Autenticação Biométrica Baseada em PPG e ECG utilizando Aprendizado Profundo

Eduardo T. Tristão, Kristtopher K. Coelho, Edelberto Franco Silva,
Alex B. Vieira, Michele Nogueira, José Augusto M. Nacif

***Abstract*—The Internet of Things popularization has significantly increased the requirements for transmitting and storing sensitive personal data. Consequently, these advances require strict access control policies with the need to guarantee security and privacy effectively. It is possible to find in the literature that biometric authentication based on PPG (photoplethysmography) or ECG (electrocardiogram) signals is potential support in meeting these requirements. In this sense, this article proposes a method of multimodal identification of individuals, combining both signals. Our proposal combines two cascaded convolutional neural networks, giving advances to state-of-the-art. As numerical results, the method achieves 99.34% accuracy, 93.83% precision, and 0.04% FAR in different databases.**

***Resumo*—A popularização da Internet das Coisas aumentou significativamente os requisitos para a transmissão e armazenamento de dados pessoais sensíveis. Consequentemente, esses avanços exigem políticas rígidas de controle de acesso com a necessidade de garantir segurança e privacidade de forma eficaz. É possível encontrar na literatura que a autenticação biométrica baseada em sinais de PPG (fotopletismografia) ou ECG (eletrocardiograma) são potenciais suportes no atendimento a esses requisitos. Pensando nisso, este artigo propõe um método de identificação multimodal de indivíduos, combinando ambos os sinais. Nossa proposta combina duas redes neurais convolucionais em cascata, dando avanços ao estado da arte. Como resultados numéricos, o método atinge 99,34% de acurácia, 93,83% de precisão e 0,04% FAR em diferentes bases de dados.**

***Index Terms*—Autenticação Biométrica, Sinais Biométricos, ECG, PPG, Aprendizado Profundo**

I. Introdução

NOS últimos anos experimentamos um aumento significativo do acesso à Internet. Fato proporcionado pelo avanço das tecnologias de redes de comunicação e a popularização de dispositivos móveis. Recentemente ainda destaca-se a popularização do uso dos dispositivos/coisas interconectadas, os quais compõem a *Internet of Things* (IoT) [1]. Com tais dispositivos cada vez mais presentes no cotidiano, tem-se o crescimento significativo do volume de dados sensíveis transmitidos desde os dispositivos IoT pessoais pelas redes IoT. Uma consequência desse cenário é, além da maior facilidade de comunicação e cobertura de sinal, uma maior superfície para que ataques cibernéticos sejam realizados, criando um cenário motivador para ações maliciosas na rede. Tais ataques podem compreender, por exemplo, a manipulação de dados durante sua troca entre os dispositivos, que podem ser capturados e/ou alterados.

Este trabalho foi apoiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/Brasil), pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/Brasil), processo 88887.509309/2020-00 e pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG). Eduardo Tristão, Kristtopher Coelho e José Augusto M. Nacif são do Instituto de Ciências Exatas e Tecnológicas da Universidade Federal de Viçosa, Brasil. E-mail: {eduardo.tristao, kristtopher.coelho, jnacif}@ufv.br. Alex B. Vieira e Edelberto Franco Silva são do Departamento de Ciência da Computação da Universidade Federal de Juiz de Fora, Brasil. E-mail: alex.borges@ufjf.edu.br e edelberto@ice.ufjf.br. Michele Nogueira é do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais, Brasil. E-mail: michele@dcc.ufmg.br.

Neste contexto, dois dos mais importantes pilares do desenvolvimento IoT são a segurança e a privacidade dos dados [2]–[4]. Considerando o cenário de IoT temos algumas possíveis categorias, tendo destaque para o contexto deste trabalho os dispositivos vestíveis, tais como: relógios, pulseiras e óculos inteligentes, por exemplo. Estes dispositivos são capazes de aferir e monitorar dados referentes a sinais vitais do usuário. Portanto, são de grande utilidade, principalmente, para aplicações inteligentes na área da saúde. Neste contexto específico, os dispositivos são considerados como pertencentes à classe de internet das coisas aplicadas à saúde (*Internet of Healthcare Things* – IoHT). Com o objetivo de incrementar a segurança neste ambiente, tem sido de interesse da academia o uso de dispositivos vestíveis voltados para a aferição de sinais vitais. Tornando esses dispositivos, uma fonte de dados para a aplicação de autenticação biométrica de usuários [5].

Dentre os principais sinais vitais considerados junto ao estado da arte em autenticação, os sinais Fotopletismografia (PPG) e Eletrocardiograma (ECG) possuem grande potencial de evolução [6]. Consequentemente, houve aumento em pesquisas relacionadas ao uso destes sinais como biometrias para a identificação de usuários [7]. Tais pesquisas demonstram que ambas biometrias são excelentes candidatas para identificação. Ademais, em sua grande maioria, estes estudos utilizaram aprendizado de máquina baseado em redes neurais buscando aumentar a segurança dos sistemas de autenticação. Como exemplo, destacam-se pesquisas as quais exploram as redes neurais convolucionais (*Convolutional Neural Networks* – CNNs) [8]–[16]. O principal objetivo é fornecer acesso seguro e confiável aos dispositivos e seus respectivos dados. Entretanto, apesar da boa $Acurácia$, alguns desses estudos não apresentam bons resultados ou omitem outras métricas, como $Precisão$ e Taxa de Falsa Aceitação (*False Acceptance Rate* – FAR).

Este artigo propõe a utilização de redes neurais convolucionais para a identificação biométrica de usuários. Ainda destaca-se a combinação dos sinais PPG e ECG (multimodal). Esses sinais foram usados em conjunto para proporcionar uma maior precisão. Isso é obtido porque mesmo que um sinal seja identificado erroneamente, a probabilidade de ambos serem incorretamente identificados é significativamente menor. Por-

tanto, foi usada uma CNN em cascata que explora convoluções unidimensionais para a identificação desses indivíduos. Além disso, foi realizada uma otimização dos parâmetros usados na rede. Os principais avanços em relação ao estado da arte compreendem o uso de uma nova rede neural multimodal para minimizar a FAR e aumentar a $Precisão$, mantendo alta $Acurácia$ em comparação com outros métodos. Todos os algoritmos e bases de dados encontram-se disponíveis publicamente[1].

O restante desse artigo é estruturado como segue: A Seção II apresenta o estado da arte relacionado ao uso de redes neurais convolucionais. Na Seção III detalha-se a estrutura das bases de dados e como os sinais são preparados como dados de entrada. Além disso, descreve a arquitetura e funcionamento do método proposto. A Seção IV apresenta os resultados obtidos, também os compara ao estado da arte. Por fim, a Seção V proporciona uma análise geral do artigo, destacando pontos importantes e indicando trabalhos futuros.

II. Trabalhos relacionados

Com o propósito de catalogar o estado da arte esta seção lista vários trabalhos com o objetivo de autenticar um usuário autorizado por meio de características biométricas. Recentemente, o uso de CNN se mostrou promissor para melhorar a precisão dos resultados. Em [8], é apresentado um método de autenticação de usuários baseado em fotopletismografia de pulso que utiliza aprendizado profundo (*deep learning* - DP) chamada de BiometricNet. O artigo usa da combinação de CNN com *Long and Short Term Memory - LSTM* para modelar a sequência de dados biométricos. Isso faz com que não seja necessário um pré-processamento de dados como normalmente ocorre na maioria dos trabalhos do estado da arte. Essa rede foi validada com sinais retirados de pessoas que executavam exercícios intensos. Os bons resultados mostrados nesse artigo indicam pesquisas futuras que se apoiam somente em redes neurais para a identificação dos indivíduos, sem a necessidade de extração de características. Essa mesma rede foi testada em um ambiente ambulatorial [9] mantendo ainda uma alta $Acurácia$, $Precisão$ e $Recall$.

Também usando PPG, temos [11] com um novo método de extração de características múltiplas. Nesse artigo, dado as dificuldades do uso de aprendizado profundo, é usado como classificador *Naive Bayes*. Adicionalmente, foi usado o método de *Sliding Window* para representar de outro modo os dados de PPG brutos. Isso faz com que não seja mais necessário um método de redução de ruídos. Em [10] foi usado também o mesmo método de representação do sinal bruto. Após isso foi usado uma extração de características com três camadas usando *sparse softmax vector* e K-ésimo Vizinho mais Próximo(KNN) como classificador. Já em [12], ainda explorando PPG como biometria, foi usado um método que integra *sparse representation learning* com aprendizado de máquina profundo em cascata. Advinda dessa junção, esse modelo proposto possui uma boa escalabilidade. Adicionalmente, por causa da complementaridade das múltiplas características usadas há uma melhora na identificação.

O Deep-ECG [17] explora o ECG como sinal para reconhecimento biométrico. Neste, os autores propõem o uso de redes CNN profundas (seis camadas) para extrair recursos que permitem realizar identificação de conjunto fechado de indivíduos, verificação de identidade e reautenticação periódica. O Deep-ECG analisa os complexosQRS (ondas Q, R e S que representam a despolarização ventricular) do ECG de entrada e produz características que geram dois modelos, o real e o binário. Estes modelos são combinados usando as distâncias euclidiana e de Hamming, respectivamente. Os testes ocorreram em conjuntos de amostras adquiridas em condições não controladas para identificação de conjunto fechado, verificação de identidade e reautenticação periódica. Em [13] é proposto um método baseado em CNN chamado Cascaded CNN para identificação humana pelo sinal de ECG. No geral, esta abordagem consiste em quatro etapas antes da fase de identificação. A primeira fase consiste na coleta, pré-processamento e geração dos modelos de diferentes indivíduos para registro. Em seguida, uma CNN para extrair características (F-CNN) é treinada usando multi classificação. Na terceira fase, uma segunda CNN (M-CNN) é treinada para correspondência com base nas características extraídas pela F-CNN. Por fim, as duas redes são incorporadas em cascata nos dispositivos para identificação humana. Devido a maior capacidade de generalização esta metodologia em cascata pode ser utilizada para vários grupos com indivíduos variáveis alcançando desempenho significativo. Por isso, em nossa abordagem também aplicamos o método em cascata, entretanto, alcançamos melhor desempenho uma vez que aplicamos a autenticação baseada em parâmetros multimodais.

Outro método de autenticação é apresentado em [14], em que fazem uso de uma CNN bidimensional. Para usar uma rede bidimensional, foi gerado um gráfico com o sinal ECG e então convertido para uma imagem binária é então direcionado à rede. Isso foi feito para poder usar convoluções e operações de *pool* bidimensionais se tornando melhor para o uso de CNN's tendo em vista que aumenta a localidade espacial. Com isso, foi obtido um resultado menos sensível a ruídos.

No caso de autenticação multimodal, temos [15] usando CNN como forma de identificação. O método propõe o uso da CNN para extração de características e a *Recurrent Neural Network* - RNN para classificador. Para a RNN, obteve uma significativa melhora nos resultados com o uso de *Gated Recurrent Unit* - GRU nas suas células.

Para os estágios de treino e teste das redes descritas foram usados conjuntos de dados publicamente disponíveis. Os autores de [8], [9], [15] usaram o banco de dados TROIKA [18], [19]. Já em [10]–[12] usaram CapnoBase [20] e BIDMC [21]. Apenas os autores de [22], [23] usaram somente o conjunto de dados BIDMC. Para avaliar a presente proposta, todas as bases de dados acima mencionadas foram consideradas. Cada um desses conjuntos de dados são descritos na seção III-A. Cada um dos conjuntos de dados, seus respectivos métodos e biometrias são sumarizados na Tabela I.

[1]https://github.com/lesc-ufv/Autentica-o-Multimodal

Tabela I: Dados e biometrias usados nos trabalhos

Trabalhos	Dados	Biometrias	Método
[8]	TROIKA	PPG	CNN e LSTM
[9]	TROIKA	PPG	CNN e LSTM
[11]	CapnoBase, BIDMC e MIMIC	PPG	Naive Bayes
[10]	CapnoBase e BIDMC	PPG	KNN
[12]	CapnoBase e BIDMC	PPG	
[13]	FANTASIA, CEBSDB, NSRDB, STDB, AFDB	ECG	CNN em Cascata
[14]	CYBHi [24] e PTB [25]	ECG	CNN bidimensional
[15]	TROIKA	ECG e PPG	
[22], [23]	BIDMC	ECG e PPG	CNN E RNN
Proposto	TROIKA, CapnoBase e BIDMC	ECG e PPG	CNN em Cascata

III. Metodologia

Nesta seção será descrita metodologia completa de como foi desenvolvida a proposta da pesquisa. Serão apresentadas as especificações das bases dados e como estes são preparados. Também, será detalhada a composição das redes neurais convolucionais e como elas interagem aplicadas em cascata. Ademais, também, são especificadas as configurações da máquina utilizada para treinamento e avaliação do método.

A. Base de dados

Nesta proposta serão considerados os bancos de dados TROIKA [18], [19], CapnoBase [20] e BIDMC [21], a fim de treinar, testar e validar o método. Todos os bancos de dados contém amostras referentes aos sinais fisiológicos PPG e ECG. Ambos os sinais foram extraídos à 125 Hz. Estas bases ainda disponibilizam outros sinais que não são considerados nesta pesquisa. A Tabela II apresenta a organização dos conjuntos de dados informando a quantidade de indivíduos distintos com sinais armazenados. Além disso, descreve a duração total de cada gravação de sinal. Por fim, indica o equipamento usado para a aferição dos sinais.

Tabela II: Organização dos dados

Características	CapnoBase	BIDMC	TROIKA
Quantidade de indivíduos	42	53	12
Duração (minutos)	8	8	5
Extração do PPG	Oxímetro de pulso	Oxímetro de pulso	Oxímetro de dedo
Extração do ECG	–	–	Sensor molhado no peito

Para realizar o treinamento da rede foram utilizados 70% do volume total dos dados. Do restante, 15% foram destinados a testes e os demais 15% aplicados para a validação do sinal de cada pessoa. Além disso, foram retirados 5 indivíduos de cada base BIDMC e Capnobase durante a fase de treino. Estes 10 indivíduos são utilizados com a finalidade de avaliar possíveis autenticações que partem de usuários não vinculados ao modelo. Em particular, o conjunto de dados TROIKA foi aplicado exclusivamente para o estágio de validação, consequentemente, não participou dos estágios de teste e treino.

1) Pré-processamento: Com intuito de padronizar ambos os sinais PPG e ECG presentes nas três bases, foi necessário introduzir um estágio de tratamento dos dados. O pré-processamento de dados produz uma lista de batimentos cardíacos para todos os indivíduos e seus respectivos rótulos numéricos que representam aquele indivíduo. A seguir a lista foi particionada aleatoriamente em dados para treino, teste e validação. Os rótulos dentro do conjunto de teste e validação variam de 0 a 106, onde cada valor representa um indivíduo. Os rótulos do conjunto de testes se pertencem ao intervalo entre 0 a 37 e 47 a 95. Os indivíduos rotulados entre 37 e 47 representam o conjunto de indivíduos não cadastrados. A padronização do formato dos sinais PPG e ECG ocorre igualmente para os três conjuntos de dados. As características relevantes selecionadas para representar os batimentos cardíacos são representados pelo complexo QRS para ECG e do pico mais alto para PPG para todos os indivíduos e as respectivas etiquetas de identificação. Cada batimento compreende 60 amostras antes e depois da onda R para o ECG e do maior pico do PPG. Consequentemente, são obtidas duas listas de batimentos cardíacos por indivíduo, compostas pelos sinais ECG e PPG respectivamente. No entanto, em cada conjunto de dados, as amostras têm amplitudes diferentes. Portanto, limitamos a variação da amplitude dos sinais ao intervalo entre 0 e 1 com base na Equação 1. Nesta equação, a variável min representa o menor valor, max é o maior e s é o valor do sinal variando no tempo. É importante ressaltar que a limitação da variação da amplitude dos sinais não altera as características da onda, apenas torna padrão a escala de apresentação do sinal. Deste modo, é sempre fornecida uma entrada padronizada para a rede neural. A detecção de picos foi implementada utilizando a biblioteca Scipy [26] da linguagem Python.

$$s'(t) = \frac{s(t) - min}{max - min} \tag{1}$$

B. Algoritmo

A proposta implementada neste trabalho é composta por duas CNN's em cascata, técnica inicialmente apresentada por [13]. A utilização das CNN's em cascata é inspirada pelo desempenho mencionado pelos autores. O método consiste em usar uma primeira CNN (CNN-1) treinada para identificar o rótulo dentro do conjunto de testes avaliado. Em seguida, a segunda CNN (CNN-2) recebe como entrada a saída da primeira. Neste ponto, os indivíduos analisados com objetivo de identificar se duas pessoas são ou não a mesma pessoa. Como resultado, a segunda CNN produz uma saída em dois neurônios. Tais neurônios são interpretados como a probabilidade dos indivíduos analisados serem ou não a mesma pessoa.

Para o uso de tal sistema, o usuário deve previamente cadastrar seus sinais PPG e ECG. Esses dados são processados pela CNN-1 e sua saída armazenada. Então, posteriormente, ao tentar acessar suas informações médicas, os sinais vitais PPG e ECG do usuário são capturados pelo dispositivos IoHT. O sistema então computa as saídas da CNN-1 desses novos sinais.

Feito isto, esses dados são colocados na CNN-2 junto com outra saída da CNN-1, armazenada previamente para aquele usuário. Por fim, a CNN-2 retorna a probabilidade de ser e de não ser sinais do mesmo indivíduo. Consequentemente, autorizando o acesso ao indivíduo, caso a probabilidade de ser um usuário autêntico for maior para ambos sinais ECG e PPG, ou negando, em caso contrário.

As CNN's no presente artigo são unidimensionais, portanto, para camada convolucional seus filtros terão apenas uma dimensão e para as camadas *Max-pooling* as operações de *pool* ocorreram em uma dimensão. Para cada camada convolucional e camada densa, utiliza-se uma função de ativação ReLU. Para cada iteração do caminho de dados com um *Max-pooling* o tamanho da entrada é reduzido pela metade. Para as camadas convolucionais, tal tamanho é reduzido em dois ou três, visto que há ausência de *padding*. A implementação das CNN's ocorreu com auxílio da biblioteca TensorFlow [27] juntamente com a biblioteca Keras [28], ambas para linguagem de programação Python. De modo complementar, também foram utilizadas outras bibliotecas, como Pandas [29], Numpy [30], Scipy [26] e Collections a fim de facilitar o processo de construção do algoritmo.

Com o intuito de atingir uma maior exatidão na identificação de usuários, foram realizadas avaliações com diferentes parâmetros. Entre elas destacam-se o tamanho do filtro *Kernel* para cada camada, quantidade de filtros, quantidade de camadas, quantidade de neurônios na camada totalmente conectada no fim da rede, quantidade de camadas de *max-pooling* e tamanho do *pool* usado em cada uma delas. Então, a arquitetura da rede neural proposta foi derivada da melhor configuração resultante. A seguir, serão detalhadas as melhores configurações dos parâmetros para cada uma das CNN's.

1) CNN-1: Nessa seção é descrita a primeira rede neural do nosso modelo em cascata, a CNN-1. A entrada consiste nos dados pré-processados, como descrito na Seção III-A1. A camada de entrada é composta por 120 neurônios dos quais recebem a amostra referente a um batimento cardíaco. Esta rede produz a saída referente ao rótulo daquele indivíduo. Estes rótulos variam de 1 a 95. Destes, dez etiquetas não foram treinadas. Além deles, também são considerados os usuários do TROIKA, os quais não possuem uma possível etiqueta de saída, uma vez que, as etiquetas de seus respectivos indivíduos variam de 96 a 107.

Apesar de cada neurônio de saída desta rede neural representar um indivíduo do conjunto de treino, não necessariamente a sua entrada precisa ser do conjunto teste. Para demonstrar isso, 22 indivíduos foram retirados dessa etapa de treino. Portanto, não possuem suas etiquetas treinadas ou mesmo, sequer foram adicionadas à rede neural.

A arquitetura da CNN-1 é ilustrada na Figura 1. As camadas convolucionais são denominadas Conv, as camadas de *pool* máximo *Max-pooling*. Também foi aplicado o otimizador Adam [28] com uma taxa de aprendizado de 0.001 e uma quantidade de épocas igual a 30. Os tamanhos de cada camada, a quantidade de filtros e respectivos tamanhos da saída são sumarizados na Tabela III para os sinais de PPG e ECG. Os valores diferem conforme o sinal avaliado devido a característica de cada camada. Em virtude da natureza das redes neurais totalmente conectadas, a entrada de dados dessa camada precisa ser formatada em apenas um vetor. Portanto, entre o Max-Pooling-3 e a rede inteiramente conectada, a saída é transformada do formato de matriz (10x64 para ECG e 10x32 para PPG) para um vetor contíguo com 640 e 320 valores respectivamente. Posteriormente este vetor é conectado à camada seguinte.

2) CNN-2: Com o objetivo de identificar se dois batimentos que passaram pela CNN-1 são ou não do mesmo indivíduo sem a necessidade de ter participado do seu estágio de treino desenvolvemos a CNN-2. Então, seguindo a lógica em cascata, a entrada da CNN-2 recebe como entrada os valores de saída da CNN-1 de ambos indivíduos avaliados e então retorna nos dois neurônios de saída valores que são interpretados como a probabilidade de ser ou não a mesma pessoa. A rede neural decide se são ou não a mesma pessoa de acordo com o maior valor entre os dois neurônios de saída. Logo após, é feito um E lógico entre o resultado para PPG e para ECG.

Sua arquitetura se compõe de modo similar a CNN-1 e é ilustrada na Figura 2, tendo um *pool* máximo e uma camada convolucional a menos. Foi usado filtros de tamanho 2 para as camadas convolucionais e para o *pool* das camadas de *Max-pooling*. A rede inteiramente conectada ao final possui 128 neurônios, a camada de entrada possui 2x95(saída de dois indivíduos da CNN-1) neurônios e a camada de saída possui dois neurônios. A saída de cada camada da CNN-2 é detalhada na Tabela IV. De modo similar ao que acontece com a CNN-2, na transição de uma CNN para uma rede neural inteiramente conectada, é necessário transformar a saída da CNN em formato matriz(64x21 para ECG e 32x21 para PPG) em um vetor contíguo com 1344 e 672 valores respectivamente. Durante o estágio de treino, foi utilizado o otimizador Adam, 30 épocas e uma taxa de aprendizado de 0.0001. Para tal, seria gerado um grande volume de dados ao usar todos os resultados do conjunto de treino da CNN-1. Então, para o treinamento dessa rede o conjunto foi dividido em seções contendo cinco batimentos totalizando em 41 para PPG e 60 para ECG por indivíduo.

C. Hardwares e Métricas

Os experimentos foram conduzidos utilizando uma máquina com 64 GB de memória RAM, processador Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz. A máquina ainda contém uma GPU Tesla K40 com 12 GB de memória GDDR5, o que permite melhorar a eficiência do treinamento do modelo. Para avaliar o desempenho da proposta consideramos as métricas referente as taxas de falsa aceitação (FAR – *False Accept Rate*) e falsa rejeição (FRR – *False Reject Rate*), $Acurácia$, $Precisão$ e $Recall$, sendo as mais usadas na literatura. As definições formais das métricas são dadas pelas equações 2 a 6

$$FAR = \frac{FP}{FP + TN} \tag{2}$$

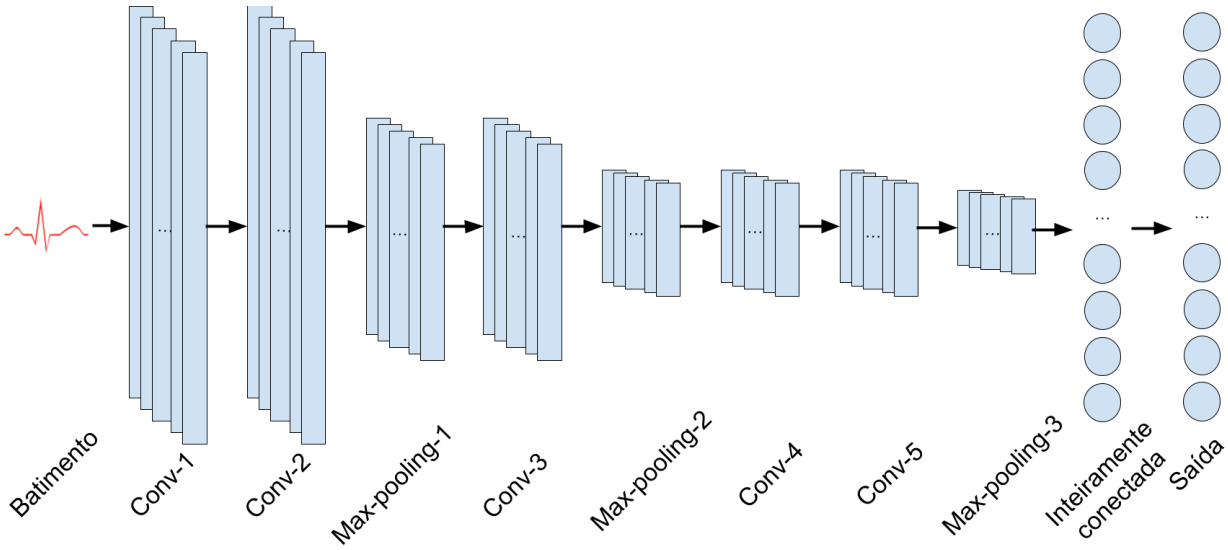

Figura 1: Representação da arquitetura da CNN-1

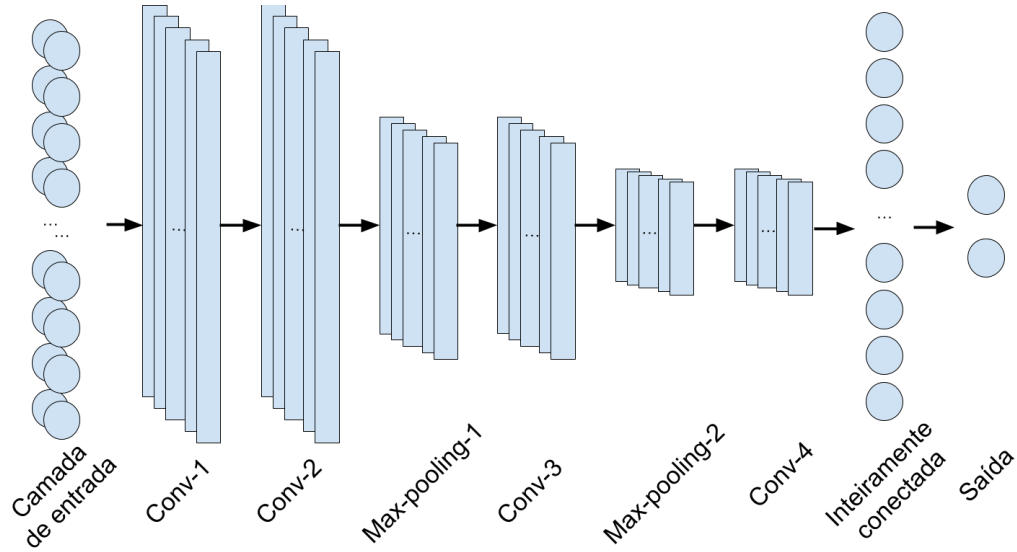

Figura 2: Representação da arquitetura da CNN-2

Tabela III: Detalhes dos valores de cada camada da CNN-1.

Camada	Tamanho		Quantidade de filtros		Tamanho da saída	
	PPG	**ECG**	**PPG**	**ECG**	**PPG**	**ECG**
Entrada	120	120	-	-	120	120
Conv-1	4	4	32	64	32x117	64x117
Conv-2	4	4	32	64	32x114	64x114
Max-pooling-1	2	2	-	-	32x57	64x57
Conv-3	3	3	32	64	32x55	64x55
Max-pooling-2	2	2	-	-	32x27	64x27
Conv-4	4	4	32	64	32x24	64x24
Conv-5	4	4	32	64	32x21	64x21
Max-pooling-3	2	2	-	-	32x10	64x10
Inteiramente conectada	128	256	-	-	128	256
Saída	95	95	-	-	95	95

Tabela IV: Saídas de cada camada da CNN-2

Camada	Tamanho da saída	
	ECG	PPG
Entrada	2x95	2x95
Conv-1	64x94	32x94
Conv-2	64x93	32x93
Max-pooling-1	64x46	32x46
Conv-3	64x45	32x45
Max-pooling-2	64x22	32x22
Conv-4	64x21	32x21
Inteiramente conectada	128	128
Saída	2	2

$$FRR = \frac{FN}{FN + TP} \tag{3}$$

$$Precisão = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$Acurácia = \frac{FP + FN}{N} \tag{6}$$

As variáveis FP, TP, FN e TN representam o total de falsos positivos, verdadeiros positivos, falsos negativos e verdadeiros negativos respectivamente e N determina o número total de amostras.

IV. Resultados e Discussão

O desempenho do método de autenticação apresentado neste artigo foi comparado com os resultados apresentados pelos demais trabalhos que representam o estado da arte. As métricas $Acurácia$, $Precisão$ e $Recall$ são sumarizados nas Tabelas V e VI separadas pelos conjuntos de dados usados.

Ao examinar os valores de referência citados, é possível observar que a técnica proposta por este trabalho obteve $Acurácia$ superior à maioria, utilizando as mesmas bases de dados avaliadas de modo isolado, seguindo o estado da arte [8]–[12], [15]. É importante ressaltar que o método proposto ainda obteve uma $Acurácia$ de 99,62% ao utilizar os sinais presentes em todas as três bases de dados. Ou seja, treinando e validando o modelo em dados de diferentes bases em conjunto. Portanto, esta foi a acurácia geral obtida pelo modelo proposto. Consequentemente, utilizar todos os conjuntos de dados proporcionou uma acurácia superior quando comparado aos conjuntos de dados isolados. Outros trabalhos [13], [14], usaram bases de dados as quais possuíam somente um dos dois sinais PPG ou ECG. Por este motivo não puderam ser comparados.

Ao considerar os demais trabalhos, ressalta-se que a $Acurácia$, apesar de uma importante métrica para verificar a validade do método, ela sozinha pode não ser a mais relevante, uma vez que representa a quantidade de acertos pelo total de predições. Portanto, existindo uma quantidade de rejeições significativamente maior que o número de aceitações, pode-se obter uma $Acurácia$ alta mesmo que não tão segura ou precisa.

Tabela V: Precisão, Acurácia e Recall do estado da arte usando o TROIKA

Trabalho	$Precisão$ (%)	$Recall$ (%)	$Acurácia$ (%)
[8]	89	84	96
[9]	67	86	96
[15]	94	93	94
Proposto	93,83	35,26	94,12

Tabela VI: Acurácia do estado da arte usando os conjuntos de dados CapnoBase e BIDMC

Trabalho	$Acurácia$ (%)	
	CapnoBase	BIDMC
[11]	98,65	97,76
[10]	99,92	99,95
[12]	99,88	99,12
Proposto	99,34	99,12

De modo contrário, a $Precisão$ e a FAR, representam uma medida mais eficiente do ponto de vista da segurança. Dado que a $Precisão$ representa, dentre aqueles que foram julgados como verdadeiros, quantos deveriam ser de fato preditos como tal. A FAR representa dentre aqueles que deveriam ser falsos, quais foram preditos erroneamente. Embora muito relevantes para a segurança do modelo, muitos dos estudos apresentados não exibem tais métricas. A proposta deste artigo, avaliada sob todos os conjuntos de dados, apresenta uma $Precisão$ de 93,85%. Este valor é superior aos informados pelos demais trabalhos que compartilham tal métrica, exceto em [15]. Portanto, este indício sugere que redes neurais submetidas a mais de uma biometria podem apresentar maior segurança.

Ao avaliar a métrica FAR, o método proposto também se destaca. Ele apresenta uma taxa de falsa aceitação equivalente à 0,042%. Isso significa que aproximadamente 4 a cada 10.000, dentre aqueles inválidos, conseguem penetrar a segurança do modelo apresentado. Em caráter comparativo, o único trabalho que apresentou tal métrica é [14]. Neste, o valor da FAR representa 0,81%, ou seja, mais de $20X$ maior. No que diz respeito a FRR e o $Recall$, o modelo proposto não obteve destaque. Entretanto, essas métricas são menos significativas no quesito segurança. Elas ilustram o quão bem o modelo é capaz de julgar corretamente aqueles indivíduos dos quais deveriam ser de fato aprovadas.

No método proposto quando o usuário válido tenta se autenticar, algumas vezes seria considerado como inválido devido ao baixo $Recall$ e alto FRR, mas ele poderá repetir a tentativa de autenticação. Para o caso de usuários inválidos o oposto acontece. Como as métricas FAR e $Precision$ indicam, um usuário que não pode ter acesso ao sistema terá seu acesso provavelmente negado, mesmo após múltiplas tentativas. Ao utilizar dois sinais é muito improvável que ele consiga acesso tornando o sistema mais robusto a tentativas de acessos não autorizados. Esta proposta garante que os usuários que não devem ter acesso sejam negados. Portanto, com o uso desse método, há uma maior garantia de manter informações sensíveis desses dispositivos IoHT seguras. Esses resultados demonstram que a rede proposta possui uma segurança superior às apresentadas no estado da arte.

V. Conclusão

Em suma, o artigo trouxe avanços na segurança dos dispositivos IoT por meio de CNNs adaptadas para a autenticação multimodal de sinais PPG e ECG. A validação do método foi feita em três diferentes conjuntos de dados, TROIKA, CapnoBase e BIDMC. Foi alcançada uma $Acurácia$ geral alta de 99,62% em relação à outras pesquisas do estado da arte. Ainda mais, mostra outros bons valores como um baixo FAR 0,04% e uma alta $Precisão$ de 93,85%. Isso mostra que a rede é mais robusta que os trabalhos relacionados a ataques de invasores desconhecidos, indicando que autenticações multimodais podem ser mais seguras. Em trabalhos futuros, pretende-se explorar a combinação entre PPG e/ou ECG e outros sinais com características biométricas como por exemplo o eletroencefalograma (EEG).

Referências

[1] S. Singh and N. Singh, "Internet of things (iot): Security challenges, business opportunities amp; reference architecture for e-commerce," in *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)*, 2015, pp. 1577–1581.

[2] M. M. Dhanvijay and S. C. Patil, "Internet of things: A survey of enabling technologies in healthcare and its applications," *Computer Networks*, vol. 153, pp. 113–131, 2019.

[3] K. K. Coelho, M. Nogueira, M. C. Marim, E. F. Silva, A. B. Vieira, and J. A. M. Nacif, "Lorena: Low memory symmetric-key generation method for based on group cryptography protocol applied to the internet of healthcare things," *IEEE Access*, vol. 10, pp. 12 564–12 579, 2022.

[4] K. Coelho, D. Damião, G. Noubir, A. Borges, M. Nogueira, and J. Nacif, "Cryptographic algorithms in wearable communications: An empirical analysis," *IEEE Communications Letters*, vol. 23, no. 11, pp. 1931–1934, 2019.

[5] W. Yang, S. Wang, N. M. Sahri, N. M. Karie, M. Ahmed, and C. Valli, "Biometrics for internet-of-things security: A review," *Sensors*, vol. 21, no. 18, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/18/6163

[6] E. Cerqueira, P. Resque, I. Medeiros, L. Bastos, A. Santos, T. Tavares, D. Rosário, A. Santos, and M. Nogueira, "Autenticação usando sinais biométricos: Fundamentos, aplicações e desafios," *Sociedade Brasileira de Computação*, 2019.

[7] A. S. Rathore, Z. Li, W. Zhu, Z. Jin, and W. Xu, "A survey on heart biometrics," *ACM Comput. Surv.*, vol. 53, no. 6, dec 2020.

[8] L. Everson, D. Biswas, M. Panwar, D. Rodopoulos, A. Acharyya, C. H. Kim, C. Van Hoof, M. Konijnenburg, and N. Van Helleputte, "Biometricnet: Deep learning based biometric identification using wrist-worn ppg," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018, pp. 1–5.

[9] D. Biswas, L. Everson, M. Liu, M. Panwar, B.-E. Verhoef, S. Patki, C. H. Kim, A. Acharyya, C. Van Hoof, M. Konijnenburg, and N. Van Helleputte, "Cornet: Deep learning framework for ppg-based heart rate estimation and biometric identification in ambulant environment," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 2, pp. 282–291, 2019.

[10] J. Yang, Y. Huang, F. Huang, and G. Yang, "Photoplethysmography biometric recognition model based on sparse softmax vector and k-nearest neighbor," *Journal of Electrical and Computer Engineering*, vol. 2020, p. 9653470, Oct 2020.

[11] J. Yang, Y. Huang, R. Zhang, F. Huang, Q. Meng, and S. Feng, "Study on ppg biometric recognition based on multifeature extraction and naive bayes classifier," *Scientific Programming*, vol. 2021, p. 5597624, May 2021.

[12] C. Liu, Y. Huang, F. Huang, and J. Yu, "Multifeature deep cascaded learning for ppg biometric recognition," *Scientific Programming*, vol. 2022, p. 7477746, Mar 2022.

[13] Y. Li, Y. Pang, K. Wang, and X. Li, "Toward improving ecg biometric identification using cascaded convolutional neural networks," *Neurocomputing*, vol. 391, pp. 83–95, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231220300485

[14] M. Hammad, S. Zhang, and K. Wang, "A novel two-dimensional ecg feature extraction and classification algorithm based on convolution neural network for human authentication," *Future Generation Computer Systems*, vol. 101, pp. 180–196, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X18329923

[15] C. Yaacoubi, R. Besrour, and Z. Lachiri, "A multimodal biometric identification system based on ecg and ppg signals," in *Proceedings of the 2nd International Conference on Digital Tools & Uses Congress*, ser. DTUC '20. New York, NY, USA: Association for Computing Machinery, 2020. [Online]. Available: https://doi.org/10.1145/3423603.3424053

[16] S. Bianco and P. Napoletano, "Biometric recognition using multimodal physiological signals," *IEEE Access*, vol. 7, pp. 83 581–83 588, 2019.

[17] R. Donida Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, "Deep-ecg: Convolutional neural networks for ecg biometric recognition," *Pattern Recognition Letters*, vol. 126, pp. 78–85, 2019, robustness, Security and Regulation Aspects in Current Biometric Systems. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865518301077

[18] Z. Zhang, Z. Pi, and B. Liu, "TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 522–531, feb 2015. [Online]. Available: https://doi.org/10.1109%2Ftbme.2014.2359372

[19] B. L. Zhilin Zhang, Zhouyue Pi, "Ieee signal processing cup 2015: Heart rate monitoring during physical exercise using wrist-type photoplethysmographic (ppg) signals," https://web.archive.org/web/20201009233959/https://sites.google.com/site/researchbyzhang/ieeespcup2015, august 2014, acessado em 05/2022.

[20] W. Karlen, "CapnoBase IEEE TBME Respiratory Rate Benchmark," 2021. [Online]. Available: https://doi.org/10.5683/SP2/NLB8IT

[21] M. A. F. Pimentel, A. E. W. Johnson, P. H. Charlton, D. Birrenkott, P. J. Watkinson, L. Tarassenko, and D. A. Clifton, "Toward a robust estimation of respiratory rate from pulse oximeters," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1914–1923, 2017.

[22] L. Bastos, T. Tavares, D. Rosário, E. Cerqueira, A. Santos, and M. Nogueira, "Double authentication model based on ppg and ecg signals," in *2020 International Wireless Communications and Mobile Computing (IWCMC)*, 2020, pp. 601–606.

[23] L. D. L. BASTOS, "Sistema de identificacao duplo (sid) de usuarios atraves dos biosinais fotopletismograma e eletrocardiograma," Master's thesis, Universidade Federal do Pará, 2020.

[24] H. P. da Silva, A. Lourenço, A. Fred, N. Raposo, and M. A. de Sousa, "Check your biosignals here: A new dataset for off-the-person ecg biometrics," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 2, pp. 503–514, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260713003891

[25] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das internet," pp. 317–318, Jan. 1995.

[26] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[27] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et al.*, "{TensorFlow}: a system for {Large-Scale} machine learning," in *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, 2016, pp. 265–283.

[28] F. Chollet *et al.*, "Keras," 2015. [Online]. Available: https://github.com/fchollet/keras

[29] W. McKinney *et al.*, "Data structures for statistical computing in python," in *Proceedings of the 9th Python in Science Conference*, vol. 445. Austin, TX, 2010, pp. 51–56.

[30] C. R. Harris, K. J. Millman, S. J. Van Der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith *et al.*, "Array programming with numpy," *Nature*, vol. 585, no. 7825, pp. 357–362, 2020.