

Universidade Federal do Maranhão - UFMA
Centro de Ciências Exatas e Tecnológicas - CCET
Programa de Pós-Graduação em Engenharia de Eletricidade - PPEE
Laboratório de Processamento da Informação Biológica - PIB

Alfredo Costa Oliveira Junior

Classificação de Fibrilação Atrial utilizando Curtose

São Luís - MA

2017

Universidade Federal do Maranhão - UFMA
Centro de Ciências Exatas e Tecnológicas - CCET
Programa de Pós-Graduação em Engenharia de Eletricidade - PPEE
Laboratório de Processamento da Informação Biológica - PIB

Alfredo Costa Oliveira Junior

Classificação de Fibrilação Atrial utilizando Curtose

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Eletricidade da UFMA, como registro para obtenção do grau de Mestre em Engenharia de Eletricidade.

Orientador: Allan Kardec Duailibe Barros Filho

São Luís - MA

2017

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Núcleo Integrado de Bibliotecas/UFMA

Costa Oliveira Junior, Alfredo.

Classificação de Fibrilação Atrial utilizando Curtose /
Alfredo Costa Oliveira Junior. - 2017.

57 f.

Orientador(a): Allan Kardec Duailibe Barros Filho.

Dissertação (Mestrado) - Programa de Pós-graduação em
Engenharia de Eletricidade/ccet, Universidade Federal do
Maranhão, São Luís - MA, 2017.

1. Apoio ao diagnóstico. 2. Estatística de alta
ordem. 3. Estrutura do sinal de ECG. 4. Sinais esparsos.
I. Kardec Duailibe Barros Filho, Allan. II. Título.

Alfredo Costa Oliveira Junior

Classificação de Fibrilação Atrial utilizando Curtose

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Eletricidade da UFMA, como registro para obtenção do grau de Mestre em Engenharia de Eletricidade.

Trabalho aprovado. São Luís - MA, 16 de fevereiro de 2017:

**Prof. Dr. Allan Kardec Duailibe Barros
Filho**
DEE/UFMA

**Prof. Dr. Francisco das Chagas de
Souza**
DEE/UFMA

Prof. Dr. Fausto Lucena de Oliveira
Universidade CEUMA

São Luís - MA
2017

*Este trabalho é dedicado à Deus e a minha família: Alfredo, Josicicleide e Allas Jony.
Também dedico aos meus amigos e a todas as crianças adultas que, quando
pequenas, sonharam em se tornar cientistas.*

AGRADECIMENTOS

Agradeço ao Professor Allan Kardec por todos os questionamentos e motivações no processo de pesquisa.

Aos amigos do Lab. PIB, Jonathan Queiroz, José de Ribamar, Marcus Vinícius, Luís Oliveira, Caio, Gean, Letícia, Carlos, prof. Ewaldo Eder, Prof. André Cavalcante, Aline, Guilherme e Marta por todos os momentos de luta, aprendizado e companheirismo.

*“Não vos amoldeis às estruturas deste mundo,
mas transformai-vos pela renovação da mente,
a fim de distinguir qual é a vontade de Deus:
o que é bom, o que Lhe é agradável, o que é perfeito.
(Bíblia Sagrada, Romanos 12, 2)*

RESUMO

A Fibrilação atrial (FA) é uma das arritmias cardíacas mais comuns em todo o mundo. Por isso, amplos são os esforços para implementar sistemas que apoiem o diagnóstico de FA. A principal forma não invasiva de avaliar a saúde cardíaca, é através da análise do sinal de eletrocardiograma (ECG), o qual representa a atividade elétrica do músculo cardíaco, e possui marcações temporais características: as ondas P, Q, R, S e T. Alguns autores utilizaram técnicas de filtragem, análise estatística e até redes neurais para detectar FA com base no intervalo RR, que é dado pela diferença temporal entre os picos da onda R. Entretanto, a análise do intervalo RR permite avaliar apenas as variações que ocorrem na onda R do sinal de ECG, não permitindo avaliar, por exemplo, as alterações na onda P, provocadas pela FA. Diante disso, propõe-se caracterizar a amplitude do sinal de ECG, a fim de classificar pacientes com FA e saudáveis. Na metodologia proposta, o sinal de ECG, foi analisado por meio das seguintes estatísticas: variância, assimetria e curtose. Para avaliar o classificador proposto, usou-se sinais obtidos das bases de dados MIT-BIH Atrial Fibrillation e MIT-BIH Normal Sinus Rhythm referentes aos pacientes com FA e com ritmo cardíaco normal, respectivamente. Dentre as estatísticas analisadas, a curtose foi a que apresentou resultados superiores em termos de sensibilidade ($Se = 100\%$), especificidade ($Sp = 88,33\%$) e acurácia ($Ac = 91,33\%$). Esses resultados são de se esperar pelo fato de que a curtose é uma medida de não-gaussianidade e que o sinal de ECG possui distribuição esparsa. A metodologia proposta também requer um número menor de etapas de pré-processamento, e sua simplicidade permite implementações em sistemas embarcados que apoiarão o diagnóstico clínico.

Palavras-chave: Estrutura do sinal de ECG, sinais esparsos, estatística de alta ordem, apoio ao diagnóstico.

ABSTRACT

Atrial fibrillation(AF) is one of the most common cardiac arrhythmias worldwide. Thus, there are ample efforts to implement AF diagnosis systems. The main noninvasive way to assess cardiac health is through electrocardiogram (ECG) signal analysis, which represents the electrical activity of the cardiac muscle, and has characteristic temporal markings: P, Q, R, S and T waves. Some authors use filtering techniques, statistical analysis and even neural networks for detecting AF based on the RR interval, that is given by the temporal difference between the peaks of the R wave. However, analyses of the RR interval allows for evaluating changes occurring only in the R wave of the ECG signal, not allowing to assess, for example, variations in the P wave provoked by the AF. In face of that, we propose characterize the ECG signal amplitude aiming at classifying both healthy and AF patients. The ECG signal was analyzed in the proposed methodology through the following statistics: variance, asymmetry, and kurtosis. Herein, we use the MIT-BIH Atrial Fibrillation and MIT-BIH Normal Sinus Rhythm database signals to evaluate AF and normal heartbeat intervals. Our study shown that kurtosis outperformed variance and asymmetry with respect to sensibility ($Se = 100\%$), specificity ($Sp = 88.33\%$) and accuracy ($Ac = 91.33\%$). The results were expected since kurtosis is a non-Gaussian measure and the ECG signal has sparse distribution. The proposed methodology also requires a lower number of pre-processing stages, and its simplicity allows for implementations in imbedded systems supporting the clinical diagnosis.

Keywords: Structure of the ECG signal, Sparse signals, Statistical high order, Diagnostic support.

LISTA DE ILUSTRAÇÕES

<p>Figura 1 – Comparativo entre ciclo cardíaco e sinal de eletrocardiograma (ECG). Em que a onda P representa o impulso elétrico que causa a contração dos átrios (Fase 1); o complexo QRS (formado por três ondas: Q, R e S) representa o impulso da contração dos ventrículos (Fase 2); e a onda T corresponde à recuperação elétrica dos ventrículos quando estes voltam para repouso (Fase 3). Fonte: Adaptado de http://cienciacognitiva.planetaclix.pt/ECG.html.</p>	16
<p>Figura 2 – Atividade elétrica de um coração saudável e um coração com FA. (a) Coração e ECG de um paciente saudável. (b) Coração e ECG de paciente com FA.</p>	16
<p>Figura 3 – Comparação entre sinais de ECG com FA e Ritmo Sinusal Normal (RSN). (a) Exemplo de registros de ECG de paciente saudável (RSN). (b) Exemplo de registro de ECG de paciente com fibrilação atrial.</p>	20
<p>Figura 4 – Comparação entre sinais de ECG com FA e Ritmo Sinusal Normal (RSN). (a) Histograma da base de dados MIT-BIH RSN. (b) Histograma da base de dados MIT-BIH FA.</p>	25
<p>Figura 5 – Metodologia proposta para classificação de FA. (a) Aquisição dos sinais de ECG da BD MIT-BIH AF e MIT-BIH NSR. (b) Na etapa de pré-processamento, acontece a detecção dos picos R. (c) Os sinais de ECG são janelados e caracterizados estatisticamente. (d) Os registros de ECG com FA e saudáveis são classificados.</p>	28
<p>Figura 6 – Representação esquemática do algoritmo de detecção de picos R. Fonte: Adaptado de Lucena (2006).</p>	30
<p>Figura 7 – Representação esquemática do intervalo RR, de modo que $RR = [b_1, b_2, \dots, b_m]$, e $b_m = t_m - t_{m-1}$ representa o m-ésimo batimento e t_m é o instante de tempo do m-ésimo pico R. Fonte: Adaptado de Lucena (2006).</p>	31
<p>Figura 8 – Ilustração do primeiro segmento do sinal de ECG, ou seja, $x(t)$ nos instantes $t_1 = 98ms$ e $t_0 = 350ms$.</p>	31
<p>Figura 9 – Exemplo de análise e caracterização de um registro de ECG. Em que e_n é a n-ésima estimação (caracterização) através de alguma estatística (tais como, média, variância, assimetria ou curtose) e E é a estimativa ou caracterização do sinal completo através de alguma estatística, como a média e a mediana.</p>	33
<p>Figura 10 – Resultado para o classificador proposto. (a) Utilizando a variância. (b) Utilizando a assimetria. (c) Utilizando a curtose.</p>	36

Figura 11 – Resultados do classificador proposto. Em azul a caracterização para os sinais de ECG saudáveis. Em vermelho a caracterização dos sinais de ECG com FA. Em preto tem-se o limiar e intervalo de confiança representando a função de decisão do método de classificação proposto.	37
Figura 12 – Ilustração de um boxplot, em que (A) representa um valor atípico (ou outlier, valor que está além do traço superior ou inferior); (B) representa o valor máximo (traço superior); (C) representa o terceiro quartil; (D) representa o Segundo quartil; (E) representa o primeiro quartil; (F) representa o valor mínimo (traço inferior).	48
Figura 13 – Dispersão dos grupos de pacientes saudáveis e com FA. a) Análise baseada apenas nos batimentos (ver Eq. 4.1). b) Análise utilizando a variância. c) Análise utilizando a assimetria. d) Análise utilizando a curtose.	50
Figura 14 – Resultado para o classificador proposto utilizando a média.	51

LISTA DE TABELAS

Tabela 1 – Lista dos registros das bases de dados adotadas (MOODY; MARK, 1983; GOLDBERGER et al., 2000).	28
Tabela 2 – Análise da dispersão dos grupos de pacientes saudáveis (com ritmo sinusal normal) e com FA em diferentes estatísticas (variância, assimetria e curtose), em que o limiar superior é o máximo valor de caracterização (E) do grupo, limiar inferior é o menor valor de caracterização do grupo e o Δ é a diferença entre as medianas do grupo de pessoas saudáveis e do grupo com FA.	35
Tabela 3 – Resultados do método de classificação proposto utilizando variância, assimetria e curtose para um grupo de 36 indivíduos, sendo 18 saudáveis e 18 com FA.	37
Tabela 4 – Comparação de desempenho avaliadas para a base de dados MIT-BIH Atrial Fibrillation (AFDB). AFDB ₁ corresponde a base de dados (BD) sem os registros 00735 e 03665, AFDB ₂ a BD sem 04936 e 05091, e AFDB _b corresponde a BD sem os registros 04746, 04908, 05261, 07879, 07910, 00735 e 03665.	39

LISTA DE ABREVIATURAS E SIGLAS

DB	Database
ECG	Eletrocardiograma
FA	Fibrilação atrial
KS	Kolmogorov–Smirnov
IC	insuficiência cardíaca
MIT	Massachusetts Institute of Technology
NSR	Normal Sinus Rhythm (rítmo cardíaco de uma pessoa saudável)

SUMÁRIO

1	INTRODUÇÃO	15
2	OBJETIVOS	19
2.1	Objetivo geral	19
2.2	Objetivos específicos	19
3	FIBRILAÇÃO ATRIAL	20
3.1	Introdução	20
3.2	Causas e condições associadas	21
3.2.1	Causas reversíveis de FA	21
3.2.2	FA sem doença cardíaca associada	22
3.2.3	Situações associadas à FA	22
3.2.4	FA associada à doença cardíaca	22
3.2.5	FA familiar	22
3.3	Diagnóstico clínico	23
3.4	Técnicas de detecção de FA	23
3.4.1	Esperança e momentos	23
3.4.2	Histograma	25
3.4.3	Teste Kolmogorov–Smirnov	25
3.4.4	Entropia	26
4	MATERIAIS E MÉTODOS	28
4.1	Base de Dados	28
4.2	Detecção de picos R e caracterização do sinal de ECG pela amplitude	29
4.3	Janelamento do sinal de ECG	32
4.4	Classificador de FA (C_{FA_m})	33
4.5	Avaliação	34
5	RESULTADOS	35
5.1	Análise Estatística da amplitude do sinal de ECG	35
5.2	Classificação da FA	37
6	DISCUSSÃO	38
7	CONCLUSÃO	40
8	CONTRIBUIÇÃO E PUBLICAÇÃO DE RESULTADOS	41

9	TRABALHO FUTUROS	42
	Referências	43
	APÊNDICES	47
	APÊNDICE A – BOXPLOT	48
A.1	Determinação dos outliers ou valores atípicos	49
	APÊNDICE B – DISPERSÃO DOS GRUPOS SAUDÁVEL E COM FA	50
	APÊNDICE C – CLASSIFICADOR PROPOSTO UTILIZANDO A MÉ-	
	DIA	51
	APÊNDICE D – CÓDIGO EM MATLAB	52

1 INTRODUÇÃO

As doenças cardíacas são a primeira causa de morte nos países desenvolvidos. Todos os anos, cerca de 17 milhões de pessoas no mundo morrem de doenças relacionadas ao coração, em particular devido a ataques cardíacos e acidentes vasculares cerebrais (WHO et al., 2016).

Estima-se que, no Brasil, mais de 300 mil pessoas por ano sejam acometidas de morte súbita devido às arritmias cardíacas (SOBRAC, 2016). Deste total, apenas 1,5% dos pacientes que sofrem da doença recebem os cuidados necessários, isso falando de cidade como São Paulo, uma das metrópoles mais bem equipadas no país.

A arritmia cardíaca são alterações elétricas que provocam modificações no ritmo das batidas do coração. As arritmias cardíacas podem ser classificadas, conforme (LUCIANO et al., 2011), de acordo com a sua origem, como arritmias ventriculares (originadas nos ventrículos) ou arritmias supra-ventriculares (originadas em áreas do coração acima dos ventrículos, habitualmente nos átrios). Podem igualmente ser classificadas de acordo com o seu efeito sobre a frequência cardíaca, em que a bradicárdia indica uma frequência cardíaca inferior a 60 batimentos por minuto e a taquicárdia indica uma frequência cardíaca superior a 100 batimentos por minuto (LUCIANO et al., 2011).

A atividade elétrica do músculo cardíaco pode ser analisada de forma não invasiva por meio do eletrocardiograma (ECG), ilustrado na Figura 1. Há cinco ondas que correspondem aos impulsos das atividades elétricas do coração. A onda P representa o impulso elétrico que causa a contração dos átrios (Fase 1). O complexo QRS (formado por três ondas: Q, R e S) representa o impulso da contração dos ventrículos (Fase 2) e a onda T corresponde à recuperação elétrica dos ventrículos quando estes voltam para repouso (Fase 3). Para fins de avaliação clínica, as deformações dessas ondas representam informação essencial do estado de saúde do coração (ASHLEY; NIEBAUER, 2004).

Um das arritmias mais prevalentes em todo o mundo, principalmente, quando se fala de população idosa, é a fibrilação atrial (FA) (SOBRAC, 2014). A FA é uma arritmia supraventricular em que ocorre uma completa desorganização na atividade elétrica atrial, fazendo com que os átrios percam sua capacidade de contração, não gerando sístole atrial (conforme Figura 2, em que Figura 2a ilustra a atividade elétrica de um coração saudável e Figura 2b ilustra um atividade elétrica de um coração com FA). Essa desorganização elétrica é tamanha que inibe o nó sinusal enquanto a FA persistir.

A FA é uma arritmia que tem características muito peculiares com relação as

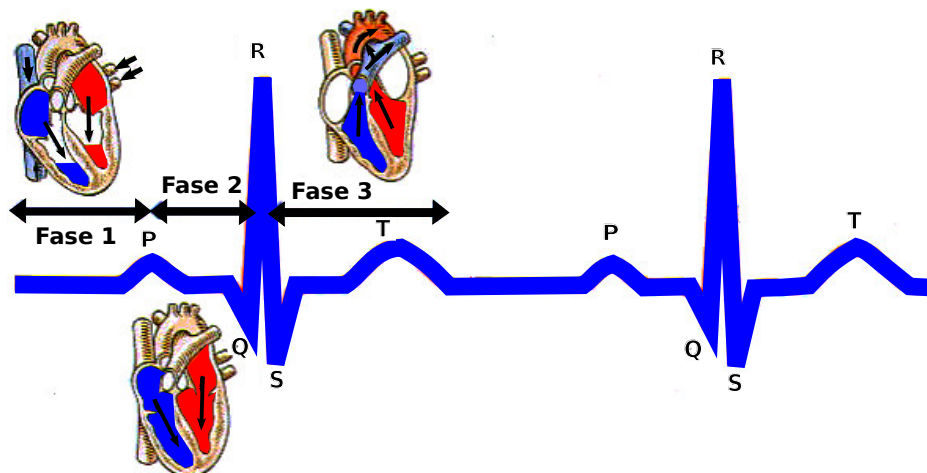


Figura 1 – Comparativo entre ciclo cardíaco e sinal de eletrocardiograma (ECG). Em que a onda P representa o impulso elétrico que causa a contração dos átrios (Fase 1); o complexo QRS (formado por três ondas: Q, R e S) representa o impulso da contração dos ventrículos (Fase 2); e a onda T corresponde à recuperação elétrica dos ventrículos quando estes voltam para repouso (Fase 3). Fonte: Adaptado de <http://cienciacognitiva.planetaclix.pt/ECG.html>.

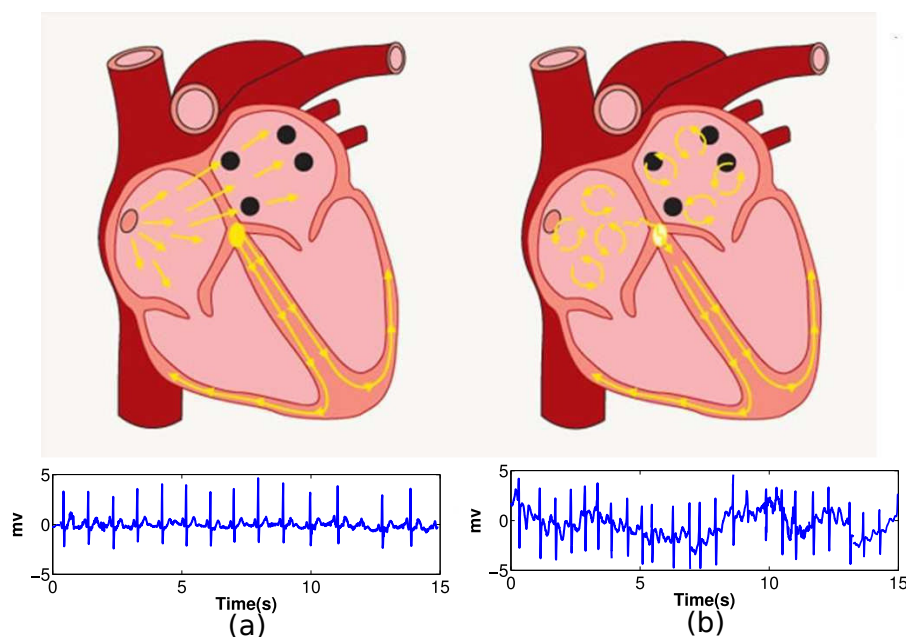


Figura 2 – Atividade elétrica de um coração saudável e um coração com FA. (a) Coração e ECG de um paciente saudável. (b) Coração e ECG de paciente com FA.

demais arritmias. A FA atinge cerca de 1,5% a 2% da população mundial e a tendência é aumentar esse percentual devido à maior longevidade da população, tendo em vista que essa doença é mais prevalente nas pessoas idosas (MAGALHÃES et al., 2016).

Segundo estimativas da Sociedade Brasileira de Arritmias Cardíacas, em 2016, cerca de 5% a 10% dos brasileiros tiveram esse tipo de arritmia, que é a segunda maior causa de mortes em todo mundo (SOBRAC, 2016).

Estima-se que a FA seja responsável por 33% de todas as internações por arritmias (VIANA; WHITAKER, 2009). O envelhecimento populacional e o avanço da medicina permitiram o aumento da prevalência de cardiopatias crônicas na população, aumentando expressivamente o número de internações hospitalares por FA nos últimos 20 anos (FRIBERG et al., 2003; SAAD; JUNIOR; CAMANHO, 2012). Estima-se que no Brasil existam em torno de 1,5 milhões de pacientes com FA e que essa população se correlacione com a pirâmide etária (SAAD; JUNIOR; CAMANHO, 2012).

A média de idade dos pacientes com FA é de 75 anos sendo que 70% da população com FA está na faixa etária dos 65 aos 85 anos. A prevalência de FA ajustada para a idade é maior em homens do que em mulheres, entretanto, acima dos 75 anos, 60% dos pacientes com FA são mulheres.

A FA compromete os átrios porque na sua ocorrência há uma completa desorganização da atividade elétrica atrial, fazendo com que os átrios deixem de se contrair, passando a vibrarem descompasadamente com relação aos ventrículos.

No eletrocardiograma, a ausência de despolarização atrial organizada, devido a FA, reflete-se com a substituição das ondas P, características do ritmo sinusal, por um tremor de alta frequência da linha de base do eletrocardiograma que varia em sua forma e amplitude (MAGALHÃES et al., 2016; SAAD; JUNIOR; CAMANHO, 2012), conforme Figura 2b.

Devido ao número de casos clínicos, a AF tem despertado grande interesse nas diversas áreas do conhecimento, tornando o seu estudo cada vez mais importante. Vários autores (Andrius Petrėnas, Marozas Vaidotas, 2015; CARVALHO et al., 2012; LEE et al., 2013; LOGAN; HEALEY, 2005; ZHOU et al., 2014) utilizaram técnicas de filtragem, análise estatística e até redes neurais aplicadas ao sinal de ECG, a fim de caracterizar e detectar essa doença. Tais trabalhos lidam com a detecção de FA em segmentos do sinal de ECG, assumindo que o doente já foi diagnosticado com a doença, de modo que o objetivo é o acompanhamento clínico.

Neste trabalho propõe-se um método para auxiliar o diagnóstico preditivo de FA, isto é, na fase de predição da doença, como um método computacional para apoio a tomada de decisão do médico.

A metodologia proposta em (OLIVEIRA et al., 2016; OLIVEIRA et al., 2017) se baseia na análise estatística da amplitude do sinal de ECG utilizando a curtose. A curtose é uma estatística de alta ordem, que mede a não-gaussianidade de um conjunto de dados e é sensível a pequenas variações (SHARMA; DANDAPAT; MAHANTA, 2010).

Autores, como (CHANG et al., 2010; COUCEIRO et al., 2015), utilizaram a curtose, para caracterizar o intervalo RR ($RR = [RR_1, RR_2, RR_3, \dots, RR_n]$, em que RR_n é o intervalo de tempo entre o pico da onda R_n e o pico da onda R_{n+1}). No entanto, utilizar o intervalo RR se mostra ineficiente em avaliar as alterações que acontecem nas demais ondas do sinal de ECG, as quais são Q, R, S e T (ASHLEY; NIEBAUER, 2004), e tais alterações podem estar relacionadas em boa parte dos casos com o mal funcionamento do coração (ZIMERMAN et al., 2009). Diante disso, a metodologia proposta detecta a presença ou ausência de FA com base nas alterações que acontecem na amplitude do sinal de ECG. A análise dos resultados é realizada com base em sinais de ECG em pacientes saudáveis e com FA. Já a avaliação da metodologia se dá pelas métricas de especificidade, sensibilidade e acurácia.

2 OBJETIVOS

2.1 Objetivo geral

Desenvolver um classificador de fibrilação atrial utilizando a curtose a fim de apoiar o diagnóstico clínico.

2.2 Objetivos específicos

- Caracterizar sinais de ECG de pacientes saudáveis e com fibrilação atrial utilizando curtose;
- Propor um classificador de sinais de ECG que diferencie pacientes saudáveis de pacientes com fibrilação atrial;
- Avaliar a metodologia proposta a partir das métricas de sensibilidade, especificidade e acurácia.

3 FIBRILAÇÃO ATRIAL

Nas próximas subseções serão apresentados os seguintes tópicos, respectivamente: introdução sobre a fibrilação atrial, causas e condições associadas, diagnóstico clínico e, finalmente, técnicas de detecção de FA.

3.1 Introdução

A fibrilação atrial (FA) é uma arritmia supraventricular que se caracteriza pela perda da capacidade de contração atrial, isto é, perda da sístole atrial (SAAD; JUNIOR; CAMANHO, 2012; JANUARY et al., 2014; ZIMERMAN et al., 2009). Durante 30s, período médio referente a um episódio de FA, o sinal de FA é caracterizado como um sinal de frequência rápida e irregular (HURST et al., 1964).

A FA está entre as principais causas de morte no mundo provocadas por doenças cardíacas. A incidência dessa doença aumenta significativamente com o aumento da idade e sua prevalência é de 1,5% a 2,5% da população mundial e, geralmente, são pessoas na faixa etária acima de 75 anos (FERREIRA; SANTOS, 2015; SAAD; JUNIOR; CAMANHO, 2012; ZIMERMAN et al., 2009). Estima-se que a FA é responsável por 33% de todas as hospitalizações por arritmia no mundo, tornando-se de grande importância o desenvolvimento de técnicas e métodos que auxiliam no seu diagnóstico (JANUARY et al., 2014; ZIMERMAN et al., 2009).

A atividade elétrica do músculo cardíaco pode ser analisada de forma não invasiva por meio do ECG. Há três estruturas que correspondem aos impulsos das atividades elétricas do coração: onda P, complexo QRS e onda T (ASHLEY; NIEBAUER, 2004). A onda P representa o impulso elétrico que causa a contração dos átrios. Com a FA, o ciclo cardíaco normal é alterado, tal alteração se manifesta no sinal de ECG pela alteração da onda P, e ondulação da linha de base. A Figura 3 ilustra um registro de ECG de paciente saudável e de um paciente com FA.

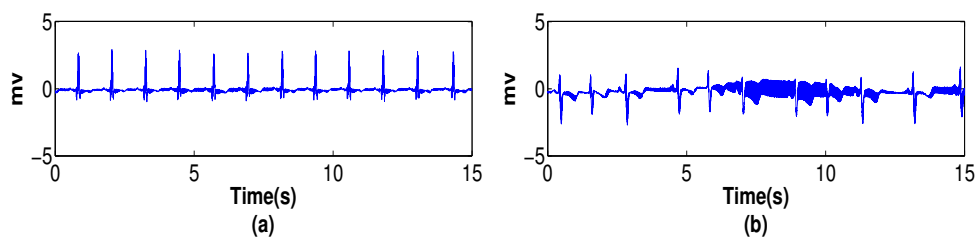


Figura 3 – Comparação entre sinais de ECG com FA e Ritmo Sinusal Normal (RSN). (a) Exemplo de registros de ECG de paciente saudável (RSN). (b) Exemplo de registro de ECG de paciente com fibrilação atrial.

Vários trabalhos propuseram detectores de FA baseados na análise das irregularidades do intervalo RR (irregularidade do intervalo RR, que é uma série temporal formado pelos tempos entre os sucessivos picos R ou batimentos, definida como $RR = [RR_1, RR_2, RR_3, \dots, RR_n]$, em que RR_n é o intervalo de tempo entre o pico da onda R_n e o pico da onda R_{n+1}).

Petrènas et al. (2015) propuseram um detector de FA baseado na teoria de entropia, a fim de apoiar o acompanhamento clínico e obtiveram 97,1% e 98,3% de sensibilidade e especificidade, respectivamente (Andrius Petrènas, Marozas Vaidotas, 2015). No trabalho de Huag et al. (2011), através da análise do histograma, do desvio padrão e teste Kolmogorov-Smirnov do intervalo RR, foi proposto um limiar para detecção de FA com sensibilidade 96,1% e especificidade de 97,9% (HUANG et al., 2011). Logan e Healey (2005) propuseram um detector de FA robusto a artefatos de movimentação para apoio do telemonitoramento de pacientes. O detector de Logan e Healey teve sensibilidade e especificidade de 96% e 89%, respectivamente (LOGAN; HEALEY, 2005). Tatenó e Glass (2001) apresentaram dois métodos de detecção de FA: um baseado no coeficiente de variação e o outro baseado no teste Kolmogorov-Smirnov para o intervalo RR. Esse último método foi o que apresentou melhores resultados com sensibilidade e especificidade, respectivamente, de 94,4% e 97,2% (TATENO; GLASS, 2001).

3.2 Causas e condições associadas

Vários são os fatores que aumentam o risco de FA, além dos clássicos, como hipertensão, diabetes, doença valvar, infarto do miocárdio e insuficiência cardíaca (IC) (WANG et al., 2003); há também destaque para a influência da obesidade (FROST; HUNE; VESTERGAARD, 2005) e presença de Apneia Obstrutiva do Sono (AOS) (GAMI et al., 2007). Além dessas influências, há também, conforme (SAAD; JUNIOR; CAMANHO, 2012), um importante preditor para o desenvolvimento da FA, que é o tamanho dos átrios. Isso porque, para um mesmo comprimento de onda, quanto maior a quantidade de tecido disponível, maior a elevação da incidência de FA na IC (LI et al., 1999).

Nas subseções a seguir serão apresentadas com mais detalhes algumas das principais causas e condições associadas a FA.

3.2.1 Causas reversíveis de FA

FA pode ser originada durante situações transitórias, pela ingestão de álcool, durante cirurgias, choque elétrico, infarto do miocárdio, embolia pulmonar, entre outras doenças pulmonares (SAAD; JUNIOR; CAMANHO, 2012). Em tais situações, ainda

conforme (SAAD; JUNIOR; CAMANHO, 2012), o tratamento eficaz das causas bases pode ser a única medida necessária para que se restaure o ritmo cardíaco normal.

3.2.2 FA sem doença cardíaca associada

Aproximadamente 25% dos casos de FA ocorrem em pacientes jovens sem cardiopatia estrutural demonstrável (SAAD; JUNIOR; CAMANHO, 2012). Entretanto, pode vir a surgir uma anomalia cardíaca não detectável no primeiro momento, podendo ser detectada com o tempo (BRUGADA et al., 1997).

No caso de idosos, a própria perda de complacência muscular e consequente enrijecimento, podem ser suficientes para permitir o surgimento de FA, e não necessariamente está relacionada a presença de cardiopatia; assim como uma cardiopatia pode não está relacionada com a FA (CAMM et al., 2010; SAAD; JUNIOR; CAMANHO, 2012).

3.2.3 Situações associadas à FA

De acordo com (FROST; HUNE; VESTERGAARD, 2005), a obesidade tem sido considerado um importante fator de risco associado a FA. Alguns trabalhos demonstraram haver correlação direta entre o tamanho do átrio esquerdo e o índice de massa corporal (FOX et al., 2004; WANG; PARISE; LEVY, 2005). Tais estudos demonstram ligação entre obesidade, FA e acidente vascular cerebral (AVC), além de que a perda de peso pode reduzir os riscos relacionados a FA.

3.2.4 FA associada à doença cardíaca

As principais doenças cardíacas associadas a FA são: doença valvular cardíaca, insuficiência cardíaca (IC), doença arterial coronariana e hipertensão arterial sistêmica (MAGALHÃES et al., 2016).

Além disso, a FA pode ser associada a outras arritmias como flutter atrial, bigeminismo atrial, bigeminismo ventricular, trigeminismo ventricular, taquicardia ventricular, taquiarritmias supraventriculares, bradicardia sinusal com ritmo idioventricular. Essa associação é devido a uma relação de causa e efeito entre elas ou, simplesmente, porque os átrios estão doentes e isso possibilita a ocorrência de tais arritmias, sem interdependência entre elas (MOODY; MARK, 1983).

3.2.5 FA familiar

Em algumas famílias essa herança parece estar relacionada a mutações genéticas em sítios cromossômicos específicos (SAAD; JUNIOR; CAMANHO, 2012),

entretanto, há casos em que a FA, estão relacionadas à hipertensão arterial sistêmica, diabetes ou IC de caráter familiar (ELLINOR et al., 2003; DARBAR et al., 2003).

3.3 Diagnóstico clínico

O diagnóstico da arritmia pode ser confirmado pelo registro da FA através do ECG basal, sistema Holter-24h, teste ergométrico, monitor de eventos ou eletrogramas intracavitários dos dispositivos implantáveis (CAMM et al., 2000). Em alguns casos a FA pode ser um achado de exame complementar e o paciente pode estar assintomático. A presença de fatores clínicos como: idade avançada, hipertensão arterial sistêmica (HAS), diabetes mellitus (DM), insuficiência cardíaca congestiva (ICC), infarto agudo do miocárdio (IAM), valvopatia (reumática ou não) e cirurgia cardiotorácica estão associados à maior prevalência de FA (MAGALHÃES et al., 2016).

Fatores de risco como obesidade, síndrome metabólica, apneia do sono, estatura elevada e excesso de atividade física também são considerados fatores para a ocorrência de FA (CAMM et al., 2010; MAGALHÃES et al., 2016; SAAD; JUNIOR; CAMANHO, 2012).

Para o diagnóstico de síndrome metabólica são necessários pelo menos três dos cinco fatores: obesidade visceral, hipertensão arterial, hipertrigliceridemia, níveis baixos de colesterol HDL (sigla inglesa para Low Density Lipoproteins, que significa lipoproteínas de densidade baixa) e glicemia de jejum $> 110mg/dl$ (SAAD; JUNIOR; CAMANHO, 2012). O estado pró-inflamatório, o estresse oxidativo e a dilatação do átrio esquerdo são fatores comuns da FA, induzindo o processo de remodelamento atrial (MAGALHÃES et al., 2016; SAAD; JUNIOR; CAMANHO, 2012).

3.4 Técnicas de detecção de FA

Nessa subseção serão apresentadas as principais técnicas de detecção de FA utilizadas na literatura: esperança e momentos, análise de histograma, teste Kolmogorov–Smirnov e entropia. Todas as técnicas tem em comum a análise estatística no intervalo RR.

3.4.1 Esperança e momentos

Seja $g(\mathbf{x})$ qualquer quantidade derivada do vetor aleatório \mathbf{x} . A quantidade $g(\mathbf{x})$ pode ser um vetor escalar, ou mesmo uma matriz. A esperança de $g(x)$ é denotada por $E\{g(x)\}$ e definida por

$$E\{g(x)\} = \int_{-\infty}^{\infty} \mathbf{g}(\mathbf{x})p_{\mathbf{x}}d\mathbf{x} \quad (3.1)$$

A integral da Eq. 3.1 é calculada sobre todos os componentes de \mathbf{x} . A operação de integração é aplicada separadamente a cada componente do vetor ou elemento da matriz, produzindo como resultado outro vetor ou matriz do mesmo tamanho.

A esperança tem algumas propriedades fundamentais (PAPOULIS; PILLAI, 2002, cap. 8):

- Linearidade: Seja $\mathbf{x}_i, i = 1, \dots, m$ um conjunto de diferentes vetores aleatórios e $a_i, i = 1, \dots, m$ alguns coeficientes escalares não aleatórios. Então,

$$E\left\{\sum_{i=1}^m a_i \mathbf{x}_i\right\} = \sum_{i=1}^m a_i \mathbf{x}_i, \quad (3.2)$$

- Transformação linear: Seja \mathbf{x}_i um vetor aleatório m -dimensional, e \mathbf{A} e \mathbf{B} matrizes não aleatórias $k \times m$ e $m \times l$, respectivamente. Então,

$$E\{\mathbf{A}\mathbf{x}\} = \mathbf{A}E\{\mathbf{x}\}, \quad E\{\mathbf{x}\mathbf{B}\} = E\{\mathbf{x}\}\mathbf{B}, \quad (3.3)$$

- Invariância da transformação: Seja $\mathbf{y} = \mathbf{g}(\mathbf{x})$ uma função vetorial (função a valores vetoriais) de um vetor aleatório \mathbf{x} . Então,

$$\int_{-\infty}^{\infty} \mathbf{y} p_{\mathbf{y}} d\mathbf{y} = \int_{-\infty}^{\infty} \mathbf{g}(\mathbf{x}) p_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}. \quad (3.4)$$

Portanto, $E\{\mathbf{y}\} = E\{\mathbf{g}(\mathbf{x})\}$, mesmo que as integrações sejam realizadas através de diferentes funções de densidade de probabilidade.

Momentos de uma variável aleatória \mathbf{x} são as esperanças típicas utilizadas para caracterizar \mathbf{x} . Eles são obtidos quando $\mathbf{g}(\mathbf{x})$ consiste de produtos de componentes de \mathbf{x} . Em particular, o primeiro momento de um vetor aleatório \mathbf{x} é chamado de vetor média $\mathbf{m}_{\mathbf{x}}$. É definido como a esperança de \mathbf{x} :

$$\mathbf{m}_{\mathbf{x}} = E\{\mathbf{x}\} = \int_{-\infty}^{\infty} \mathbf{x} p_{\mathbf{x}} d\mathbf{x} \quad (3.5)$$

Cada componente de m_{x_i} de um vetor aleatório $\mathbf{m}_{\mathbf{x}}$ é dado por

$$m_{x_i} = E\{x_i\} = \int_{-\infty}^{\infty} x_i p_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} = \int_{-\infty}^{\infty} x_i p_{x_i}(x_i) dx_i \quad (3.6)$$

em que $p_{x_i}(x_i)$ é a densidade marginal da i -ésima componente x_i de \mathbf{x} . Isso ocorre porque as integrais sobre todos os outros componentes de \mathbf{x} se reduzem à unidade devido à definição da densidade marginal (PAPOULIS; PILLAI, 2002, cap. 4).

Autores, como (TATENO; GLASS, 2001; LOGAN; HEALEY, 2005; HUANG et al., 2011), utilizaram o intervalo RR como variável aleatório \mathbf{x} e aplicaram as definições de momentos de primeira e segunda ordem para propor detectores de FA.

3.4.2 Histograma

Histograma é uma representação gráfica (um gráfico de barras verticais ou barras horizontais) da distribuição de frequências de um conjunto de dados (CASTA-NHEIRA, 2008, cap. 2). O histograma pode ser um gráfico por valores absolutos ou frequência relativa ou densidade.

A frequência absoluta (f_i) é o número de observações correspondente a cada classe. A frequência absoluta é, geralmente, chamada apenas de frequência. Já a frequência relativa (f_{ri}) é o quociente entre a frequência absoluta da classe correspondente e a soma das frequências (total observado), isto é,

$$f_{ri} = \frac{f_i}{\sum_j^n f_i} \quad (3.7)$$

onde n representa o número total de observações.

No caso do histograma ser um gráfico por densidade, a frequência relativa do intervalo i , é representada pela área de um retângulo que é colocado acima do ponto médio da classe i . Consequentemente, a área total do histograma (igual a soma das áreas de todos os retângulos) será igual a 1.

A Figura 4 ilustra o histograma de dois conjuntos de dados, um composto por sinais de pacientes saudáveis (Figura 4a) e outro com pacientes com FA (Figura 4b).

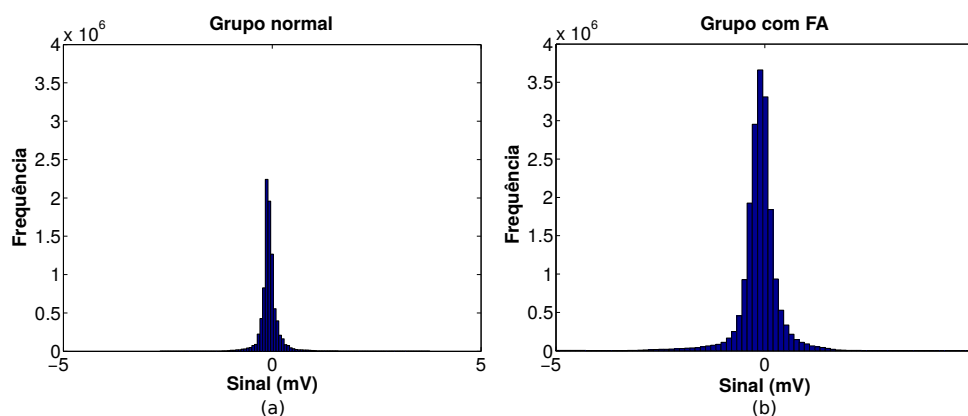


Figura 4 – Comparação entre sinais de ECG com FA e Ritmo Sinusal Normal (RSN). (a) Histograma da base de dados MIT-BIH RSN. (b) Histograma da base de dados MIT-BIH FA.

Pela Figura 4, nota-se que ambos os grupos possuem distribuição não-gaussiana, o que já foi mostrado em (LUCENA et al., 2011). Entretanto, o grupo de pacientes saudáveis possui uma distribuição esparsa maior do que a de pacientes com FA.

3.4.3 Teste Kolmogorov–Smirnov

O teste de Kolmogorov-Smirnov (KS) destina-se a averiguar se uma amostra pode ser considerada como proveniente de uma população com uma determinada

distribuição (TATENO; GLASS, 2001; HUANG et al., 2011). O teste é particularmente indicado para distribuições contínuas.

O Teste KS considera as seguintes hipóteses (PAPOULIS; PILLAI, 2002, cap. 8):

- H_0 : a população tem uma determinada distribuição D ;
- H_1 : a população não tem a determinada distribuição D .

Este teste observa a máxima diferença absoluta entre a função de distribuição acumulada assumida para os dados, na pratica considera-se como referência a Normal (TATENO; GLASS, 2001; HUANG et al., 2011), e a função de distribuição empírica dos dados. Como critério, comparamos esta diferença com um valor crítico, para um dado nível de significância.

Seja a variável aleatória x cuja função de distribuição acumulada contínua F_x desconhecida. A estatística utilizada para o teste é

$$D_n = \sup |F(x) - F_n(x)| \quad (3.8)$$

A função da Equação 3.8 corresponde a distância máxima vertical entre os gráficos de $F(x)$ e $F_n(x)$ sobre a amplitude dos possíveis valores de x , em que: $F(x)$ representa a função de distribuição acumulada assumida para os dados e $F_n(x)$ representa a função de distribuição acumulada empírica dos dados.

Autores como (TATENO; GLASS, 2001; HUANG et al., 2011) utilizaram a teoria de análise de histograma e teste Kolmogorov–Smirnov para analisar o intervalo RR e propor detectores de FA.

3.4.4 Entropia

Entropia (H), segundo (PAPOULIS; PILLAI, 2002, cap. 14), é a medida de incerteza de uma variável aleatória x dada por

$$H(x) = -\sum_i P(x = a_i) \log(P(x = a_i)) \quad (3.9)$$

em que a_i são os possíveis valores de x . Dependendo de qual é a base do logaritmo, diferentes unidades de entropia são obtidas. Geralmente, o logaritmo de base 2 é utilizado.

Seja uma função f definida por

$$f(p) = -p \log p, \quad \text{for } 0 \leq p \leq 1 \quad (3.10)$$

Essa é uma função não negativa que é zero para $p = 0$ e para $p = 1$, e positiva para os valores de p entre 0 e 1. Usando essa função, a entropia pode ser escrita como

$$H(\mathbf{x}) = -\sigma_i f(P(\mathbf{x} = a_i)) \quad (3.11)$$

A entropia de uma variável aleatória pode ser interpretada como o grau de de informação que a observação de uma variável fornece. Quanto mais aleatória, isto é, imprevisível e desestruturada for a variável, maior é a entropia.

Desse modo, a entropia é pequena para valores de probabilidade $P(\mathbf{x} = a_i)$ iguais a 0 e 1, e a entropia é grande se a probabilidade estiver entre 0 e 1.

Autores como (DASH et al., 2009; Andrius Petrėnas, Marozas Vaidotas, 2015; ZHOU et al., 2014) propuseram detectores de FA baseado na teoria de entropia.

4 MATERIAIS E MÉTODOS

Nesta seção, são apresentados os passos de pré-processamento e o algoritmo de detecção de picos R e intervalos RR, bem como a base de dados. São apresentadas também a caracterização do sinal de ECG e a classificação da FA. A Figura 5 ilustra o esquema da metodologia classificação de FA proposta.

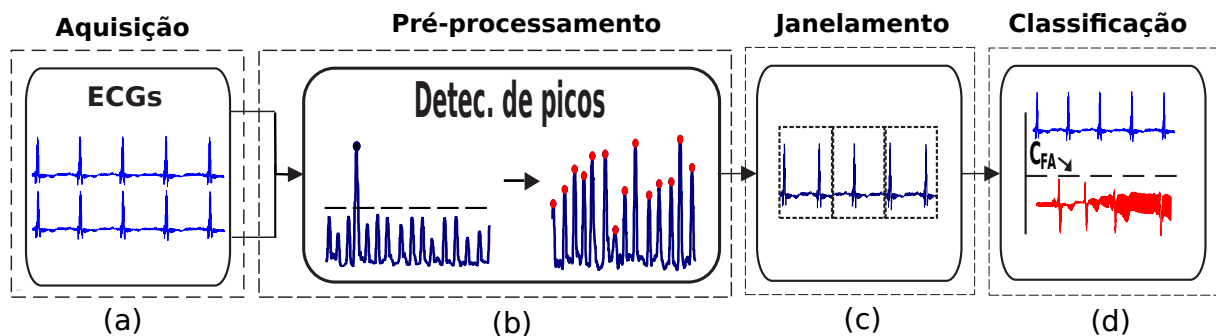


Figura 5 – Metodologia proposta para classificação de FA. (a) Aquisição dos sinais de ECG da BD MIT-BIH AF e MIT-BIH NSR. (b) Na etapa de pré-processamento, acontece a detecção dos picos R. (c) Os sinais de ECG são janelados e caracterizados estatisticamente. (d) Os registros de ECG com FA e saudáveis são classificados.

4.1 Base de Dados

Nesta seção, são apresentadas as bases de dados (DB) MIT-BIH¹ AF (MOODY; MARK, 1983) e MIT-BIH Normal Sinus Rhythm (GOLDBERGER et al., 2000), sendo 18 registros com ritmo sinusal normal (banco de dados MIT-BIH NSR) e 25 com FA (MIT-BIH AF), como mostrado na Tabela 1.

Tabela 1 – Lista dos registros das bases de dados adotadas (MOODY; MARK, 1983; GOLDBERGER et al., 2000).

Base de dados	Registros
MIT-BIH AF	00735; 03665; 04015; 04043; 04048; 04126; 04746; 04908; 04936; 05091; 05121; 05261; 06426; 06453; 06995; 07162; 07859; 07879; 07910; 08215; 08219; 08378; 08405; 08434; 08455.
MIT-BIH NSR	16265; 16272; 16273; 16420; 16483; 16539; 16773; 16786; 16795; 17052; 17453; 18177; 18184; 19088; 19090; 19093; 19140; 19830.

¹ Massachusetts Institute of Technology - Beth Israel Hospital

A base de dados MIT-BIH de Fibrilação Atrial (FA) contém 319 episódios de fibrilação atrial. Esses episódios representam 94,99 horas do total de 260 horas ou 36,53% da BD MIT-BIH AF. As gravações individuais têm aproximadamente 10 horas de duração, de 25 indivíduos numa frequência de amostragem de 250 Hz com resolução de 12 bits numa gama de ± 10 milivolts. Cada registro da base de FA contém notas sobre os tipos de ritmos: fibrilação atrial (FA), flutter atrial (FLA), ritmo AVJ (ritmo atrioventricular juncional) e demais ritmos (OR, termo usado para indicar outros ritmos). Registos 00735 e 03665 onde estão excluídos da simulação uma vez que não estão disponíveis na base de dados.

A base de dados MIT-BIH NSR contém 18 gravações de ECG com frequência de amostragem de 128Hz. Cada registro possui aproximadamente 24 horas de duração, com um total de quase 2 milhões de batimentos. Os sinais incluídos nesta base de dados não tiveram arritmias significativas, e são bem adequados para a avaliação da especificidade do detector. Para facilitar a comparação, o desempenho do método proposto foi avaliado nas combinações comumente usadas das bases de dados MIT-BIH AF e MIT-BIH NSR. Nesse trabalho foram utilizados registros, de ambas as bases, e o download desses registros foi realizado em <https://physionet.org/cgi-bin/atm/ATM>.

Entretanto, nesse trabalho ao invés das 10h de gravação para cada registro da BD de FA e as 24h para os registros da BD de pacientes saudáveis, utiliza-se apenas 1h de gravação para ambas as bases de dados. Os registros 04746, 04908, 05261, 07879 da base de dados MIT-BIH AF foram desconsiderados por não possírem episódios de FA durante 1h de gravação disponibilizada pela Physionet² (detentora dos direitos sobre AFDB e NSRDB) e o registro 07910 também foi desconsiderado por possuir menos de 5% de episódios para 1h de gravação. Além disso, por recomendação da Physionet, foram desconsiderados os registros 00735 e 03665 por não possuírem amostras válidas.

4.2 Detecção de picos R e caracterização do sinal de ECG pela amplitude

Para detectar os picos R, foi utilizado o algoritmo de (PAN; TOMPKINS, 1985) devido a sua acurácia de 99,3% em detectar os complexos QRS da base de dados Arritmia do MIT-BIH (MOODY; MARK, 1983).

O algoritmo de (PAN; TOMPKINS, 1985) é baseado na representação esquemática ilustrada na Figura 6. Entretanto, tal algoritmo pode ser descrito através das seguintes etapas:

- **Passo 1:** utiliza-se um filtro passa-alta tipo FIR de ordem 60 com frequência de corte em 40 Hz usando uma janela de Hamming para eliminar os artefatos

² <https://physionet.org/cgi-bin/atm/ATM>

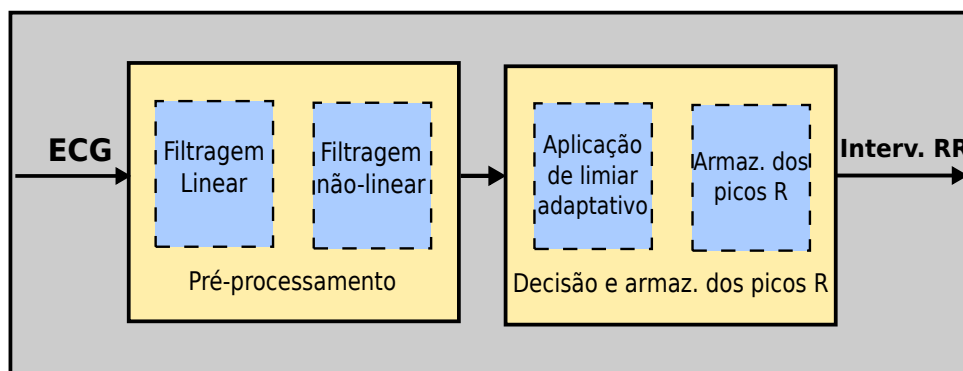


Figura 6 – Representação esquemática do algoritmo de detecção de picos R. Fonte: Adaptado de Lucena (2006).

indesejados de alta frequência, tais como a interferência de 60 Hz.

- **Passo 2:** utiliza-se um filtro digital passa-alta tipo FIR de alta ordem com frequência de corte em 4 Hz para eliminar os artefatos indesejados de baixa frequência, tais como respiração, linha de base e artefatos de movimento. Nesse momento o sinal de ECG está limitado em banda entre 4 - 40 Hz.
- **Passo 3:** Diferencia-se o sinal de ECG do passo 2 para obter informações sobre a inclinação.
- **Passo 4:** Elevamos ao quadrado o sinal do passo 3 (processamento não linear). Isto causa uma elevação da amplitude do sinal.
- **Passo 5:** Usa-se uma janela deslizante (W) com um detector de limiar adaptativo. A princípio o limiar adaptativo é ajustado em $\frac{1}{3}$ do valor máximo de W , mas tende a variar com o deslocamento de uma janela incremental. Assim o algoritmo definirá como valor máximo do complexo QRS os valores acima do limiar pré-definido. Após a detecção do ponto máximo, localização temporal da onda R, tem-se um salto em W definido pelo intervalo mínimo da duração do ciclo cardíaco. Este valor é baseado nos limites fisiológicos do ECG.
- **Passo 6:** Armazena-se os momentos de pico ($R[n]$) para se calcular os intervalos RR, isto é, $RR[n] = RR[n + 2] - R[n]$.

A partir dos picos R, pode-se construir o intervalo RR. Pode-se definir o intervalo RR como a coleção de batimentos,

$$RR = [b_1, b_2, b_3, \dots, b_m] \quad (4.1)$$

em que b_m representa o m -ésimo batimento. A Figura 7 ilustra o intervalo RR descrito em termo de seus batimentos.

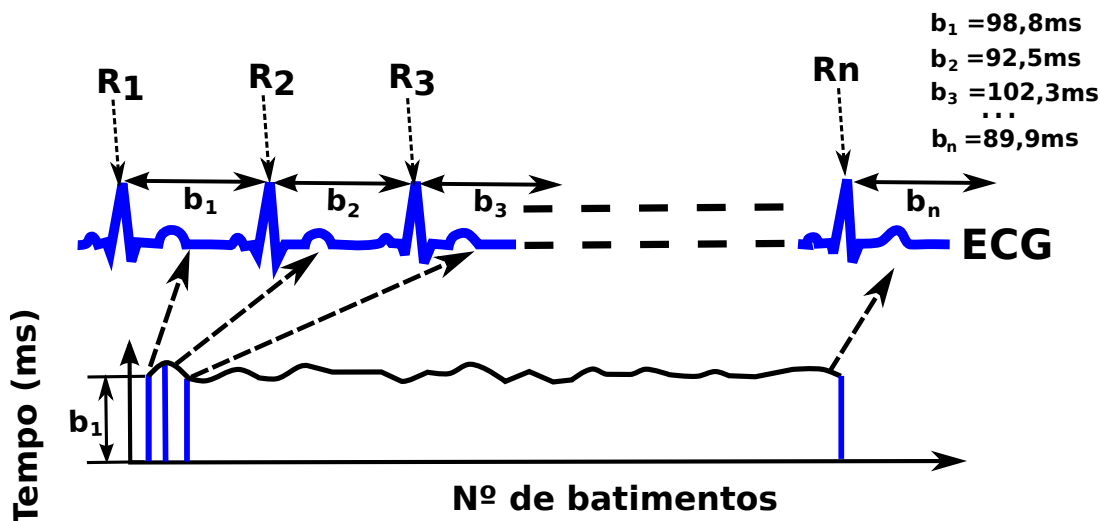


Figura 7 – Representação esquemática do intervalo RR, de modo que $RR = [b_1, b_2, \dots, b_m]$, e $b_m = t_m - t_{m-1}$ representa o m -ésimo batimento e t_m é o instante de tempo do m -ésimo pico R. Fonte: Adaptado de Lucena (2006).

A análise temporal do ECG, denominada de intervalo RR, permite avaliar as irregularidades das ondas R. Entretanto, através da análise baseada na amplitude do sinal de ECG (OLIVEIRA et al., 2016; OLIVEIRA et al., 2017), pode-se analisar as deformações em todas as ondas do ECG (onda P, Q, R, S e T) através de n estatísticas diferentes. A Figura 8 ilustra o sinal de ECG com suas marcações temporais características (onda P, complexo QRS e onda T) e em destaque, tracejado em preto, o segmento do sinal de ECG analisado na amplitude, $x(t)$, nos instantes t_1 e t_0 .

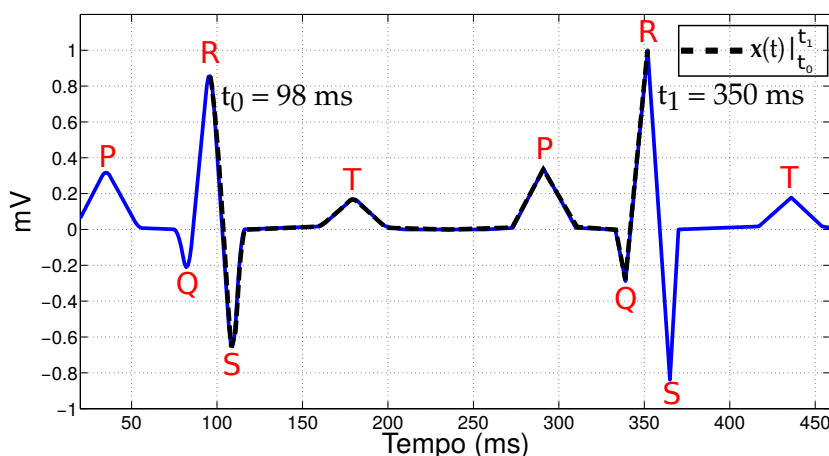


Figura 8 – Ilustração do primeiro segmento do sinal de ECG, ou seja, $x(t)$ nos instantes $t_1 = 98ms$ e $t_0 = 350ms$.

Considerando o sinal de ECG, $x(t)$, e o i -ésimo batimento (b_i), pode-se definir

$$\mu_i = \frac{1}{n} \sum_{j=1}^n x_j, \tag{4.2}$$

em que μ_i é a média do i -ésimo batimento e n é o número de amostras entre duas ondas R consecutivas do i -ésimo batimento. Desse modo, o mesmo i -ésimo batimento, poderá ser caracterizado por k -ésimas estatísticas diferentes, dentre elas estão:

$$\sigma_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (x_j - \mu_i)^2} \quad (4.3)$$

$$\sigma_i^2 = \frac{1}{n} \sum_{j=1}^n (x_j - \mu_i)^2 \quad (4.4)$$

$$\gamma_i = \frac{1}{n} \sum_{j=1}^n (x_j - \mu_i)^3 \quad (4.5)$$

$$\kappa_i = \frac{1}{n} \sum_{j=1}^n (x_j - \mu_i)^4 - 3\left(\frac{1}{n} \sum_{j=1}^n (x_j - \mu_i)^2\right)^2 \quad (4.6)$$

em que σ_i , σ_i^2 , γ_i e κ_i são, respectivamente, o desvio padrão, a variância, a assimetria e a curtose do i -ésimo batimento.

Diante disso, a análise estatística (caracterização) utilizando a amplitude do sinal de ECG possibilitará a criação de novos métodos de classificação da FA.

O método de classificação proposto utiliza limiar de decisão, assim como parte dos trabalhos aqui citados (Andrius Petrėnas, Marozas Vaidotas, 2015; ZHOU et al., 2014; LOGAN; HEALEY, 2005; TATENO; GLASS, 2001; HUANG et al., 2011; LIAN; WANG; MUESSIG, 2011). Embora, esses métodos usem algum tipo de limiar de decisão, eles não possuem embasamento estatístico que garanta um intervalo de confiança. O desafio é, então, desenvolver limiar de decisão que considere um intervalor de confiança. Esse trabalho propõem um limiar baseados na dispersão dos grupos de FA. Assim, o limiar irá variar conforme a dispersão do grupo de pacientes com FA.

Para assegurar os resultados e a generalização do método, propõem-se estabelecer um intervalo de confiança (IC) de 99,73%, ou seja, baseado em três vezes o desvio padrão (CAMPBELL; SWINSCOW, 2011, cap. 4).

4.3 Janelamento do sinal de ECG

O ideal é que a análise do ECG fosse feita batimento a batimento, no entanto, cada episódio de FA pode conter até 5 vezes mais batimentos por segundo que uma pessoa saudável (ZIMERMAN et al., 2009), aumentando o custo computacional. O aumento do custo computacional pode inviabilizar a implementação em sistemas embarcados, fazendo com que estes métodos nunca se tornem realmente métodos que apoiem o diagnóstico de cardiopatias como a FA. A fim de facilitar a implementação

da metodologia em dispositivos embarcados, como smartphones, alguns trabalhos propõem o janelamento do sinal de ECG (LEE et al., 2013; TATENO; GLASS, 2001). A janela mais utilizada na literatura é a de 60 batimentos, que corresponde a média de batimentos de uma pessoa saudável (60 batimentos por minuto) (JANUARY et al., 2014; ZIMERMAN et al., 2009).

A Figura 9 ilustra a caracterização de um registro de ECG (E).

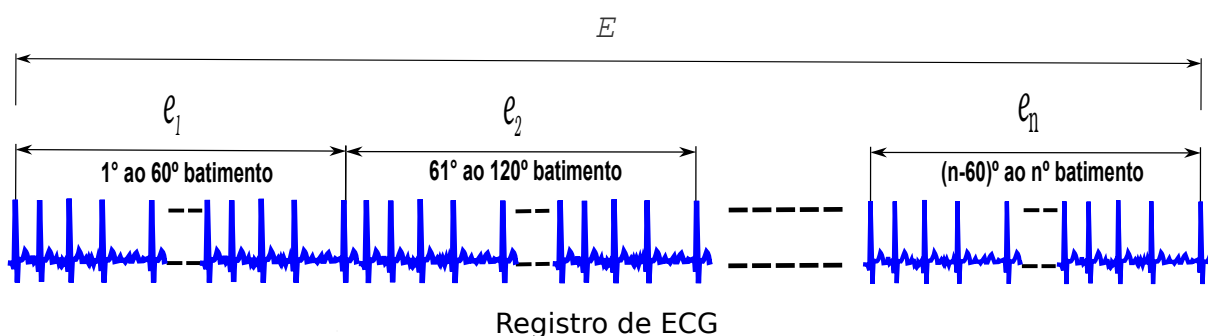


Figura 9 – Exemplo de análise e caracterização de um registro de ECG. Em que e_n é a n -ésima estimativa (caracterização) através de alguma estatística (tais como, média, variância, assimetria ou curtose) e E é a estimativa ou caracterização do sinal completo através de alguma estatística, como a média e a mediana.

Conforme ilustrado em Figura 9, considerando que cada segmento de ECG entre dois picos R consecutivos, representa um batimento cardíaco, é realizado um janelamento de 60 batimentos e cada janela é caracterizada utilizando uma estatística (média, variância, assimetria ou curtose). Posteriormente, utilizando estatísticas, tais como a média e mediana, o registro inteiro de ECG (1h de gravação) é caracterizado (E). Dentre as estatísticas avaliadas (média e mediana) para caracterizar o registro de ECG (E), a mediana foi a que permitiu maior separação, isso é de se esperar tendo em vista que a média é sensível a outliers.

4.4 Classificador de FA (C_{FA_m})

O classificador proposto, C_{FA_m} , é baseado na variabilidade populacional do grupo de FA e a partir dessa variabilidade, propôs-se um limiar de decisão, que é definido na Equação 4.7, em que os registros de ECG que atendem a esse limiar são classificados com FA.

$$C_{FA_m} = \tilde{X}_{FA} \pm 3\sigma_{FA}, \quad (4.7)$$

onde \tilde{X}_{FA} é a média das caracterizações dos registros da Base de Dados (BD) de FA e σ_{FA} é o desvio padrão das caracterizações dos registros da BD de FA. Para o classificador proposto, a caracterização de cada registro de ECG (E) foi realizada por meio da mediana.

4.5 Avaliação

Os resultados são avaliados em quatro categorias:

- Verdadeiro positivo (VP), em que um registro com FA é classificado com FA;
- Verdadeiro negativo (VN), em que um registro saudável é classificado como saudável;
- Falso negativo (FN), para esta categoria, um registro com FA é classificado como saudável;
- Falso positivo (FP), em que um registro saudável é classificado com FA.

Como métrica para avaliação da fase de classificação, usamos a *sensibilidade*, que reflete como a metodologia adotada é efetiva em corretamente identificar os VPs, *especificidade*, refletindo como a metodologia adotada é eficaz na correta identificação dos VNs e *precisão* indicando como a metodologia é eficaz na correta realização do diagnóstico.

A sensibilidade e especificidade são definidas, respectivamente,

$$VP/(VP + FN),$$

e

$$VN/(VN + FP),$$

A acurácia é definida como

$$(VP + VN)/n,$$

em que n é a quantidade de registros da BD.

5 RESULTADOS

As próximas subseções tratam da análise estatística da amplitude do ECG e os resultados do classificador proposto, respectivamente.

5.1 Análise Estatística da amplitude do sinal de ECG

Para a etapa de simulações, foram utilizado os registros da Tabela 1, desconsiderando os registros 04746, 04908, 05261, 07879, 07910, 00735 e 03665, conforme Subseção 4.1, e também foram retirados os batimentos ectópicos (Andrius Petrėnas, Marozas Vaidotas, 2015).

A partir de então, foram investigadas algumas estatísticas para o classificador proposto, a fim de avaliar qual permitiria uma maior separação entre os grupos de estudo (pacientes com FA e saudáveis). Para este trabalho, o classificador proposto, C_{FA_m} , foi analisado utilizando a variância, assimetria e curtose. A partir dessas estatísticas, cada registro de ECG foi caracterizado (conforme Seção 4.3) e, em seguida, foi analisada a dispersão de cada grupo de estudo. Como critério para avaliar a separabilidade entre os grupos, utilizou-se o valor de Δ , que é a diferença entre as medianas dos grupos com FA e pessoas saudáveis. A Tabela 2 apresenta o resumo da análise estatística para o classificador proposto, C_{FA_m} .

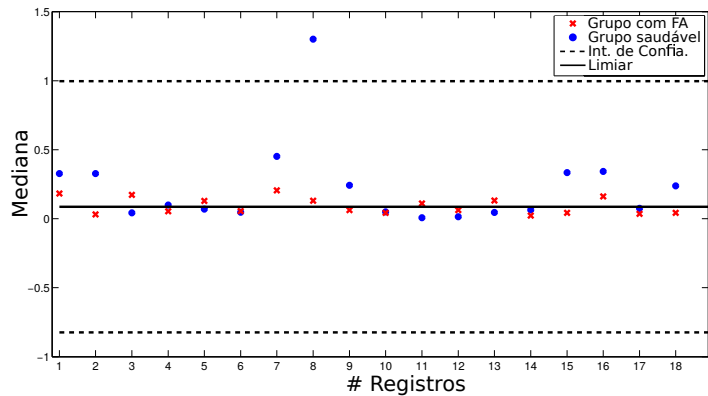
Tabela 2 – Análise da dispersão dos grupos de pacientes saudáveis (com ritmo sinusal normal) e com FA em diferentes estatísticas (variância, assimetria e curtose), em que o limiar superior é o máximo valor de caracterização (E) do grupo, limiar inferior é o menor valor de caracterização do grupo e o Δ é a diferença entre as medianas do grupo de pessoas saudáveis e do grupo com FA.

	Análise	Saudáveis	FA	 \Delta
Batimentos (b_n)	Limiar Superior	2,0625	2,3600	0,104
	Mediana	1,3120	1,4160	
	Limiar Inferior	0,6	0,3640	
Variância	Limiar Superior	0,2054	1,3018	0,024
	Mediana	0,062	0,086	
	Limiar Inferior	0,022	0,007	
Assimetria	Limiar Superior	4,9496	1,6949	2,7776
	Mediana	2,3241	-0,4535	
	Limiar Inferior	-0,9574	-2,3424	
Curtose	Limiar Superior	37,2921	10,2829	12,9484
	Mediana	19,6150	6,6666	
	Limiar Inferior	8,1533	5,6508	

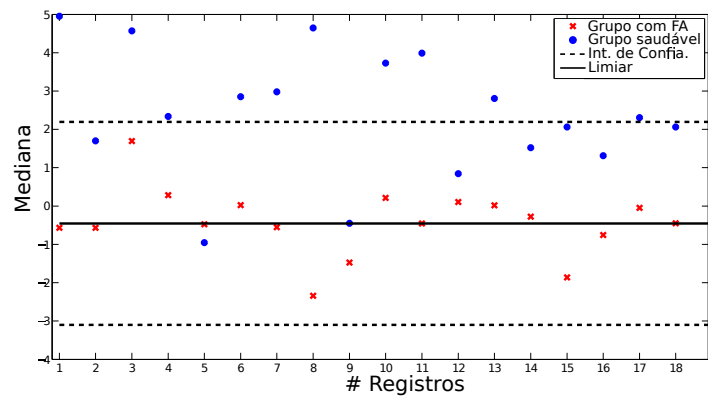
Os resultados apresentados na Tabela 2 foram baseados em um janelamento

do sinal de ECG de 60 batimentos, tendo em vista que 60 bpm é a média de frequência cardíaca de uma pessoa normal (JANUARY et al., 2014; ZIMERMAN et al., 2009).

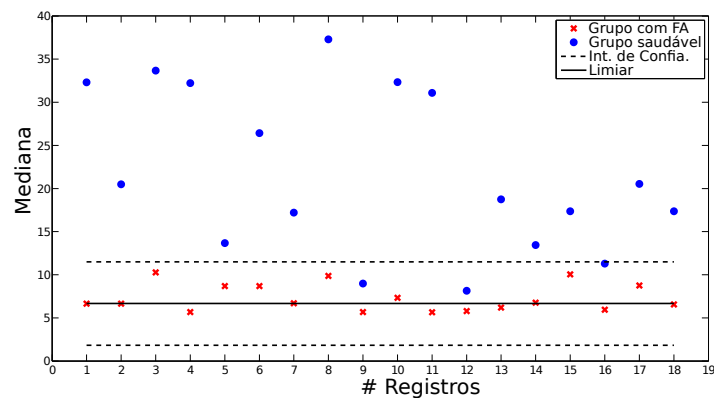
As análises ilustradas na Figura 10 são referentes ao classificador proposto utilizando, respectivamente, análise estatística através da variância, assimetria e curtose, e corroboram com a análise apresentada na Tabela 2.



(a)



(b)



(c)

Figura 10 – Resultado para o classificador proposto. (a) Utilizando a variância. (b) Utilizando a assimetria. (c) Utilizando a curtose.

5.2 Classificação da FA

Os resultados obtidos pelo método do classificador proposto, C_{FA_m} , utilizando variância, assimetria e curtose são apresentados na Tabela 3, e avaliados em relação a sensibilidade, especificidade e acurácia.

Tabela 3 – Resultados do método de classificação proposto utilizando variância, assimetria e curtose para um grupo de 36 indivíduos, sendo 18 saudáveis e 18 com FA.

C_{FA_m}	Se%	Sp%	Ac%
Variância	100	5	52,7
Assimetria	100	55,55	77,77
Curtose	100	83,33	91,66

A Figura 11 ilustra a região de aceitação, para o intervalo de confiança de 99,73% (CAMPBELL; SWINSCOW, 2011, cap. 4), do classificador utilizando a curtose, o melhor resultado conforme Tabela 3.

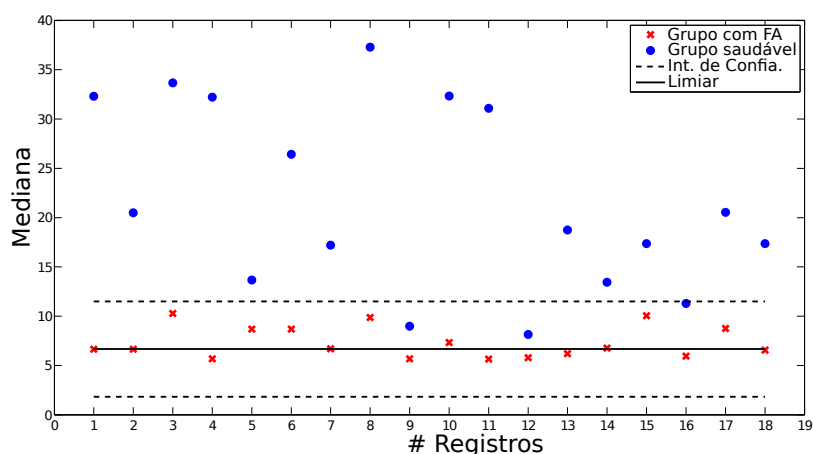


Figura 11 – Resultados do classificador proposto. Em azul a caracterização para os sinais de ECG saudáveis. Em vermelho a caracterização dos sinais de ECG com FA. Em preto tem-se o limiar e intervalo de confiança representando a função de decisão do método de classificação proposto.

Os códigos das simulações realizadas para esse trabalho se encontram disponíveis no Apêndice D.

6 DISCUSSÃO

Os resultados apresentados na Tabela 2, que são baseado no critério da diferença entre as medianas dos grupos de estudo, Δ , mostram que a curtose permitiu maior separabilidade dos grupos, obtendo um $\Delta = 12,9484$, superior aos valores de Δ das demais estatísticas analisadas. Isso é de se esperar tendo em vista que os sinais de ECG e complexos QRS são formados por componentes esparsas, tendo uma distribuição não-gaussiana (LUCENA et al., 2011).

Os resultados obtidos pelo o classificador proposto, C_{FA_m} , foram avaliados em função da sensibilidade, especificidade e acurácia, para um intervalo de confiança de 99,73% (ver Tabela 3). Os resultados da Tabela 3 mostram que a curtose foi a estatística que permitiu maior acurácia ($Ac = 91,66\%$) para o classificador proposto.

Ainda pela Tabela 3, os resultados referentes a variância e assimetria não apresentaram acurácia superior aos resultados utilizando a curtose (ilustrados na Figura 10). A Figura 11 ilustra os resultados do classificador proposto, C_{FA_m} , que obteve sensibilidade de 100%, especificidade de 83,33%, e acurácia de 91,67%. Nota-se também que o classificador ilustrado na Figura 11 apresenta uma dispersão menor para o grupo FA do que para o grupo de saudáveis, o que corroborará com a Tabela 2. Além da variância, assimetria e curtose, foi realizado teste com a média (ver Apêndice C), a qual não permitiu a separação entre os grupos de pacientes com FA e saudáveis.

Na Tabela 4, a metodologia proposta é comparada com os melhores desempenhos na literatura com relação à sensibilidade, especificidade e acurácia. Em todas as comparações (conforme Tabela 4), o método de classificação, C_{FA_m} , apresenta sensibilidade (100%) superior ao demais métodos, ou seja, o método consegue reconhecer os doentes de forma mais eficiente. No entanto, a capacidade do método proposto, de reconhecer os não doentes é inferior em grande parte aos métodos da literatura.

Investigações utilizando o classificador proposto também foram realizadas com o intervalo RR (ver Apêndice B). No entanto, nenhuma das estatísticas utilizadas (variância, assimetria e curtose) permitiu a separação dos grupos.

O método da avaliação da amplitude do sinal permite extrair uma maior quantidade de informação do ECG, e isso possibilitou a simplicidade do classificador proposto, baseando-se em um limiar que considera a distribuição estatística do grupo de FA. Diferentemente do que é feito na literatura (Andrius Petrėnas, Marozas Vaidotas, 2015; DASH et al., 2009; TATENO; GLASS, 2001), o classificador proposto, C_{FA_m} , leva em consideração um intervalo de confiança de 99,73%.

O C_{FA_m} também tem a vantagem com relação as demais metodologias da

Tabela 4 – Comparação de desempenho avaliadas para a base de dados MIT-BIH Atrial Fibrillation (AFDB). AFDB₁ corresponde a base de dados (BD) sem os registros 00735 e 03665, AFDB₂ a BD sem 04936 e 05091, e AFDB_b corresponde a BD sem os registros 04746, 04908, 05261, 07879, 07910, 00735 e 03665.

Método	Ano	Base de Dados	Se %	Sp %	Ac %
C_{FA_m}	2016	AFDB _b	100	91,66	83,33
R. Couceiro et al.	2015	AFDB	88,50	92,90	90,60
Petrenas et al.	2015	AFDB	97,1	98,3	-
Zhou et al.	2014	AFDB	97,53	98,26	98,16
Jinseok Lee et al.	2013	AFDB	97,26	95,91	96,14
Carvalho et al.	2012	AFDB ₁	93,8	96,1	-
Huang et al.	2011	AFDB	96,1	98,1	-
Lian et al.	2011	AFDB	95,8	96,4	-
Pei-Chann C. et al.	2010	TAFGH	68,49	96,00	75,51
Dash et al.	2009	AFDB ₂	94,4	95,1	99,1
Logan et al.	2005	AFDB	96	89	-
Tateno et al.	2001	AFDB	94,4	97,2	-

Tabela 4 de não só avaliar a deformação das ondas R do ECG, mas também avaliar as deformações nas demais ondas do ECG. Tal característica, pode ser utilizada para investigar outras doenças cardíacas. Entretanto, a curtose, que é utilizada no classificador proposto, tem como problema a sensibilidade a outliers (PAPOULIS; PILLAI, 2002).

7 CONCLUSÃO

Nesse trabalho, foi apresentado um classificador de FA baseado na variabilidade da amplitude do sinal de ECG. A principal diferença deste método, foi de que, em vez de considerar as irregularidades das séries temporais RR (intervalo RR), analisou-se as alterações e deformações que ocorrem na amplitude do sinal de ECG, como, por exemplo, na onda P, ocasionadas pela FA, permitindo obter mais informação do sinal cardíaco. Além disso, o classificador, se diferencia dos demais classificadores desenvolvidos pela literatura (Andrius Petrėnas, Marozas Vaidotas, 2015; HUANG et al., 2011; LIAN; WANG; MUESSIG, 2011; LOGAN; HEALEY, 2005; ZHOU et al., 2014; TATENO; GLASS, 2001), pelo fato dele ser baseado em um intervalo de confiança de 99,73% e também por considerar em seu estudo registros de ECG, contendo episódios de FA, com apenas 1h de gravação.

O fato do classificador proposto avaliar as deformações nas ondas do ECG, torna possível prever, não só se o paciente tem FA, mas também se ele é saudável ou não.

O classificador poderá ser utilizado em estudos epidemiológicos e no apoio a tomada de decisão em práticas clínicas, com uma taxa de acerto de até 91,67% do diagnóstico. Os resultados apresentados nesse trabalho estão publicados em (OLIVEIRA et al., 2016; OLIVEIRA et al., 2017).

Além disso, a simplicidade do método proposto e a velocidade de processamento permitirá aplicações em sistemas embarcados, similar ao Holter 24h. No entanto, um sistema baseado no método proposto será capaz de não só registrar a atividade elétrica do coração e suas variações, mas também fornecer o diagnóstico da FA.

A implementação da metodologia proposta se mostra relevante para países de grande extensão territorial, como o Brasil, em que a concentração de médicos é desproporcional a média nacional de 1,84 médico para cada 1000 habitantes (BRASIL, 2010). Também se torna relevante, tendo em vista que tal sistema poderá apoiar o diagnóstico remoto, como por exemplo, em locais de difícil acesso e onde não existem médicos especializados.

8 CONTRIBUIÇÃO E PUBLICAÇÃO DE RESULTADOS

- Publicações:

- OLIVEIRA, A; QUEIROZ, J. A.; LUCENA, F.; BARROS, A. K., Method to characterize and classify atrial fibrillation based on kurtosis. 2016 IEEE International Conference on Industrial Technology (ICIT), Taipei, Taiwan, 2016, pp. 1028-1032. doi: 10.1109/ICIT.2016.7474895.
- OLIVEIRA, A; QUEIROZ, J. A.; LUCENA, F.; BARROS, A. K.; SILVA, L. C. O; Lopes, M. V. Classification of Atrial Fibrillation Based on ECG Signal Amplitude. In: 2017 IEEE International Conference on Biomedical and Health Informatics, Resumo expandido publicado em Anais da conferência, 2017.

- Artigo aceito:

- OLIVEIRA, A; QUEIROZ, J. A.; LUCENA, F.; BARROS, A. K.; SILVA, L. C. O; Lopes, M. V. Atrial Fibrillation Classification Based on Kurtosis. International Conference on Applied Mathematics and Computer Science (ICAMCS), 2017.

9 TRABALHO FUTUROS

Como trabalhos futuros pretende-se subamostrar os registros de ECG da Base de Dados (BD) de FA¹, para a partir de então poder fazer uma comparação dos resultados do classificador com a amostragem igual a da BD de pacientes saudáveis. Isso porque, supõe-se que a amostragem do registro influencia na etapa de caracterização do mesmo, pois quanto mais ou menos amostras naquele registro, tem-se mais ou menos informação para se caracterizar estatisticamente.

Também pretende-se aprimorar o método de classificação proposto de modo que se aumente a acurácia do classificador. Posteriormente, pretende-se implementar o classificador proposto em dispositivos embarcados, como smartphones.

¹ A BD de FA possui 256Hz, já a BD de pacientes saudáveis possui 128Hz; conforme Subseção 4.1

REFERÊNCIAS

- Andrius Petrėnas, Marozas Vaidotas, L. S. Low-complexity detection of atrial fibrillation in continuous long-term monitoring. *Computers in Biology and Medicine*, p. 1–8, 2015. Citado 7 vezes nas páginas 17, 21, 27, 32, 35, 38 e 40.
- ASHLEY, E. A.; NIEBAUER, J. *Cardiology explained*. [S.l.]: Remedica, 2004. Citado 3 vezes nas páginas 15, 18 e 20.
- BRASIL, M. d. S. *Sistema de Informações de Recursos Humanos para o SUS (SIRH) e base demográfica do IBGE*. 2010. Brasília: Ministério da Saúde. Citado na página 40.
- BRUGADA, R. et al. Identification of a genetic locus for familial atrial fibrillation. *New England Journal of Medicine*, Mass Medical Soc, v. 336, n. 13, p. 905–911, 1997. Citado na página 22.
- CAMM, A. J. et al. Guidelines for the management of atrial fibrillation. *European heart journal*, Eur Soc Cardiology, p. ehq278, 2010. Citado 2 vezes nas páginas 22 e 23.
- CAMM, A. J. et al. Don't you agree, or what part of the problem don't you understand? *Journal of Interventional Cardiac Electrophysiology*, Springer, v. 4, n. 4, p. 559–560, 2000. Citado na página 23.
- CAMPBELL, M. J.; SWINSCOW, T. D. V. *Statistics at square one*. [S.l.]: John Wiley & Sons, 2011. Citado 2 vezes nas páginas 32 e 37.
- CARVALHO, P. de et al. Model-based atrial fibrillation detection. In: *ECG signal processing, classification and interpretation*. [S.l.]: Springer, 2012. p. 99–133. Citado na página 17.
- CASTANHEIRA, N. P. *Estatística aplicada a todos níveis*. [S.l.]: Editora Ibpex, 2008. Citado na página 25.
- CHANG, P.-C. et al. Atrial fibrillation analysis based on blind source separation in 12-lead ecg. *World Academy of Science, Engineering and Technology*, 2010. Citado na página 18.
- COUCEIRO, R. et al. Detection of atrial fibrillation using 12-lead ecg for mobile applications. *International Conference on Biomedical and Health Informatics*, 2015. Citado na página 18.
- DARBAR, D. et al. Familial atrial fibrillation is a genetically heterogeneous disorder. *Journal of the American College of Cardiology*, Am Coll Cardio Found, v. 41, n. 12, p. 2185–2192, 2003. Citado na página 23.
- DASH, S. et al. Automatic real time detection of atrial fibrillation. *Annals of biomedical engineering*, Springer, v. 37, n. 9, p. 1701–1709, 2009. Citado 2 vezes nas páginas 27 e 38.
- ELLINOR, P. T. et al. Locus for atrial fibrillation maps to chromosome 6q14–16. *Circulation*, Am Heart Assoc, v. 107, n. 23, p. 2880–2883, 2003. Citado na página 23.

FERREIRA, J. P.; SANTOS, M. Heart failure and atrial fibrillation: from basic science to clinical practice. *International journal of molecular sciences*, Multidisciplinary Digital Publishing Institute, v. 16, n. 2, p. 3133–3147, 2015. Citado na página 20.

FOX, C. S. et al. Parental atrial fibrillation as a risk factor for atrial fibrillation in offspring. *Jama*, American Medical Association, v. 291, n. 23, p. 2851–2855, 2004. Citado na página 22.

FRIBERG, J. et al. Rising rates of hospital admissions for atrial fibrillation. *Epidemiology*, LWW, v. 14, n. 6, p. 666–672, 2003. Citado na página 17.

FROST, L.; HUNE, L. J.; VESTERGAARD, P. Overweight and obesity as risk factors for atrial fibrillation or flutter: the danish diet, cancer, and health study. *The American journal of medicine*, Elsevier, v. 118, n. 5, p. 489–495, 2005. Citado 2 vezes nas páginas 21 e 22.

GAMI, A. S. et al. Obstructive sleep apnea, obesity, and the risk of incident atrial fibrillation. *Journal of the American College of Cardiology*, Am Coll Cardio Found, v. 49, n. 5, p. 565–571, 2007. Citado na página 21.

GOLDBERGER, A. L. et al. Physiobank, physiokit, and physionet components of a new research resource for complex physiologic signals. *Circulation*, Am Heart Assoc, v. 101, n. 23, p. 215–220, 2000. Citado 2 vezes nas páginas 11 e 28.

HUANG, C. et al. A novel method for detection of the transition between atrial fibrillation and sinus rhythm. *IEEE Transactions on Biomedical Engineering*, IEEE, v. 58, n. 4, p. 1113–1119, 2011. Citado 5 vezes nas páginas 21, 24, 26, 32 e 40.

HURST, J. W. et al. Management of patients with atrial fibrillation. *The American journal of medicine*, Elsevier, v. 37, n. 5, p. 728–741, 1964. Citado na página 20.

JANUARY, C. T. et al. 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: Executive Summary A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *Journal of the American College of Cardiology*, v. 64, n. 21, p. 2246–2280, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.jacc.2014.03.021>>. Citado 3 vezes nas páginas 20, 33 e 36.

LEE, J. et al. Atrial Fibrillation Detection using a Smart Phone. *International Journal of Cardiology*, v. 60, n. 1, p. 1177–1180, 2013. ISSN 1558-2531. Citado 2 vezes nas páginas 17 e 33.

LI, D. et al. Promotion of atrial fibrillation by heart failure in dogs atrial remodeling of a different sort. *Circulation*, Am Heart Assoc, v. 100, n. 1, p. 87–95, 1999. Citado na página 21.

LIAN, J.; WANG, L.; MUESSIG, D. A simple method to detect atrial fibrillation using rr intervals. *The American journal of cardiology*, Elsevier, v. 107, n. 10, p. 1494–1497, 2011. Citado 2 vezes nas páginas 32 e 40.

LOGAN, B.; HEALEY, J. Robust detection of atrial fibrillation for a long term telemonitoring system. In: IEEE. *Computers in Cardiology*, 2005. [S.l.], 2005. p. 619–622. Citado 5 vezes nas páginas 17, 21, 24, 32 e 40.

LUCENA, F. et al. Statistical coding and decoding of heartbeat intervals. *PloS one*, Public Library of Science, v. 6, n. 6, p. e20227, 2011. Citado 2 vezes nas páginas 25 e 38.

LUCIANO, P. M. et al. Atendimento de arritmia cardíaca em emergência de hospital universitário terciário. *Rev Bras Cardiol*, v. 24, n. 4, p. 225–232, 2011. Citado na página 15.

MAGALHÃES, L. et al. li diretrizes brasileiras de fibrilação atrial. *Arquivos Brasileiros de Cardiologia*, SciELO Brasil, v. 106, n. 4, p. 1–22, 2016. Citado 4 vezes nas páginas 16, 17, 22 e 23.

MOODY, G. B.; MARK, R. G. A new method for detecting atrial fibrillation using rr intervals. *Computers in Cardiology*, v. 10, n. 1, p. 227–230, 1983. Citado 4 vezes nas páginas 11, 22, 28 e 29.

OLIVEIRA, A. et al. Method to characterize and classify atrial fibrillation based on kurtosis. *2016 IEEE International Conference on Industrial Technology (ICIT)*, p. 1028–1032, March 2016. Citado 3 vezes nas páginas 17, 31 e 40.

OLIVEIRA, A. et al. Classification of Atrial Fibrillation Based on ECG Signal Amplitude. *In: 2017 IEEE International Conference on Biomedical and Health Informatics*, Resumo expandido publicado em Anais da conferência, 2017. Citado 3 vezes nas páginas 17, 31 e 40.

PAN, J.; TOMPKINS, W. J. A real-time QRS detection algorithm. *IEEE transactions on bio-medical engineering*, v. 32, n. 3, p. 230–236, 1985. ISSN 0018-9294. Citado na página 29.

PAPOULIS, A.; PILLAI, U. *Probability, random variables, and stochastic processes*. 4. ed. [S.I.]: McGraw-Hill, 2002. ISBN 0073660116. Citado 3 vezes nas páginas 24, 26 e 39.

SAAD, E. B.; JUNIOR, L. A. I.; CAMANHO, L. E. Tratamento da fibrilação atrial: o estado da arte. *J. bras. med*, p. 7–18, 2012. Citado 5 vezes nas páginas 17, 20, 21, 22 e 23.

SHARMA, L.; DANDAPAT, S.; MAHANTA, A. ECG signal denoising using higher order statistics in wavelet subbands. *Biomedical Signal Processing and Control*, Elsevier, v. 5, n. 3, p. 214–222, 2010. Citado na página 17.

SOBRAC, S. *29 DE SETEMBRO - DIA MUNDIAL DO CORAÇÃO*. 2014. Disponível em: <<http://www.sobrac.org/publico-geral/wp-content/uploads/2014/04/SOBRAC-Release-Arritmologista-Dia-Mundial-do-Cora+%C2%BA+%C3%BAo-Setembro-20141.pdf>>. Citado na página 15.

SOBRAC, S. *Sociedade Brasileira de Arritmias Cardíacas*. 2016. Disponível em: <<http://www.sobrac.org>>. Citado 2 vezes nas páginas 15 e 17.

TATENO, K.; GLASS, L. Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and Δ RR intervals. *Medical and Biological Engineering and Computing*, Springer, v. 39, n. 6, p. 664–671, 2001. Citado 7 vezes nas páginas 21, 24, 26, 32, 33, 38 e 40.

VIANA, R. A. P. P.; WHITAKER, I. Y. *Enfermagem em terapia intensiva: práticas e vivências*. [S.l.]: Artmed Editora, 2009. Citado na página 17.

WANG, T.; PARISE, H.; LEVY, D. Obesity and the risk of new-onset atrial fibrillation. *ACC Current Journal Review*, No longer published by Elsevier, v. 14, n. 3, p. 45, 2005. Citado na página 22.

WANG, T. J. et al. Temporal Relations of Atrial Fibrillation and Congestive Heart Failure and Their Joint Influence on Mortality The Framingham Heart Study. *Circulation*, Am Heart Assoc, v. 107, n. 23, p. 2920–2925, 2003. Citado na página 21.

WHO, W. H. O. et al. World health statistics 2016: monitoring health for the sdgs, sustainable development goals. World Health Organization (WHO), 2016. Citado na página 15.

ZHOU, X. et al. Automatic online detection of atrial fibrillation based on symbolic dynamics and shannon entropy. *Biomedical engineering online*, BioMed Central Ltd, v. 13, n. 1, p. 18, 2014. Citado 4 vezes nas páginas 17, 27, 32 e 40.

ZIMERMAN, L. et al. Diretrizes brasileiras de fibrilação atrial. *Arq. bras. cardiol*, v. 92, n. 6 supl. 1, p. 1–42, 2009. Citado 5 vezes nas páginas 18, 20, 32, 33 e 36.

Apêndices

APÊNDICE A – BOXPLOT

O boxplot é um sumário gráfico da distribuição de uma amostra que exibe sua forma, tendência central e variabilidade. Ele é construído com base no resumo dos cinco números, constituído por:

- Valor mínimo;
- Primeiro quartil (Q1);
- Segundo quartil (Q2);
- Terceiro quartil (Q3);
- Valor máximo;

A Figura 12 ilustra um típico gráfico de boxplot.

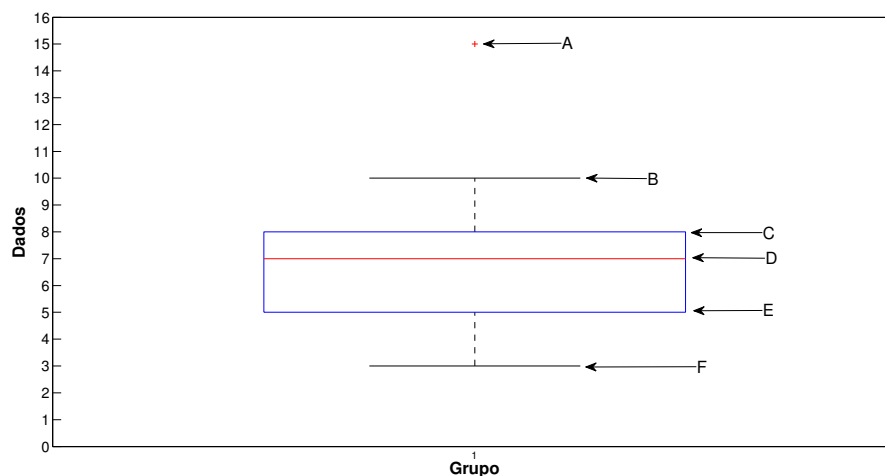


Figura 12 – Ilustração de um boxplot, em que (A) representa um valor atípico (ou outlier, valor que está além do traço superior ou inferior); (B) representa o valor máximo (traço superior); (C) representa o terceiro quartil; (D) representa o Segundo quartil; (E) representa o primeiro quartil; (F) representa o valor mínimo (traço inferior).

O boxplot da Figura 12 é formado por uma caixa construída paralelamente ao eixo da escala dos dados (pode ser horizontal ou vertical). Essa caixa vai desde o primeiro quartil até o terceiro quartil e nela traça-se uma linha na posição da mediana. Essa caixa, que descreve os 50% centrais da distribuição, é comum a todas as variantes do boxplot.

Entretanto, o que vem a ser os quartis? Dada uma escala de mensuração dos dados, os quartis são valores nessa escala que dividem o conjunto de dados em quatro partes, todas elas com o mesmo número de observações. Isso significa que 25% das observações são menores que o primeiro quartil, 50% são menores que o segundo quartil e 75% são menores que o terceiro quartil. Quando se fala de maior ou menor observação se trata de escala, de ordem. Assim, para calcular os quartis, tem-se que ordenar os dados.

A.1 Determinação dos outliers ou valores atípicos

A regra que é geralmente adotada para identificação dos valores atípicos (outliers) se baseia na amplitude interquartil AIQ, definida como a distância entre o primeiro e o terceiro quartis:

$$AIQ = Q3 - Q1 \quad (A.1)$$

Note que AIQ é o comprimento da caixa. Quaisquer valores abaixo de $Q1$ ou acima de $Q3$ por mais de $1,5 \times AIQ$ serão considerados valores atípicos e terão tratamento especial no boxplot. Assim, serão valores atípicos os valores x tais que

$$x < Q1 - 1,5AIQ \quad (A.2)$$

ou

$$x > Q3 + 1,5AIQ \quad (A.3)$$

Os valores que se encontram entre $Q1 - 1,5AIQ$ e $Q3 + 1,5AIQ$ são chamados valores adjacentes e sua representação se completa traçando uma linha que vai de $Q1$ até o menor valor adjacente (isto é, o valor mínimo dos dados, excluídos os valores atípicos) e outra que vai de $Q3$ até o maior valor adjacente (isto é, o valor máximo dos dados, excluídos os valores atípicos).

APÊNDICE B – DISPERSÃO DOS GRUPOS SAUDÁVEL E COM FA

A Figura 13 ilustra a dispersão dos grupos de pacientes saudáveis e com FA, e baseado no classificador proposto, definido em 4.7, comparou-se as estatísticas variância, assimetria e curtose com uma análise similar baseada, simplesmente, nos batimentos cardíacos (ver Eq. 4.1).

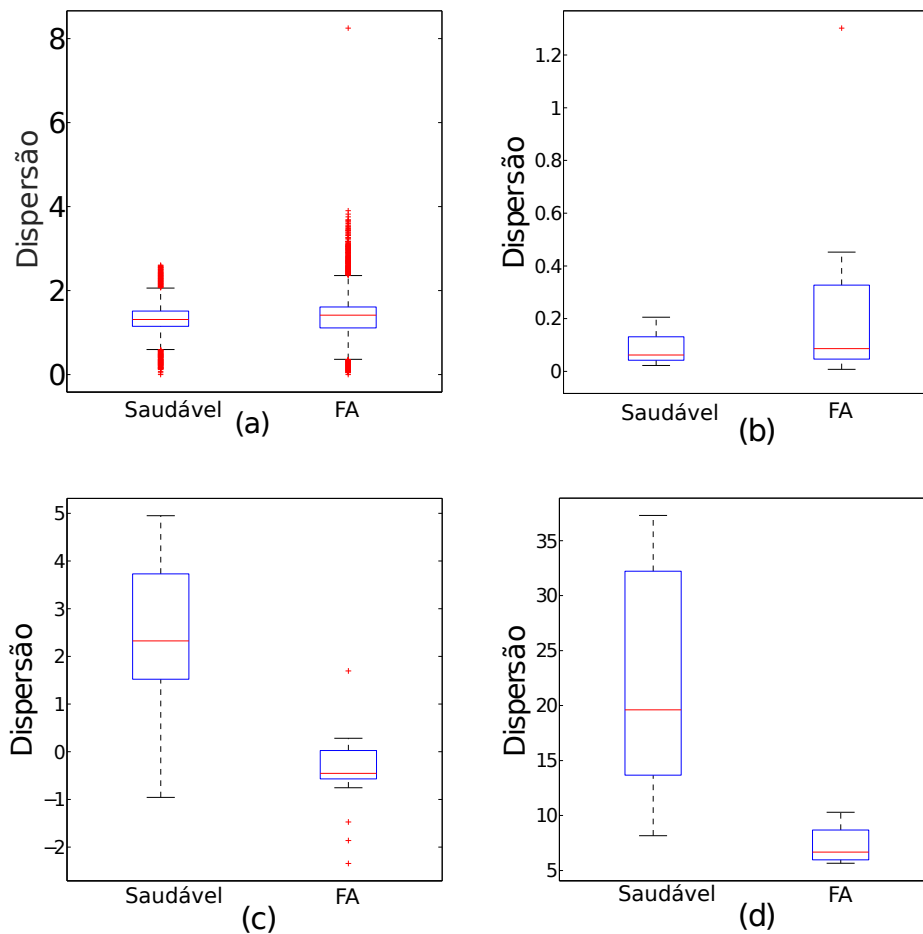


Figura 13 – Dispersão dos grupos de pacientes saudáveis e com FA. a) Análise baseada apenas nos batimentos (ver Eq. 4.1). b) Análise utilizando a variância. c) Análise utilizando a assimetria. d) Análise utilizando a curtose.

É possível observar pela Figura 13 (pelo critério da diferença entre a mediana dos grupos de estudo, descrito por Δ) que a curtose supera as estatísticas variância e assimetria (skewness) com $\Delta \cong 12,94$.

Qualquer dúvida sobre o gráfico boxplot, consultar Apêndice A.

APÊNDICE C – CLASSIFICADOR PROPOSTO UTILIZANDO A MÉDIA

A análise ilustrada na Figura 14 é referente ao classificador proposto utilizando a média.

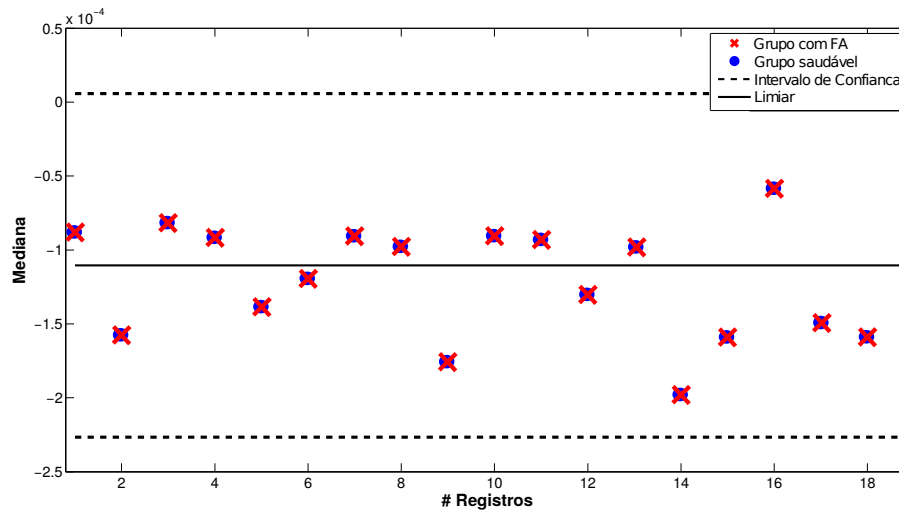


Figura 14 – Resultado para o classificador proposto utilizando a média.

Os resultados utilizando a média não foram tão significativos com relação as demais estatísticas utilizadas nesse trabalho (principalmente, assimetria e curtose), tendo em vista que a média é uma medida de tendência central que é sensível a outliers.

APÊNDICE D – CÓDIGO EM MATLAB

• Script principal

```

1 %% UNIVERSIDADE FEDERAL DO MARANHAO
2 %% MESTRADO DE ENGENHARIA DE ELETRICIDADE
3 %% LABORATORIO DE PROCESSAMENTO DA INFORMACAO BIOLOGICA
4 %% ALFREDO COSTA OLIVEIRA JUNIOR
5 %% VERSION: 1.0
6
7 %%          CODIGO : CLASSIFICACAO DE FIBRILACAO ATRIAL
8
9 %% Referencias:
10 % OLIVEIRA, A. et al. Method to characterize and classify atrial ...
    fibrillation based on kurtosis. p. 1028 1032 , March 2016.
11 % OLIVEIRA, A. et al. Classification of Atrial Fibrillation ...
    Based on ECG Signal Amplitude. 2017.
12
13 clc
14 clear all
15
16 %% Inicializacao de variaveis
17 flagFigura = 0;
18 kurtosisRR = [];
19 energiaRR = [];
20 varRR = [];
21 skewnessRR = [];
22 meanRR = [];
23
24 qrs_amplitude = [];
25 qrs_index = [];
26 histogramas = [];
27 ecgs = [];
28 result = [];
29
30 % Carrega variavel listaRegistros que tem a lista de registros das
31 % bases de dados da MIT-BIH FA e MIT-BIH NSR
32 load('listaRegistros'); % matriz com [ 'nome_registro', ...
    'max_tempo', 'freq_amostragem' ]
33
34 DataInput = listaRegistros; % Variavel auxiliar
35 diretorio = 'basesmit/'; % nome do diretorio que contem as ...
    bases: MIT-BIH FA e MIT-BIH NSR
36

```

```

37 [linhas colunas] = size(DataInput); % guarda a qtd de linhas e ...
    qtd de colunas da matriz DataInput
38
39 for k = 1:linhas
40 % define e abri arquivos .m das bases de dados (BD) MIT-BIH FA e ...
    MIT-BIH NSR
41 % registros de 1 a 18 sao da BD MIT-BIH FA e de 19 a 36 sao da ...
    BD MIT-BIH NSR
42 local = cell2mat(strcat(diretorio, DataInput(k,1)))
43 load(local(1:end)); % vetor com os valorse do sinal de ECG
44
45 minmax = str2num(cell2mat(DataInput(1,2))); % obtem a qtd de ...
    tempo maximo daquele registro
46
47 min = [0 minmax]; % intervalo de tempo de captura: [0 66 minutos]
48
49 Fs = str2num(cell2mat(DataInput(k,3))); % frequencia de amostragem
50
51
52 signal = signal(:); % padroniza forma de vetor signal
53
54 % normalizacao do sinal e limitando o registro de ECG a apenas ...
    66 minutos
55 ecg = signal(1:min(2)*Fs*60)/(200*max(signal(1:min(2)*Fs*60))); ...
    % utiliza esse valor 200 por recomendacao da Physionet
56
57 [ecg_m, qrs_amplitude,qrs_index,delay] = detection_peaksR(ecg, ...
    min, Fs, flagFigura);
58
59
60 %% JANELA DE 60 BATIMENTOS
61 % Caracterizacao por janela (e)
62 kurtosisRR = [ kurtosisRR; kurtosis_intervRRwindow( ecg, ...
    qrs_index, 60, 28) ];
63 % skewnessRR = [skewnessRR; skewness_intervRRwindow( ecg, ...
    qrs_index, 60, 28 )];
64 % varRR      = [ varRR; var_intervRRwindow( ecg, qrs_index, 60, ...
    28) ];
65
66 clear signal ecg % apaga o valor das variaveis ecg e signal
67
68 % armazena os valores das medianas da estatistica desejada ...
    (caracterizacao do registro: E)
69 if(k==linhas-18)
70     medianKurtosisRRFA      = median(kurtosisRR'); %% mediana
71     % medianSkewnessRRFA    = median(skewnessRR'); %% mediana
72     % medianVarianceRRFA    = median(varRR'); %% mediana

```

```

73     clear kurtosisRR
74     kurtosisRR = []; % inicializando variavel
75 end
76 if(k==linhas)
77     medianKurtosisRRNormal = median(kurtosisRR');
78     % medianSkewnessRRNormal = median(skewnessRR'); %% mediana
79     % medianVarianceRRNormal = median(varRR'); %% mediana
80 end
81
82 end
83
84 %% variaveis auxiliares: substituir pela estatistica desejada: ...
    medianKurtosis, medianSkewness, medianVariance
85 medianStatisticRRFA = medianKurtosisRRFA;
86 medianStatisticRRNormal = medianKurtosisRRNormal;
87
88 %%
89 figure,
90 p1 = plot(medianStatisticRRFA, 'ro', 'LineWidth', 5);
91 hold on;
92 p2 = plot(medianStatisticRRNormal, 'bx', 'LineWidth', 5);
93 title('Error using kurtosis method');
94
95 %% Construcao do limiar e intervalo de confianca
96 mediaTotal = mean(medianStatisticRRFA); % media das medianas ...
    (media das caracterizacoes do grupo de FA)
97 desvio_af = std(medianStatisticRRFA); % desvio padrao do grupo ...
    de FA
98 faixa = (3*desvio_af); % 3 desvios padroes (intervalo de ...
    confianca de 99,73%)
99 y1 = (mediaTotal+faixa)*ones(1,23);
100 y2 = (mediaTotal-faixa)*ones(1,23);
101 hold on; p3 = plot(y1, '--k')
102 hold on; p4 = plot(y2, '--k')
103 x = 1:18;
104 y = mediaTotal*ones(1,23);
105 hold on; p5 = plot(y, 'k');
106 legend([p1,p2,p3,p5], 'Grupo com FA', 'Grupo normal', 'Intervalo de ...
    Confianca', 'Limiar');

```

- **Função que janelas o sinal de ECG e o caracteriza utilizando a curtose**

```

1 function [ kurtosisRR ] = kurtosis_intervRRwindow( ecg_m, ...
    qrs_i_raw, windows, qtd )
2 %% function [kurtosisRR]=kurtosis_intervRRwindow(ecg_m, ...
    qrs_i_raw, windows, qtd)

```

```
3 % Characterization algorithm based on ECG signal amplitude using ...
   kurtosis.
4
5 %% Inputs
6 % ecg_m : raw ecg vector signal 1d signal
7 % qrs_i_raw : indice dos picos R
8 % windows : lenght of window in beat
9 % qtd: amount of window
10 %% Outputs
11 % kurtosisRR : vector of windows characterized by kurtosis
12 %% Author : Hooman Sedghamiz
13 % Universidade Federal do Maranhao
14 % email : alfredo.coj@gmail.com
15 % github.com/alfredocoj
16
17 % Any direct or indirect use of this code should be referenced
18 % Copyright 2017
19 %% References :
20 % OLIVEIRA, A. et al. Method to characterize and classify atrial ...
   fibrillation based on kurtosis. 2016 IEEE International ...
   Conference on Industrial Technology (ICIT), p. 1028 1032 , ...
   March 2016.
21 %OLIVEIRA, A. et al. Classification of Atrial Fibrillation Based ...
   on ECG Signal Amplitude. In: 2017 IEEE International ...
   Conference on Biomedical and Health Informatics, Resumo ...
   expandido publicado em Anais da conferncia, 2017.
22
23 % Inicializacao
24 kurtosisRR = [];
25
26 positions = qrs_i_raw';
27 %N = round(length(qrs_i_raw)/windows);
28 N = qtd;
29 aux=windows;
30 i=1;
31
32 % janelamento e caracterizacao o sinal de ECG
33 for data=1:windows:length(positions)
34     if(N ≠ 0 && (aux ≤ length(positions)) )
35         kurtosisRR(i) = ...
           kurtosis(ecg_m(positions(data):positions(aux)));
36         N=N-1;
37         aux=aux+windows;
38         i=i+1;
39     end
40 end
41
```



```
42 end
```

- Demais métodos, bases de dados e arquivos necessários para a execução desse código em matlab estão disponíveis em: <github.com/alfredocoj/dissertacao>.