

Universidade Federal do Ceará
Pró-Reitoria de Pesquisa e Pós-Graduação
PIBIC 2023/2024 - Edital Nº 02/2023

Deteccção de Insuficiência Cardíaca em Pacientes com Doença de Chagas a partir de Algoritmos de Extração Automática de Parâmetros e Aprendizado de Máquina em Sinais ECG

Resumo

A doença de Chagas (DC), também conhecida como tripanossomíase americana, ocorre em mais de 21 países, principalmente na América Latina (no Brasil e no México). O número de infectados gira de 6 a 7 milhões de pessoas. A disseminação da doença geralmente ocorre quando as fezes dos insetos da subfamília Triatominae, portadores do protozoário *Trypanosoma cruzi*, entram em contato com o sangue do paciente. A infecção, então, ocorre em duas fases: aguda e crônica. De acordo com a Organização Mundial de Saúde (OMS), durante a fase crônica, 30% dos pacientes desenvolvem algum tipo de cardiomiopatia, como arritmias cardíacas ou insuficiência cardíaca progressiva, que, se não tratadas, podem levar o paciente chagásico à Morte Súbita Cardíaca (MSC). Embora alguns desenvolvam apenas defeitos de condução e anormalidades segmentares leves da motilidade da parede, outros desenvolvem sintomas graves de insuficiência cardíaca (IC), fenômenos tromboembólicos e arritmias ventriculares com risco de vida. A IC é uma das principais causas de morte de pacientes com DC, no entanto, medicamentos de baixo custo são capazes de reduzir a mortalidade e melhorar a qualidade de vida de portadores de IC. Devido à falta de serviços terciários fora dos centros urbanos, uma ferramenta de diagnóstico automático baseada no ECG, que é um exame relativamente simples, tem potencial de melhorar a capacidade de diagnóstico da IC. Neste contexto, o presente projeto propõe utilizar algoritmos de processamento digital de sinais, extração automática de parâmetros, modelos clássicos e modernos de aprendizado de máquina para análise de desempenho de diferentes abordagens computacionais para detecção de IC em pacientes com DC a partir de sinais ECG de longa duração (Holter). Para tanto, contaremos com um repositório de dados de sinais ECG e de dados clínicos de 219 pacientes portadores de cardiomiopatia chagásica fornecido pelo Hospital Universitário Clementino Fraga Filho (HUCFF/UFRJ).

1. Introdução

A doença de Chagas é um problema de saúde crônico em toda a América Latina, onde se estima que em torno de 8 milhões de pessoas estejam infectadas [1,2]. A doença é causada por um parasita intracelular, o *Trypanosoma cruzi*, sendo uma das principais causas de cardiomiopatias na América Latina. O parasita é transmitido ao ser humano principalmente pelo contato com excrementos de insetos triatomíneos hematófagos infectados, embora mecanismos não-vetoriais de transmissão também sejam reconhecidas como transfusão de sangue e congênito (da mãe para o feto) [3]. Enquanto manifestações digestivas da doença de Chagas (megaesôfago e megacólon) ocorrem em torno de 10% dos indivíduos infectados, a cardiomiopatia crônica, afetando de 30 a 40% dos infectados, é certamente a manifestação mais grave e frequente, a qual acarreta diversas complicações: insuficiência cardíaca, aneurismas ventriculares, tromboembolismo e arritmias severas. Tais complicações podem, por sua vez, ocasionar a morte súbita cardíaca. A patogênese

do comprometimento cardíaco na Doença de Chagas ainda não está totalmente compreendida, e quatro mecanismos principais foram propostos, a saber: 1) dano ao miocárdio decorrente da ação parasitária; 2) lesão miocárdica imunomediada; 3) distúrbios microvasculares e 4) disautonomia cardíaca. Se não tratada, a doença de Chagas pode ser fatal, na maioria dos casos devido ao comprometimento cardíaco. A morte súbita cardíaca está relacionada a 55% - 65% das mortes em doença de Chagas, podendo ser a primeira manifestação da doença ou seu evento terminal [4]. A morte súbita é definida como um colapso abrupto com registro de perda de sinais vitais [5].

O sinal ECG devido às suas características e facilidade de aquisição, é uma ferramenta adequada para o diagnóstico de distúrbios da atividade cardíaca. Constitui-se em um método de diagnóstico de baixo custo e não-invasivo, fundamental para monitoramento de doenças cardíacas congênitas ou adquiridas [6]. Arritmias cardíacas severas podem ser fatais, de forma que a detecção precoce de alterações eletrocardiográficas tem potencial para evitar desfechos adversos [7]. Um sinal ECG típico consiste em uma série temporal resultante dos processos de despolarização e repolarização elétrica do tecido cardíaco [8]. Estes eventos são responsáveis pela geração de ondas características, espaçadas uma das outras por intervalos específicos. São denominadas pelas seguintes letras: P, Q, R, S, e T. As ondas Q, R e S formam o complexo QRS, o qual está relacionado ao processo de despolarização ventricular. Três intervalos importantes para fins de diagnóstico são o intervalo PR

(do início da onda P ao início do complexo QRS), o intervalo QT (do início do complexo QRS ao final da onda T) e o intervalo ST (do final do complexo QRS ao início da onda T). Alterações detectáveis no sinal ECG causadas pela doença de Chagas podem ocorrer em até 68,5% dos pacientes infectados [9,10]. As extrassístoles ventriculares e os distúrbios de condução são as alterações mais frequentes detectáveis no sinal ECG em pacientes chagásicos. As principais causas de morte nesses pacientes são a morte súbita cardíaca, a insuficiência cardíaca e o acidente vascular encefálico [11,12,13].

A Variabilidade da Frequência Cardíaca (VFC), representada pelas alterações nos intervalos entre batimentos, tem sido aplicada já há muito tempo como marcador não-invasivo da modulação autonômica cardíaca. Vários estudos já relataram alterações em índices da VFC em pacientes chagásicos, confirmando que a disautonomia cardíaca está presente na patogênese da doença de Chagas. No entanto, ainda há controvérsia sobre se a disautonomia cardíaca em pacientes com doença de Chagas é um mecanismo primário ou secundário a outro dano cardíaco [14,15].

Numerosos algoritmos, especialmente no campo do Aprendizado de Máquina, têm sido desenvolvidos com a finalidade de realização de predições a partir dos batimentos cardíacos [16,17]. A aplicabilidade desses algoritmos está não apenas no diagnóstico automático de doenças cardíacas através do processamento do sinal ECG, mas também para acompanhamento da evolução de tais patologias, sugerindo-se a aplicabilidade em algumas medidas terapêuticas.

Algoritmos como Floresta Aleatória (Random Forest), Redes Neurais, Máquinas de Vetores Suporte (Support Vector Machine), dentre outros, podem fornecer novos marcadores e permitir a análise do sinal ECG quase sempre sem requerer pré-processamento ou alguma informação prévia (dados clínicos). Tradicionalmente, tais algoritmos têm sido aplicados de forma bem sucedida para classificação de batimentos nas seguintes classes [18]: batimento ectópico supraventricular, batimento ectópico ventricular, fibrilação, insuficiência cardíaca, dentre outras. Já outras pesquisas consideram a aplicação de algoritmos de aprendizado de máquina para diferenciação de classes de pacientes, por exemplo, isquêmicos e não-isquêmicos e portadores ou não-portadores de doença coronariana [19,20,21].

Neste contexto, o presente trabalho propõe a investigação de métodos de extração automática de parâmetros de sinais ECG e a aplicação de algoritmos de aprendizado de máquina para o desenvolvimento de modelos de detecção de Insuficiência Cardíaca (IC), mais precisamente, a disfunção sistólica do ventrículo esquerdo, a qual pode ter seu diagnóstico confirmado através da medição da fração de ejeção do ventrículo esquerdo (FEVE) menor ou igual a 40% através do exame Ecocardiograma. Diversos atributos derivados do processamento do sinal ECG têm potencial de caracterizar estados patológicos, dentre os quais: amplitude, duração e morfologia do complexo QRS, amplitude, duração e morfologia da onda P, amplitude, duração e morfologia da

onda T, intervalos PR, QT e ST.

Dentre os classificadores potencialmente aplicáveis para a predição de eventos adversos com o uso de sinais ECG, destacamos os classificadores Bayesianos, as redes neurais MLP, as máquinas de vetores suporte (SVM) e o algoritmo de florestas aleatórias (Random Forest) [22].

2. Perguntas de Partida

- 1) É possível desenvolver um modelo de detecção automática de insuficiência cardíaca (IC), através da detecção da disfunção sistólica do ventrículo esquerdo, em pacientes chagásicos com base em algoritmos de aprendizado de máquina a partir do processamento de sinais ECG ?
- 2) Dada a disponibilidade de sinais ECG de longa duração, quais métricas derivadas dos parâmetros extraídos do processamento do ECG impactam na detecção automática de insuficiência cardíaca?
- 3) Modelos de Deep Learning que processam diretamente o sinal original produzem melhores resultados que modelos tradicionais, baseados em engenharia de parâmetros?
- 4) É possível detectar insuficiência cardíaca a partir do processamento de subintervalos de ECG, como sinais de 1 minuto, 5 minutos, 10 minutos ou 15 minutos ?

3. Hipóteses

- 1) A partir do processamento e extração de parâmetros de um sinal ECG de 24 horas, dentre os quais, as amplitudes das ondas características, os intervalos entre as ondas e os índices de variabilidade da frequência cardíaca, é possível desenvolver um modelo de detecção de insuficiência cardíaca em pacientes chagásicos a partir do uso de uma base de dados com registros de fração de ejeção do ventrículo esquerdo, métrica obtida pelo exame ecocardiograma;
- 2) Dados diferentes cenários em que diferentes tipos de dados estão disponíveis, diferentes modelos podem ser propostos;
- 3) Aplicando-se diferentes metodologias de classificação (por exemplo, classificação por paciente, classificação por sinal ECG e classificação por batimento ou segmento de sinal) é possível estratificar o peso/importância de diferentes métricas nos modelos de detecção de insuficiência cardíaca.

4. Objetivos

Objetivo Geral: Desenvolver um modelo de detecção de insuficiência cardíaca (IC) em pacientes portadores de Doença de Chagas (DC) baseado em algoritmo de aprendizado de máquina a partir do processamento do sinal ECG de 24 horas e de dados clínicos desses pacientes.

Objetivos Específicos:

- 1) Organizar uma base de dados de parâmetros extraídos dos sinais ECG (batimento a batimento) avaliados em intervalos específicos (1 minuto, 5 minutos, 10 minutos e 15 minutos) a partir de uma parceria com o Hospital Universitário Clementino Fraga Filho da Universidade Federal do Rio de Janeiro - UFRJ;
- 2) Investigar na literatura e aplicar algoritmos de processamento digital de sinais bem estabelecidos sobre toda a base de sinais ECG para extração dos parâmetros correspondentes: amplitudes e durações das ondas P, T e complexo QRS, morfologia das ondas P, T e complexo QRS, intervalos PR, QT e ST;
- 3) Selecionar e avaliar diferentes modelos de aprendizado de máquina, incluindo-se redes convolucionais para processamento dos dados em diferentes cenários: subintervalos de curta duração e sinais de longa duração;
- 4) Aplicar processos de rotulação dos sinais, a partir de parceria com pesquisadores da área de

medicina cardiologia, para marcação manual dos limites das ondas características, considerando amostras de formas de onda dos ECGs, e anotação de ocorrência de eventos importantes, tais como bradicardia e taquicardia sinusal, fibrilação atrial, bloqueios de ramo, taquicardia ventricular, bloqueio sinoatrial, dentre outros;

4) Calcular diferentes métricas de desempenho para cada modelo avaliado: acurácia, precisão, sensibilidade (recall), especificidade, AUC, matriz de confusão, escore F1;

5) Selecionar o modelo mais robusto em cada cenário e ranquear os pesos/importâncias dos diversos parâmetros que formam o vetor de entrada dos modelos de aprendizado de máquina nos diversos cenários.

5. Materiais e Métodos

Para a execução da pesquisa proposta, experimentos computacionais serão realizados com o uso de uma base de dados clínicos e de sinais ECG de pacientes portadores de doença de Chagas, em que parte desses pacientes foi a óbito por ocorrência de morte súbita cardíaca (MSC), parte foi a óbito por outras causas e parte ainda se encontra viva.

(1) Descrição do banco de dados:

O Hospital Universitário Clementino Fraga Filho da Universidade Federal do Rio de Janeiro (HUCFF-UFRJ) é um centro de referência regional no Brasil. A população de pacientes inclui residentes urbanos na cidade do Rio de Janeiro, que são atendidos independentemente das condições financeiras. Os pacientes encaminhados para este hospital entre 1990 e 2017 nos quais a CCC (cardiomiopatia chagásica crônica) foi verificada com base em critérios diagnósticos validados foram solicitados a cooperar em um programa prospectivo de acompanhamento clínico protocolado, compreendendo uma avaliação sistemática dos parâmetros cardíacos. Em resumo, o diagnóstico de CCC foi estabelecido com base na presença de pelo menos dois ensaios sorológicos positivos para anticorpos para *Trypanosoma cruzi* (imunofluorescência indireta, hemaglutinação indireta ou teste de imunoabsorção enzimática) juntamente com anormalidades ecocardiográficas (diâmetro diastólico final do VE (ventrículo esquerdo) > 55 mm, disfunção global do VE ou anormalidade do movimento da parede regional) ou alterações características no ECG de 12 derivações (bloqueio avançado de ramo direito, geralmente com hemibloqueio anterior esquerdo). Indivíduos que não consentiram ao programa de acompanhamento clínico protocolado, aqueles com episódio prévio de parada cardíaca súbita, aqueles com marcapasso cardíaco implantado e aqueles com qualquer outro tipo de cardiomiopatia, bem como aqueles que foram submetidos anteriormente a tratamento específico para infecção chagásica, foram excluídos do estudo.

O protocolo foi apreciado pelo comitê de ética do HUCFF-UFRJ, o qual abdicou da necessidade de permissão por escrito sob o número 45360915.1.1001.5262, de acordo com as normas atualmente aplicadas pelo Comitê Nacional de Ética em Pesquisa e com os princípios descritos na Declaração de Helsinque.

A partir do momento do diagnóstico, os pacientes foram avaliados sistematicamente no HUCFF-UFRJ a cada 6-8 meses. Cada avaliação do estudo compreendeu o seguinte: histórico clínico e exame completo, hematologia de rotina, bioquímica sérica e urinária, eletrocardiografia de 12 derivações, monitoramento Holter 24 horas, teste ergométrico e ecocardiografia detalhada.

Dados sobre ocorrências de morte súbita cardíaca foram coletados a partir de registros hospitalares; entrevistas com médicos, testemunhas e/ou parentes próximos. Em casos de morte, realizaram-se consultas de dados de mortalidade de 1.985.735 registros do Sistema de Informações sobre Mortalidade, disponível gratuitamente, no Ministério da Saúde (SIM-Sistema de Informação de Mortalidade). O tempo de acompanhamento foi atribuído como o número de dias entre o Holter e o evento de morte súbita (MS).

A morte súbita cardíaca (MSC) é descrita como um colapso repentino com perda documentada de sinais vitais, que pode resultar em tentativas de restaurar a circulação [1]. Apenas é considerada de origem cardíaca após serem descartadas as possibilidades de morte súbita devido à doença não

cardíaca aguda, doença vascular não cardíaca, overdose de drogas, doença terminal ou condições metabólicas. O evento de morte súbita é comprovado por dois médicos independentes. Neste banco de dados, assume-se que a morte súbita foi causada por arritmias cardíacas. Assume-se também que todos os pacientes receberam o tratamento mais adequado, com base em diretrizes, ajustadas individualmente para a CCC, disponíveis na época do estudo.

Os sinais a serem utilizados na pesquisa foram adquiridos a partir de 308 pacientes com doença de Chagas, coletados entre 1992 e 2013. Quanto ao gênero, os sinais pertencem a 180 (58,4%) mulheres e 128 (41,6%) homens. Em relação à idade, os sinais pertencem a 4 (~ 1,3%) pacientes entre 15 e 20 anos, 22 (~ 7,1%) pacientes entre 20 e 30 anos, 77 (~ 25%) pacientes entre 30 e 40 anos, 106 (~ 34,4%) pacientes entre 40 e 50 anos, 95 (~ 30,8%) pacientes entre 50 e 60 anos, e 12 (~ 3,9%) pacientes entre 60 e 70 anos. De todos os sinais analisados, apenas 1 pertencia a um paciente que apresentava algum outro tipo de doença cardíaca. Em relação à classe alvo, 181 (~ 48%) dos sinais foram obtidos de pacientes com fração de ejeção reduzida do ventrículo esquerdo, o que caracteriza insuficiência cardíaca, e 196 (~ 52%) dos sinais foram obtidos de pacientes sem a referida disfunção.

Em relação ao escore Rassi, que avalia o risco de mortalidade em 5 a 10 anos em pacientes com doença de Chagas, 58 (~ 18,8%) sinais são de pacientes com alto risco de morte, 75 (~ 24,4%) pertencem a pacientes com risco intermediário de morte e 175 (~ 56,8%) sinais pertencem a pacientes com baixo risco de morte.

(2) Algoritmos para processamento digital de sinais e extração de parâmetros:

Conforme sugerido por pesquisadores médicos especialistas do Hospital Universitário Clementino Fraga Filho da Universidade Federal do Rio de Janeiro (HUCFF-UFRJ), serão investigados três diferentes cenários quanto à extensão do processamento do exame Holter de cada paciente com doença de Chagas: (i) extração de parâmetros em toda a extensão de aproximadamente 24 horas; (ii) extração de parâmetros diferenciada em dois segmentos de aproximadamente 12 horas (8h da manhã até 8h da noite/8h da noite até 8h da manhã do dia seguinte); e (iii) extração de parâmetros apenas no segmento final de sinal ECG de quatro horas de duração, correspondente ao intervalo das 4 horas da manhã às 8 horas da manhã. A justificativa específica para este terceiro cenário é que exatamente neste período, em que o paciente monitorado encontrava-se no processo de despertar, acontecem processos de modulação do sistema nervoso autônomo que podem favorecer o desencadeamento de arritmias malignas e, por conseguinte, a ocorrência de morte súbita. Uma vez obtidos os segmentos de sinal, algoritmos de pré-processamento para eliminação de interferências de alta e baixa frequência são aplicados, e em sequência, algoritmos para detecção dos complexos QRS [21,26], segmentação ou delineamento dos complexos QRS [22,26], modelagem morfológica dos complexos QRS [28], detecção e delineamento das ondas P e T [23,24], e modelagem morfológica das ondas P e T [23,24]. Como frutos da aplicação de todos os algoritmos citados de pré-processamento e extração de parâmetros, obteremos para cada ciclo cardíaco de cada sinal ECG processado os seguintes atributos: duração e amplitude da onda P; duração e amplitude do complexo QRS; duração e amplitude da onda T; atributos morfológicos da onda P; atributos morfológicos do complexo QRS e atributos morfológicos das ondas P e T. Os atributos morfológicos de cada onda característica consistem em parâmetros de modelos matemáticos, tais como funções Gaussiana e Chapéu-Mexicano, que buscam aproximar o formato das formas de onda.

(3) Pré-processamento dos atributos obtidos e aplicação de modelos de aprendizado de máquina:

Uma vez obtidos os atributos de caracterização de cada ciclo cardíaco de cada sinal ECG processado, serão aplicados algoritmos de redução de dimensionalidade e também remoção de outliers. Serão propostos três cenários de classificação: (i) classificação por janela de sinal de tamanho 1 minuto; (ii) classificação por janela de sinal de tamanho 5 minutos; e (iii) classificação por janela de sinal de tamanho 10 minutos. Cada janela de sinal, em todos os três cenários, é uma amostra independente, cujos atributos podem ser métricas estatísticas calculadas a partir dos atributos dos batimentos contidos no sinal ou as próprias amplitudes do sinal na janela. Então, serão organizados os conjuntos de treinamento e teste, considerando que cada janela de sinal (cenário 2) será a amostra ou indivíduo de

entrada. Serão aplicados, então, os algoritmos classificador bayesiano, redes neurais MLP, máquina de vetores suporte (SVM), florestas aleatórias (Random Forest) e redes convolucionais. O objetivo de classificação de cada um dos modelos de aprendizado de máquina é prever se um dado intervalo de sinal, extraídos de um dado sinal ECG e caracterizados por um dado conjunto de atributos, associa-se ou não a um paciente com insuficiência cardíaca;

(4) Processo de rotulação dos sinais ECG:

Os sinais ECG do repositório cedido pelo Hospital Universitário Clementino Fraga Filho não se encontram rotulados, nem com relação às marcações das delimitações das ondas características, nem com relação a eventos adversos ocorrendo ao longo dos registros, como arritmias malignas. Portanto, por ocasião do presente projeto, também será construído uma interface de rotulação que também aplicará algoritmos de aprendizagem de máquina não-supervisionados para prover identificação de marcações de início, pico e final das ondas características, para um dado número amostrado de ciclos cardíacos, bem como a identificação da ocorrência de eventos importantes ao longo de intervalos selecionados dos sinais ECG, tais como: QRS alargado, ectopias ventriculares polimórficas, fibrilação atrial de baixa resposta ventricular, bradicardia sinusal, bloqueios Atrioventriculares;

(5) Aplicação de modelos de aprendizado de máquina para detecção de eventos adversos ao longo dos sinais ECG:

Com o processo de rotulação, descrito na etapa 4, finalizado, passaremos a construção de modelos de aprendizado de máquina para a identificação das ocorrências acima descritas, a partir da extração de parâmetros e dos datasets obtidos a partir da organização dos parâmetros em correspondentes grupos de intervalos.

6. Dados Preliminares

Preliminarmente, dispomos das informações constantes no banco de dados disponibilizado pelo Hospital Universitário Clementino Fraga Filho da Universidade Federal do Rio de Janeiro. Este banco de dados contém informações clínicas e sinais ECG de exames Holter de pacientes portadores de doença de Chagas. Dentre as informações clínicas disponibilizadas para cada paciente, destacam-se, dentre outras, data de nascimento, data de cada exame Holter realizado, sexo, estado de nascimento, índice de massa corpórea, se possui outra cardiomiopatia ou não, escore Rassi, dados de exame de ecocardiograma e se apresentou ou não evento de MSC.

No total, temos 308 pacientes, alguns dos quais possuem um exame Holter, e outros possuem mais de um exame Holter, totalizando 385 registros de sinais ECG.

7. Referências Bibliográficas

- [1] A. Rassi Jr., A. Rassi, J.A. Marin-Neto, Chagas disease, *Lancet Lond. Engl.*, vol. 375, pp. 1388-1402, 2010.
- [2] A. Rassi Jr., A. Rassi, S.G. Rassi, Predictors of mortality in chronic Chagas disease: a systematic review of observational studies, *Circulation*, vol. 115, pp. 1101-1108, 2007.
- [3] M.V. Simões, M.M.D. Romano, A. Schmidt, K.S.M. Martins, J.A. Marin-Neto, Chagas disease cardiomyopathy, *Int. J. Cardiovasc. Sci.*, vol. 31, pp. 173-189, 2018.
- [4] J.A. Marin-Neto, E. Cunha-Neto, B.C. Maciel, M.V. Simões, Pathogenesis of chronic Chagas heart disease, *Circulation*, vol. 115, pp. 1109-1123, 2007.
- [5] A. C. J. de Souza, G. Salles, A. M. Hasslocher-Moreno, A. S. de Sousa, P. E. A. A. do Brasil, R.M. Saraiva, and S. S. Xavier, Development of a risk score to predict sudden death in patients with Chagas' heart disease, *International journal of cardiology*, vol. 187, pp. 700-704, 2015.
- [6] A. A. Fossa, T. Wisialowski, K. Crimin, E. Wolfgang, J.-P. Couderc, M. Hinterseer, S. Kaab, W. Zareba, F. Badilini, and N. Sarapa, Analyses of dynamic beat-to-beat QT-TQ interval (ECG

restitution) changes in humans under normal sinus rhythm and prior to an event of torsades de pointes during QT prolongation caused by sotalol, *Annals of non-invasive electrocardiology*, vol. 12, no. 4, pp. 338-348, 2007.

[7] A. A. Fossa, M. Zhou, A. Robinson, J. Purkayastha, and P. Martin, Use of ECG restitution (beat-to-beat QT-TQ interval analysis) to assess arrhythmogenic risk of QTc prolongation with guanfacine, *Annals of Noninvasive Electrocardiology*, vol. 19, no. 6, pp. 582-594, 2014.

[8] P. E. Lazzerini, F. Laghi-Pasini, I. Bertolozzi, G. Morozzi, S. Lorenzini, A. Simpatico, E. Selvi, M. R. Bacarelli, F. Finizola, F. Vanni, et al., Systemic inflammation as a novel QT-prolonging risk factor in patients with torsades de pointes, *Heart*, vol. 103, no. 22, pp. 1821-1829, 2017.

[9] E. Roggero, A. R. Pérez, N. Pollachini, S. R. Villar, J. Wildmann, H. Besedovsky, and A. del Rey, The sympathetic nervous system affects the susceptibility and course of trypanosoma cruzi infection, *Brain, behavior, and immunity*, vol. 58, pp. 228-236, 2016.

[10] G. A. Ng, A. Mistry, X. Li, F. S. Schlindwein, and W. B. Nicolson, Lifemap: towards the development of a new technology in sudden cardiac death risk stratification for clinical use, *Ep Europace*, vol. 20, no. F12, pp. f162-f170, 2018.

[11] K. E. Brack, J. Winter, and G. A. Ng, Mechanisms underlying the autonomic modulation of ventricular fibrillation initiation tentative prophylactic properties of vagus nerve stimulation on malignant arrhythmias in heart failure, *Heart failure reviews*, vol. 18, no. 4, pp. 389-408, 2013.

[12] G. A. Ng, Vagal modulation of cardiac ventricular arrhythmia, *Experimental physiology*, vol. 99, no. 2, pp. 295-299, 2014.

[13] Z. M. Cucunuba, O. Okuwoga, M.-G. Basáñez, and P. Nouvellet, Increased mortality attributed to chagas disease: a systematic review and meta-analysis, *Parasites & vectors*, vol. 9, no. 1, p. 42, 2016.

[14] M. C. P. Nunes, A. Beaton, H. Acquatella, C. Bern, A. F. Bolger, L. E. Echeverria, W. O. Dutra, J. Gascon, C. A. Morillo, J. Oliveira-Filho, et al., Chagas cardiomyopathy: an update of current clinical knowledge and management: a scientific statement from the American heart association, *Circulation*, vol. 138, no. 12, pp. e169-e209, 2018.

[15] C. P. Dobson, A. Kim, and M. Haigney, QT variability index, *Progress in cardiovascular diseases*, vol. 56, no. 2, pp. 186-194, 2013.

[16] E. M. Cherry, F. H. Fenton, and R. F. Gilmour Jr, Mechanisms of ventricular arrhythmias: a dynamical systems-based perspective, *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 302, no. 12, pp. H2451-H2463, 2012.

[17] W. B. Nicolson, G. P. McCann, M. I. Smith, A. J. Sandilands, P. J. Stafford, F. S. Schlindwein, N. J. Samani, and G. A. Ng, Prospective evaluation of two novel ECG-based restitution biomarkers for prediction of sudden cardiac death risk in ischaemic cardiomyopathy, *Heart*, vol. 100, no. 23, pp. 1878-1885, 2014.

[18] J. Ramírez, M. Orini, A. Mincholé, V. Monasterio, I. Cygankiewicz, A. Bayés de Luna, J. P. Martínez, E. Pueyo, and P. Laguna, T-wave morphology restitution predicts sudden cardiac death in patients with chronic heart failure, *Journal of the American Heart Association*, vol. 6, no. 5, p. e005310, 2017.

[19] A. Rassi Jr, A. Rassi, W. C. Little, S. S. Xavier, S. G. Rassi, A. G. Rassi, G. G. Rassi, A. Hasslocher-Moreno, A. S. Sousa, and M. I. Scanavacca, Development and validation of a risk score for predicting death in Chagas' heart disease, *New England Journal of Medicine*, vol. 355, no. 8, pp. 799-808, 2006.

[20] G. J. Volpe, H. T. Moreira, H. S. Trad, K. C. Wu, M. F. Braggion-Santos, M. K. Santos, B. C. Maciel, A. Pazin-Filho, J. A. Marin-Neto, J. A. Lima, et al., Left ventricular scar and prognosis in chronic chagas cardiomyopathy, *Journal of the American College of Cardiology*, vol. 72, no. 21, pp. 2567-2576, 2018.

[21] G. Salles, S. Xavier, A. Sousa, A. Hasslocher-Moreno, and C. Cardoso, Prognostic value of QT interval parameters for mortality risk stratification in Chagas' disease: results of a long-term follow-up study, *Circulation*, vol. 108, no. 3, pp. 305-312, 2003.

[22] T. Senra, B. M. Ianni, A. C. Costa, C. Mady, M. Martinelli-Filho, R. Kalil-Filho, and C. E. Rochitte, Long-term prognostic value of myocardial fibrosis in patients with chagas cardiomyopathy, *Journal*

of the American College of Cardiology, vol. 72, no. 21, pp. 2577-2587, 2018.

- [23] A. Brayner, J. Maria Monteiro, and J. Paulo Madeiro, PeMo-EC: An Intelligent, Pervasive and Mobile Platform for ECG Signal Acquisition, Processing, and Pre-Diagnostic Extraction, in *Intelligent Pervasive Computing Systems for Smarter Healthcare*. Wiley Online Library, 2019, pp. 37-65.
- [24] J. P. do Vale Madeiro, P. C. Cortez, J. L. Salinet Jr, R. C. Pedrosa, J. M. da Silva Monteiro Filho, and A. R. A. Brayner, Classical and modern features for interpretation of ECG signal, in *Developments and Applications for ECG Signal Processing*. Elsevier, 2019, pp. 1-28.
- [25] J. P. do Vale Madeiro, J. M. da Silva Monteiro Filho, and P. R. F. Rodrigues, Techniques for QRS complex detection, in *Developments and Applications for ECG Signal Processing*. Elsevier, 2019, pp. 89-118.
- [26] J. P. do Vale Madeiro, J. M. da Silva Monteiro Filho, and P. R. F. Rodrigues, Delineation of QRS complex: Challenges for the development of widely applicable algorithms, in *Developments and Applications for ECG Signal Processing*. Elsevier, 2019, pp. 119-139.
- [27] J. P. do Vale Madeiro, E. M. B. E. dos Santos, P. C. Cortez, J. H. da Silva Felix, and F. S. Schlindwein, Evaluating Gaussian and Rayleigh-based mathematical models for T and P-waves in ECG, *IEEE Latin America Transactions*, vol. 15, no. 5, pp. 843-853, 2017.
- [28] J. P. do Vale Madeiro, P. C. Cortez, J. M. da Silva Monteiro Filho, and P. R. F. Rodrigues, Mathematical modeling of T-wave and P-wave: A robust alternative for detecting and delineating those waveforms, in *Developments and Applications for ECG Signal Processing*. Elsevier, 2019, pp. 141-167.
- [29] J. P. Madeiro, W. B. Nicolson, P. C. Cortez, J. A. Marques, C. R. Vazquez-Seisdedos, N. Elangovan, G. A. Ng, and F. S. Schlindwein, New approach for T-wave peak detection and T-wave end location in 12-lead paced ECG signals based on a mathematical model, *Medical Engineering & physics*, vol. 35, no. 8, pp. 1105-1115, 2013.
- [30] J. P. Madeiro, P. C. Cortez, J. A. Marques, C. R. Seisdedos, and C. R. Sobrinho, An innovative approach of QRS segmentation based on first-derivative, Hilbert and Wavelet transforms, *Medical Engineering & Physics*, vol. 34, no. 9, pp. 1236-1246, 2012.
- [31] J. P. Madeiro, P. C. Cortez, F. I. Oliveira, and R. S. Siqueira, A new approach to QRS segmentation based on wavelet bases and adaptive threshold technique, *Medical Engineering and Physics*, vol. 29, no. 1, pp. 26-37, 2007.
- [32] J. P. do Vale Madeiro, J. A. L. Marques, T. Han, and R. C. Pedrosa, Evaluation of mathematical models for QRS feature extraction and QRS morphology classification in ECG signals, *Measurement*, vol. 156, p. 107580, 2020.
- [33] N. M. M. Nascimento, L. B. Marinho, S. A. Peixoto, J. P. do Vale Madeiro, V. H. C. de Albuquerque, and P. P. Rebouças Filho, Heart arrhythmia classification based on statistical moments and structural co-occurrence, *Circuits, Systems, and Signal Processing*, pp. 1-20, 2019.
- [34] L. B. Marinho, N. de MM Nascimento, J. W. M. Souza, M. V. Gurgel, P. P. Rebouças Filho, and V. H. C. de Albuquerque, A novel electrocardiogram feature extraction approach for cardiac arrhythmia classification, *Future Generation Computer Systems*, vol. 97, pp. 564-577, 2019.
- [35] A. Lyon, A. Mincholé, J. P. Martínez, P. Laguna, and B. Rodriguez, Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances, *Journal of The Royal Society Interface*, vol. 15, no. 138, p. 20170821, 2018.
- [36] Y. Özbay, R. Ceylan, and B. Karlik, A fuzzy clustering neural network architecture for classification of ECG arrhythmias, *Computers in Biology and Medicine*, vol. 36, no. 4, pp. 376-388, 2006.
- [37] H. Kawazoe, Y. Nakano, H. Ochi, M. Takagi, Y. Hayashi, Y. Uchimura, T. Tokuyama, Y. Watanabe, H. Matsumura, S. Tomomori, et al., Risk stratification of ventricular fibrillation in brugada syndrome using non-invasive scoring methods, *Heart Rhythm*, vol. 13, no. 10, pp. 1947-1954, 2016.
- [38] R. Bailon, J. Mateo, S. Olmos, P. Serrano, J. García, A. Del Río, I. Ferreira, and P. Laguna, Coronary artery disease diagnosis based on exercise electrocardiogram indexes from

repolarisation, depolarisation and heart rate variability, Medical and Biological Engineering and Computing, vol. 41, no. 5, pp. 561-571, 2003.

[39] P. R. F. Rodrigues, J. M. da Silva Monteiro Filho, and J. P. do Vale Madeiro, The issue of automatic classification of heartbeats, in Developments and Applications for ECG Signal Processing. Elsevier, 2019, pp. 169-193.

[40] S. Consoli, D. R. Recupero, and M. Petkovic, Data Science for Healthcare. Springer, 2019.

[41] V. N. Vapnik, An overview of statistical learning theory, IEEE Transactions on Neural Networks, vol. 10, no. 5, pp. 988-999, 1999.

[42] L. Breiman, Random forests, Machine learning, vol. 45, no. 1, pp.5-32, 2001.

8. Plano de Atividades

| Mês | Bolsista 1 | Bolsista 2 |
|-----|---|---|
| 1 | Revisão bibliográfica acerca de métodos de extração de atributos de sinais ECG com foco na detecção de insuficiência cardíaca | Desenvolvimento de sistema com interface visual para rotulação |
| 2 | Seleção de algoritmos de detecção de pontos fiduciais do sinal ECG e de análise da variabilidade da frequência cardíaca | Desenvolvimento de sistema com interface visual para rotulação |
| 3 | Implementação de algoritmos no ambiente Python | Revisão bibliográfica acerca de métodos de aprendizagem não-supervisionada |
| 4 | Experimentos computacionais sobre a base de sinais de pacientes portadores de doença de Chagas | Implementação de algoritmo de clustering sobre intervalos de sinal ECG do repositório disponibilizado ao projeto |
| 5 | Experimentos computacionais sobre a base de sinais de pacientes portadores de doença de Chagas | Implementação de algoritmo de clustering sobre intervalos de sinal ECG do repositório disponibilizado ao projeto |
| 6 | Construção de datasets com métricas extraídas de sinal variando-se o comprimento do sinal para cada dataset: 1 minuto, 5 minutos, 10 minutos | Testes de validação do processo de recomendação de intervalos para rotulação a partir de aprendizagem não-supervisionada |
| 7 | Implementação e teste dos classificadores: regressão logística, KNN, Random Forest, Redes Neurais MLP, SVM, Naive Bayes, Gradient Boosting, Extreme Gradient Boosting | Execução de rotulação a partir da execução do sistema desenvolvido por pesquisadores da área de medicina cardiológica |
| 8 | Implementação e teste dos classificadores: regressão logística, KNN, Random Forest, Redes Neurais MLP, SVM, Naive Bayes, Gradient Boosting, Extreme Gradient Boosting | Execução de rotulação a partir da execução do sistema desenvolvido por pesquisadores da área de medicina cardiológica |
| 9 | Testes de desempenho dos classificadores e comparação entre os diferentes cenários | Implementação de algoritmos de extração de parâmetros do ECG |
| 10 | Implementação de modelos que utilizam o próprio sinal ECG cru como entrada: CNN-1D e redes LSTM | Implementação de classificadores para detecção de eventos adversos junto aos sinais ECG: regressão logística, KNN, Random Forest, Redes Neurais MLP, SVM, Naive Bayes, Gradient Boosting, Extreme Gradient Boosting |

| | | |
|----|---|---|
| 11 | Implementação de modelos que utilizam o próprio sinal ECG cru como entrada: CNN-1D e redes LSTM | Implementação de classificadores para detecção de eventos adversos junto aos sinais ECG: regressão logística, KNN, Random Forest, Redes Neurais MLP, SVM, Naive Bayes, Gradient Boosting, Extreme Gradient Boosting |
| 12 | Redação de artigo científico e relatório de atividades | Redação de artigo científico e relatório de atividades |