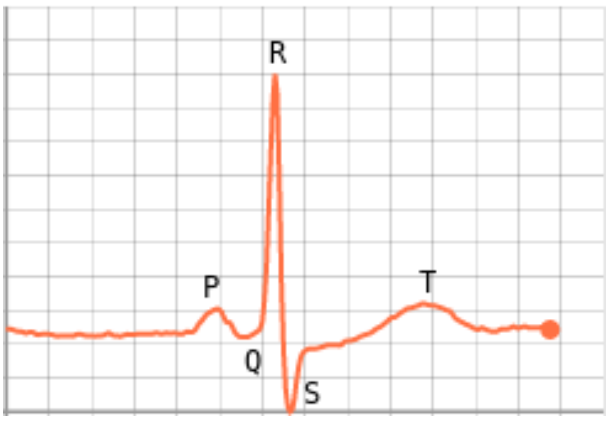


Figura 4.8: Ilustração das principais partes de um batimento cardíaco: uma onda P corresponde a contração atrial, seguida por um complexo QRS representando a contração dos ventrículos, seguido por uma onda T do relaxamento dos ventrículos.

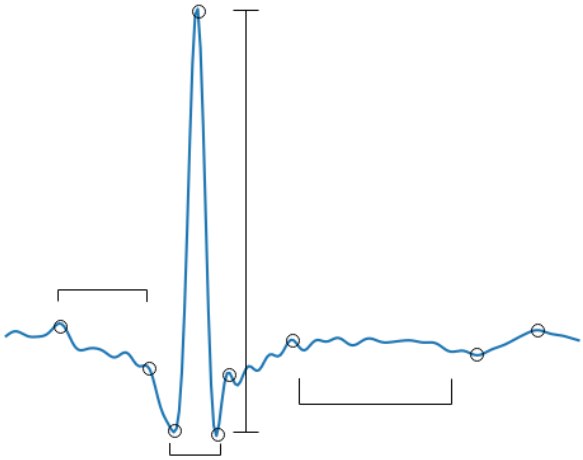


Figura 4.9: Exemplos de características fiduciais.

indivíduos de uma base foram utilizados nos experimentos. Isso dificulta a comparação dos trabalhos e a análise da capacidade de generalização dos descritores encontrados na literatura. Por isso, neste trabalho também foi feito um estudo sobre os métodos, desafios, problemas, e foi proposto um novo descritor para ECG. Como esse não é o foco deste trabalho, todas as informações sobre tal estudo e sobre o novo descritor encontram-se no Apêndice A deste documento.

4.2.1 Aquisição de Amostra

Para a aquisição de amostras de ECG, foi utilizado o PhysioNet kit (GOLDBERGER et al., 2000). A plataforma contém diversas bases de dados biométricos, dentre eles o ECG[3]. Para cada base há gravações de ECGs contendo os valores do sinais elétricos medidos, assim como informações sobre a mesma. A Figura 4.10 mostra um exemplo de um trecho de ECG extraído de uma das bases. Devido a variedade de canais, caso mais de uma base de dados seja utilizada, elas devem compartilhar um mesmo canal.

4.2.2 Pré-processamento e Descrição da Amostra

Como dito anteriormente, através do Physionet é possível extrair o ECG de gravações. Para segmentar os batimentos cardíacos em uma gravação, foi utilizado a ferramenta Biosppy (CARREIRAS et al., 2015–). Ela permite não só segmentar uma gravação em batimentos cardíacos como também filtrar o sinal, reduzindo ruídos causados por interferências durante a gravação do ECG. Após filtrado e segmentado, cada batimento cardíaco é linearmente interpolado para um comprimento de exatamente 256 valores e sua amplitude é normalizada para o intervalo de 0 a 1 utilizando os valores mínimo e máximo. O resultado dessas etapas são exibidos na Figura 4.11. De maneira similar ao processo adotado por faces, o batimento normalizado é utilizado como entrada para uma CNN descritora que gera um vetor de 84 posições. Os detalhes sobre o treinamento da

[3]https://www.physionet.org/about/database/#ecg

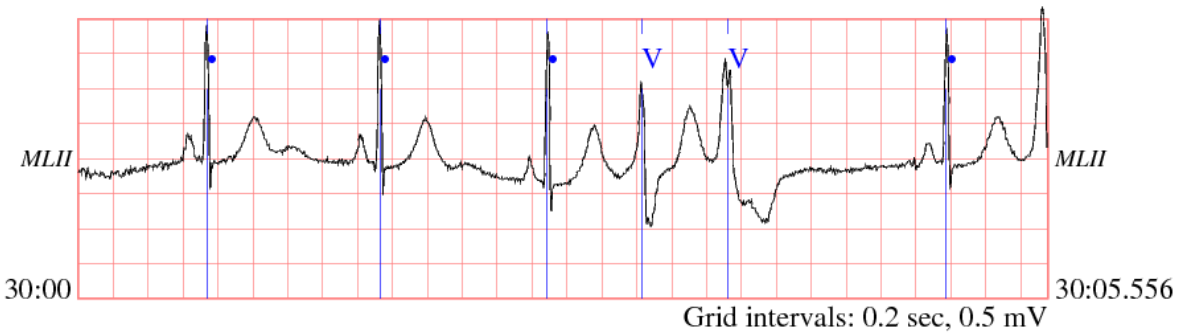
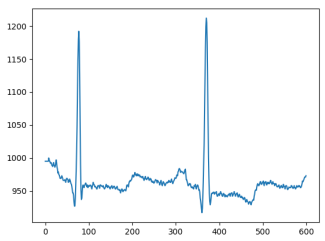
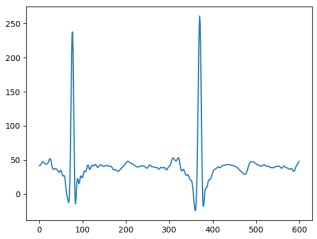


Figura 4.10: Exemplo de um trecho de um ECG extraído de uma gravação.

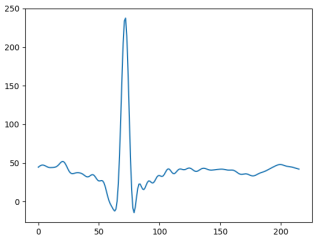
CNN descritora são encontrados no Apêndice A.



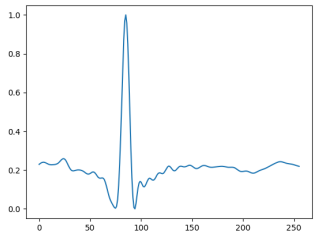
(a) sinal original



(b) sinal filtrado



(c) batimento segmentado



(d) batimento normalizado

Figura 4.11: Exemplo de cada etapa de pré-processamento para ECG.

Capítulo

5

O MÉTODO PROPOSTO

Para este trabalho foi escolhida a técnica "Possibilistic C-Means (PCM)", pois a mesma é a que mais se adequa a proposta deste trabalho, visto que o problema de autenticação contínua pode ser modelado como um problema de dois agrupamentos, onde um agrupamento representa o grupo genuíno e o outro o invasor, e não necessita de treinamento. O foco deste trabalho está no algoritmo de autenticação contínua em si, de maneira que ele possa ser visto como uma *black box*, funcionando com diferentes biometrias.

No Capítulo 3.1.2 foram definidas diferentes maneiras de se calcular o valor de $\eta_{genuíno}$. Todas foram implementadas para avaliar qual é a melhor. Para fins de nomenclatura, para cada forma de calcular o valor de $\eta_{genuíno}$, o método proposto foi divido da seguinte forma:

- **PCM**: utiliza a Equação 3.8;

$$\eta_{genuíno} = DIST\_MAX - dist(c_{genuíno}, c_{invasor})$$

- **PCM D2**: utiliza a Equação 3.9, o que reduz a distância máxima pela metade;

$$\eta_{genuíno} = DIST\_MAX * 0.5 - dist(c_{genuíno}, c_{invasor})$$

- **PCM D3**: utiliza a Equação 3.10, o que reduz a distância máxima em um terço;

$$\eta_{genuíno} = DIST\_MAX * 0.333 - dist(c_{genuíno}, c_{invasor})$$

- **PCM ZT**: utiliza a Equação 3.11, o que reduz a distância máxima pela distância entre última amostra e a de login;

$$\eta_{genuíno} = DIST\_MAX - dist(c_{genuíno}, c_{invasor}) - dist(c_{genuíno}, z_t)$$

Na etapa de login, um vetor de características é armazenado como o modelo do usuário, que servirá como protótipo do grupo genuíno durante todo o acesso. Após o login, as últimas 10 amostras (faces ou Eletrocardiograma (ECG)s) descritas são mantidas como o histórico das observações. O algoritmo PCM é executado neste histórico e os graus de pertinências dos grupos genuíno e invasor da última observação são calculados como apresentado na subseção 3.1.2. O Algoritmo 4 ilustra o funcionamento para o PCM padrão.

Algoritmo 4 Algoritmo de fusão de similaridades utilizando PCM

armazena a amostra de login;
$X =$ últimas n amostras em $\mathcal{Z}_t$;
$P(x = seguro|\mathcal{Z}_0) = 1$;
$P(x = invadido|\mathcal{Z}_0) = 0$;
repita
  se z_t = *nova amostra válida* **então**
    $X = \{X - z_v\} \cup z_t$, sendo z_v a observação mais antiga em X;
    $c_{genuíno} =$ amostra de login;
    $c_{invasor} = max_{x_i \in X}(dist(x_i, c_{genuino}))$;
    $\eta_{genuíno} = DIST\_MAX - dist(c_{genuíno}, c_{invasor})$;
    $\eta_{invasor} = dist(c_{genuíno}, c_{invasor})$;
    Atualize cada pertinência $w_{i,j}$ de $x_i \in X$ utilizando: $w_{ij} = \frac{1}{1+(\frac{dist(x_i,c_j)}{\eta_j})^{\frac{1}{m-1}}}$;
    repita
      $c_{invasor} = \frac{\sum_{i=1}^{n}(w_{iinvasor})^m x_i}{\sum_{i=1}^{n}(w_{iinvasor})^m}$;
      $\eta_{invasor} = dist(c_{genuíno}, c_{invasor})$;
      $\eta_{genuíno} = DIST\_MAX - dist(c_{genuíno}, c_{invasor})$;
      Atualize cada pertinência $w_{i,j}$ de $x_i \in X$ utilizando:
      $w_{ij} = \frac{1}{1+(\frac{dist(x_i,c_j)}{\eta_j})^{\frac{1}{m-1}}}$;
    até $c_{invasor}$ *não mude*;
    $P(z_t \mid x = seguro) = w_{t,genuíno}$;
    $P(z_t \mid x = invadido) = w_{t,invasor}$;
  fim
  $P(x = seguro|\mathcal{Z}_t) \propto P(z_t|x = seguro) + 2^{\frac{u-t}{K}} \times P(x = seguro|\mathcal{Z}_u)$;
  $P(x = invadido|\mathcal{Z}_t) \propto P(z_t|x = invadido) + 2^{\frac{u-t}{K}} \times P(x = invadido|\mathcal{Z}_u)$;
  $P_{seguro} = \frac{2^{-\frac{\Delta t}{K}} \times P(x=seguro|\mathcal{Z}_t)}{P(x=seguro|\mathcal{Z}_t)+P(x=invadido|\mathcal{Z}_t)}$;
até $P_{seguro} < limiar$;

A distância utilizada para comparar os descritores foi a distância cosseno, a mesma utilizada para comparar os descritores de Wu et al. (2015), descrita na Equação 5.1. Sendo assim, $DIST\_MAX = 2$ na Equação 3.8.

$$distância\_cosseno(x, y) = 1 - similaridade\_cosseno(x, y) \tag{5.1}$$

$$similaridade\_cosseno(x, y) = \frac{x \cdot y}{||x||||y||} \tag{5.2}$$

Para os parâmetros do PCM utilizados, o valor de fuzzificação escolhido foi 1.5, o tamanho do histórico de observações como 10 e o número máximo de iterações para convergência do PCM foi 100. O valor de fuzzificação de 1.5 é recomendado pelos autores do PCM, Krishnapuram e Keller (KRISHNAPURAM; KELLER, 1996) como valor mais adequado para a função de pertinência. Como o histórico de amostras é pequeno, o algoritmo converge em média em até 5 iterações, portanto um máximo de 100 iterações é mais do que o suficiente para esse contexto. Na Figura 5.1 podemos ver o resumo do método proposto.

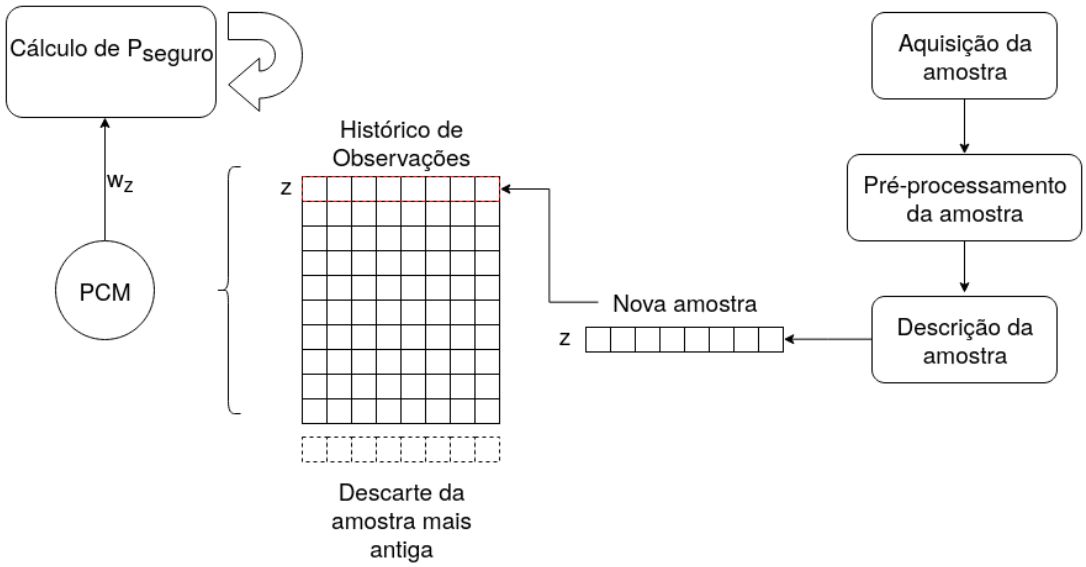


Figura 5.1: Ilustração do método de autenticação contínua utilizando PCM.

Após construído o sistema de autenticação contínua utilizando PCM, foram feitos experimentos, comparando-o com o método estado-da-arte CDF, proposto por Pamplona et al. (2013). Os experimentos e resultados são descritos no capítulo a seguir.

Capítulo 6

EXPERIMENTOS

Os experimentos foram realizados para a nova proposta utilizando o PCM e para o método estado-da-arte proposto por Pamplona et al. (2013) descrito na Seção 2.2.

6.1 FACE

6.1.1 Bases de Dados

Para os experimentos com faces foi criada uma base de dados. A base de faces é composta por 7 gravações (ou vídeos) de 7 sujeitos diferentes capturados em 2D, Near Infrared (NIR) e 3D utilizando o sensor Microsoft Kinect One, onde cada gravação tem duração média de 40 minutos e foram capturados a uma taxa de 15 quadros por segundo. Desta forma temos um cenário com longa duração, sem restrições sobre como os usuários deveriam usar o sistema.

Essas gravações foram utilizados como entrada para o sistema de autenticação contínua. Para cada gravação, foi concatenada ao seu final o início de cada outra gravação para simular uma situação em que o usuário genuíno deixa o sistema e um invasor assume, como ilustrado na Figura 6.1. No total foram 42 simulações de ataque.

6.1.2 Constantes do método Cumulative Distribution Function (CDF)

Como o método de Pamplona et al. (2013) utiliza funções de distribuição acumulada (Cumulative Distribution Function (CDF)) e necessita dos parâmetros μ e σ (média e desvio padrão) para os estados seguro e invadido, os mesmos tiveram que ser calculados.

Para faces, foram utilizadas 4 bases de dados diferentes. A primeira é composta por 97 estudantes e docentes da Universidade Federal da Bahia, com um total de 5565 imagens capturadas em 2D, 3D e NIR. A segunda é a CASIA NIR-VIS 2.0 (CASIA) (LI et al., 2013), que contém 17.580 imagens de 725 sujeitos em 2D e NIR. A terceira é a Face Recognition Grand Challenge (FRGC), composta de 4.950 imagens registradas de 3D e 2D de 556 sujeitos. A quarta base utilizada é a Labeled Faces in the Wild (LFW), com mais de 13.000 imagens de 1.680 sujeitos capturados em 2D. A primeira base é muito

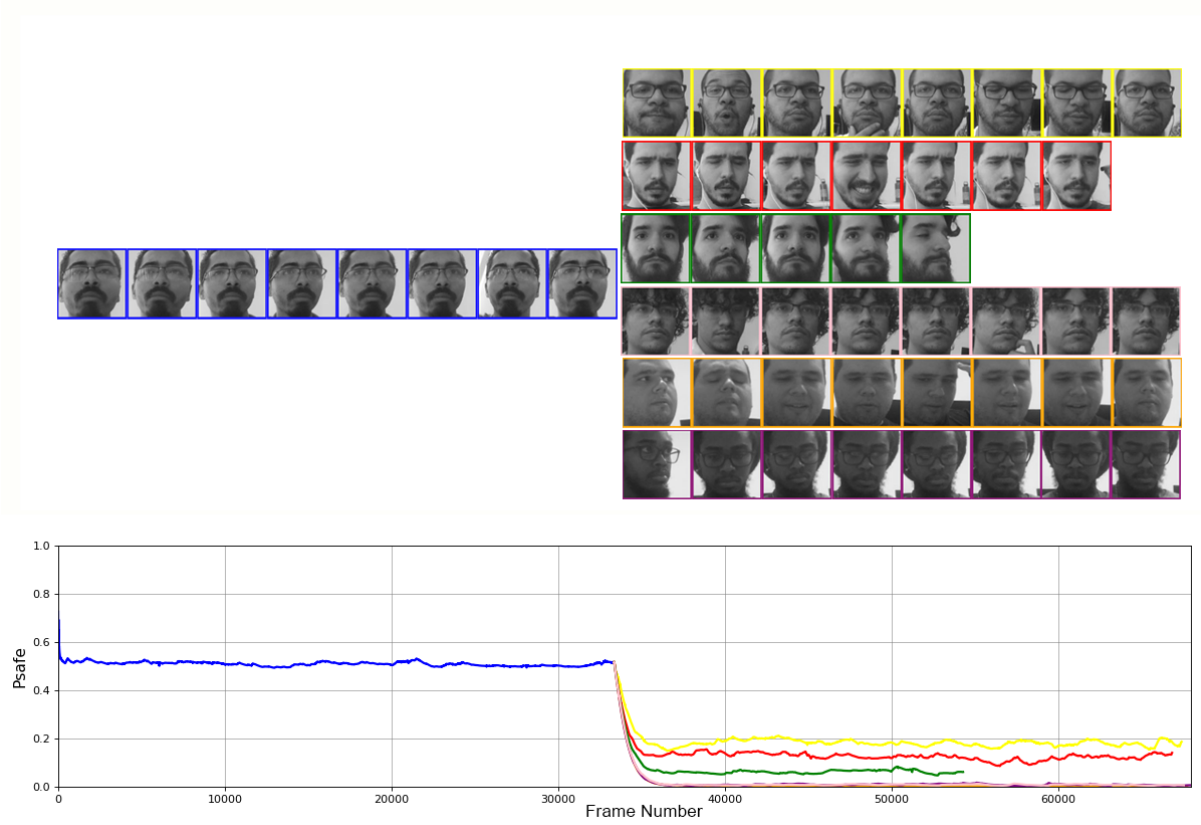


Figura 6.1: Ilustração dos valores de P_{seguro} ao longo do tempo. A linha azul são os valores de P_{seguro} quando o usuário genuíno estava utilizando o sistema. As outras linhas representam quando um invasor utilizava o sistema. As faces com contorno azul representam exemplos de quadros onde o usuário genuíno utilizava o sistema. As outras representam os invasores.

controlada, com restrição de pose e expressão facial. As outras bases foram escolhidas, pois são as mais utilizadas na literatura para avaliação de reconhecimento facial utilizando 2D, NIR e 3D. A CASIA possui menos variação de pose e de expressão facial do que a FRGC. A LFW é uma base sem restrições, sendo utilizada como *benchmark* para reconhecimento facial 2D.

As bases possuem diferentes níveis de variações intraclasse (*e.g.* variações de pose e expressão facial de uma mesma pessoa) e extraclasse (*e.g.* variações entre faces de pessoas diferentes), e não necessariamente representam os desafios de um cenário real de autenticação contínua. Essas variações resultam em diferentes parâmetros de CDF.

Os parâmetros foram calculados utilizando a distância cosseno e são exibidos na Tabela 6.1.

Base	Tipo	$\mu_{genuíno}$	$\sigma_{genuíno}$	$\mu_{invasor}$	$\sigma_{invasor}$
UFBA	2D	0.2741	0.1268	0.9158	0.1137
	NIR	0.2789	0.1384	0.7837	0.1360
	3D	0.3976	0.1521	0.4932	0.1359
CASIA	2D	0.0914	0.0985	0.9089	0.1442
	NIR	0.0950	0.1222	0.8325	0.1481
FRGC	2D	0.2190	0.1080	0.9550	0.1062
	3D	0.1641	0.1008	0.5406	0.1356
LFW	2D	0.3731	0.1274	0.9906	0.0929

Tabela 6.1: Parâmetros CDF obtidos de diferentes bases de dados para serem utilizados no método de Pamplona et al. (2013) para modalidade de biometria facial.

6.1.3 Resultados

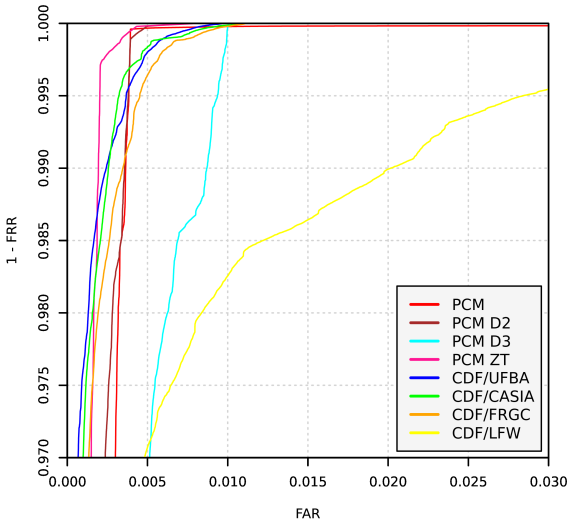
Para cada uma das biometrias, foram avaliados o método de Pamplona et al. (2013) baseado em CDFs utilizando todos os possíveis parâmetros da Tabela 6.1 e o método Possibilistic C-Means (PCM) com suas variações, para cada sessão de acesso. Os valores de P_{seguro} de todas as sessões em cada um desses testes foram compilados em uma curva de característica de operação do receptor (Receiver Operating Characteristic (ROC)) e os resultados são exibidos na Figura 6.2.

As curvas mostram quão boa é a separação entre os valores de P_{seguro} do usuário genuíno e dos invasores. Como pode ser observado, o método proposto é comparável aos melhores resultados de faces 2D, melhor para 3D e pior para NIR. A Tabela 6.2 mostra os valores Equal Error Rate (EER) de cada uma dessas curvas para ilustrar melhor a observação anterior.

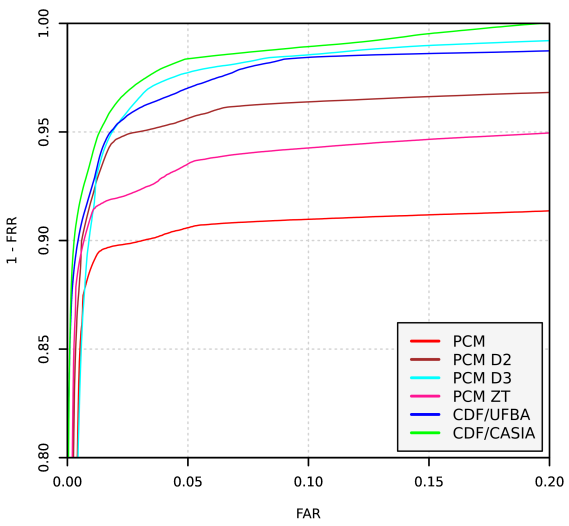
Método	Treinamento	2D	NIR	3D
PCM	-	0.0038	0.0906	**0.0304**
PCM D2	-	0.0038	0.0455	0.0476
PCM D3	-	0.0088	0.0318	0.0518
PCM ZT	-	**0.0023**	0.0405	0.0616
CDF	UFBA	0.0040	0.0360	0.0956
CDF	CASIA	0.0035	**0.0282**	-
CDF	FRGC	0.0046	-	0.0634
CDF	LFW	0.0141	-	-

Tabela 6.2: EERs para autenticação contínua nas 4 modalidades para o método PCM e o método de Pamplona et al. (2013) baseado em CDF. Em negrito os melhores resultados.

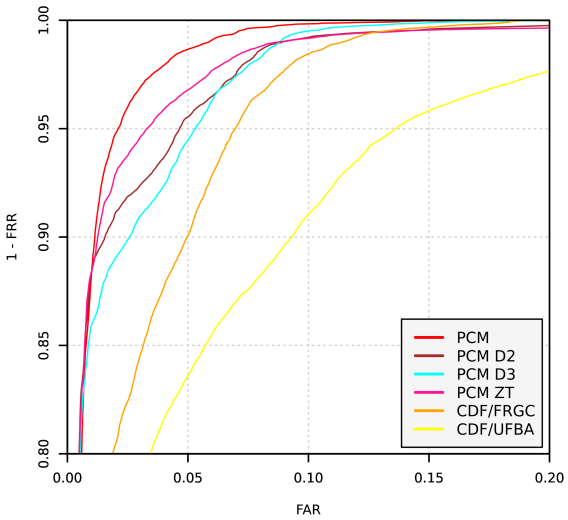
Para uma avaliação qualitativa, as Figuras 6.3 a 6.5 mostram os valores de P_{seguro} ao longo do tempo para cada um dos experimentos, comparando o método PCM (D3) com o melhor resultado do método baseado em CDF. O PCM D3 foi escolhido para essa análise, pois na média dos EERs ele é o melhor. Devido ao grande número de



(a) 2D



(b) NIR

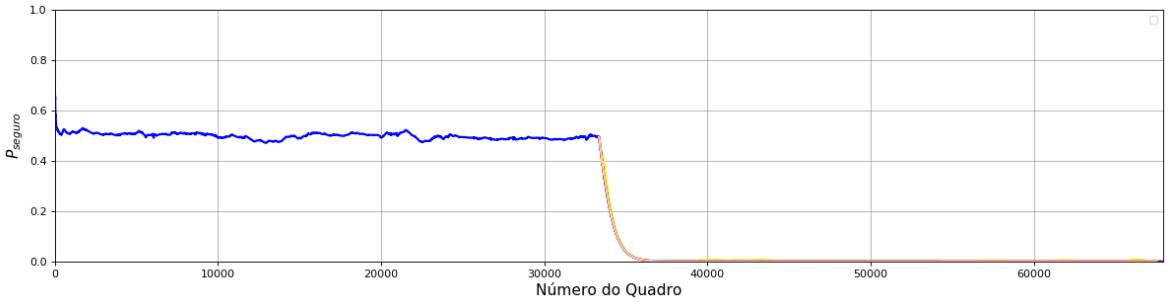


(c) 3D

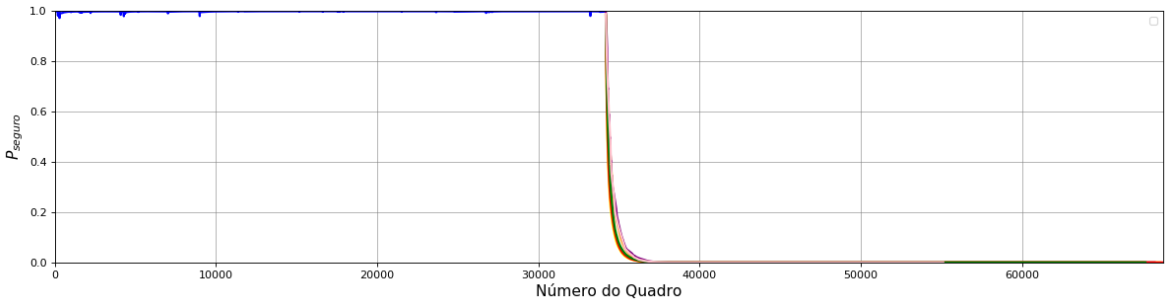
Figura 6.2: Curvas ROC de autenticação contínua usando PCM e CDF para cada modalidade. Quanto mais próxima do canto superior esquerdo, melhor o método.

experimentos, apenas o melhor e o pior caso de cada algoritmo são exibido nas Figuras 6.3 a 6.5. O melhor caso é aquele com menores variações intraclasse/interclasse e cuja diferença entre o valor de P_{seguro} genuíno e P_{seguro} de invasores seja a maior possível. O pior caso é justamente o contrário.

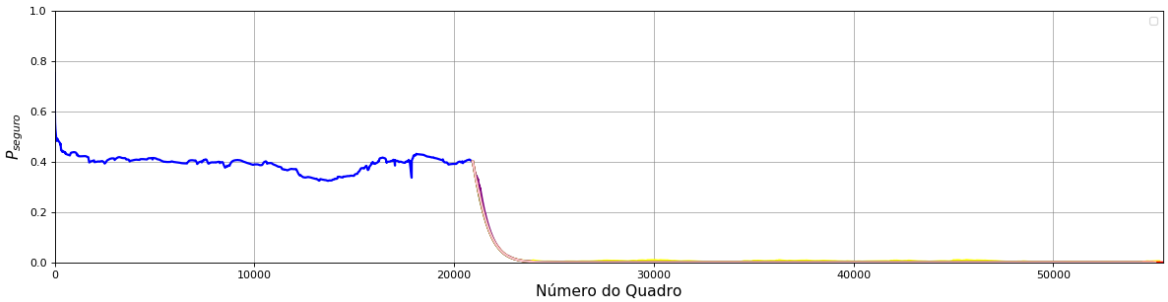
Podemos observar na Figura 6.3 que ambos os métodos possuem comportamento similar. Durante a utilização do sistema pelo usuário genuíno, o valor de P_{seguro} se mantém constante (*e.g.* $P_{seguro} \simeq 1$ para o método CDF e $P_{seguro} \simeq 0.5$ para o método



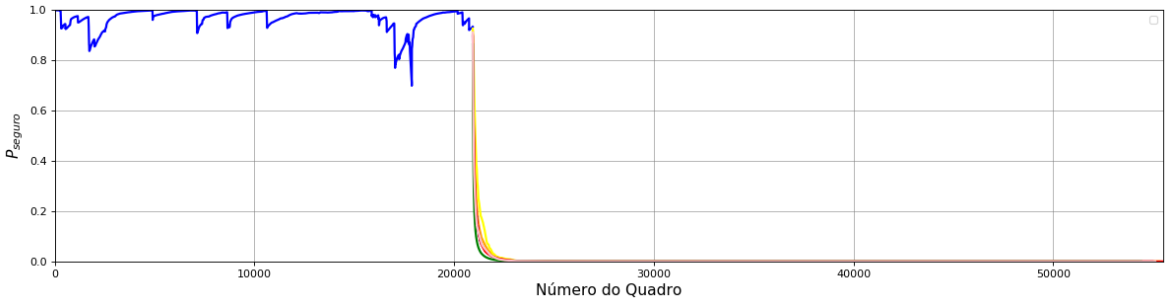
(a) Melhor caso para PCM D3



(b) Melhor caso para CDF

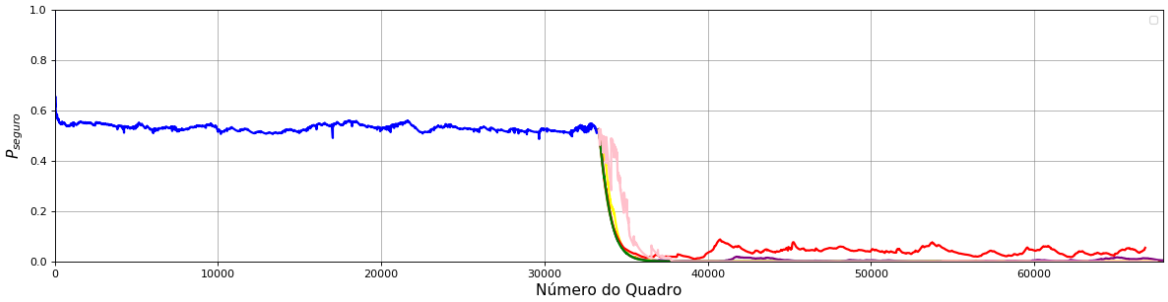


(c) Pior caso para PCM D3

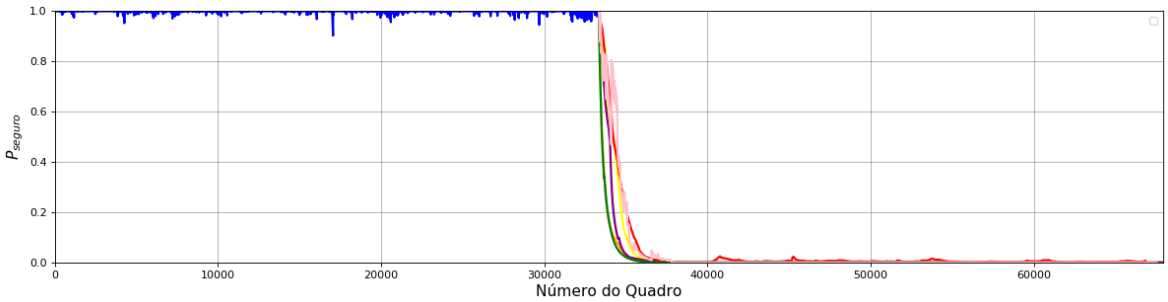


(d) Pior caso para CDF

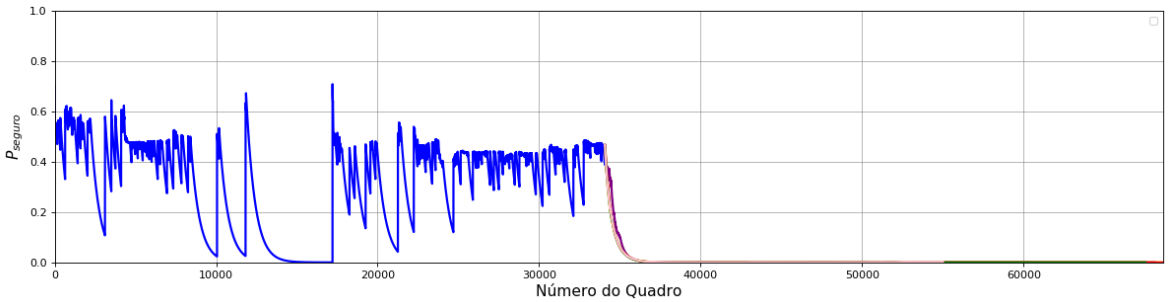
Figura 6.3: Valores de P_{seguro} ao longo do tempo para modalidade de faces em 2D. O pior caso de ambos pertence ao mesmo sujeito. Os valores de CDF foram obtidos usando os parâmetros da CASIA/2D (vide Tabela 6.1).



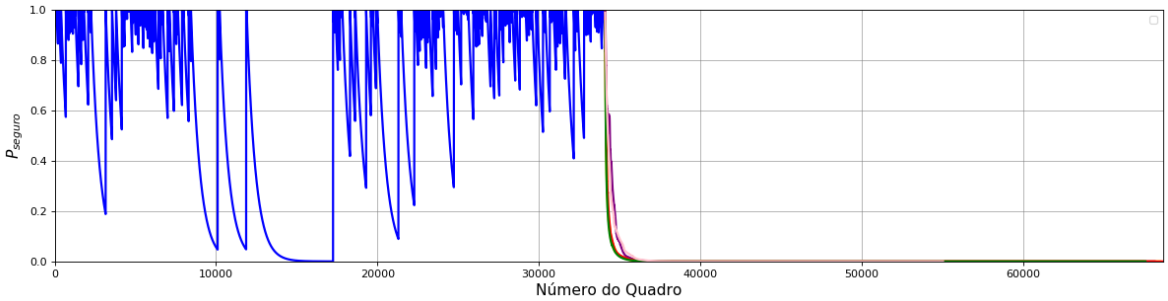
(a) Melhor caso para PCM D3



(b) Melhor caso para CDF

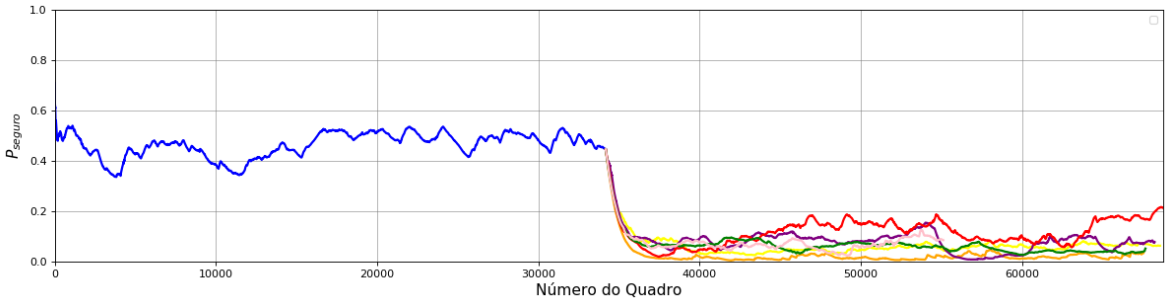


(c) Pior caso para PCM D3

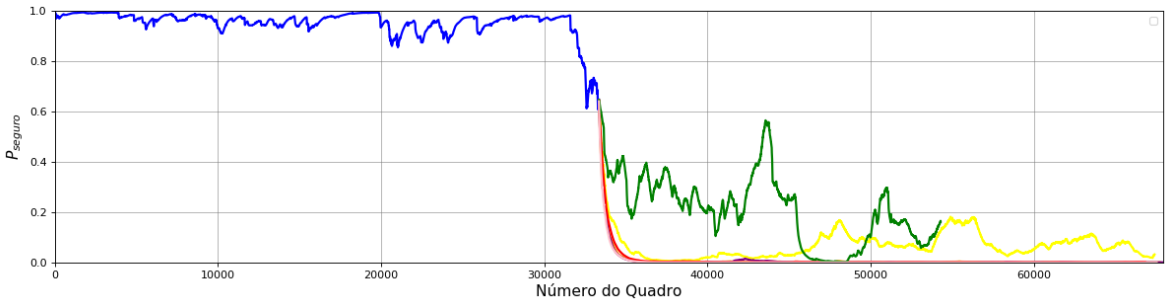


(d) Pior caso para CDF

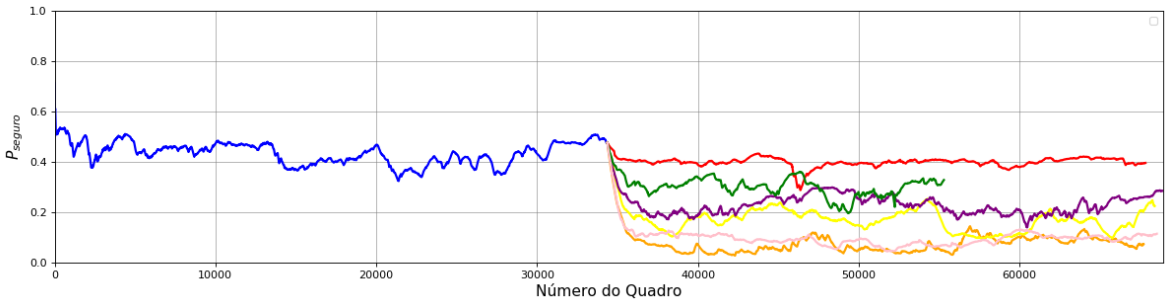
Figura 6.4: Valores de P_{seguro} ao longo do tempo para modalidade de faces em NIR. Neste exemplo, os sujeitos são os mesmos para ambos os casos. Os valores de CDF foram obtidos usando os parâmetros da CASIA/NIR (vide Tabela 6.1).



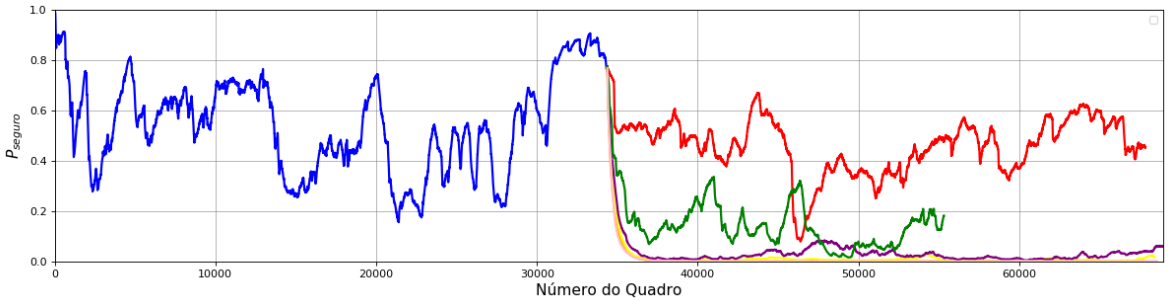
(a) Melhor caso para PCM D3



(b) Melhor caso para CDF



(c) Pior caso para PCM D3



(d) Pior caso para CDF

Figura 6.5: Valores de P_{seguro} ao longo do tempo para modalidade de faces em 3D. O pior caso de ambos pertence ao mesmo sujeito. Os valores de CDF foram obtidos usando os parâmetros da FRGC/3D (vide Tabela 6.1).

PCM) e quando um invasor assume o sistema, o valor de P_{seguro} cai rapidamente para próximo de 0. Entretanto, é possível perceber que os valores de P_{seguro} para o usuário genuíno no método PCM oscilam menos. Isso demonstra uma maior robustez do método a variações intraclasse.

Na Figura 6.4 podemos observar que a estabilidade durante o usuário genuíno do método PCM é mantida (exceto no pior caso, que é ruim para ambos), porém quando o invasor assume, a queda do valor de P_{seguro} para o invasor não é tão rápida, como pode ser observado na curva rosa da Figura 6.4a. Isso é esperado, pois a rede neural de Wu et al. (2015) foi treinada para reconhecer faces 2D. Sendo assim, quando utilizada para descrever faces em infravermelho, o descritor possui menos separabilidade entre a classe genuína e invasora. Por outro lado, o método CDF possui informação prévia da CASIA sobre a separabilidade de faces em NIR. Além disso, para a detecção facial foi utilizado a Multi-Task Convolutional Neural Networks (MTCNN) (ZHANG et al., 2016b), treinada para faces 2D. Apesar de ter um bom desempenho detectando faces em NIR, ela falha para as faces mais difíceis desta modalidade (*e.g.* com maior variação intraclasse, como expressões faciais ou variações de pose). Isso faz com que os experimentos em NIR contenham faces com menos variações, tornando-os mais controlados, o que se aproxima do cenário da CASIA, melhorando os resultados do método CDF. Os piores casos da Figura 6.4 são atípicos e trata-se justamente de uma gravação onde o detector facial falha muitas vezes por quadros consecutivos. Isso faz com que o valor de P_{seguro} caia naturalmente (PAMPLONA et al., 2013) e provoque os picos observados nas Figuras 6.4c e 6.4d para ambos os métodos.

Na Figura 6.5 são mostrados os resultados para a modalidade de 3D. A estabilidade intraclasse do método PCM para o melhor e o pior caso é afetada, mas se mantém consistente, ao contrário do método CDF. Porém, a queda no valor de P_{seguro} para os invasores é bem menor do que nas outras modalidades. Os resultados para ambos os métodos são esperados, pois como dito anteriormente, o descritor (WU et al., 2015) foi treinado para faces 2D e faces em 3D são muito diferentes de faces 2D, logo sua capacidade discriminativa cai bastante e ambos os métodos não conseguem distinguir tão bem os genuínos dos invasores.

6.2 Eletrocardiograma (ECG)

6.2.1 Bases de Dados

Para os experimentos com ECG foi utilizada a base MIT-BIH Arrhythmia Database (MITDB) (MOODY; MARK, 2001). A base MITDB contém 48 gravações de ECGs de 47 sujeitos, com duração de meia hora cada. As gravações foram feitas utilizando dois canais (Modified Limb Lead I (MLII) e Modified Limb V1 (MLV1)). Os registros foram obtidos entre os anos de 1975 e 1979. Essa base de dados foi especialmente designada para o desenvolvimento de detectores de arritmias e contém gravações de pessoas saudáveis e de pessoas com problemas cardíacos. O objetivo desse trabalho é analisar o comportamento dos métodos de autenticação contínua de maneira não excludente, e por isso, as gravações de pessoas que sofrem com arritmias cardíacas não foram descartadas. Os

únicos registros descartados foram os de número 202, 102 e 104, pois o primeiro pertence a um sujeito com duas gravações e os últimos dois não apresentam o canal MLII utilizado pela Convolutional Neural Network (CNN) descritora. No total foram utilizadas 45 gravações de 45 sujeitos, uma gravação por sujeito. De maneira análoga a faces, para cada sessão de acesso começando por uma gravação, foi concatenado ao seu final o início de cada outro gravação para simular uma situação em que o usuário genuíno deixa o sistema e um invasor assume.

6.2.2 Constantes do método CDF

Para ECG, como a CNN descritora utiliza o canal MLII, nós escolhemos o conjunto de validação da base Long-Term ST Database (LTST) (JAGER et al., 2003) no treinamento da CNN para calcular os parâmetros CDF. A base LTST contém 86 gravações de 80 sujeitos que possuem algumas variações no segmento ST dos batimentos cardíacos. Cada gravação contém entre 21 e 24 horas de duração. Apenas 16 gravações possuem o canal MLII, e durante o treinamento da rede neural que descreve os batimentos cardíacos, apenas 4 gravações foram utilizadas para o conjunto de validação e esse subconjunto foi utilizado para cálculo dos parâmetros de CDF para ECG. Apenas este subconjunto foi utilizado, pois assim, garante-se que indivíduos utilizados no treino ou no teste não sejam utilizados para calcular os parâmeros de CDF, o que em cenários reais é o mais provável de se acontecer. Mais detalhes sobre o processo de treinamento da rede são mostrados no Apêndice A.

Os parâmetros foram calculados utilizando a distância cosseno e são exibidos na Tabela 6.3.

Base	**Tipo**	$\mu_{genuíno}$	$\sigma_{genuíno}$	$\mu_{invasor}$	$\sigma_{invasor}$
LTST	ECG	0.0997	0.1180	0.5550	0.1722

Tabela 6.3: Parâmetros CDF obtidos na base de dados LTST para serem utilizados no método de Pamplona et al. (2013) para a modalidade de ECG.

6.2.3 Resultados

De maneira similar a faces, os valores de P_{seguro} de todas as sessões em cada um desses testes para ECG foram compilados em uma curva ROC e os resultados são exibidos na Figura 6.6. As curvas ROC mostram que todas as variações do PCM são melhores que o método CDF. A Tabela 6.4 mostra os valores EER de cada uma dessas curvas.

A Figura 6.7 mostra a análise qualitativa dos experimentos, seguindo o mesmo protocolo utilizado para faces.

Observando a Figura 6.7 podemos perceber que ECG é a modalidade mais difícil para ambos os métodos. Existem muito mais indivíduos, mais variações intra e inter classes, especialmente devido ao fato da base de dados conter casos de arritmias. Os batimentos cardíacos anormais reduzem o desempenho dos sistemas. Isso reforça a hipótese de que a biometria ECG é bastante difícil, tanto no reconhecimento (como apontado no Apêndice

Método	Treinamento	ECG
PCM	-	0.1349
PCM D2	-	**0.0938**
PCM D3	-	0.0951
PCM ZT	-	0.1170
CDF	LTST	0.1260

Tabela 6.4: EERs para autenticação contínua utilizando ECG. Em negrito os melhores resultados.

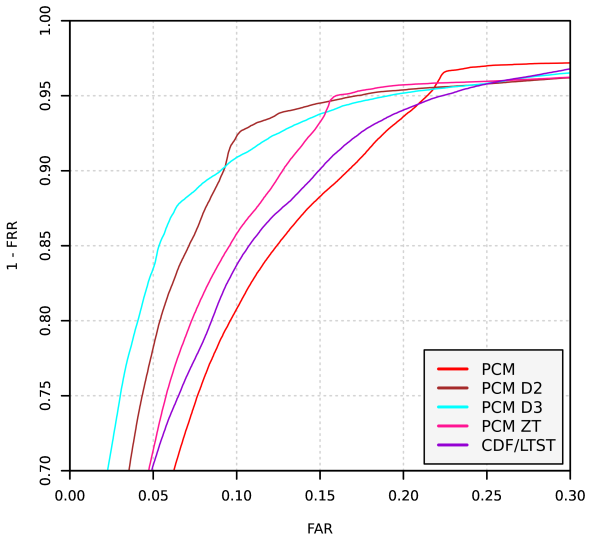
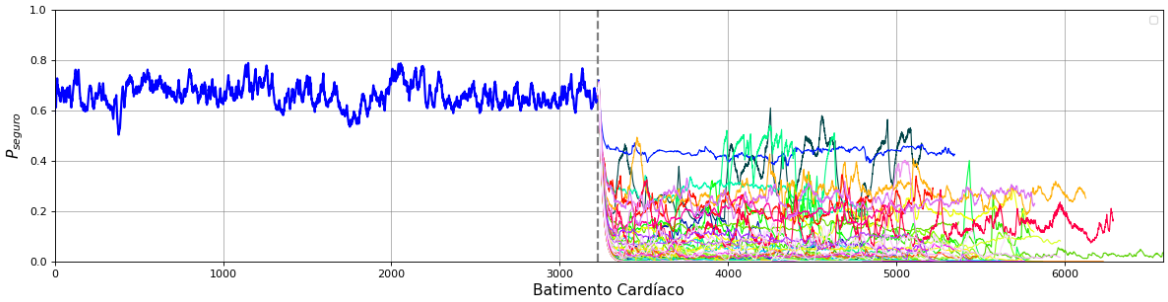


Figura 6.6: Curvas ROC de autenticação contínua usando PCM e CDF para modalidade ECG.

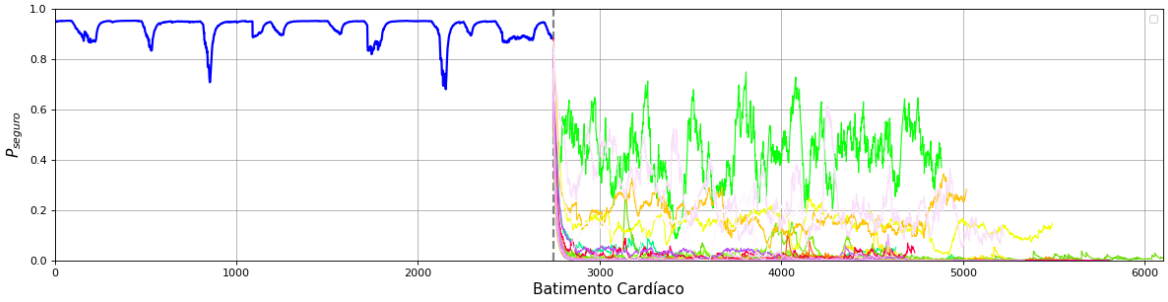
A) quanto na autenticação contínua. No pior caso, podemos observar que o valor de P_{safe} fica praticamente zerado para o usuário genuíno. Isso acontece porque algumas gravações da MITDB começam com um batimento cardíaco anormal. Tal batimento é utilizado como amostra de login e influencia o restante da gravação para ambos os métodos.

6.3 DISCUSSÃO

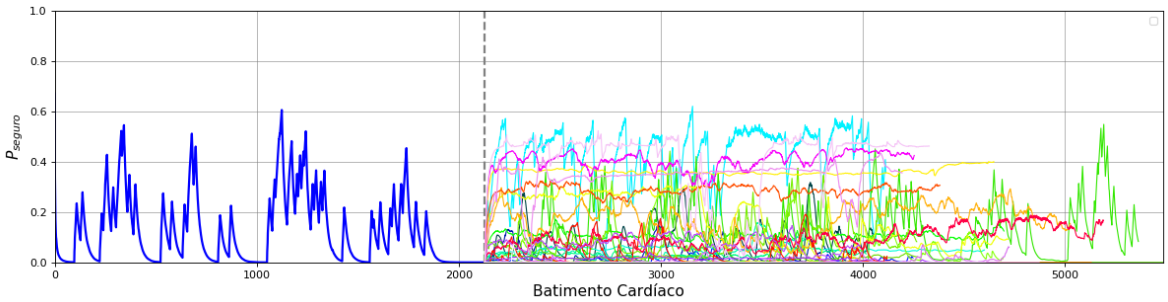
Apesar do método PCM D3 não ter os melhores resultados para cada biometria, ele é balanceado ao longo dos resultados. Como o valor de $\eta_{invasor}$ é fixo para todas as variações do PCM, o que determina o seu desempenho final é o valor de $\eta_{genuíno}$. Para entender como o valor de $\eta_{genuíno}$ afeta o desempenho, para cada variante do PCM, a Figura 6.8 coloca o centroide invasor na menor distância da amostra de login, cuja pertinência ao grupo genuíno se torne 0. Podemos observar que o PCM D3 tem uma bom desempenho enquanto as amostras invasoras estejam a pelo menos 0.666 de distância da amostra



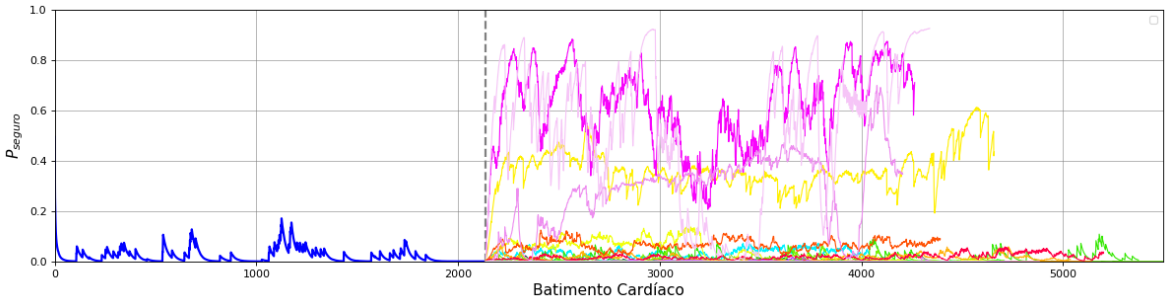
(a) Melhor caso para PCM D3



(b) Melhor caso para CDF



(c) Pior caso para PCM D3



(d) Pior caso para CDF

Figura 6.7: Valores de P_{seguro} ao longo do tempo para modalidade de ECG. O pior caso de ambos pertence ao mesmo sujeito. Os valores de CDF foram obtidos usando os parâmetros da LTST (vide Tabela 6.3).

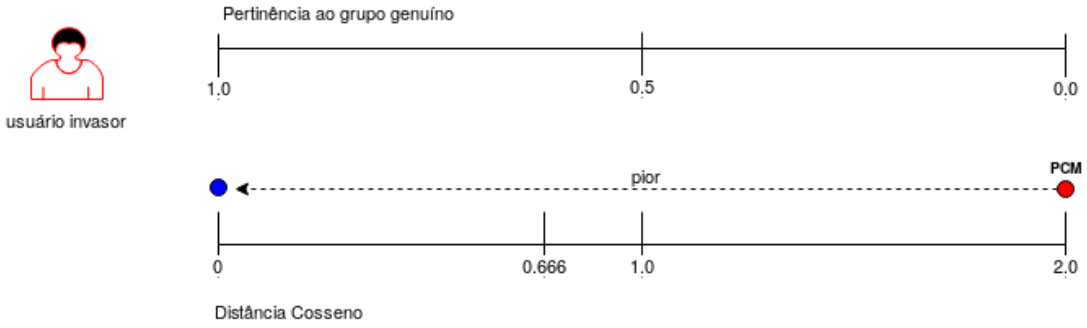
de login. Em outras palavras, ele possue uma margem de segurança melhor: qualquer descritor que coloque amostras invasoras a 0.666 de distância da amostra de login, fará com que a pertinência ao grupo genuíno seja igual a 0, o que não acontece com outros métodos.

Em contrapartida, isso traz uma desvantagem para o PCM D3: ele é mais afetado por variações intraclasse. Como o intervalo de distâncias que define a pertinência ao grupo genuíno é menor (vai de 0 a 0.666), pequenas variações intraclasse reduzem a pertinência ao grupo genuíno quando o usuário permitido está utilizando o sistema. Em adicional, caso o descritor não deixe as amostras genuínas tão próximas umas das outras, o PCM D3 terá desempenho pior do que os outros métodos.

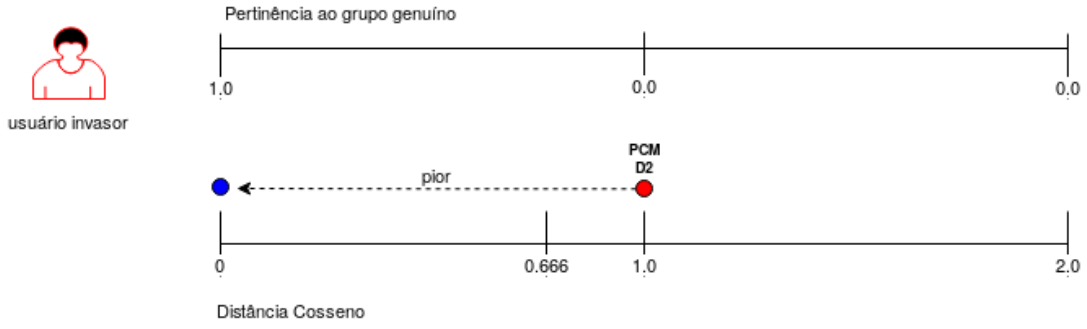
Em resumo, quando comparado com os outros métodos, o PCM D3 possui melhor desempenho quando o descritor não consegue separar bem amostras interclasse, e possui pior desempenho quando o descritor não consegue unir bem amostras intraclasse. Apesar do PCM ZT não ser incluído nessa comparação, podemos afirmar que trata-se de uma técnica intermediária entre o PCM e o PCM D2, desde que $0 < dist(c_{genuíno}, z_t) < 1$, o que é bastante razoável.

Os resultados dos experimentos indicam que na média, os descritores agrupam bem amostras genuínas e não separam tão bem as invasoras, o que favorece o PCM D3. Por exemplo, ele possui o pior EER para faces 2D, cujo descritor foi treinado com milhões de faces de milhares de pessoas diferentes, visto que as distâncias intraclasses são mínimas. Entretanto, o descritor é muito bom, e essa piora é praticamente insignificante. Por outro lado, quando analisamos o desempenho para ECG, temos um descritor treinado com um número muito pequeno de usuários e com muitas amostras dos mesmos. Assim, o descritor não consegue separar muito bem amostras de pessoas diferentes, e o método ganha muito mais desempenho quando comparamos ao PCM que supõe que amostras de pessoas diferentes usem quase o máximo da métrica.

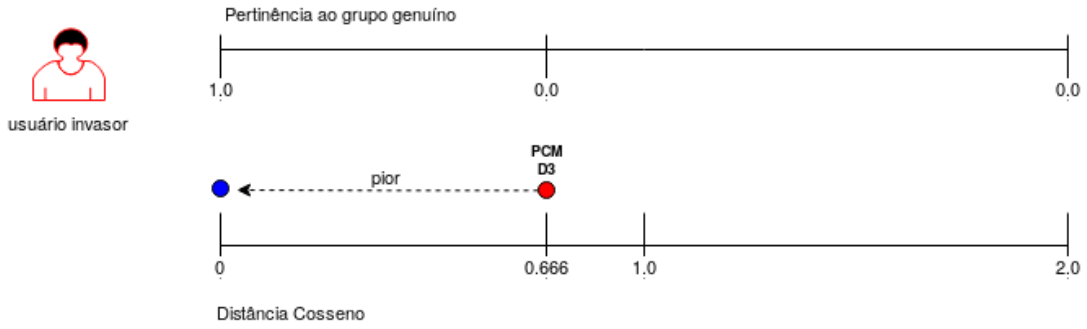
Por fim, o PCM não precisa de treinamento, se adequando a qualquer mudança de característica biométrica. Neste trabalho foram utilizadas várias bases de dados para o cálculo dos parâmetros de CDF, mas em uma aplicação real não é possível escolher os parâmetros que dão o melhor resultado, pois o conjunto de teste seria o próprio sistema em execução.



(a) PCM



(b) PCM D2



(c) PCM D3

Figura 6.8: Relação entre desempenho e distâncias. O círculo azul escuro representa a amostra de login. O círculo vermelho representa o centróide invasor. Para cada ilustração, supomos que o centroide invasor está na distância mínima para que a pertinência ao grupo genuíno seja igual a 0 (que é o desejável). Por exemplo, em (c), o centroide invasor está a 0.666 de distância da amostra de login. Com isso, temos que $\eta_{genuíno} = 2 * 0.333 - 0.666 = 0$. Com $\eta_{genuíno} = 0$, a pertinência ao grupo genuíno é 0. A medida que a distância for menor que 0.666, o valor de $\eta_{genuíno}$ começa a aumentar, e consequentemente, aumenta a pertinência ao grupo genuíno.

Capítulo 7

CONCLUSÃO

Após uma análise na literatura de detecção de anomalias, o algoritmo Possibilistic C-Means (PCM) foi selecionado como mais promissor para o contexto de autenticação contínua, incluindo diferentes variações. Após implementado e comparado em 4 modalidades biométricas diferentes com o estado-da-arte, os resultados mostram que o novo método é ligeiramente superior ao estado-da-arte, porém sem a necessidade de calcular parâmetros sob bases de dados previamente.

As diferenças entre os resultados do estado-da-arte revelam que em cenários reais, nem sempre a biometria irá seguir aquela distribuição de probabilidade previamente calculada. O novo método também apresenta grande estabilidade durante a utilização do usuário genuíno. Isso permite que possa ser definido um limiar para diferentes modalidades e o usuário genuíno terá menor probabilidade de ter o acesso negado ao sistema devido a variações intra-classe. Quando um invasor assume o sistema o valor de P_{seguro} cai e também fica estável, o que dificultaria um invasor conseguir ultrapassar o limiar, mesmo que por um breve instante.

O fato de nenhum dos métodos ter o melhor desempenho individual para cada biometria também demonstra que o problema de autenticação contínua é muito mais difícil do que é reportado na literatura. Após isolarmos o problema, tornando-o invariante à biometria utilizada, podemos constatar que desenvolver um método robusto ao tipo de biometria é bastante difícil, pois em geral, tentativas de melhorar o método para um caso/biometria específica, pode causar uma queda no desempenho para outro caso/biometria.

7.1 TRABALHOS FUTUROS

O PCM pode ser melhorado, pois por mais estável que seja o valor de P_{seguro}, idealmente ele deve ser próximo de 1 para o usuário genuíno e próximo de 0 para o invasor. O valor de P_{seguro} em 0.5 durante o usuário genuíno revela que as amostras tem pertinências iguais tanto ao grupo genuíno quanto ao invasor. Isso se deve ao fato de que a pertinência ao grupo invasor nunca será próxima de 0. As amostras sempre terão alta pertinência ao

grupo invasor, pois elas estão sempre próximas de seu centroide e o valor de $\eta_{invasor}$ não é baixo o suficiente para reduzir a pertinência de tais amostras ao grupo invasor. Na prática, o que controla o valor de P_{seguro} é a pertinência ao grupo genuíno.

Possíveis soluções para este problema são formas melhores de definir $\eta_{genuíno}$ e $\eta_{invasor}$ ou até mesmo reduzir a influência no cálculo de P_{seguro} da pertinência ao grupo invasor e aumentar a influência da pertinência ao grupo genuíno, visto que na prática, a pertinência ao grupo genuíno é a que varia muito e controla o valor de P_{seguro}. Também é possível utilizar um método híbrido que utiliza o PCM e o método CDF em conjunto para calcular o valor de P_{seguro}.

Com o avanço de *deep learning* e do aumento das bases de dados, em um futuro próximo, pode-se utilizar a ideia deste trabalho em conjunto com redes neurais, fazendo com que elas analisem um histórico de amostras e aprendam a inferir sobre a segurança do sistema.

O resultado deste trabalho foi aceito em 2018 no evento IEEE International Conference on Fuzzy Systems. O código utilizado neste trabalho está disponível em ⟨https://github.com/ivision-ufba/pcm-continuous-authentication⟩.

REFERÊNCIAS BIBLIOGRÁFICAS

ACAR, A. et al. Waca: Wearable-assisted continuous authentication. In: IEEE. *2018 IEEE Security and Privacy Workshops (SPW)*. [S.l.], 2018. p. 264–269.

AGRAFIOTI, F.; HATZINAKOS, D. Ecg based recognition using second order statistics. In: IEEE. *Communication Networks and Services Research Conference*. [S.l.], 2008. p. 82–87.

AGRAFIOTI, F.; HATZINAKOS, D. Ecg biometric analysis in cardiac irregularity conditions. *Signal, Image and Video Processing*, p. 3(4):329–343, 2009.

AHONEN, T.; HADID, A.; PIETIKÄINEN, M. Face recognition with local binary patterns. In: SPRINGER. *European conference on computer vision*. [S.l.], 2004. p. 469–481.

ALBRECHT, P. *ST segment characterization for long term automated ECG analysis*. Tese (Doutorado) — Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 1983.

ALESKEROV, E.; FREISLEBEN, B.; RAO, B. Cardwatch: A neural network based database mining system for credit card fraud detection. In: IEEE. *Computational Intelligence for Financial Engineering (CIFEr), 1997., Proceedings of the IEEE/IAFE 1997*. [S.l.], 1997. p. 220–226.

ALTINOK, A.; TURK, M. Temporal integration for continuous multimodal biometrics. In: CITESEER. *Proceedings of the Workshop on Multimodal User Authentication*. [S.l.], 2003.

ANSCOMBE, F. J.; GUTTMAN, I. Rejection of outliers. *Technometrics*, Taylor & Francis, Ltd., American Statistical Association, American Society for Quality, v. 2, n. 2, p. 123–147, 1960. ISSN 00401706. Disponível em: ⟨http://www.jstor.org/stable/1266540⟩.

BALTRUŠAITIS, T.; ROBINSON, P.; MORENCY, L.-P. Openface: an open source facial behavior analysis toolkit. In: IEEE. *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. [S.l.], 2016. p. 1–10.

BARBELLO, B. Continuous user authentication on mobile devices: Recent progress and remaining challenges. 2016.

BELGACEM, N. et al. Person identification system based on electrocardiogram signal using labview. *International Journal on Computer Science and Engineering*, Citeseer, v. 4, n. 6, p. 974, 2012.

BELHUMEUR, P.; HESPANHA, J.; KRIEGMAN, D. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. In: BUXTON, B.; CIPOLLA, R. (Ed.). *Computer Vision — ECCV '96*. Springer Berlin Heidelberg, 1996, (Lecture Notes in Computer Science, v. 1064). p. 43–58. ISBN 978-3-540-61122-6. Disponível em: ⟨http://dx.doi.org/10.1007/BFb0015522⟩.

BEZDEK, J. C.; EHRLICH, R.; FULL, W. Fcm: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, Elsevier, v. 10, n. 2-3, p. 191–203, 1984.

BIEL, L. et al. Ecg analysis: a new approach in human identification. *IEEE Transactions on Instrumentation and Measurement*, v. 50, n. 3, p. 808–812, 2001.

BOUMBAROV, O.; VELCHEV, Y.; SOKOLOV, S. Personal biometric identification based on ecg features. *J Inf Technol Control*, v. 3, n. 4, p. 11–8, 2008.

BOUSSELJOT, R.; KREISELER, D.; SCHNABEL, A. Nutzung der ekg-signaldatenbank cardiodat der ptb über das internet. *Biomedizinische Technik/Biomedical Engineering*, Walter de Gruyter, Berlin/New York, v. 40, n. s1, p. 317–318, 1995.

BREUNIG, M. M. et al. Lof: identifying density-based local outliers. In: ACM. *ACM sigmod record*. [S.l.], 2000. v. 29, n. 2, p. 93–104.

BRIDGES, S. M.; VAUGHN, R. B. et al. Fuzzy data mining and genetic algorithms applied to intrusion detection. In: *Proceedings of 12th Annual Canadian Information Technology Security Symposium*. [S.l.: s.n.], 2000. p. 109–122.

BYERS, S.; RAFTERY, A. E. Nearest-neighbor clutter removal for estimating features in spatial point processes. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 93, n. 442, p. 577–584, 1998.

CAO, Q. et al. Vggface2: A dataset for recognising faces across pose and age. In: IEEE. *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. [S.l.], 2018. p. 67–74.

CARREIRAS, C. et al. *BioSPPy: Biosignal Processing in Python*. 2015–. [Online; accessed ¡today¿]. Disponível em: ⟨https://github.com/PIA-Group/BioSPPy/⟩.

CHAN, A. D. et al. Person identification using electrocardiograms. In: IEEE. *Electrical and Computer Engineering, 2006. CCECE'06. Canadian Conference on*. [S.l.], 2006. p. 1–4.

CHAN, A. D. et al. Wavelet distance measure for person identification using electrocardiograms. *IEEE transactions on instrumentation and measurement*, IEEE, v. 57, n. 2, p. 248–253, 2008.

CHANDOLA, V.; BANERJEE, A.; KUMAR, V. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, ACM, v. 41, n. 3, p. 15, 2009.

COUTINHO, D. P. et al. Novel fiducial and non-fiducial approaches to electrocardiogram-based biometric systems. *IET biometrics*, IET, v. 2, n. 2, p. 64–75, 2013.

CROUSE, D. et al. Continuous authentication of mobile user: Fusion of face image and inertial measurement unit data. In: IEEE. *Biometrics (ICB), 2015 International Conference on*. [S.l.], 2015. p. 135–142.

DAHIA, G.; SANTOS, M.; SEGUNDO, M. P. A study of cnn outside of training conditions. In: *2017 IEEE International Conference on Image Processing (ICIP)*. [S.l.: s.n.], 2017. p. 3820–3824.

DAMOUSIS, I. G.; TZOVARAS, D.; BEKIARIS., E. Unobtrusive multimodal biometric authentication: the humabio project concept. *EURASIP Journal on Advances in Signal Processing*, v. 2008, p. 110:1–110:11, 2008.

DUGELAY, J. L. et al. Recent advances in biometric person authentication. In: *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*. [S.l.: s.n.], 2002. v. 4, p. IV–4060–IV–4063.

FATEMIAN, S. Z.; HATZINAKOS, D. A new ecg feature extractor for biometric recognition. In: IEEE. *Digital Signal Processing, 2009 16th International Conference on*. [S.l.], 2009. p. 1–6.

FENG, H.; FAWAZ, K.; SHIN, K. G. Continuous authentication for voice assistants. In: ACM. *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*. [S.l.], 2017. p. 343–355.

FENU, G.; MARRAS, M.; BORATTO, L. A multi-biometric system for continuous student authentication in e-learning platforms. *Pattern Recognition Letters*, Elsevier, v. 113, p. 83–92, 2018.

FERNANDES, S.; BALA, J. Performance analysis of pca-based and lda-based algorithms for face recognition. *International Journal of Signal Processing Systems*, v. 1, n. 1, p. 1–6, 2013.

FIERREZ, J. et al. Benchmarking touchscreen biometrics for mobile authentication. *IEEE Transactions on Information Forensics and Security*, IEEE, v. 13, n. 11, p. 2720–2733, 2018.

FLIOR, E.; KOWALSKI, K. Continuous biometric user authentication in online examinations. In: IEEE. *Information Technology: New Generations (ITNG), 2010 Seventh International Conference on*. [S.l.], 2010. p. 488–492.

FLIOR, E.; KOWALSKI, K. Continuous biometric user authentication in online examinations. *Seventh International Conference on Information Technology: New Generations,*, 2010.

GASCON, H. et al. Continuous authentication on mobile devices by analysis of typing motion behavior. In: *Sicherheit*. [S.l.: s.n.], 2014. p. 1–12.

GHOFRANI, N.; BOSTANI, R. Reliable features for an ecg-based biometric system. In: IEEE. *Biomedical Engineering (ICBME), 2010 17th Iranian Conference of*. [S.l.], 2010. p. 1–5.

GOLDBERGER, A. et al. Physiobank, physiotoolkit, and physionet : Components of a new research resource for complex physiologic signals. v. 101, p. E215–20, 07 2000.

GRUBBS, F. E. Sample criteria for testing outlying observations. *The Annals of Mathematical Statistics*, JSTOR, p. 27–58, 1950.

GRUBBS, F. E. Procedures for detecting outlying observations in samples. *Technometrics*, Taylor & Francis, v. 11, n. 1, p. 1–21, 1969.

GUNETTI, D.; PICARDI, C. Keystroke analysis of free text. *ACM Trans. Inf. Syst. Secur*, 2005.

GUTTORMSSON, S. E. et al. Elliptical novelty grouping for on-line short-turn detection of excited running rotors. *IEEE Transactions on Energy Conversion*, IEEE, v. 14, n. 1, p. 16–22, 1999.

H, M.; PJ, P. Computational and performance aspects of pca-based face-recognition algorithms. *Perception*, v. 30, p. 303 – 321, 2001.

HAMDI, T.; SLIMANE, A. B.; KHALIFA, A. B. A novel feature extraction method in ecg biometrics. In: *Image Processing, Applications and Systems Conference (IPAS)*. [S.l.: s.n.], 2014. p. 1–5.

HARTIGAN, J. A.; WONG, M. A. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, JSTOR, v. 28, n. 1, p. 100–108, 1979.

HAUTAMAKI, V.; KARKKAINEN, I.; FRANTI, P. Outlier detection using k-nearest neighbour graph. In: IEEE. *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*. [S.l.], 2004. v. 3, p. 430–433.

HODGE, V.; AUSTIN, J. A survey of outlier detection methodologies. *Artificial intelligence review*, Springer, v. 22, n. 2, p. 85–126, 2004.

IDÉ, T.; PAPADIMITRIOU, S.; VLACHOS, M. Computing correlation anomaly scores using stochastic nearest neighbors. In: IEEE. *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on*. [S.l.], 2007. p. 523–528.

ISRAEL, S. A. et al. Ecg to identify individuals. *Pattern recognition*, Elsevier, v. 38, n. 1, p. 133–142, 2005.

JAGER, F. et al. Long-term st database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia. *Medical and Biological Engineering and Computing*, Springer, v. 41, n. 2, p. 172–182, 2003.

JAIN, A. K. et al. An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, v. 14, n. 1, 2004.

JANAKIRAMAN, R. et al. Using continuous face verification to improve desktop security. In: *Application of Computer Vision, 2005. WACV/MOTIONS '05 Volume 1. Seventh IEEE Workshops on*. [S.l.: s.n.], 2005. v. 1, p. 501–507.

JEKOVA, I.; BORTOLAN, G. Personal verification/identification via analysis of the peripheral ecg leads: Influence of the personal health status on the accuracy. *BioMed research international*, Hindawi, v. 2015, 2015.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

KONG, S. et al. Recent advances in visual and infrared face recognition—a review. *Computer Vision and Image Understanding*, v. 97, p. 103–135, January 2005. Issue 1.

KONG, S. G. et al. Recent advances in visual and infrared face recognition—a review. *Computer Vision and Image Understanding 97 (2005) 103–135*, v. 97, p. 103–135, April 2004. Issue 1.

KRISHNAPURAM, R.; KELLER, J. M. A possibilistic approach to clustering. *IEEE Transactions on Fuzzy Systems*, IEEE, v. 1, n. 2, p. 98–110, 1993.

KRISHNAPURAM, R.; KELLER, J. M. The possibilistic c-means algorithm: insights and recommendations. *IEEE Transactions on Fuzzy Systems*, v. 4, n. 3, p. 385–393, 1996.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, v. 60, n. 6, p. 84–90, 2017.

KUMAR, S. et al. Using continuous biometric verification to protect interactive login sessions. In: IEEE. *21st Annual Computer Security Applications Conference (ACSAC'05)*. [S.l.], 2005. p. 10–pp.

KUMAR, V. Parallel and distributed computing for cybersecurity. *IEEE Distributed Systems Online*, IEEE, v. 6, n. 10, 2005.

KUNZ, M. et al. Continuous speaker verification in realtime. *BIOSIG 2011–Proceedings of the Biometrics Special Interest Group*, Gesellschaft für Informatik eV, 2011.

LAGUNA, P. et al. A database for evaluation of algorithms for measurement of qt and other waveform intervals in the ecg. In: IEEE. *Computers in cardiology 1997*. [S.l.], 1997. p. 673–676.

LECUN, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, IEEE, v. 86, n. 11, p. 2278–2324, 1998.

LEGGETT, J. et al. Dynamic identity verification via keystroke characteristics. *International Journal of Man-Machine Studies*, Elsevier, v. 35, n. 6, p. 859–870, 1991.

LI, D.-Y.; LIAO, W.-H. Facial feature detection in near-infrared images. *Proc. of 5th International Conference on Computer Vision, Pattern Recognition and Image Processing*, 2003.

LI, S. et al. The casia nir-vis 2.0 face database. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. [S.l.: s.n.], 2013. p. 348–353.

LI, S. Z. et al. Illumination invariant face recognition using near-infrared images. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, v. 29, n. 4, p. 627 – 639, April 2007.

LI, Y. et al. Patient-specific ecg classification by deeper cnn from generic to dedicated. *Neurocomputing*, v. 314, p. 336–346, 2018.

LIAO, Y.; VEMURI, V. R. Use of k-nearest neighbor classifier for intrusion detection. *Computers & security*, Elsevier, v. 21, n. 5, p. 439–448, 2002.

LU, J.; PLATANIOTIS, K. N.; VENETSANOPOULOS, A. N. Face recognition using lda-based algorithms. *IEEE Transactions on Neural networks*, IEEE, v. 14, n. 1, p. 195–200, 2003.

MAI, V.; KHALIL, I.; MELI, C. Ecg biometric using multilayer perceptron and radial basis function neural networks. In: IEEE. *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. [S.l.], 2011. p. 2745–2748.

MOCK, K. et al. Real-time continuous iris recognition for authentication using an eye tracker. In: *Proceedings of the 2012 ACM Conference on Computer and Communications Security*. New York, NY, USA: ACM, 2012. (CCS '12), p. 1007–1009. ISBN 978-1-4503-1651-4. Disponível em: ⟨http://doi.acm.org/10.1145/2382196.2382307⟩.

MOGHADDAM, W. W. B.; PENTLAND, A. Beyond eigenfaces: Probabilistic matching for face recognition. *Automatic Face and Gesture Recognition*, p. 30 – 35, 1998.

MONACO, J. V. et al. Developing a keystroke biometric system for continual authentication of computer users. In: IEEE. *Intelligence and Security Informatics Conference (EISIC), 2012 European*. [S.l.], 2012. p. 210–216.

MONACO, J. V.; TAPPERT, C. C. The partially observable hidden markov model and its application to keystroke dynamics. *Pattern Recognition*, Elsevier, v. 76, p. 449–462, 2018.

MONDAL, S.; BOURS, P. A continuous combination of security & forensics for mobile devices. *Journal of information security and applications*, Elsevier, v. 40, p. 63–77, 2018.

MOODY, G. B.; MARK, R. G. The impact of the mit-bih arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, IEEE, v. 20, n. 3, p. 45–50, 2001.

MURMURIA, R. et al. Continuous authentication on mobile devices using power consumption, touch gestures and physical movement of users. In: SPRINGER. *International Symposium on Recent Advances in Intrusion Detection*. [S.l.], 2015. p. 405–424.

NAKANISHI, I.; BABA, S.; MIYAMOTO, C. Eeg based biometric authentication using new spectral features. In: *Intelligent Signal Processing and Communication Systems, 2009. ISPACS 2009. International Symposium on*. [S.l.: s.n.], 2009. p. 651–654.

NIINUMA, K.; PARK, U.; JAIN, A. Soft biometric traits for continuous user authentication. *Information Forensics and Security, IEEE Transactions on*, v. 5, n. 4, p. 771–780, Dec 2010.

PAMPLONA, M. S. et al. Continuous 3d face authentication using rgb-d cameras. *Computer Vision and Pattern Recognition Workshops (CVPRW)*, p. 64 – 69, June 2013.

PATCHA, A.; PARK, J.-M. An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer networks*, Elsevier, v. 51, n. 12, p. 3448–3470, 2007.

PINTO, J. R. et al. Towards a continuous biometric system based on ecg signals acquired on the steering wheel. *Sensors*, v. 17, n. 10, 2017.

PLATANIOTIS, K. N.; HATZINAKOS, D.; LEE, J. K. Ecg biometric recognition without fiducial detection. In: IEEE. *Biometric Consortium Conference, 2006 Biometrics Symposium: Special Session on Research at the*. [S.l.], 2006. p. 1–6.

PORTNOY, L.; ESKIN, E.; STOLFO, S. Intrusion detection with unlabeled data using clustering. In: CITESEER. *In Proceedings of ACM CSS Workshop on Data Mining Applied to Security (DMSA-2001*. [S.l.], 2001.

POURBABAEE, B. et al. Deep convolutional neural network for ecg-based human identification. *CMBES Proceedings*, v. 41, 2018.

RABINER, L. R. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, Ieee, v. 77, n. 2, p. 257–286, 1989.

RAMASWAMY, S.; RASTOGI, R.; SHIM, K. Efficient algorithms for mining outliers from large data sets. In: ACM. *ACM Sigmod Record*. [S.l.], 2000. v. 29, n. 2, p. 427–438.

ROY, A.; HALEVI, T.; MEMON, N. An hmm-based behavior modeling approach for continuous mobile authentication. In: IEEE. *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. [S.l.], 2014. p. 3789–3793.

SAFIE, S. I.; SORAGHAN, J. J.; PETROPOULAKIS, L. Ecg biometric authentication using pulse active width (paw). In: IEEE. *Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2011 IEEE Workshop on*. [S.l.], 2011. p. 1–6.

SAFIE, S. I.; SORAGHAN, J. J.; PETROPOULAKIS, L. Electrocardiogram (ecg) biometric authentication using pulse active ratio (par). *IEEE Transactions on Information Forensics and Security*, IEEE, v. 6, n. 4, p. 1315–1322, 2011.

SANTOS, M.; SEGUNDO, M. P. Autenticação facial contínua usando imagens de infravermelho. In: *Conference on Graphics, Patterns and Images (SIBGRAPI)*. [S.l.: s.n.], 2015. v. 28.

SASIKALA, P.; WAHIDABANU, R. Identification of individuals using electrocardiogram. *International journal of computer science and network security*, v. 10, n. 12, p. 147–153, 2010.

SCHROFF, F.; KALENICHENKO, D.; PHILBIN, J. Facenet: A unified embedding for face recognition and clustering. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 815–823.

SEGUNDO, M. P. et al. Orthogonal projection images for 3d face detection. *Pattern Recognition Letters*, v. 50, p. 72–81, 2014. Depth Image Analysis.

SHEN, T.-W.; TOMPKINS, W.; HU, Y. One-lead ecg for identity verification. In: IEEE. *Engineering in medicine and biology, 2002. 24th annual conference and the annual fall meeting of the biomedical engineering society embs/bmes conference, 2002. proceedings of the second joint*. [S.l.], 2002. v. 1, p. 62–63.

SILVA, A. da; SEGUNDO, M. P. Reconhecimento facial 2d para autenticação contínua. In: *Conference on Graphics, Patterns and Images (SIBGRAPI)*. [S.l.: s.n.], 2015. v. 28.

SIM, T. et al. Continuous verification using multimodal biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 29, n. 4, April 2007.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

SINGH, S. et al. Keystroke dynamics for continuous authentication. In: IEEE. *2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. [S.l.], 2018. p. 205–208.

SINGLA, S. K.; SHARMA, A. Ecg based biometrics verification system using labview. *Sonklanakarin Journal of Science and Technology*, v. 32, n. 3, p. 241, 2010.

SITOVÁ, Z. et al. Hmog: New behavioral biometric features for continuous authentication of smartphone users. *IEEE Transactions on Information Forensics and Security*, IEEE, v. 11, n. 5, p. 877–892, 2016.

SPENCE, C.; PARRA, L.; SAJDA, P. Detection, synthesis and compression in mammographic image analysis with a hierarchical image probability model. In: IEEE. *Mathematical Methods in Biomedical Image Analysis, 2001. MMBIA 2001. IEEE Workshop on*. [S.l.], 2001. p. 3–10.

SRIVASTAVA, N. et al. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, v. 15, p. 1929–1958, 2014. Disponível em: ⟨http://jmlr.org/papers/v15/srivastava14a.html⟩.

STEFANSKY, W. Rejecting outliers in factorial designs. *Technometrics*, Taylor & Francis Group, v. 14, n. 2, p. 469–479, 1972.

SUFI, F.; KHALIL, I. Faster person identification using compressed ecg in time critical wireless telecardiology applications. *Journal of Network and Computer Applications*, Elsevier, v. 34, n. 1, p. 282–293, 2011.

TAN, P.-N.; STEINBACH, M.; KUMAR, V. *Introduction to Data Mining, (First Edition)*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2005. ISBN 0321321367.

TANG, J. et al. Enhancing effectiveness of outlier detections for low density patterns. *Advances in Knowledge Discovery and Data Mining*, Springer, p. 535–548, 2002.

TAWFIK, M. M.; SELIM, H.; KAMAL, T. Human identification using time normalized qt signal and the qrs complex of the ecg. In: IEEE. *Communication Systems Networks and Digital Signal Processing (CSNDSP), 2010 7th International Symposium on*. [S.l.], 2010. p. 755–759.

TING, C.-M.; SALLEH, S.-H. Ecg based personal identification using extended kalman filter. In: IEEE. *Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference on*. [S.l.], 2010. p. 774–777.

TSAI, P.-W. et al. Interactive artificial bee colony supported passive continuous authentication system. *IEEE Systems Journal*, IEEE, v. 8, n. 2, p. 395–405, 2014.

WAILI, T. et al. A hasty approach to ecg person identification. In: IEEE. *2016 International Conference on Computer and Communication Engineering (ICCCE)*. [S.l.], 2016. p. 267–271.

WALD, A. *Sequential analysis*. [S.l.]: Courier Corporation, 1973.

WANG, Y. et al. Analysis of human electrocardiogram for biometric recognition. *EURASIP journal on Advances in Signal Processing*, Springer, v. 2008, n. 1, p. 148658, 2007.

WANG, Y.; PLATANIOTIS, K. N.; HATZINAKOS, D. Integrating analytic and appearance attributes for human identification from ecg signals. In: IEEE. *Biometric Consortium Conference, 2006 Biometrics Symposium: Special Session on Research at the*. [S.l.], 2006. p. 1–6.

WEI, T.; ZHIHUA, X. Infrared face recognition based on local binary pattern and multi-objective genetic algorithm. *Proceeding of the IEEE International Conference on Information and Automation Shenzhen*, p. 359 – 362, June 2011.

WU, X. et al. A light cnn for deep face representation with noisy labels. *arXiv preprint arXiv:1511.02683*, 2015.

WÜBBELER, G. et al. Verification of humans using the electrocardiogram. *Pattern Recognition Letters*, Elsevier, v. 28, n. 10, p. 1172–1175, 2007.

YAO, J.; WAN, Y. A wavelet method for biometric identification using wearable ecg sensors. In: IEEE. *Medical Devices and Biosensors, 2008. ISSS-MDBS 2008. 5th International Summer School and Symposium on*. [S.l.], 2008. p. 297–300.

YE, C.; COIMBRA, M. T.; KUMAR, B. V. Investigation of human identification using two-lead electrocardiogram (ecg) signals. In: IEEE. *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*. [S.l.], 2010. p. 1–8.

ZHANG, C. et al. Understanding deep learning requires rethinking generalization. *CoRR*, abs/1611.03530, 2016. Disponível em: ⟨http://arxiv.org/abs/1611.03530⟩.

ZHANG, K. et al. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, v. 23, n. 10, p. 1499–1503, 2016.

ZHANG, Q.; ZHOU, D.; ZENG, X. Heartid: a multiresolution convolutional neural network for ecg-based biometric human identification in smart health applications. *IEEE Access*, IEEE, v. 5, p. 11805–11816, 2017.

ZHAO, S.; GRIGAT, R.-R. An automatic face recognition system in the near infrared spectrum. *Machine Learning and Data Mining in Pattern Recognition*, v. 3587, p. 437–444, 2005.

ZHENG, Y. Near infrared face recognition using orientation-based face patterns. *Biometrics Special Interest Group (BIOSIG)*, p. 1 – 4, 2012.

ZHIHUA, X.; GUODONG, L. Weighted infrared face recognition in multiwavelet domain. *Imaging Systems and Techniques (IST)*, p. 70 – 74, 2013.

Apêndice

A

UMA ANÁLISE SOBRE ECG COMO CARACTERÍSTICA BIOMÉTRICA

A.1 INTRODUÇÃO

Após uma busca na literatura, foi possível notar que apesar dos bons resultados reportados na literatura, a maioria dos trabalhos de reconhecimento de Eletrocardiograma (ECG) apresentam o mesmo problema: o uso do mesmo conjunto de sujeitos para treinamento e teste (SUFI; KHALIL, 2011; WAILI et al., 2016; ZHANG; ZHOU; ZENG, 2017; YAO; WAN, 2008). Outros problemas comuns são o uso de bases de dados privadas (BIEL et al., 2001; ISRAEL et al., 2005; CHAN et al., 2006; BOUMBAROV; VELCHEV; SOKOLOV, 2008; SINGLA; SHARMA, 2010; TAWFIK; SELIM; KAMAL, 2010; BELGACEM et al., 2012; HAMDI; SLIMANE; KHALIFA, 2014; POURBABAEE et al., 2018), o uso de bases de dados que não estão mais disponíveis (PLATANIOTIS; HATZINAKOS; LEE, 2006; WÜBBELER et al., 2007; FATEMIAN; HATZINAKOS, 2009; GHOFRANI; BOSTANI, 2010; SAFIE; SORAGHAN; PETROPOULAKIS, 2011b, 2011a; COUTINHO et al., 2013; JEKOVA; BORTOLAN, 2015) ou deixar de fornecer informações críticas para reprodutibilidade (*e.g.* escolha de um subconjunto de uma base de dados sem especificar o critério de seleção) (SHEN; TOMPKINS; HU, 2002; WANG; PLATANIOTIS; HATZINAKOS, 2006; WANG et al., 2007; AGRAFIOTI; HATZINAKOS, 2008; CHAN et al., 2008; COUTINHO et al., 2013). Tais práticas dificultam uma análise confiável do poder de generalização dos métodos apresentados e impede uma comparação de seus resultados. Como consequência, é difícil definir o estado-da-arte do reconhecimento de pessoas baseado em ECG.

Neste trabalho, investigamos a confiabilidade do ECG para o reconhecimento biométrico. Para isso, propomos o uso de uma simples Convolutional Neural Network (CNN) adaptada de uma arquitetura bem conhecida (LECUN et al., 1998) para reconhecer indivíduos utilizando uma pequena porção do sinal de ECG: um único batimento cardíaco. Com isso, esperamos estabelecer um patamar de comparação com os resultados do estado-da-arte que possam orientar futuras comparações na literatura. Replicamos os experimentos

de outros trabalhos da literatura para mostrar que nossa CNN é comparável às abordagens existentes. Posteriormente, usamos duas bases de dados públicas para realizar uma análise mais profunda, que em nossa opinião, revela o verdadeiro potencial do ECG para fins de reconhecimento biométrico.

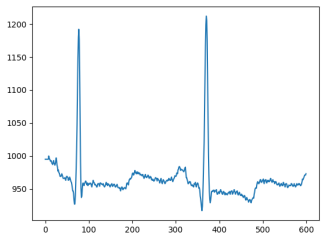
A.2 RECONHECIMENTO BASEADO EM ECG UTILIZANDO CNN

Existem diferentes maneiras de reconhecer indivíduos através de sinais de ECG. Alguns métodos usam um registro ECG inteiro ou usam heurísticas para selecionar batimentos específicos um registro de ECG (FATEMIAN; HATZINAKOS, 2009; SASIKALA; WAHIDABANU, 2010; SAFIE; SORAGHAN; PETROPOULAKIS, 2011b; POURBABAEE et al., 2018) e outros usam um único batimento cardíaco (ZHANG; ZHOU; ZENG, 2017; MAI; KHALIL; MELI, 2011). Como métodos que usam um único batimento cardíaco podem ser facilmente estendidos para usar sinais mais longos, neste trabalho focamos neste tipo de método.

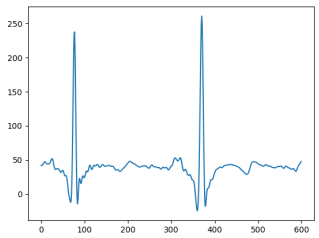
Para aquisição do sinal de ECG, usamos o kit PhysioNet (GOLDBERGER et al., 2000) para extrair os registros das bases de dados existentes. A Figura A.1a mostra uma parte de um sinal original contido em um registro de ECG. Como pode ser observado, cada registro contém vários batimentos cardíacos e o sinal pode conter ruídos. Portanto, para todos os registros de ECG, usamos o a ferramenta Biosppy (CARREIRAS et al., 2015–) tanto para filtrar os ruídos, conforme mostrado na Figura A.1b, quanto para segmentar os batimentos cardíacos existentes, conforme mostrado na Figura A.1c. Por fim, cada batimento cardíaco é interpolado linearmente para um comprimento fixo de 256 unidades e sua amplitude é normalizada no intervalo de 0 a 1 (utilizando o mínimo e o máximo), como mostrado na Figura A.1d. Após esse processo, os batimentos cardíacos pré-processados podem passar por um classificador para fins de reconhecimento.

No nosso caso, esse classificador é uma CNN, que recentemente se tornou o estado-da-arte para problemas de reconhecimento de padrões (KRIZHEVSKY; SUTSKEVER; HINTON, 2017). Consequentemente, as CNNs também atraíram a atenção de pesquisas baseadas em ECG (ZHANG; ZHOU; ZENG, 2017; LI et al., 2018). Como um batimento cardíaco é representado por um vetor unidimensional de 256 posições, usar uma arquitetura de CNN muito profunda para descrever ou classificá-los pode causar *overfit*. Para evitar esse problema, escolhemos uma arquitetura simples e bem conhecida na literatura: a LeNet-5 (LECUN et al., 1998). Como essa arquitetura foi projetada para classificar dígitos em imagens 32×32, tivemos que adaptá-la para sinais unidimensionais de batimentos cardíacos e modernizá-la com as práticas mais comuns no design de CNNs. Isso foi feito: (1) substituindo filtros 2D por filtros 1D; (2) adicionando unidades lineares retificadas, Rectified Linear Units (ReLU), como função de ativação após cada camada oculta; e (3) adicionando uma camada de *dropout* de 50% após a primeira camada densa para melhorar generalização (SRIVASTAVA et al., 2014). A tabela A.1 resume a arquitetura utilizada.

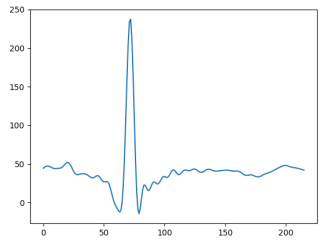
No treinamento, usamos o otimizador Adam (KINGMA; BA, 2014) para minimizar o função de perda de entropia cruzada *softmax*. Cada experimento foi repetido para diferentes taxas de aprendizado inicial que variam de 10^{-1} a 10^{-4} e diferentes tamanhos



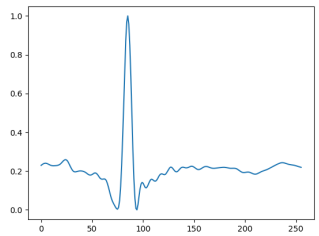
(a) sinal original



(b) sinal filtrado



(c) batimento segmentado



(d) batimento normalizado

Figura A.1: Exemplo da etapa de pré-processamento adotada em nossa abordagem de reconhecimento baseado em ECG: (a) o sinal original passa por (b) filtragem de ruído, (c) segmentação de batimentos cardíacos e (d) normalização em termos de comprimento e amplitude.

de *batch* que variam de 2^3 a 2^8. O treinamento é realizado por 500 épocas com um declínio linear da taxa de aprendizado de 10^{-1} ou até o desempenho no conjunto de validação parar de melhorar após 10 épocas (*early stopping*). Depois de encontrarmos os melhores hiperparâmetros para um experimento (taxa de aprendizado inicial e tamanho do *batch*), combinamos os conjuntos de treinamento e validação para otimizar ainda mais a CNN até que a acurácia do treinamento + validação esteja acima de 98% durante 10 épocas consecutivas.

A.3 EXPERIMENTOS

Os experimentos foram divididos em duas partes. Primeiro, comparamos o desempenho da nossa CNN com outros trabalhos da literatura usando seus protocolos de experimentação (Seção A.3.2). Depois, propomos um novo protocolo para avaliar ao máximo a capacidade de generalização de nosso método em cenários que estão mais próximos da vida real (Seção A.3.8). Todas as bases de dados usadas em nossos experimentos são brevemente descritas a seguir.

A.3.1 Bases de dados

Para os experimentos, selecionamos as bases de dados mais conhecidas da literatura: a QT (LAGUNA et al., 1997), MIT-BIH Arrhythmia Database (MITDB) (MOODY;

#	Type	Input	Filter size	Stride	Dropout	Output
1	Convolution + ReLU	256×1	$5 \times 1 \times 6$	1		256×6
	Max pooling	256×6	2×1	2		128×6
2	Convolution + ReLU	128×6	$5 \times 1 \times 16$	1		128×16
	Max pooling	128×16	2×1	2		64×16
	Flattening	64×16				1024
3	Dense + ReLU	1024			50%	120
4	Dense + ReLU	120				84
5	Dense + Softmax	84				# subjects

Tabela A.1: Descrição da arquitetura da CNN unidimensional baseada na LeNet-5 (LECUN et al., 1998).

MARK, 2001), a MIT-BIH Normal Sinus Rhythm Database (NSRDB) (GOLDBERGER et al., 2000), a MIT-BIH Noise Stress Test Database (STDB) (ALBRECHT, 1983) e a Long-Term ST Database (LTST) (JAGER et al., 2003). A maioria delas foi produzida pelo Laboratório de Arritmia do MIT-BIH. Apesar da grande utilização da base PTB Diagnostic ECG Database (PTB) (BOUSSELJOT; KREISELER; SCHNABEL, 1995) na literatura, até onde sabemos, os trabalhos reportados neste trabalho usaram uma versão desatualizada da base de dados ou selecionou um subconjunto de pacientes saudáveis sem especificar quais são. Devido a esses problemas de reprodutibilidade, a PTB foi não utilizada neste trabalho.

- A QT (LAGUNA et al., 1997) é composta por registros de outras bases de dados. Ela consiste em 105 trechos de quinze minutos de ECG de dois canais, que foram escolhidos para incluir uma ampla variedade nos complexos QRS e ST-T.
- A MITDB (MOODY; MARK, 2001) contém 48 trechos de meia hora, oriundos de registros ambulatoriais de ECG, obtidos de 47 indivíduos estudados entre 1975 e 1979. As gravações apresentam dois canais. Esta base de dados foi especialmente desenvolvida para desenvolver detectores de arritmia.
- A NSRDB (GOLDBERGER et al., 2000) inclui 18 gravações de ECG com longa duração, de diferentes sujeitos sem arritmias significativas.
- A STDB (ALBRECHT, 1983) inclui 28 gravações de ECG de comprimentos variados, e foram obtidos durante testes de esforço.
- A LTST (JAGER et al., 2003) contém 86 gravações longas de ECG de 80 sujeitos escolhidos para exibir uma variedade de eventos de alterações no segmento ST. Cada gravação dura entre 21 e 24 horas e contém dois ou três canais de ECG.

A.3.2 Comparação com trabalhos da literatura

Para fazer uma comparação justa, pesquisamos na literatura por trabalhos que utilizaram bases de dados públicas e que possuem uma descrição clara sobre seus experimentos, o que nos permite replicar as mesmas condições de teste. Muitos trabalhos pesquisados não atendiam a esses requisitos e foram descartados (SUFI; KHALIL, 2011; WAILI et al., 2016; ZHANG; ZHOU; ZENG, 2017; YAO; WAN, 2008; BIEL et al., 2001; ISRAEL et al., 2005; CHAN et al., 2006; BOUMBAROV; VELCHEV; SOKOLOV, 2008; SINGLA; SHARMA, 2010; TAWFIK; SELIM; KAMAL, 2010; BELGACEM et al., 2012; HAMDI; SLIMANE; KHALIFA, 2014; POURBABAEE et al., 2018; PLATANIOTIS; HATZINAKOS; LEE, 2006; WÜBBELER et al., 2007; FATEMIAN; HATZINAKOS, 2009; GHOFRANI; BOSTANI, 2010; SAFIE; SORAGHAN; PETROPOULAKIS, 2011b, 2011a; COUTINHO et al., 2013; JEKOVA; BORTOLAN, 2015; SHEN; TOMPKINS; HU, 2002; WANG; PLATANIOTIS; HATZINAKOS, 2006; WANG et al., 2007; AGRAFIOTI; HATZINAKOS, 2008; CHAN et al., 2008). Nosso objetivo neste experimento é mostrar que uma simples CNN pode desempenhar tão bem ou até melhor do que outros trabalhos, utilizando os mesmos protocolos de experimentação. Isso se deve ao fato das CNNs funcionarem muito bem para reconhecimento biométrico quando todos os sujeitos do conjunto de teste também estão no conjunto de treinamento. A Tabela A.2 resume todas as comparações entre nosso método e outros trabalhos da literatura. Uma demonstração visual dos protocolos de avaliação é ilustrada na Figura A.2, que mostra todos os batimentos cardíacos utilizados durante o treinamento de dois sujeitos (o mais fácil e o mais difícil) para cada protocolo. Uma discussão detalhada sobre cada comparação é apresentada abaixo.

A.3.3 Ting e Salleh (2010)

Usaram um filtro estendido de Kalman para reconhecimento de pessoas usando ECG. Eles treinaram um modelo dinâmico de ECG para cada sujeito e a amostra de teste é rotulada de acordo com o modelo com a maior proximidade. Os experimentos deles usam um único canal Modified Limb Lead I (MLII), de 15 gravações da base QT. Os dados foram divididos em dois conjuntos, um contendo apenas batimentos cardíacos normais e outro com normais e anormais. Somente o primeiro minuto de cada registro é utilizado, com os 30 primeiros segundos usados para treinamento e os 30 segundos restantes para teste. Nós utilizamos 10% dos dados de treinamento para o conjunto de validação e os melhores resultados foram obtidos ao usar uma taxa de aprendizado inicial de 10^{-3} e um tamanho de *batch* de 8. No primeiro conjunto (apenas batimentos normais), alcançamos uma acurácia de 99,639% versus 87,5% para Ting e Salleh (2010). No segundo conjunto (normais + anormais), alcançamos 97,8448% de acurácia versus 61,5%. O baixo número de sujeitos e o uso de batimentos cardíacos consecutivos facilitam a memorização dos dados de treinamento pela CNN, tornando o protocolo de avaliação muito fácil para métodos do estado-da-arte de reconhecimento biométrico. Como exibido na Figura A.2f, mesmo para sujeitos mais difíceis, os picos R estão no mesmo local e o sinal não apresenta tantos ruídos quanto em outros protocolos.

Trabalho	Método	Protocolo (por pessoa)	Base	Resultado reportado	Nosso resultado
[1]	Extended Kalman filter	1º minuto (normal); 30s para treino/30s para teste	QT	87.5%	**99.639%**
		1º minuto (normal e anormal); 30s para treino/30s para teste	QT	61.5%	**97.8448%**
[2]	Wavelet transform, auto-correlation, CNN	500 trechos aleatórios de 2s; 250 para treino/250 para teste	NSRDB	95.1%	**98.7333%**
			MITDB	91.1%	**97.5733%**
			STDB	90.3%	**95.9429%**
[3]	23 pontos do complexo QRS, MLP	18 complexos QRS aleatórios; 12 para treino/6 para teste	NSRDB	**99.69%**	93.5185%
[4]	Wavelet transform, ICA, PCA, SVM	gravação de 30 minutos; primeiros 5min para treino restante para teste	MITDB	92.34%	**92.9102%**
		trechos de 5 min a cada 20 min nas primeiras 2 horas; primeiro trecho para treino restante para teste	NRSDB	88.20%	**96.3703%**
		trechos de 5 min a cada 2h nas primeiras 22 horas; primeiro trecho para treino restante para teste	NRSDB	**79.94%**	72.9820%

Tabela A.2: Resumo das comparações realizadas neste trabalho. Valores em negrito mostram o método com melhor desempenho para cada experimento. [1] (TING; SALLEH, 2010), [2] (ZHANG; ZHOU; ZENG, 2017), [3] (MAI; KHALIL; MELI, 2011), [4] (YE; COIMBRA; KUMAR, 2010)

A.3.4 Zhang, Zhou e Zeng (2017)

Usaram *wavelet transform* para decompor uma janela de ECG filtrada e pré-processada em várias *wavelets*, depois aplicaram a auto-correlação nelas e treinaram uma CNN (muito maior que a nossa) para cada uma (ZHANG; ZHOU; ZENG, 2017). A saída dessas CNNs são unidas por camadas densas cuja saída é um vetor de probabilidades para cada sujeito. Seus experimentos consistem na seleção aleatória de 500 janelas de 2 segundos cada, sendo 250 para treinamento e 250 para teste. Como o principal objetivo em usar uma janela de 2 segundos é incluir pelo menos um batimento cardíaco em cada janela, nós simplesmente selecionamos 500 batimentos cardíacos aleatórios. Embora isso não seja idêntico ao protocolo utilizado, não foi obtida nenhuma vantagem sobre eles, pois no final acabamos usando uma janela de ECG mais curta. Eles repetiram esse experimento em 8 bases de dados diferentes, mas apenas três delas estavam entre as selecionadas por nós: NSRDB, MITDB e STDB. Tais bases de dados contém indivíduos saudáveis, indivíduos com arritmias cardíacas e indivíduos se exercitando, respectivamente. Para todos os experimentos, 10% dos dados de treinamento são usado para validação e os melhores hiper-parâmetros foram uma taxa inicial de aprendizado de 10^{-3} e um tamanho de *batch* de 128.

Para a NSRDB, alcançamos uma acurácia de 98,7333% versus 95,1% para Zhang, Zhou e Zeng (2017). Embora a descrição da NSRDB afirme que contém registros sem arritmias significativas, podemos ver na Figura A.2c de que este protocolo é muito mais difícil do que o de Ting e Salleh (2010) (Figura A.2f). A principal razão por trás disso é a seleção aleatória de batimentos cardíacos, o que reduz a proximidade temporal e

consequentemente aumentam as variações intraclasses.

Para a MITDB, o registro 202 foi descartado, pois os registros 201 e 202 são do mesmo indivíduo e todos os outros indivíduos têm apenas um registro. Além disso, os registros 102 e 104 também foram descartados porque não possuem o canal usado pelos autores (MLII). Obtivemos uma acurácia de 97,5733% versus 91,1%. Como pode ser visto na Figura A.2e, a presença de batimentos cardíacos anormais na MITDB aumentam ainda mais as variações intraclasses.

Para STDB, alcançamos 95,9429% de acurácia versus 90,3%. Como seus registros de ECG foram capturados durante testes de esforço, isso representa um cenário de reconhecimento mais caótico, como pode ser observado na Figura A.2g. No entanto, a acurácia obtida é muito alta, revelando que mesmo este protocolo não é desafiador o suficiente.

Embora Zhang, Zhou e Zeng (2017) também use CNNs, nosso método é mais simples e ainda alcança uma acurácia mais alta em todos os casos. Isso sugere que deixar a CNN aprender as características discriminantes diretamente do batimento cardíaco é melhor do que usar descritores *handcrafted* como entrada para a CNN.

A.3.5 Mai, Khalil e Meli (2011)

Detectam 23 pontos ao longo do complexo QRS como características discriminativas de ECG e utilizam um Multi Layer Perception (MLP) para classificação. Em seus experimentos, eles extraíram aleatoriamente 324 complexos QRS de 18 indivíduos da NSRDB. O conjunto de treinamento contém 216 complexos QRS (12 por sujeito) e o conjunto de testes possui 108 complexos QRS (6 por sujeito). Como cada batimento cardíaco contém um único complexo QRS, usamos 324 batimentos cardíacos em nossa replicação dos experimentos. Novamente, embora essa não seja uma reprodução exata, Mai, Khalil e Meli (2011) opcionalmente descartam outras partes do batimento cardíaco que não sejam complexos QRS. Uma taxa de aprendizado inicial de 10^{-3} e um tamanho de *batch* de 8 foram definidos usando 10% dos dados de treinamento para validação. Alcançamos 93,5185% de acurácia versus 99,69% para Mai, Khalil e Meli (2011). A combinação de poucos dados de treinamento e alta variação intraclasse, como mostrado a Figura A.2b, diminuíram o desempenho da nossa abordagem (de 98,7333% ao usar 250 batimentos cardíacos por sujeito para treinamento, para 93,5185% ao usar apenas 12). Esse cenário favoreceu a abordagem de Mai, Khalil e Meli (2011), que tira proveito do conhecimento de especialistas para lidar com dados escassos. Esse resultado pode sugerir, no entanto, que o complexo QRS é mais confiável para reconhecimento do que usar todo o sinal de batimento cardíaco.

A.3.6 Ye, Coimbra e Kumar (2010)

Concatenaram *wavelets* e coeficientes de Independent Component Analysis (ICA), utilizando Principal Component Analysis (PCA) para redução de dimensionalidade, e em seguida, aplicando Support Vector Machine (SVM) para classificar um sinal de entrada como um dos possíveis sujeitos. Comparamos nosso método com seus resultados em duas bases de dados: MITDB e NSRDB. Embora eles também mostrem resultados para LTST, a versão reportada não está mais disponível e não conseguimos reproduzir esta parte dos

seus experimentos. Como eles relatam resultados para diferentes canais de ECG, sempre relatamos o melhor resultado para um único canal. Como os protocolos de Ye, Coimbra e Kumar (2010) incluem muito mais dados para treinamento, usamos 20% deles para validação. Para ambas MITDB e NSRDB, os hiperparâmetros encontrados foram uma taxa inicial de aprendizado de 10^{-3} e um tamanho de *batch* de 32.

Para a MITDB, os primeiros 5 minutos de cada registro foram usados para treinamento e o restante foi usado para teste. Atingimos 92,9102% versus 92,34% para Ye, Coimbra e Kumar (2010). Ao observar as Figuras A.2e e A.2d, podemos ver que os dados de treinamento disponíveis são muito semelhantes. No entanto, esse protocolo possui muito mais dados de teste, o que introduz mais variação intraclasse e reduz a acurácia.

Para a NSRDB, a base foi dividida em dois cenários diferentes: curto prazo e longo prazo. Para o cenário de curto prazo, trechos de 5 minutos foram extraídos a cada 20 minutos durante as primeiras duas horas de cada registro, sendo o primeiro trecho usado para treinamento e o restante para teste. Para o cenário de longo prazo, foram extraídos trechos de 5 minutos a cada 2 horas durante as primeiras 22 horas de cada registro, com o primeiro trecho sendo usado para treinamento, e o restante para teste. No cenário de curto prazo, alcançamos 96,3703% versus 88,20%. No cenário de longo prazo, alcançamos 72,9820% contra 79,94%. A Figura A.2a mostra o conjunto de treinamento para os dois cenários. Como o conjunto de treinamento é o mesmo, essa queda na acurácia de curto para longo prazo nos dois métodos indica que as variações intraclasses são muito mais altas após um longo período de tempo. Deste modo, a separação temporal pode ser tão desafiadora quanto arritmias e prática de exercícios para sistemas de reconhecimento baseados em ECG.

A.3.7 Discussão

Apesar de alguns dos protocolos apresentados serem mais desafiadores que outros, os resultados obtidos estão acima de 95% de acurácia em quase todos eles. Não acreditamos que esses resultados representem o real potencial de ECG como uma característica biométrica, principalmente porque todos eles usam o mesmo conjunto de indivíduos para treinamento e teste.

Para investigar essa questão, na subseção a seguir usamos nosso próprio protocolo, considerando a divisão de treinamento e teste sem interseção de sujeitos, e avaliação utilizando bases de dados cruzadas. No entanto, podemos concluir a partir dos experimentos prévios que nossa CNN é bem comparável aos trabalhos existentes na literatura e pode servir como base para o desempenho do reconhecimento de pessoas utilizando ECG, em um protocolo de avaliação muito mais desafiador.

A.3.8 Análise do poder de generalização do ECG

Nos experimentos anteriores, ter os mesmos sujeitos nos conjuntos de treino e teste facilita o processo de classificação e não nos fornece uma boa estimativa da real capacidade de generalização do nosso método. Para uma análise mais profunda, executamos vários experimentos sem sobreposição de sujeitos nos conjuntos de treinamento e teste, e medimos o desempenho através do Equal Error Rate (EER) e Rank-1 para identificação.

Escolhemos duas bases de dados públicas que compartilham o mesmo canal de ECG (MLII): MITDB e LTST. Descartamos os registros que não possuem o canal escolhido em ambas as bases. No total, foram 16 registros da LTST e 45 da MITDB, todos de sujeitos diferentes.

Dividimos cada base de dados de uma maneira que 50% dos sujeitos são usados para treinamento, 25 % para validação e 25 % para teste. Nenhum sujeito aparece em mais de um subconjunto. A tabela A.3 lista as *labels* de cada sujeito nesses três subconjuntos.

Base de dados	**Treinamento**	**Validação**	**Teste**
MITDB	100, 103, 106, 108, 111, 113, 115, 117, 119, 122, 124, 201, 205, 208, 210, 213, 215, 219, 221, 223, 230, 232, 234	101, 107, 112, 116, 121, 200, 207, 212, 217, 222, 231	105, 109, 114, 118, 123, 203, 209, 214, 220, 228, 233
LTST	s20071, s20091, s20101, s20121, s20131, s20141, s20221, s20231	s20061, s20111, s20201, s20211	s20011, s20051, s20081, s20241

Tabela A.3: *Labels* de cada sujeito na divisão das bases utilizadas nos experimentos.

Neste experimento, descartamos a última camada densa da CNN após o treinamento e usamos a saída CNN como um vetor descritor. A distância cosseno foi utilizada para medir a semelhança entre dois descritores de batimentos cardíacos. O treinamento e validação são similares aos utilizados na comparação com outros trabalhos. Porém, usamos a informação de EER para *early stopping*. Entretanto, como calcular a distância entre todos os pares possíveis de batimentos cardíacos seria extremamente custoso, selecionamos aleatoriamente 150 batimentos cardíacos de cada sujeito e calculamos o EER para este subconjunto. Para fins de validação, repetimos esse processo 5 vezes a cada época e utilizamos o EER médio como a medida. Para medir o desempenho no conjunto de testes, após a conclusão do treinamento, repetimos o processo acima 10 vezes e calculamos a média e o desvio padrão do EER. Além disso, calculamos o Rank-1 selecionando aleatoriamente um batimento por sujeito para compor o conjunto galeria e usamos os restantes como amostras de teste. Também calculamos a média e o desvio padrão do Rank-1. Os hiperparâmetros para a MITDB foram: taxa de aprendizado de 10^{-3} e tamanho de *batch* de 128. Já para a LTST foram 10^{-4} e 128 respectivamente.

Utilizamos quatro configurações experimentais, conforme descrito na Tabela A.4. Como alternamos qual base de dados é **A** ou **B**, temos um total de oito experimentos. No geral: no Experimento 1, treinamos e testamos na mesma base de dados; no Experimento 2, treinamos em **A** e testamos em **B**; no Experimento 3, treinamos em toda

a **A** e testamos em **B**; e no Experimento 4, treinamos em toda a **A** e testamos em toda a **B**.

	Experimento			
	1	**2**	**3**	**4**
Conjunto de Treinamento A	Treino	Treino	Treino	Treino
Conjunto de Validação A	Validação	Validação	Validação	Validação
Conjunto de Teste A	Teste	-	Treino	Treino
Conjunto de Treinamento B	-	-	-	Teste
Conjunto de Validação B	-	-	-	Teste
Conjunto de Teste B	-	Teste	Teste	Teste

Tabela A.4: Diferentes configurações para os experimentos nas duas bases de dados.

Os resultados obtidos em termos de EER são mostrados na Tabela A.5 e para Rank-1 na Tabela A.6. Ao analisar esses resultados, fica claro que a MITDB é mais desafiadora que a LTST. Os batimentos cardíacos anormais da MITDB afetam tanto o treinamento quanto o teste. Mesmo quando treinamos na MITDB, o EER para a LTST é menor que para a MITDB. Adicionar mais dados anormais para o treinamento (MITDB$^+$) não melhora os resultados para LTST. Por outro lado, usando mais dados normais para treinamento (LTST$^+$) melhora o desempenho no experimento de dados cruzados para a MITDB.

		Teste			
		MITDB	MITDB$^+$	LTST	LTST$^+$
Treinamento	MITDB	17.37%	-	14.83%	-
	MITDB$^+$	-	-	16.74%	**19.38%**
	LTST	19.12%	-	12.49%	-
	LTST$^+$	17.98%	**18.98%**	-	-

Tabela A.5: EER para os nossos experimentos. A$^+$ é a versão completa da base de dados A. O desvio padrão em todos os casos foi menor que 0.01, e por isso foi omitido nesta tabela. Valores em negrito destacam o cenário de conjunto de dados cruzados mais difícil.

No cenário mais desafiador, usamos uma base de dados inteira para treinamento e outra base inteira para teste. Estes resultados ilustram muito bem um cenário real, em que os dados para login/registro são limitados, os sujeitos são desconhecidos para o sistema e as condições de aquisição de amostras podem variar muito. Como pode ser observado, os valores obtidos para Rank-1 estão longe da precisão alcançada nos resultados das comparações com a literatura (A.3.2). A mesma CNN que superou vários trabalhos na literatura não consegue atingir taxas razoáveis de Rank-1 na maioria dos experimentos relatados na Tabela A.6. Isso confirma nossa hipótese de que os protocolos de avaliação existentes eram fáceis demais e não representavam bem o verdadeiro potencial

do ECG para fins de reconhecimento biométrico. Com esses resultados, nós acreditamos que os trabalhos atuais de ECG (incluindo este) estão longe de alcançar um desempenho aceitável, principalmente quando comparamos o ECG com outras biometrias, como faces e impressões digitais.

A.4 CONCLUSÃO

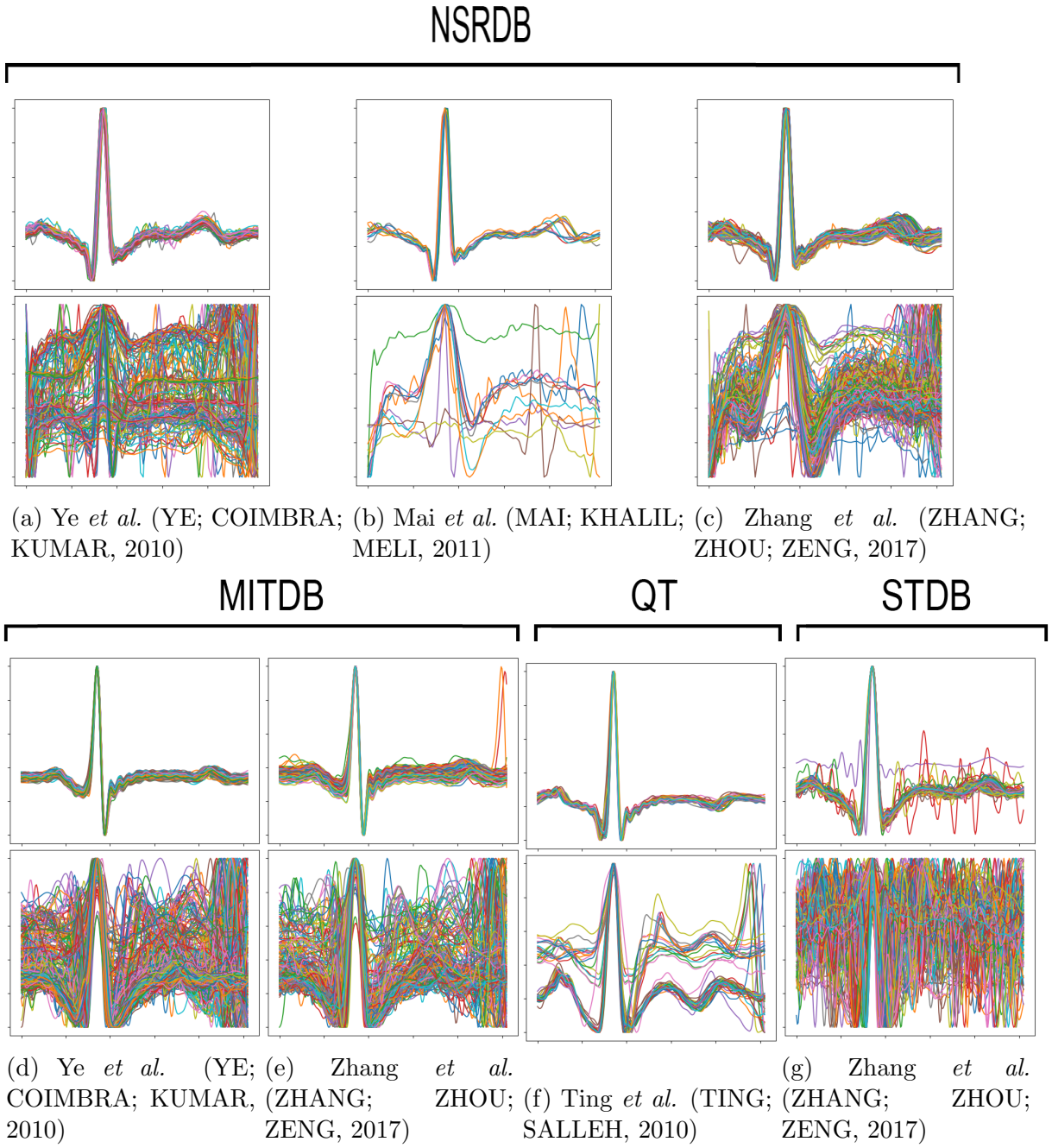
Nós apresentamos uma simples CNN para reconhecimento biométrico utilizando ECG e comparamos com diferentes trabalhos na literatura usando suas configurações experimentais. Apesar de nosso método se comparar bem a esses métodos e obter alta precisão, não acreditamos que esses resultados representassem o real potencial do ECG como uma característica biométrica, principalmente porque todos os experimentos replicados usaram o mesmo conjunto de sujeitos para treinamento e teste. Quando testamos a mesma CNN em um cenário mais difícil, os valores de EER e Rank-1 obtidos colocam o ECG muito atrás de outras biometria populares, como faces e impressões digitais. Isso mostra que as altas precisões relatadas anteriormente na literatura devem-se principalmente à facilidade dos protocolos de avaliação não condizentes com um cenário realista. Nosso trabalho indica que utilizar ECG como característica biométrica é um problema muito desafiador, mas também mostra que há muito espaço para crescer.

Propomos um protocolo de avaliação que ilustra melhor o desempenho de um método de reconhecimento de ECG usando bases de dados conhecidas e publicamente disponíveis. Os treinamentos e os testes foram realizados de maneira independente (sem interseção de sujeitos) para evitar problemas de memorização (ZHANG et al., 2016a). Os resultados para identificação foram calculados usando uma única amostra por sujeito na galeria (pior cenário). A avaliação entre bases de dados cruzadas mostrou ser um cenário mais próximo da vida real, pois adiciona variações nas condições de aquisição e um número maior de sujeitos de teste na mistura de desafios existentes (*e.g.* arritmia, exercício).

As futuras direções de pesquisa para este trabalho incluem: 1) avaliação de diferentes partes de um batimento cardíaco, especialmente o complexo QRS (MAI; KHALIL; MELI, 2011); 2) avaliar outras arquiteturas de CNN; 3) usar redes neurais recorrentes para tirar proveito do natureza temporal dos sinais de ECG; e 4) combinar batimentos cardíacos consecutivos para melhorar o desempenho sem comprometer a viabilidade.

		Teste			
		MITDB	MITDB$^+$	LTST	LTST$^+$
Treinamento	MITDB	77.63% ± 0.037	-	85.37% ± 0.035	-
	MITDB$^+$	-	-	65.03% ± 0.0432	**83.66% ± 0.0596**
	LTST	74.64% ± 0.046	-	90.64% ± 0.038	-
	LTST$^+$	65.11% ± 0.0180	**76.21% ± 0.0389**	-	-

Tabela A.6: Rank-1 para os nossos experimentos. A$^+$ é a versão completa da base de dados A. O desvio padrão em todos os casos foi menor que 0.01, e por isso foi omitido nesta tabela. Valores em negrito destacam o cenário de conjunto de dados cruzados mais difícil.



(a) Ye *et al.* (YE; COIMBRA; KUMAR, 2010)

(b) Mai *et al.* (MAI; KHALIL; MELI, 2011)

(c) Zhang *et al.* (ZHANG; ZHOU; ZENG, 2017)

(d) Ye *et al.* (YE; COIMBRA; KUMAR, 2010)

(e) Zhang *et al.* (ZHANG; ZHOU; ZENG, 2017)

(f) Ting *et al.* (TING; SALLEH, 2010)

(g) Zhang *et al.* (ZHANG; ZHOU; ZENG, 2017)

Figura A.2: Batimentos cardíacos do conjunto de treino, de bases de dados diferentes para configurações de experimentos diferentes. Cada linha colorida representa uma amostra de batimento diferente. Para cada figura acima, o grupo de batimentos cardíacos na parte superior representa um sujeito com uma pequena variação intraclasse e o grupo de batimentos cardíacos na parte inferior representa um sujeito com uma alta variação intraclasse. Melhor visualizado em cores.