

Combinações de redes neurais e discriminantes lineares para classificação de arritmias cardíacas

Natália Nagata, Renato Candido e Magno T. M. Silva

Resumo— Neste trabalho, utilizam-se redes neurais *perceptron* multicamada, redes neurais recorrentes, análise de discriminantes lineares e combinações desses métodos para classificação automática de arritmias cardíacas. A fim de se obter resultados clinicamente realistas para o diagnóstico, dados de um mesmo paciente não foram usados simultaneamente nas fases de treinamento e de teste. Os resultados indicam que a combinação da rede *perceptron* multicamada com a análise de discriminantes lineares apresenta um desempenho melhor em relação aos modelos individuais. Além disso, esse esquema combinado alcança métricas de classificação superiores às da literatura para as classes de arritmia de batimentos supraventriculares ectópicos (S) e ventriculares ectópicos (V).

Palavras-Chave— Aprendizado de máquina, redes neurais, análise de discriminante linear, eletrocardiograma, arritmias cardíacas.

Abstract— In this work, we use multilayer neural networks (MLP), recurrent neural networks, linear discriminant analysis (LDA) and combinations of these methods for automatic classification of cardiac arrhythmias. In order to obtain clinically realistic results for the diagnosis, patients records used during the training phase were not used in the test set. The results indicate that the combination of MLP with LDA presents a better performance compared to those of individual models. Furthermore, this combined scheme achieves classification metrics superior to those reported in the literature for the supraventricular ectopic beats (S) and the ventricular ectopic beats (V) arrhythmia classes.

Keywords— Machine learning, neural networks, linear discriminant analysis, electrocardiogram, cardiac arrhythmias.

I. INTRODUÇÃO

O eletrocardiograma (ECG) registra a atividade elétrica do coração por meio de um arranjo de eletrodos que captam os potenciais superficiais do corpo ao longo do tempo. É um teste efetivo de baixo custo e não invasivo que se tornou uma ferramenta padrão na identificação de doenças do coração [1]. Segundo a Organização Mundial da Saúde (OMS), as doenças cardiovasculares são a principal causa de morte no mundo [2]. Dentre essas doenças, as arritmias cardíacas são as mais comuns e a sua classificação precisa é de grande interesse em estudos biomédicos [2]. As arritmias correspondem a qualquer distúrbio na taxa, na regularidade e nos locais de origem ou condução dos impulsos elétricos cardíacos. Um importante passo no seu diagnóstico é a classificação de batimentos consecutivos do sinal de ECG [3].

A análise manual desses batimentos demanda muito tempo do cardiologista. Em alguns casos, é necessário gravar o sinal

Natália Nagata, Renato Candido, Magno T. M. Silva, Depto. de Engenharia de Sistemas Eletrônicos, Escola Politécnica, Universidade de São Paulo, São Paulo, SP, Brasil, e-mails: nagata.natalia@usp.br; renatocan@lps.usp.br; magno.silva@usp.br. Este trabalho foi financiado pela FAPESP (2019/26911-6 e 2021/02063-6) e CAPES (código de financiamento 001).

de ECG por até uma semana para identificar determinadas arritmias, possibilitando a perda de informações importantes e dificultando resultados imediatos [4], [5]. Uma alternativa é o uso de métodos computadorizados para a identificação automática das arritmias. A partir da década de 1990, surgiram vários trabalhos em que soluções baseadas em redes neurais *perceptron* multicamada (*multilayer perceptron* – MLP) [6], análise de discriminantes lineares (*linear discriminant analysis* – LDA) [7] e máquinas de vetores de suporte (*support vector machines* – SVM) [8] foram propostas para esse problema.

As altas taxas de erro desses métodos em conjunto com o crescimento da área de aprendizado de máquina [9] fizeram com que a pesquisa em automatização do diagnóstico de arritmias cardíacas atraísse novamente a atenção da comunidade científica. Trabalhos recentes consideram o uso de redes neurais convolucionais (*convolutional neural network* – CNN) [10] e redes neurais recorrentes (*recurrent neural networks* – RNN) [11] na classificação de arritmias. Porém, em vários desses trabalhos, os autores não seguem as recomendações da *Association for the Advancement of Medical Instrumentation* (AAMI) [12] e utilizam dados dos mesmos pacientes tanto no conjunto de teste, quanto no de treinamento. Com exceção de métodos direcionados para o monitoramento de pacientes específicos (*patient-specific*) [13], essa abordagem não é clinicamente realista, já que na prática o sistema será utilizado em pacientes, cujos dados não foram usados no treinamento.

Dentre os classificadores que seguem uma divisão realista dos pacientes, os trabalhos descritos resumidamente a seguir são considerados como estado da arte na literatura. Em [4], desenvolveu-se uma LDA para classificar os batimentos em um problema de cinco classes, a partir da extração características descrita aqui na Subseção IV-C. Em [14], características temporais, morfológicas e estatísticas do sinal foram utilizadas junto a um algoritmo de *sequential forward floating search* para encontrar combinações de características ótimas, com classificadores baseados em LDA e MLP. Em [15], foi proposta uma LDA com pesos e extração de características de intervalos RR e características morfológicas usando transformada de wavelet. Em [16], o ECG foi representado em três dimensões por meio do *temporal vectorcardiogram* (TVCG). Essa representação foi usada em redes complexas para extração de características, que, por sua vez, foram consideradas como entrada de um classificador SVM. Por fim, [17] aprimorou o modelo de [16] com o uso de *particle swarm optimization* para seleção de características.

Como a divisão de pacientes em treinamento e teste nem sempre é considerada, torna-se difícil avaliar e comparar os métodos de classificação de arritmias cardíacas contidos na

literatura. Além disso, a combinação de classificadores foi pouco explorada para o diagnóstico automático de arritmias seguindo as recomendações da AAMI [18].

Neste artigo, propõe-se o uso de redes neurais dos tipos MLP e RNN, do método estatístico da LDA e de combinações desses classificadores para o diagnóstico automático de arritmias. Foram considerados sinais de ECG do banco de dados *MIT-BIH Arrhythmia Database* (MITDB) [19], [20], levando-se em conta a separação dos pacientes nas fases de treinamento e teste. Dessa forma, foi possível comparar o desempenho dos diferentes classificadores propostos em termos de métricas de classificação, avaliar o efeito da combinação desses classificadores e comparar os resultados obtidos com os da literatura.

O artigo está organizado da seguinte forma. Na Seção II, descreve-se o banco de dados utilizado e as classes de arritmia consideradas. Em seguida, apresentam-se as etapas de reconhecimento do sinal de ECG anteriores à classificação. Na Seção IV, as estruturas dos classificadores são descritas em maiores detalhes. As Seções V e VI contêm respectivamente os resultados de simulação e as principais conclusões do trabalho.

II. BANCO DE DADOS

O MITDB é composto por 48 gravações ambulatoriais de duas derivações de 30 minutos de pacientes do *Boston's Beth Israel Hospital*, com anotações manuais de cada batimento feitas por cardiologistas. As derivações presentes na maioria das gravações são a bipolar II modificada (*lead A*) e a precordial V1 (*lead B*). Neste trabalho, seguiu-se a recomendação da AAMI e dividiu-se os conjuntos de treinamento e de teste conforme a separação proposta por De Chazal *et al.* [4]. A AAMI também recomenda o agrupamento das anotações em cinco classes: batimentos do nó sinoatrial (N), supraventriculares ectópicos (S), ventriculares ectópicos (V), fusão de batimentos normais e ventriculares ectópicos (F) e desconhecidos ou de marca-passo (Q). No entanto, devido à ausência de resultados promissores da classe Q, tanto na literatura quanto neste trabalho, essa classificação não foi considerada. Além desse agrupamento em um problema de quatro classes, muitos autores otimizam um problema de três classes, classificando N, S e V, e removendo a classe F devido à menor quantidade de dados [15]–[17], [21].

Para lidar com o desbalanceamento do número de dados entre as classes, os dados da classe N foram subamostrados e as funções custo foram ponderadas de forma proporcional ao inverso da quantidade de dados de cada classe. A Tabela I mostra a quantidade total disponível de dados e a quantidade utilizada no trabalho.

TABELA I: Número total de batimentos e número utilizado para os conjuntos de teste e de treinamento.

Classe	Número total de batimentos		Número utilizado	
	Conjunto de treinamento	Conjunto de teste	Conjunto de treinamento	Conjunto de teste
N	40098	40052	8492	8359
S	755	1319	755	1319
V	2538	2034	2538	2034
F	396	376	396	376

III. ETAPAS DE RECONHECIMENTO DO SINAL

Em geral, o diagnóstico a partir do sinal de ECG por técnicas computadorizadas é dividido em quatro etapas: o

pré-processamento do sinal, a segmentação, a extração das características e a classificação. A etapa de pré-processamento reduz os ruídos e artefatos provenientes de várias origens, como a interferência da rede elétrica. No entanto, essa etapa não foi considerada neste trabalho. A etapa de segmentação delimita o complexo QRS e as ondas P e T. Para isso, usou-se o método proposto em [22], baseado em transformada de wavelet e implementado em MATLAB pelo pacote *ecg-kit* [23]. Já a etapa de extração de características determina a menor quantidade de características do sinal de ECG que permite taxas de classificação aceitáveis [5]. A extração de características usada neste trabalho está descrita na Subseção IV-C. A etapa de classificação é explorada na seção seguinte.

IV. CLASSIFICADORES

Na etapa de classificação, foram considerados três classificadores: uma rede MLP, uma RNN e uma LDA. Cada método foi implementado individualmente e, posteriormente, suas saídas foram combinadas. As redes foram implementadas utilizando as bibliotecas *Tensorflow* e *Keras* [24] e a LDA, utilizando a biblioteca *scikit-learn* [25]. Na sequência, abordam-se os parâmetros e a entrada da rede MLP considerada. Detalha-se também o bloco da RNN utilizada e os seus parâmetros. Por fim, descreve-se o método de balanceamento da LDA por pesos e a extração de características usada nesse método. O número de camadas, o número de épocas e os hiperparâmetros das redes foram ajustados usando *grid search* a fim de se obter o melhor desempenho.

A. Rede neural MLP

Inspirando-se nos resultados de [26], para classificar um determinado batimento, foram utilizados também o batimento anterior e o batimento posterior. Considerou-se que cada batimento compreendia 320 amostras, totalizando 960 amostras da primeira derivação do sinal como entrada da rede.

A rede MLP utilizada é composta por duas camadas ocultas, a primeira com 32 neurônios e a segunda com 16, e funções de ativação ReLU em ambas. Na camada de saída, foram utilizados 4 neurônios com função *Softmax*. Para trabalhar com as classes desbalanceadas, usou-se a função custo de entropia cruzada categórica com pesos calculados na proporção inversa do número de dados de cada classe. Essa função é dada por

$$J_{WCE} = -\frac{1}{N} \sum_{n=1}^N \sum_{k=1}^K p_k \times y_k^{(L)}(n) \log(d_k(n)), \quad (1)$$

em que N é o número de exemplos de treinamento, K é o número de classes, p_k é o peso da classe k , $d_k(n)$ é o valor real do n -ésimo exemplo de treinamento e $y_k^{(L)}(n)$ é a saída da rede para o n -ésimo exemplo. Considerou-se ainda o algoritmo de otimização Adam [27] com $\beta_1 = 0,9$ e $\beta_2 = 0,99$, e o algoritmo de retropropagação (*backpropagation*) com passo de aprendizado $\eta = 0,001$, 250 épocas, mini-batches de tamanho $k = 2048$ e inicialização de Xavier [28] para os pesos.

B. Rede neural recorrente

A rede RNN utilizada possui uma camada com três blocos de LSTM (*Long Short-Term Memory*), o que corresponde a três passos de tempo. Como na MLP, usou-se uma entrada de

960 amostras da primeira derivação do sinal de ECG, com 320 amostras de entrada para cada passo de tempo, como ilustrado na Figura 1. A LSTM foi originalmente proposta em [29] para resolver o problema de desvanecimento do gradiente. Ela possui portas que controlam o fluxo de informação e permitem o aprendizado de dependências de longo prazo.

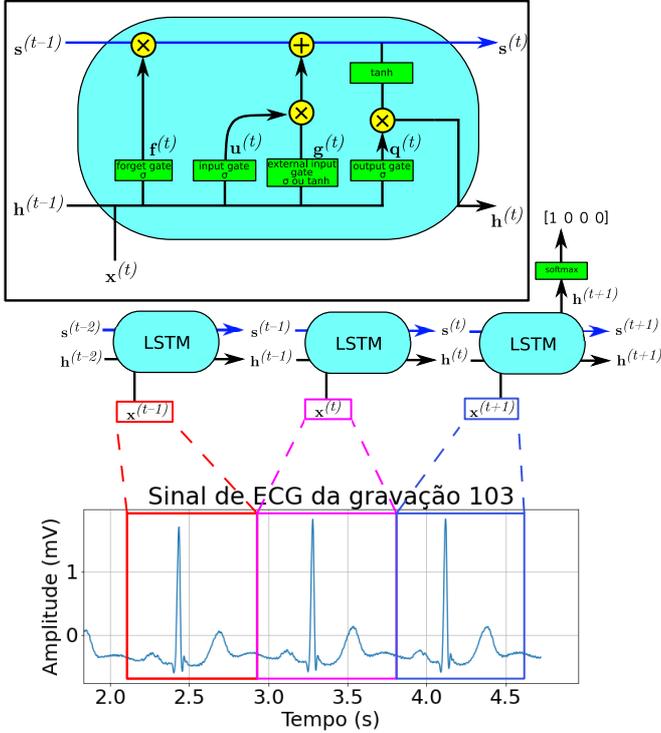


Fig. 1: Rede RNN utilizada com três passos de tempos.

O estado interno da célula LSTM (*cell internal state*) $s_i^{(t)}$ é representado pela linha horizontal azul no topo do bloco LSTM da Figura 1. Essa unidade é controlada pela porta *forget gate* $f_i^{(t)}$, para cada passo de tempo t e unidade i da camada oculta, sendo o número total de unidades igual a 72 na rede considerada. A *forget gate* usa uma função de ativação sigmoideal tendo como saída um valor entre 0 e 1 que controla o fluxo de informação do estado $s_i^{(t-1)}$ para o estado $s_i^{(t)}$.

A entrada $x_i^{(t)}$ de um bloco LSTM fornece a saída $u_i^{(t)}$ que pode ser “acumulada” no estado se a porta de entrada externa (*external input gate*) $g_i^{(t)}$ permitir. A porta de saída (*output gate*) $q_i^{(t)}$ controla a ocorrência ou não de saída na célula. O estado $s_i^{(t)}$ é atualizado a cada passo de tempo, e, após passar por uma tangente hiperbólica e ser multiplicado pela porta de saída, fornece o vetor de estado oculto $h_i^{(t)}$ da célula. No último passo de tempo, esse vetor é conectado a uma camada com 4 neurônios de saída e função *Softmax*.

Na RNN considerada neste trabalho, usou-se a função custo de entropia cruzada categórica com pesos da Equação (1), o algoritmo de otimização Adam com $\beta_1 = 0,9$ e $\beta_2 = 0,99$, e o algoritmo de retropropagação com passo de aprendizado $\eta = 0,001$, 100 épocas, mini-batches de tamanho $k = 32$, inicialização de Xavier para os pesos da entrada e inicialização ortogonal para os pesos recorrentes.

C. Análise de discriminante linear (LDA)

A LDA é uma técnica bem conhecida para extração de características, redução de dimensionalidade e classificação,

sendo uma generalização do discriminante linear de Fisher [30] para K classes. Na LDA, uma entrada \mathbf{x} de um espaço D -dimensional é projetada para um espaço $(K - 1)$ -dimensional por meio de $K - 1$ funções discriminantes $y_k = \mathbf{w}_k^T \mathbf{x}$, com $k = 1, \dots, K - 1$. Essas funções podem ser agrupadas em um vetor \mathbf{y} , e \mathbf{w}_k podem ser agrupados como colunas de uma matriz \mathbf{W} de dimensão $D \times (K - 1)$, com $\mathbf{y} = \mathbf{W}^T \mathbf{x}$. Após a projeção, é possível classificar a entrada como pertencente à classe C_k por meio da comparação com um limiar [31]. Denotando a matriz de covariância entre classes e a matriz de covariância dentro de cada classe por \mathbf{S}_B e \mathbf{S}_W , respectivamente, a função custo para o caso de $K = 2$ é dada por $J(\mathbf{w}) = (\mathbf{w}^T \mathbf{S}_B \mathbf{w}) / (\mathbf{w}^T \mathbf{S}_W \mathbf{w})$. Essa função custo busca maximizar a separação entre classes ao mesmo tempo em que reduz a sobreposição entre elas, fornecendo a melhor direção de projeção dos dados. Para $K > 2$ classes, a solução que maximiza $J(\mathbf{w})$ é obtida quando \mathbf{W} é composta pelos autovetores associados aos maiores autovalores de $\mathbf{S}_W^{-1} \mathbf{S}_B$ [31].

Para levar em conta o desbalanceamento da quantidade de dados das classes, calculam-se \mathbf{S}_B e \mathbf{S}_W com pesos p_k na proporção inversa ao número de dados das classes, ou seja,

$$\mathbf{S}_W = \sum_{k=1}^K p_k \sum_{n \in C_k} (\mathbf{x}_n - \mathbf{m}_k)(\mathbf{x}_n - \mathbf{m}_k)^T \quad \text{e} \quad (2)$$

$$\mathbf{S}_B = \sum_{k=1}^K p_k n_k (\mathbf{m}_k - \mathbf{m})(\mathbf{m}_k - \mathbf{m})^T, \quad (3)$$

em que n_k é o número de dados na classe C_k , \mathbf{m}_k é a média desses dados, e \mathbf{m} é a média ponderada

$$\mathbf{m} = \frac{\sum_{k=1}^K p_k \sum_{n \in C_k} \mathbf{x}_n}{\sum_{k=1}^K p_k n_k}. \quad (4)$$

Os vetores de entrada de cada batimento cardíaco utilizado na LDA são compostos por 10 amostras do complexo QRS, 9 amostras entre o fim do QRS e o fim da Onda T, informações dos intervalos RR, da duração do QRS e da Onda T e a presença da Onda P, como proposto por [4] e ilustrado na Figura 2. Os intervalos RR usados foram os intervalos entre o batimento atual e o anterior (*Pre-RR*), entre o batimento atual e o posterior (*Post-RR*), o intervalo médio de toda a gravação de um paciente (*Average RR*) e o intervalo médio local entre 10 batimentos adjacentes (*Local average RR*), totalizando 26 características extraídas para cada batimento.

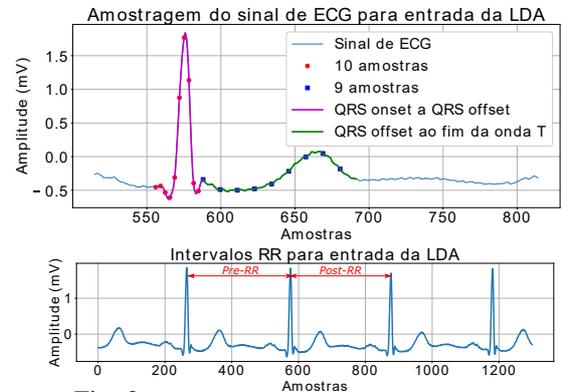


Fig. 2: Extração de características da LDA.

Para cada derivação (*lead A* e *lead B*), foi considerado um classificador. As saídas desses classificadores foram combinadas para fornecer a saída global como explicado na Subseção IV-D abaixo. A LDA usada possui um estimador

de covariância por máxima verossimilhança, probabilidades marginais iguais entre as classes, ou seja, de $1/4$, e a menor redução de dimensionalidade possível ($K - 1 = 3$).

D. Combinação de classificadores

Em geral, classificadores usados em uma aplicação específica atingem diferentes graus de sucesso, apoiados em conjuntos de características distintos. A combinação de classificadores em um único sistema de reconhecimento de arritmias ajuda a integrar o conhecimento adquirido pelos diferentes modelos [32]. Em [4], as configurações de classificadores combinados obtiveram maiores acurácias em todos os resultados em comparação ao uso de apenas um classificador. Neste trabalho, usou-se a operação de multiplicação elemento a elemento proposta em [4] para combinar os classificadores.

Seja $P_m(k | \mathbf{x})$ a probabilidade de uma entrada \mathbf{x} pertencer a uma classe k fornecida como uma posição do vetor de saída do m -ésimo classificador, o vetor final de saída de um modelo combinado para a entrada \mathbf{x} é dado por [4]

$$P(k | \mathbf{x}) = \frac{\prod_{m=1}^M P_m(k | \mathbf{x})}{\sum_{l=1}^K \prod_{m=1}^M P_m(l | \mathbf{x})}, \quad (5)$$

sendo $K = 4$ o número de classes e M o número de classificadores. A classificação global é obtida pela escolha da classe com a maior probabilidade condicional resultante dessa operação. Essa combinação fornece resultados mais confiáveis em probabilidades que foram estimadas como zero, reduzindo o erro total das redes e a incidência de falsos negativos.

V. RESULTADOS E DISCUSSÃO

Foram feitas simulações considerando a MLP, a RNN e a LDA individualmente e as combinações MLP-LDA, MLP-RNN, RNN-LDA e MLP-RNN-LDA. Para cada simulação, calcularam-se as métricas de Acurácia geral $Acc = (VP + VN)/(VP + VN + FP + FN)$ e de Sensibilidade $Se = VP/(VP + FN)$ e Precisão $P = VP/(VP + FP)$ para cada classe, sendo VP , VN , FP e FN as quantidades de verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos, respectivamente. Seguindo a recomendação da AAMI [4], [12], os falsos positivos devido à classe F foram desconsiderados para a classe V. Calcularam-se também as métricas de $F1$ -score $F1 = 2 \times Se \times P/(Se + P)$ para cada classe e as métricas gerais de $macro$ - $F1$ -score $mF1 = \sum_{k=1}^K F1_k/K$ e de $F1$ -score ponderado para 3 e 4 classes $wF1 = \sum_{k=1}^K n_k F1_k / \sum_{k=1}^K n_k$, em que n_k é o número de elementos da classe k da Tabela I. Os resultados para cada simulação estão apresentados na Tabela II, junto a outros resultados da literatura retirados diretamente dos trabalhos originais. Esses trabalhos também utilizam o banco de dados MITDB e seguem as recomendações da AAMI. Os dois maiores valores das simulações para cada métrica são apresentados em negrito.

Analisando-se os classificadores propostos neste trabalho individualmente, é possível observar que a MLP apresenta as métricas gerais de Acc e de $wF1$ maiores do que a RNN e a LDA, além de $F1$ maiores para as classe N, S e V. Para essas classes, a MLP também atinge os valores de $F1$ usuais da literatura, porém, para a classe S, obtém um valor de $F1$ de 57,4%, que é maior do que o valor obtido em todos os outros trabalhos considerados. Para a classe V, o valor de $F1$

é de 76,2%, que se encontra na média dos demais trabalhos. Com relação à classe N, o desempenho da MLP é um pouco pior. Observa-se que o melhor classificador individual para a classe F é a LDA, com a maior métrica de Se de 83,8%, e de $F1$ de 29,7%. Nota-se que as entradas propostas por [4] e utilizadas na LDA auxiliam a identificação da classe F.

Os resultados dos classificadores combinados foram melhores do que os dos classificadores individuais, em termos das métricas Acc e $wF1$. Comparando-se a combinação RNN-LDA com a RNN individual, por exemplo, aumentou-se a Acc de 74,4% da RNN para 81,7%, e o valor de $wF1$ de 75,1% da RNN para 82,3%. Além disso, o valor de $F1$ de todas as classes foi maior na combinação da RNN-LDA do que na RNN e na LDA separadamente.

Observando-se os valores em negrito, é possível perceber que a MLP-LDA e a MLP-RNN-LDA apresentaram os melhores desempenhos gerais dentre todas as simulações. Combinando-se a MLP com a LDA, a Acc de 77,2% da MLP aumentou para 84,2% e o valor de $wF1$ aumentou de 78,3% para 84,5%. Porém, ao realizar-se a combinação MLP-RNN-LDA, não é possível notar variação significativa das métricas em relação ao modelo MLP-LDA, com o valor de Acc aumentando apenas 0,2% e de $wF1$ aumentando 0,3%. Assim, a contribuição da RNN na combinação não foi relevante, além de levar a um aumento de 113476 parâmetros no modelo. Portanto, decidiu-se comparar o modelo MLP-LDA com os resultados da literatura. A matriz de confusão obtida por esse modelo está na Tabela III.

A combinação MLP-LDA obtém valores de $F1$ para as classe S e V de 61,9% e 86,6% respectivamente, maiores do que todos os outros da literatura. Além disso, obtém o valor de $F1$ de N de 90,8%, com uma diferença de 5,7% do maior valor dos demais trabalhos. No entanto, o valor de $F1$ da classe F é menor do que os de [4] e [14], que consideram o problema de quatro classes. Em relação ao desempenho geral, o valor de $wF1$ foi menor dos que os demais, com uma diferença máxima de 6,8%, no caso de quatro classes. Considerando três classes, essa diferença passa a ser de 7,1%. Dentre os classificadores que usaram as quatro classes, a diferença máxima desconsiderando a classe F foi de 4,7%.

Apesar dos valores de $F1$ maiores nas classes S e V, a combinação MLP-LDA obteve $wF1$ menor devido ao peso da classe N durante a ponderação no cálculo dessa métrica. Neste trabalho, não foram considerados todos os batimentos da classe N, pois o desbalanceamento dos dados na proporção original prejudicaria muito o treinamento dos classificadores, mesmo com a correção realizada pelos pesos na função custo. Avaliando-se a métrica $mF1$, ao atribuir pesos iguais às métricas $F1$ de todas as classes, o método proposto alcança valores maiores do que os relatados nos demais trabalhos.

VI. CONCLUSÕES

A classificação de arritmias cardíacas por métodos computadorizados é uma área com muitas possibilidades de melhoria. Neste trabalho, avaliou-se o efeito de combinar os classificadores MLP, RNN e LDA. Dentre os classificadores individuais, a MLP apresentou os melhores resultados para as classes N, S e V. Além disso, verificou-se que as combinações ajudam

TABELA II: Métricas (%) dos resultados das simulações e comparação com os valores da literatura. Os dois maiores resultados para cada métrica calculada com os resultados dos métodos propostos estão em negrito.

Métodos propostos	Acc	N			S			V			F			wF1 4 classes	wF1 3 classes	mF1
		Se	P	F1												
MLP	77,2	82,4	88,7	85,4	57,2	57,6	57,4	82,1	71,1	76,2	4,0	2,8	3,3	78,3	80,7	73,0
RNN	74,4	79,3	84,4	81,8	55,4	47,7	51,2	77,5	71,1	74,2	15,2	12,3	13,6	75,1	77,0	69,1
LDA	69,9	79,2	93,3	85,7	46,3	31,0	37,1	44,1	70,6	54,3	83,8	18,1	29,7	73,4	74,8	59,0
MLP-LDA	84,2	90,4	91,1	90,8	56,8	68,0	61,9	89,3	84,1	86,6	15,4	12,6	13,9	84,5	86,8	79,8
MLP-RNN	78,8	84,3	87,1	85,7	56,8	57,7	57,2	83,9	73,4	78,3	4,52	5,7	5,0	78,8	81,2	73,7
RNN-LDA	81,7	89,6	89,4	89,5	55,7	56,1	55,9	73,6	86,1	79,4	40,7	23,5	29,8	82,3	83,9	74,9
MLP-RNN-LDA	84,4	91,1	89,2	90,1	57,2	71,2	63,5	87,9	85,1	86,5	12,2	13,3	12,7	84,2	86,5	80,0

Métodos da literatura	Acc	N			S			V			F			wF1 4 classes	wF1 3 classes	mF1
		Se	P	F1	Se	P	F1	Se	P	F1	Se	P	F1			
Chazal [4]	85,9	86,9	99,2	92,6	75,9	38,5	51,1	77,7	81,9	79,7	89,4	8,6	15,7	89,7	90,3	74,5
Mar [†] [14]	90,0	89,6	99,1	94,11	83,2	33,5	47,8	86,8	75,9	81,0	61,1	16,6	26,1	91,0	91,5	74,3
Llamedo [‡] [21]	93,0	95,0	98,0	96,5	77,0	39,0	51,8	81,0	87,0	83,9	–	–	–	–	93,9	77,4
Lin [†] [15]	90,8	91,6	99,3	95,3	81,4	31,6	45,5	86,2	73,7	79,5	–	–	–	–	92,4	73,4
Garcia [†] [16]	91,0	95,0	96,0	95,5	30,0	26,0	27,9	85,0	66,0	74,3	–	–	–	–	91,6	65,9
Luz [‡] [17]	92,4	94,0	98,0	96,0	62,0	53,0	57,2	87,3	59,4	70,7	–	–	–	–	92,9	74,6

[†] Esses resultados foram obtidos com uma rede MLP. [‡] Autores otimizaram os métodos para o problema de três classes: N, S e V.

TABELA III: Matriz de confusão obtida para MLP-LDA.

Classes Verdadeiras		Classes Preditas			
		N	S	V	F
	N	7557	317	161	324
	S	370	749	183	17
	V	127	28	1817	62
	F	240	7	71	58

efetivamente a melhorar o desempenho geral do classificador, aumentando todas as métricas (Se , P , $F1$) das classes N, S e V e as métricas Acc , $wF1$ e $mF1$. Isso é verificado pelos valores em negrito da Tabela II. Destaca-se ainda que, para as classes S e V, consideradas como principais pela maioria dos trabalhos e pelas recomendações da AAMI, a combinação MLP-LDA atingiu valores de $F1$ -score melhores do que todos os classificadores concebidos como estado da arte, além de um valor de $mF1$ maior do que todos os demais. Assim, foi possível balancear os falsos alarmes (Pe) e o bom desempenho na detecção (Se) de cada classe de forma clinicamente relevante.

REFERÊNCIAS

- [1] L. B. Marinho *et al.*, “A novel electrocardiogram feature extraction approach for cardiac arrhythmia classification,” *Future Generation Computer Systems*, vol. 97, pp. 564–577, 2019.
- [2] S. K. Berkaya *et al.*, “A survey on ECG analysis,” *Biomedical Signal Processing and Control*, vol. 43, pp. 216–235, 2018.
- [3] S. I. Niwas, R. S. S. Kumari, and V. Sadasivam, “Artificial neural network based automatic cardiac abnormalities classification,” in *Proc. of ICCIMA’05*. IEEE, 2005, pp. 41–46.
- [4] P. de Chazal, M. O’Dwyer, and R. B. Reilly, “Automatic classification of heartbeats using ECG morphology and heartbeat interval features,” *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1196–1206, 2004.
- [5] N. K. Dewangan and S. P. Shukla, “ECG arrhythmia classification using discrete wavelet transform and artificial neural network,” in *Proc. of RTEICT*. IEEE, 2016, pp. 1892–1896.
- [6] L. Edenbrandt, B. Devine, and P. W. Macfarlane, “Neural networks for classification of ECG ST-T segments,” *J. Electrocardiol.*, vol. 25, pp. 167–173, 1992.
- [7] Yun-Chi Yeh, Wen-June Wang, and Che Wun Chiou, “Cardiac arrhythmia diagnosis method using linear discriminant analysis on ECG signals,” *Measurement*, vol. 42, pp. 778–789, 2009.
- [8] M. H. Song *et al.*, “Support vector machine based arrhythmia classification using reduced features,” *Int. J. Control, Automation, and Systems*, vol. 3, pp. 571–579, Dec. 2005.
- [9] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
- [10] U. R. Acharya *et al.*, “A deep convolutional neural network model to classify heartbeats,” *Computers in Biology and Medicine*, vol. 89, pp. 389–396, 2017.
- [11] G. Wang *et al.*, “A global and updatable ECG beat classification system based on recurrent neural networks and active learning,” *Information Sciences*, vol. 501, pp. 523–542, 2019.
- [12] Association for the Advancement of Medical Instrumentation (AAMI), “ANSI/AAMI EC57:1998(R)2008 - Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms,” *American National Standards Institute, Arlington, VA, USA*, 2008.
- [13] S. Kiranyaz, T. Ince, and M. Gabbouj, “Real-time patient-specific ECG classification by 1-D convolutional neural networks,” *IEEE Trans. Biomed. Eng.*, vol. 63, pp. 664–675, 2015.
- [14] T. Mar *et al.*, “Optimization of ECG classification by means of feature selection,” *IEEE Trans. Biomed. Eng.*, vol. 58, pp. 2168–2177, 2011.
- [15] C. Lin and C. Yang, “Heartbeat classification using normalized RR intervals and morphological features,” *Mathematical Problems in Engineering*, 2014.
- [16] G. Garcia *et al.*, “Improving automatic cardiac arrhythmia classification: Joining temporal-VCG, complex networks and SVM classifier,” in *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2016, pp. 3896–3900.
- [17] G. Garcia *et al.*, “Inter-patient ECG heartbeat classification with temporal VCG optimized by PSO,” *Scientific Reports*, vol. 7, pp. 1–11, 2017.
- [18] E. J. S. Luz *et al.*, “ECG-based heartbeat classification for arrhythmia detection: A survey,” *Computer Methods and Programs in Biomedicine*, vol. 127, pp. 144–164, 2016.
- [19] A. L. Goldberger *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, pp. e215–e220, Jun. 13, 2000.
- [20] G. B. Moody and R. G. Mark, “The impact of the MIT-BIH arrhythmia database,” *IEEE Eng. Med. Biol. Mag.*, vol. 20, pp. 45–50, 2001.
- [21] M. Llamedo and J. P. Martínez, “Heartbeat classification using feature selection driven by database generalization criteria,” *IEEE Trans. Biomed. Eng.*, vol. 58, pp. 616–625, 2011.
- [22] J. P. Martínez *et al.*, “A wavelet based ECG delineator: evaluation on standard databases,” *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 570–581, 2004.
- [23] A. Demski and M. L. Soria, “ecg-kit: a Matlab toolbox for cardiovascular signal processing,” *J. Open Research Software*, vol. 4, 2016.
- [24] Martín Abadi *et al.*, “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015, Software available from tensorflow.org.
- [25] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [26] N. Nagata, R. Candido, and M. T. M. Silva, “Classification of arrhythmias using multilayer perceptron neural networks,” in *28th USP International Symposium of Undergraduate Research*, 2020.
- [27] D. P. Kingma, and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [28] X. Glorot, and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” *Proc. of AISTATS*, p. 249–256, 2010.
- [29] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, pp. 1735–1780, 1997.
- [30] R. A. Fisher, “The statistical utilization of multiple measurements,” *Annals of Eugenics*, vol. 8, pp. 376–386, 1938.
- [31] C. M. Bishop, *Pattern recognition and machine learning*, Springer, 2006.
- [32] S. Osowski, T. Markiewicz, and L. T. Hoai, “Recognition and classification system of arrhythmia using ensemble of neural networks,” *Measurement*, vol. 41, pp. 610–617, 2008.