

Novas Ferramentas Para Análise do Controle Autônomo

Daniel Lucas Ferreira e Almeida, João Luiz Azevedo de Carvalho (orientador)

Departamento de Engenharia Elétrica (ENE/FT)

Universidade de Brasília

Brasília – DF, Brasil

daniel.lucas@msn.com, joaoluiz@pgea.unb.br

Abstract — The analysis of heart rate variability (HRV) aids in the diagnosis of various diseases related to the malfunction of the autonomic nervous system. Traditional approaches for analysis of HRV require the signal to be reasonably stationary during the period of observation. This is not possible when analyzing long duration signals. Detrended fluctuation analysis (DFA) is robust to this issue, as it removes external interferences (“trends”) and considers only intrinsic characteristics which are present throughout the signal. We recently introduced MatLab software for DFA of HRV (BIOSIGNALS 2010:225-229). We now present several improvements to this tool, which may improve the estimation of DFA coefficients. We use the new software to analyze noise signals (white noise signals, pink noise signals and Brownian noise signals) and show that the improvements makes the estimation of DFA coefficients more accurate. Then, we use it to analyze HRV signals from normal subjects, athletes, Chi meditators, and volunteers diagnosed with sleep apnea or epilepsy. The results show that it is possible to clearly distinguish individuals from these groups based on DFA coefficients calculated using the improved implementation of our software.

Keywords — heart rate variability, HRV, detrended fluctuations analysis, DFA, autonomic nervous system.

Resumo — A análise da variabilidade da frequência cardíaca (VFC) auxilia o diagnóstico de várias doenças relacionadas ao mau funcionamento do sistema nervoso autônomo. Abordagens tradicionais para análise da VFC precisam garantir que o sinal apresente um comportamento estacionário durante o período de observação. Isso se torna inviável ao se analisar sinais de longa duração. A análise das flutuações destendenciadas (DFA) torna-se uma solução para esse problema, uma vez que remove interferências externas (“tendências”) e considera apenas as características intrínsecas que estão presentes em todo o sinal. Nós recentemente introduzimos um software em ambiente MATLAB que realiza a DFA para sinais de VFC (BIOSIGNALS 2010:225-229). Este trabalho apresenta diversas melhorias a essa ferramenta a fim de melhorar a acurácia do cálculo dos coeficientes da DFA. A nova ferramenta é utilizada para analisar sinais aleatórios (ruído branco, ruído rosa e ruído Browniano), o que mostra que as melhorias tornaram o cálculo dos coeficiente mais preciso. Então, a ferramenta é usada para analisar sinais de VFC de pessoas normais, atletas, praticantes de meditação Chi e voluntários diagnosticados com apnéia noturna ou epilepsia. Os resultados mostram que é possível distinguir claramente os indivíduos desses grupos baseado nos coeficientes gerados pela DFA.

Palavras-chave — variabilidade da frequência cardíaca, VFC, análise das flutuações destendenciadas, DFA, sistema nervoso autônomo.

I. INTRODUÇÃO

Exames clínicos são ferramentas de extrema importância para o diagnóstico das mais diversas condições patológicas dos pacientes. Tais exames muitas vezes são baseados em sinais biológicos, cuja análise pode detectar irregularidades no funcionamento das funções vitais do ser humano. Por essa razão, ferramentas que permitam ou auxiliem esse tipo de análise clínica são importantes.

Neste trabalho, estivemos em contato desde o início com a análise de flutuações destendenciadas (do inglês *detrended fluctuation analysis*, ou DFA), proposta inicialmente por Peng *et al.* para analisar as sequências de nucleotídeos em cadeias de DNA [1]. Pouco tempo depois, o mesmo grupo usou uma abordagem parecida para a análise de sinais de sinais de variabilidade da frequência cardíaca (do inglês *heart rate variability*, ou HRV) [2]. Essa análise teria como objetivo identificar deficiências no controle simpático-parassimpático sobre o ritmo cardíaco, uma vez que tais características podiam ser extraídas desses sinais, como comprovou o trabalho de Malick e Camm [3]. Outros trabalhos posteriores vieram propor softwares que realizam tal análise de forma consistente [4,5] e, com base nesses trabalhos, trazemos agora novas abordagens que possam tornar a análise DFA mais viável e precisa.

Nesta abordagem apresentada, alguns conceitos importantes relacionados ao HRV e à análise DFA são brevemente comentados (seções I.A, I.B e I.C) e a metodologia usada na pesquisa é explicada (seção II). A pesquisa e os resultados obtidos da mesma estão contidos na seção III e, em seguida, são apresentadas as conclusões e considerações finais (seção IV), além da bibliografia (seção V).

A. O sinal RR

A análise da variabilidade da frequência cardíaca permite verificar o comportamento do sistema autônomo no controle das diversas atividades vitais do corpo humano. A forma como a frequência cardíaca varia pode caracterizar a presença de diversas patologias e a detecção das mesmas é importante para a qualidade de vida do indivíduo.

O sinal de HRV, porém, não é um sinal linear e também não apresenta estacionariedade. O fato de o sinal não ser linear sugere o uso de técnicas não-lineares na sua análise. O fato de ser não-estacionário implica que as características estatísticas do sinal variam ao longo do mesmo, ou seja, diferentes trechos do sinal de HRV irão apresentar diferentes valores para os parâmetros estatísticos.

O sinal RR é utilizado na análise da HRV, pois a onda R é a que apresenta maior facilidade de detecção no traçado do eletrocardiograma (ECG). O intervalo RR é obtido medindo-se o intervalo entre ondas R consecutivas (Figura 1). A detecção das ondas R é discutida em diversos trabalhos. O software ECGLab [4], por exemplo, realiza essa operação e o cálculo dos intervalos RR.

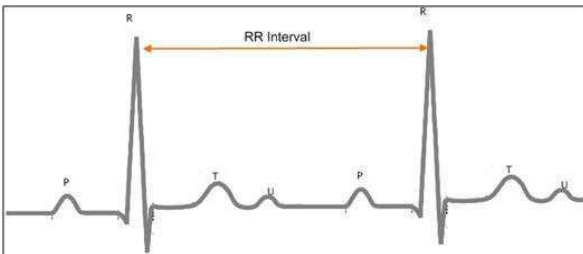


Figura 1. Intervalo RR

B. DFA - Cálculo e Aplicações

Neste trabalho, focaremos apenas na análise destendenciada do sinal de HRV. A análise das flutuações destendenciadas é vantajosa com relação às técnicas mais tradicionais (análises temporal, frequencial e geométrica) porque considera toda a característica do sinal, ou seja, a característica de não-estacionariedade do sinal não influencia o resultado da DFA. Dessa forma, a análise DFA permite encontrar correlações de longo prazo do

sinal, distinguindo as características intrínsecas do sinal daquelas que são geradas por ruídos e/ou outras interferências externas.

O uso da DFA na análise de sinais não-estacionários foi inicialmente proposta por Peng *et al.* [2]. Essa análise gera dois coeficientes, α_1 e α_2 , que representam a correlação de curto e longo prazo, respectivamente, do sinal destendenciado. Essa análise foi validada por Peng *et al.* ao mostrar que essa análise distinguia sinais com características patológicas de sinais de indivíduos saudáveis.

O cálculo da DFA irá gerar, portanto, os dois coeficientes que irão caracterizar o sinal de HRV e as influências do sistema nervoso sobre o ritmo cardíaco. O cálculo tradicional desses coeficientes para um sinal $RR(n)$ baseia-se nos seguintes passos [2,5]:

1. Obter o sinal $y(n) = \sum_{\eta=0}^n [RR(\eta) - \overline{RR}(\eta)]$, onde $\overline{RR}(\eta)$ é o valor médio dos intervalos RR.
2. O “sinal integrado”, $y(n)$, é segmentado em múltiplas janelas de comprimento l_k . Para cada janela, faz-se um ajuste linear pelo método dos mínimos quadrados. O “sinal de tendências”, $y_k(n)$, é uma aproximação por segmentos de reta de $y(n)$, obtida substituindo as amostras de $y(n)$ pelos valores obtidos por meio do ajuste linear em cada janela.
3. Calcula-se o erro de aproximação, $e_k(n)$, denominado “sinal destendenciado”, como: $e_k(n) = y(n) - y_k(n)$.
4. Calcula-se a raiz da média quadrática do erro de aproximação, isto é: $E_k = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} e_k^2(n)}$, onde N é a quantidade de intervalos RR contidos no sinal $RR(n)$.
5. Repetem-se os passos 2 a 4 para diversos valores de l_k .
6. Obtêm-se a curva $f(x)$, onde $x = \log_{10}(l_k)$, como sendo $f(x) = \log_{10}(E_k)$. Como E_k tem relação aproximadamente exponencial com l_k , então $f(x)$ terá relação aproximadamente linear com x .
7. Assim, faz-se um ajuste linear de $f(x)$ pelo método dos mínimos quadrados, para $4 \leq l_k \leq 16$. O coeficiente angular obtido será chamado α_1 .
8. Repete-se o passo 7 para $16 \leq l_k \leq N$. O coeficiente angular obtido será chamado α_2 .

Os passos acima são implementados no software já existente desenvolvido por Leite *et al.* [5], cuja interface gráfica é mostrada na Figura 2. Nessa

interface, é possível ver o sinal RR, o sinal integrado, o sinal de tendências, o sinal destendenciado, a curva $f(x)$ (com os ajustes associados a α_1 e α_2) e os valores dos coeficientes α_1 e α_2 .

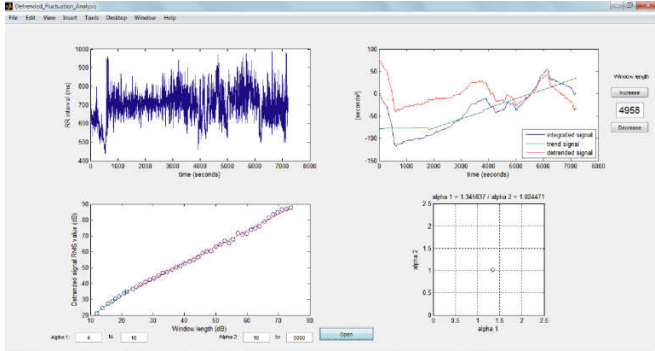


Figura 2. Interface gráfica do software que implementa a DFA, disponível em: <http://pgea.unb.br/~joaoluiz/>

Nosso trabalho consistiu em aumentar a acurácia desse software, garantindo mais precisão nos resultados da DFA. Os métodos utilizados para isso serão detalhados na seção III.A.

C. Batimentos ectópicos e falhas na detecção do sinal RR

O sinal RR muitas vezes apresenta falhas na sua detecção ou batimentos anormais com relação ao ritmo cardíaco monitorado. Tais anormalidades não devem ser consideradas na análise, pois não constituem características do sistema nervoso. Os batimentos ectópicos são batimentos que apresentam uma duração menor do que os outros, que podem ser seguidos por um intervalo compensatório de maior duração. Esses batimentos são decorrentes da estrutura cardíaca e não do controle nervoso. Portanto, estão fora do propósito da análise DFA.

A remoção desses batimentos (Figura 3) pode ser feita utilizando a ferramenta ECGLab desenvolvida por Carvalho *et al.* [4]. Em alguns casos, porém, a forma mais confiável de realizar essa retirada é pela análise do sinal RR, identificando possíveis intervalos de batimentos ectópicos. A retirada dos batimentos irregulares, portanto, é parte essencial do pré-processamento dos sinais que irão passar pela análise DFA.

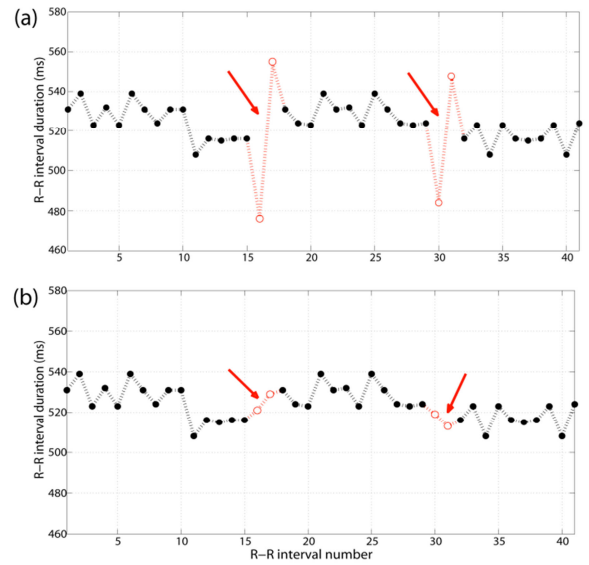


Figura 3. Exemplo da retirada dos ectópicos: (a) sinal com os batimentos ectópicos (indicados por setas); e (b) sinal após a remoção dos batimentos ectópicos e interpolação por splines cúbicas.

II. METODOLOGIA

Para o desenvolvimento do trabalho, utilizamos o software MATLAB. Inicialmente, foram feitos alguns exercícios na área de processamento e análise de sinais a fim de aumentar a familiaridade com a plataforma e com o software já existente [4,5].

Em seguida, começamos a analisar as deficiências do código original. A primeira tarefa realizada foi uma nova abordagem no cálculo das janelas usadas no cálculo da DFA. O tamanho das janelas l_k escolhido deveria ser tal que os valores de $x = \log_{10}(l_k)$ ficassem igualmente espaçados.

Mesmo com a escolha ótima dos valores de l_k , ainda é observada não uniformidade no espaçamento dos valores de x . Assim, fizemos uma alteração na forma como era feito o cálculo dos coeficientes α_1 e α_2 : ao invés de uma interpolação simples, usamos uma interpolação ponderada pelas distâncias entre os valores de x .

Outra modificação proposta foi com relação ao uso das amostras que excediam ou que faltavam ao se dividir o sinal em janelas.

Além disso, começamos a trabalhar recentemente numa nova abordagem para o cálculo da DFA, fazendo com que o sinal de tendências não apresente descontinuidades devido à aproximação com segmentos de reta.

III. PESQUISA E RESULTADOS OBTIDOS

A. Melhorias propostas ao software original

Nesta seção, as melhorias implementadas no software serão apresentadas de forma mais detalhada.

1. Cálculo dos comprimentos das janelas

Como explicado anteriormente, o sinal destendenciado é calculado para diferentes valores de janela l_k . Nós desejamos uma relação entre $f(x) = \log_{10}(E_k)$ e $x = \log_{10}(l_k)$ o mais próxima possível da linearidade, para que os valores dos coeficientes sejam obtidos da regressão linear dos pontos obtidos.

Para que o espaçamento no eixo x seja uniforme, os comprimentos de janela, l_k , devem ser calculados fazendo-se $l'_k = 10^{ks}l_0$ e, a seguir, arredondando o resultado. O espaçamento horizontal dos valores de $f(x)$ é $\Delta f_k = \log_{10}(l_k/l_{k-1})$. A Tabela 1 ilustra o efeito da escolha do passo s . Com s grande, teríamos poucos valores de $f(x)$ para $l_k \leq 16$, o que pioraria a precisão no cálculo de α_1 . Contudo, com s pequeno, são obtidos valores repetidos de l_k . Os valores repetidos poderiam ser descartados, mas note que os espaçamentos para valores pequenos de k seriam bem maiores que para valores grandes de k . O espaçamento deve ser aproximadamente igual a s para todos os valores de k , portanto essa maior não-uniformidade também pode reduzir a precisão no cálculo de α_1 . Com o valor ótimo de s , temos o “melhor dos dois mundos”.

Na implementação desenvolvida por Leite *et al.* [5], usou-se $s = 0,0703$, independentemente do menor comprimento de janela (l_0) avaliado. Se for utilizado $l_0 < 4$, o valor de s precisaria ser aumentado para garantir que não seriam obtidos valores repetidos de l_k . Em nossa implementação, o passo ótimo é calculado em função de l_0 . Isto não só resolve esse problema, como também possibilita uma maior precisão no cálculo de α_2 . Pois, usando $l_0 > 4$, o valor de s utilizado será menor, o que fornecerá mais valores de $f(x)$, melhorando, assim, a qualidade da regressão linear.

Tabela 1: Espaçamento horizontal dos valores de $f(x)$, para diferentes valores de passo e com $4 \leq l_k \leq 15$.

k	$s = 0,10$ (grande)			$s = 0,05$ (pequeno)			$s = 0,0703$ (ótimo)		
	l'_k	l_k	Δf_k	l'_k	l_k	Δf_k	l'_k	l_k	Δf_k
0	4,00	4	—	4,00	4	—	4,00	4	—
1	5,04	5	0,10	4,49	4	0	4,70	5	0,10
2	6,34	6	0,08	5,04	5	0,10	5,53	6	0,08
3	7,98	8	0,12	5,65	6	0,08	6,50	7	0,07
4	10,05	10	0,10	6,34	6	0	7,64	8	0,06
5	12,65	13	0,11	7,11	7	0,07	8,98	9	0,05
6				7,98	8	0,06	10,56	11	0,09
7				8,95	9	0,05	12,42	12	0,04
8				10,05	10	0,05	14,60	15	0,10
9				11,27	11	0,04			
10				12,65	13	0,07			
11				14,19	14	0,03			

O valor ótimo para o passo s foi calculado de modo iterativo, dependendo do comprimento da primeira janela l_0 , para que não existam valores repetidos para o comprimento das janelas após o arredondamento. O código implementado calcula todos os valores de l_k entre l_0 e l_{stop} , para um determinado valor do passo s . Caso existam valores repetidos para o comprimento das janelas, o código incrementa o passo e recalcula os valores de l_k . Esse processo é repetido até que se encontre o passo mínimo ($s_{\text{ótimo}}$) que permita a distinção entre todos os valores de comprimento l_k , garantindo uma excelente precisão.

Fazer a repetição desse processo de l_0 a N , resultaria em tempo desnecessário de processamento, uma vez que a partir de um valor l_{stop} , os valores de l_k não se repetem para qualquer valor de $s < s_{\text{ótimo}}$. O valor de l_{stop} depende do valor inicial l_0 e foi calculado de forma empírica para $1 \leq l_0 < 10^4$. A partir disso, montamos uma tabela de consulta com esses valores (Figura 4), a qual é consultada antes de realizar o processo iterativo.

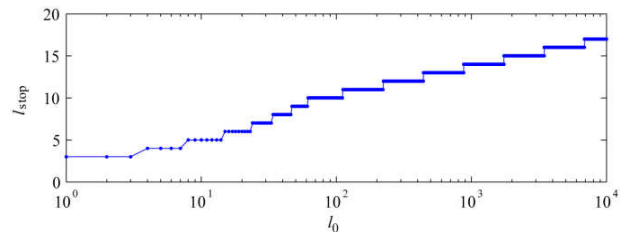


Figura 4. Valores de l_{stop} em função de l_0 , calculados empiricamente.

O cálculo do passo ótimo para diferentes valores de l_k torna-se relevante uma vez que a nova versão do software utiliza dois valores de passo distintos: o primeiro para realizar as operações das janelas a curto prazo (valores típicos de l_k entre 4 e 16) e o outro para as janelas a longo prazo (valores típicos de l_k entre 16 e N). O passo ótimo para um $l_0 = 4$ é $0,0703$ e para $l_0 = 16$ é $0,0216$. Como esperado, quanto maior o l_0 menor o passo $s_{\text{ótimo}}$. Isso garante um maior número de janelas no cálculo de α_2 .

2. Regressão Linear Ponderada

Os coeficientes α_1 e α_2 correspondem aos coeficientes angulares das retas resultantes da regressão linear realizada entre os dados de $f(x)$, como explicado anteriormente. No software original, a regressão realizada é simples e não considera o espaçamento não uniforme de x .

Contudo, percebe-se, pela análise da Tabela 1, que os valores de l_k não são igualmente espaçados após o arredondamento de l'_k . Essa discrepância afeta principalmente os menores valores de comprimento de janela, comprometendo o cálculo de α_1 . A fim de amenizar essa irregularidade, propusemos uma regressão linear ponderada, onde os pesos, w_k , são calculados com base no espaçamento entre os valores de x , desta forma:

$$w_k = (\Delta f_k + \Delta f_{k+1})/2.$$

3. Solução para as amostras faltantes ou excedentes: superposição de janelas

Quando sinal a ser analisado é subdividido em janelas de comprimento l_k , é muito provável que nem todas as amostras do sinal se enquadrem em alguma janela. Matematicamente, é possível que $N / l_k \notin \mathbf{Z}$. No software original, a última janela apresenta um número de amostras menor que l_k . Outra alternativa seria descartar as amostras que entrariam nessa última janela. De ambas as formas, porém, a precisão no cálculo de $e_k(n)$ e de E_k seria afetada.

De modo a tentar corrigir essa imprecisão, buscamos fazer com que todas as janelas apresentassem o mesmo comprimento l_k , sem que nenhuma amostra fosse descartada. Para que isso

fosse feito, usamos do artifício mostrado na Figura 5: superposição de janelas.

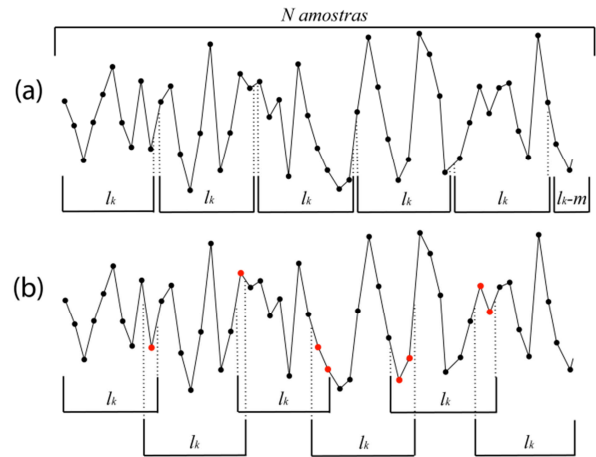


Figura 5. Uniformizando o tamanho das janelas para o cálculo do sinal das tendências: (a) na implementação original sobravam $l_k - m$ amostras quando N/l_k ; e (b) na nova implementação, onde foi utilizada superposição entre as janelas para contornar esse problema.

Apenas para que possamos realizar uma comparação, na Figura 6 temos o resultado do sinal de tendências e do sinal destendenciado utilizando a abordagem do software original (Figura 6a) e a abordagem apresentada neste tópico (Figura 6b).

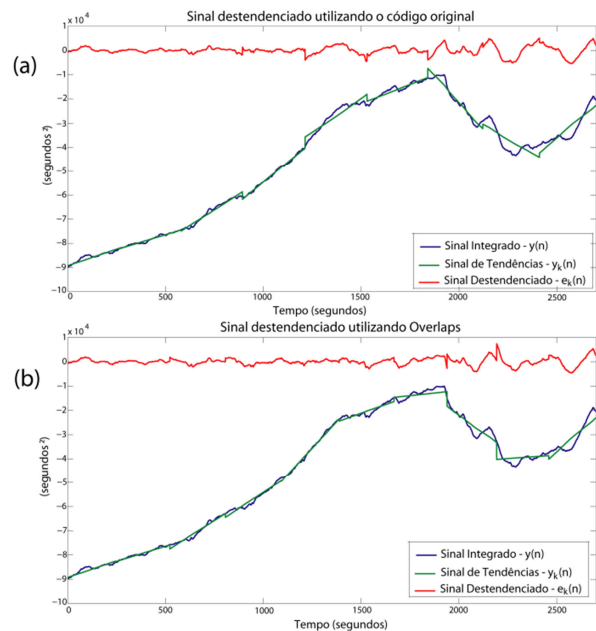


Figura 6. Sinais original, de tendências e destendenciado, para as abordagens (a) do software original e (b) utilizando superposição de janelas. Os gráficos correspondem a $l_k = 437$.

4. Solução para as amostras faltantes ou excedentes: superposição de l_k-1 amostras

Recentemente, começamos a trabalhar numa nova abordagem para o cálculo do sinal de tendências. Ao invés de calcular todo o sinal $y_k(n)$ com um tamanho de janela l_k , e, a partir do sinal $y_k(n)$ obtido, fazer o cálculo de $e_k(n)$, nós propomos fazer isso amostra por amostra.

Consideremos o ponto $y(n')$. Na janela com tamanho centrada em n' — $[y(n' - l_k/2), y(n' + l_k/2 - 1)]$, caso l_k seja par; ou $[y(n' - l_k/2), y(n' + l_k/2)]$, caso seja ímpar —, nós realizamos uma regressão linear simples, de modo que possamos, através dessa aproximação linear, calcular o valor de $y_k(n')$. O valor, nesse ponto, do erro será dado por $e_k(n') = y(n') - y_k(n')$. O sinal destendenciado $e_k(n)$ será composto por todos os valores pontuais calculados dessa forma.

O raciocínio acima vale para todas as amostras do sinal que não se encontram nas extremidades, ou seja, que estão dentro do intervalo $[l_k/2 + 1, N - l_k/2]$. No caso das amostras que estão fora desse intervalo, nós usamos, para o cálculo de $e_k(n)$, as l_k primeiras amostras, para $1 \leq n \leq l_k/2$; ou as últimas l_k amostras, para $N - l_k/2 < n \leq N$.

Para fins de comparação, na Figura 7 temos o resultado do sinal de tendências e do sinal destendenciado utilizando a abordagem do software original (Figura 7a) e a abordagem apresentada neste tópico (Figura 7b).

B. *Validação das alterações usando sinais de ruído*

Foram usados três tipos de sinais de ruídos para a validação das alterações realizadas: ruído branco, ruído rosa ($1/f$) e ruído Browniano ($1/f^2$). O coeficiente α correspondente a cada um desses ruídos é, respectivamente, 0,5, 1,0 e 1,5 [2].

Foram gerados 250 sinais de ruído de cada tipo (com 5000 amostras cada) utilizando o software MATLAB. Os sinais de ruído branco, $h(n)$, foram gerados usando a função *rand*. Os sinais de ruído rosa, $r(n)$, foram obtidos fazendo $r(n) = \mathcal{F}^{-1}\{\sqrt{f} \cdot H(f)\}$, onde \mathcal{F}^{-1} denota uma transformada discreta de Fourier inversa e $H(f) = \mathcal{F}\{h(n)\}$ é a transformada discreta de

Fourier do respectivo sinal de ruído branco. Já os sinais de ruído Browniano, $b(n)$, foram gerados a partir da integração dos sinais de ruído branco a partir do filtro digital integrador:

$$b(n) = b(n-1) + h(n).$$

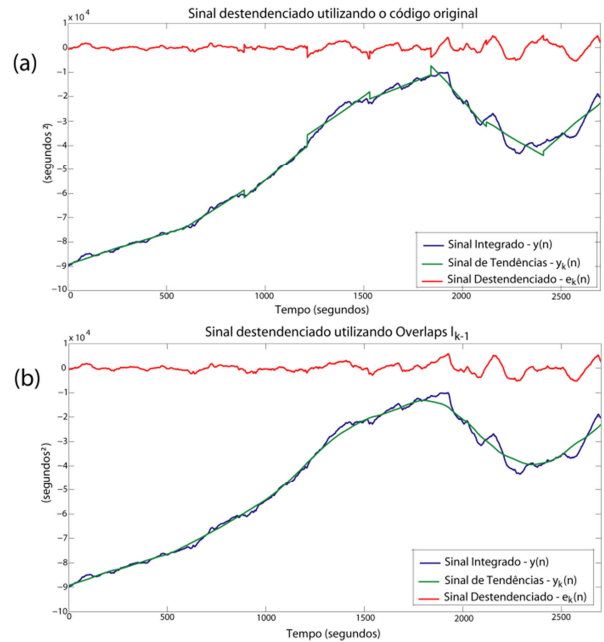


Figura 7. Sinais originais, de tendências e destendenciado para as abordagens (a) do software original e (b) utilizando a abordagem dos *overlaps* l_k-1 . O valor de l_k usado foi de 437.

Para validar as modificações feitas, usamos o seguinte procedimento: fizemos alterações isoladas no código original e, em seguida, calculamos os valores dos coeficientes α_1 e α_2 para cada alteração proposta, individualmente. Em seguida, unimos as alterações mais relevantes ou que apresentaram os melhores resultados isoladamente, e recalculamos os coeficientes para essa nova situação. Esses valores encontrados serão comparados com os valores teóricos para cada tipo de ruído.

Na Figura 8, vê-se, graficamente, os valores dos coeficientes para cada situação isolada: código original, passo ótimo único (mesmo passo foi usado para as janelas pequenas, l_k de 4 a 16, e grandes, l_k de 16 a 5000), dois passos ótimos distintos (um para as janelas menores e outro para as maiores), regressão linear ponderada para o cálculo dos coeficientes, cálculo do sinal destendenciado utilizando superposição e superposição l_k-1 . Na Tabela 2 encontram-se os valores médios e os desvios padrão para os coeficientes α_1 e α_2 calculados para cada conjunto de sinais.

Tabela 2: Valores médios dos coeficientes α_1 e α_2 para as diferentes situações propostas no novo software. Além do valor do desvio padrão de cada uma dessas médias.

Alterações	Alfa 1						Alfa 2					
	Ruído Branco		Ruído Rosa		Ruído Browniano		Ruído Branco		Ruído Rosa		Ruído Browniano	
	\bar{X}	σ^2	\bar{X}	σ^2	\bar{X}	σ^2	\bar{X}	σ^2	\bar{X}	σ^2	\bar{X}	σ^2
Original	0,590	0,012	1,029	0,016	1,502	0,016	0,492	0,029	0,981	0,042	1,478	0,054
Único Passo Ótimo	0,595	0,011	1,031	0,016	1,502	0,016	0,491	0,030	0,980	0,043	1,477	0,056
Dois Passos Ótimos	0,586	0,011	1,027	0,015	1,501	0,016	0,491	0,029	0,981	0,042	1,478	0,054
Regressão Ponderada	0,594	0,011	1,032	0,016	1,503	0,016	0,492	0,029	0,981	0,042	1,478	0,054
Overlaps	0,590	0,012	1,029	0,016	1,501	0,016	0,490	0,029	0,977	0,042	1,473	0,053
Overlaps l_{k-1}	0,564	0,011	1,038	0,013	1,525	0,012	0,484	0,030	0,975	0,042	1,475	0,052

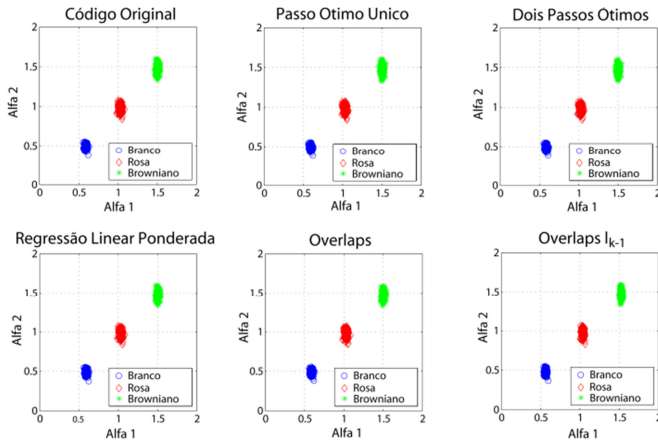


Figura 8. Valores dos coeficientes α_1 e α_2 para as diferentes alterações propostas para o novo software.

Analisando os resultados mostrados na Tabela 2, percebemos que a grande melhora encontrada foi no cálculo do α_1 para sinais de ruído branco, usando a abordagem de superposição l_{k-1} . Outro aspecto que podemos notar, foram valores menores para a medida de dispersão para essa mesma abordagem, o que pode indicar que essa análise consiga identificar de forma melhor as similaridades entre sinais de uma mesma categoria. Para as demais abordagens, não houve alteração relevante.

De uma forma geral, a abordagem que praticamente não apresenta nenhuma vantagem com relação à original é a que usa apenas um passo ótimo. Dessa forma, fizemos combinações entre as demais técnicas da seguinte forma: Teste 1, que irá compreender as técnicas de dois passos ótimos, da regressão linear ponderada e do cálculo do sinal destendenciado usando superposição l_{k-1} ; e o Teste 2, que inclui dois passos ótimos, a regressão linear ponderada e o cálculo do sinal destendenciado usando superposição.

Na figura 9, encontram-se os valores para os coeficientes α_1 e α_2 calculados nas duas situações distintas supracitadas: Teste 1 e o Teste 2. Os

valores das médias e dos desvios padrões para os coeficientes encontram-se na Tabela 3.

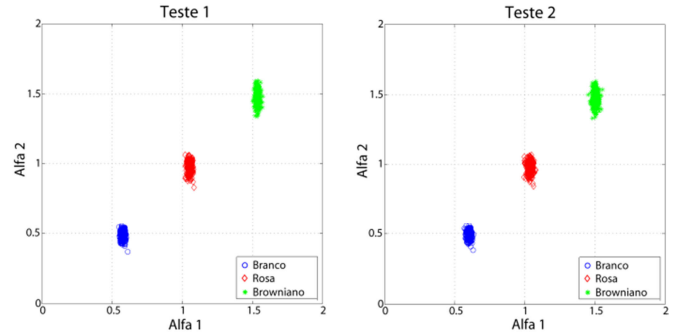


Figura 9. Valores dos coeficientes α_1 e α_2 para duas situações distintas: Teste 1, que compreende a análise DFA com dois passos distintos, regressão linear ponderada e uso de superposição; e Teste 2, que compreende a análise DFA com dois passos distintos, regressão linear ponderada e superposição l_{k-1} .

Novamente, analisando a Tabela 3, percebemos uma grande melhora no valor do coeficiente α_1 para os sinais de ruído branco no Teste 1, que inclui a técnica de utilização de superposição l_{k-1} , bem como menores valores para as medidas de dispersão para o mesmo caso. Percebe-se certa piora no valor do mesmo coeficiente para os ruídos Brownianos. No entanto, os valores médios ainda estão muito próximos para as três abordagens analisadas. Para os outros casos, os valores são praticamente os mesmos.

C. Utilização do software modificado para identificação de grupos patológicos

Os coeficientes de correlação α_1 e α_2 obtidos da análise DFA deveriam indicar as características intrínsecas do sinal. Portanto, sinais de HRV obtidos de pacientes com alguma anormalidade relacionada ao sistema nervoso autônomo devem

Tabela 3: Valores médios dos coeficientes α_1 e α_2 para comparação entre diferentes situações propostas no novo software. Além do valor do desvio padrão de cada uma dessas médias.

Alterações	Alfa 1						Alfa 2					
	Ruído Branco		Ruído Rosa		Ruído Browniano		Ruído Branco		Ruído Rosa		Ruído Browniano	
	\bar{X}	σ^2	\bar{X}	σ^2	\bar{X}	σ^2	\bar{X}	σ^2	\bar{X}	σ^2	\bar{X}	σ^2
Original	0,590	0,012	1,029	0,016	1,502	0,016	0,492	0,029	0,981	0,042	1,478	0,054
Teste 1	0,577	0,011	1,052	0,013	1,537	0,012	0,484	0,030	0,976	0,042	1,476	0,052
Teste 2	0,603	0,011	1,036	0,015	1,503	0,015	0,490	0,029	0,977	0,041	1,473	0,052

apresentar características diferentes de sinais obtidos de pacientes completamente normais.

Com o propósito de atestar essas diferenças, usamos o novo software para analisar sinais obtidos da base de dados do Physiobank (<http://www.physionet.org/physiobank/>) [6, 7, 8]. Os sinais selecionados são de onze indivíduos normais, 8 mulheres e 3 homens, em período de repouso [6]; nove atletas de elite, 3 mulheres e 6 homens, em repouso [6]; oito praticantes de meditação Chi, 5 mulheres e 3 homens, em períodos nos quais há a prática da meditação e em períodos de repouso [6]; doze indivíduos, 11 homens e 1 mulher, com quadro de apneia noturna, cujos dados foram obtidos enquanto dormiam; nove atletas de elite, 3 mulheres e 6 homens, em repouso [7]; e sete indivíduos, 5 mulheres e 2 homens, durante crises epilépticas leves [8].

Antes de realizar a análise DFA desses sinais, porém, foi necessário tratá-los com relação à existência de batimentos ectópicos ou a alguma falha na detecção das ondas R, como explicado anteriormente. Esse procedimento foi realizado com a ajuda do software ECGLab [4]. Os sinais de ruído não passaram por esse procedimento porque não constituem sinais obtidos a partir de um ECG, mas sim gerados computacionalmente.

Após esse procedimento, foi possível realizar a análise DFA e obter os valores dos coeficientes α_1 e α_2 para os diferentes grupos utilizando a DFA com superposição l_k-1 (Figura 10) e com superposição (Figura 11).

A partir dos valores dos coeficientes calculados nas duas análises, fizemos uma análise quantitativa dos mesmos.

1. Análise Quantitativa utilizando o índice Davies-Bouldin

A primeira análise quantitativa realizada foi feita com base no cálculo do índice Davies-Bouldin

(DB) entre os diferentes grupos. O índice DB é uma medida da similaridade desses grupos, baseada na dispersão e na distância entre os mesmos. O índice calculado para dois grupos, por exemplo, é o quociente entre a soma das dispersões de cada grupo e a distância entre os centroides dos mesmos [9]. Se considerarmos as medidas de dispersão como sendo os raios de círculos centrados no centroide de cada grupo, caso a soma dos dois raios seja maior que a distância entre os centroides, isso indicaria que há amostras de um dos grupos dentro do “círculo” do outro grupo, ou seja, os dois grupos não são completamente distintos. Nesse caso, o índice DB é maior que 1. Caso a distância entre os centroides seja maior que a soma da medida de dispersão usada ($DB < 1$), a diferenciação entre os grupos será clara.

Nas tabelas 4 e 5 encontram-se os valores do índice DB calculados entre os diferentes grupos para os coeficientes encontrados a partir da análise DFA usando superposição l_k-1 (Teste 1). Nas tabelas 6 e 7, estão os índices DB para os que foram calculados usando a análise DFA usando superposição (Teste 2). Apesar de ser um teste estatístico bastante limitado, o índice DB nos permite uma análise razoável dos resultados encontrados, aliado aos resultados gráficos das Figuras 10 e 11.

Fazendo uma análise dos resultados encontrados utilizando o Teste 1, percebemos que é possível distinguir alguns desses grupos: os normais dos meditadores (Figura 10a e Tabela 4), os atletas dos meditadores (Figura 10b e Tabela 4) e os normais dos apnéicos e dos epilépticos (Figura 10d e Tabela 5). Os atletas e os indivíduos normais não apresentaram distinções (Figura 10c e Tabela 4), uma vez que esses dois grupos não apresentam diferenças quanto ao funcionamento do sistema nervoso autônomo.

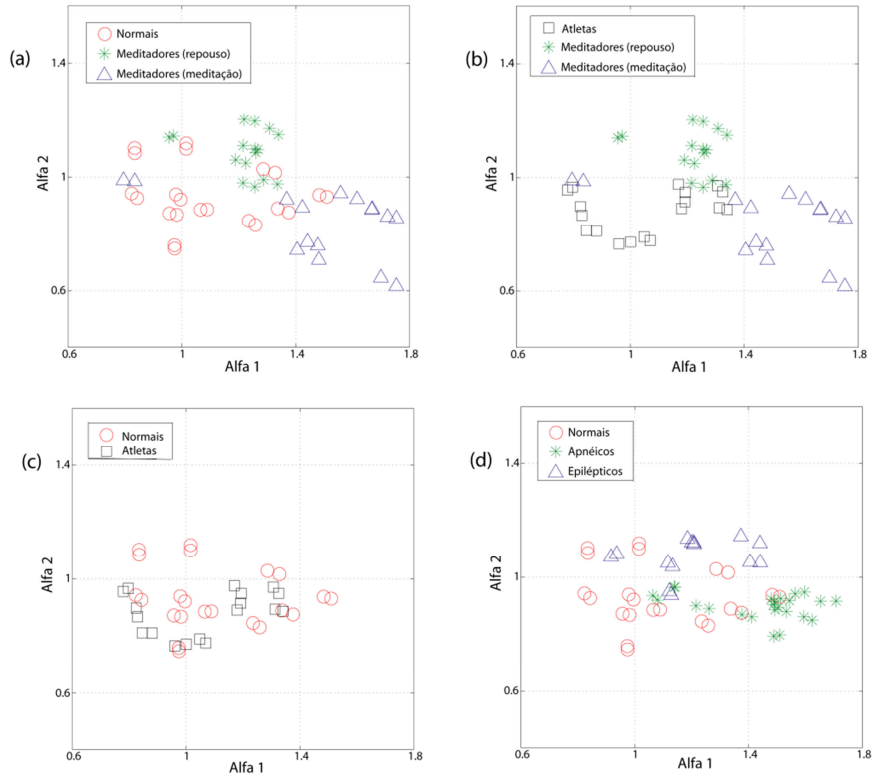


Figura 10. Comparação entre os coeficientes calculados para os diversos grupos utilizando a versão do software proposta em Teste 1: a) normais e meditadores (reposu e meditação); b) atletas e meditadores (reposu e meditação); c) normais e atletas; e d) normais,

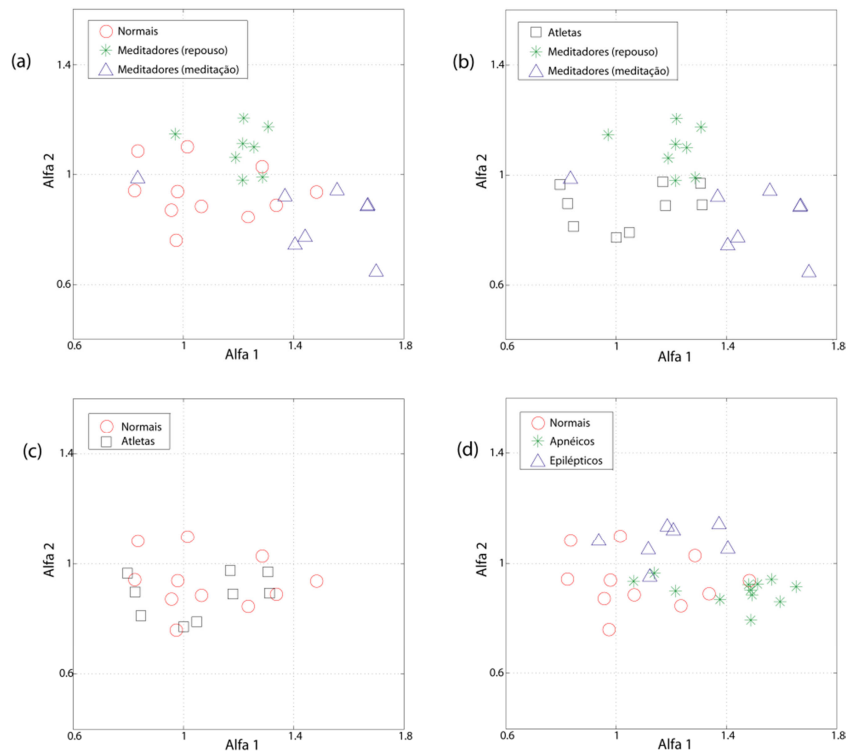


Figura 11. Comparação entre os coeficientes calculados para os diversos grupos utilizando a versão do software proposta em Teste 2: a) normais e meditadores (reposu e meditação); b) atletas e meditadores (reposu e meditação); c) normais e atletas; e d) normais,

Tabela 4: Índice Davies–Bouldin para o Teste 1: indivíduos normais, atletas e meditadores.

Grupos	Índice DB
Normais × Atletas	5,21
Normais × Meditadores (repouso)	1,35
Normais × Meditadores (meditação)	1,12
Atletas × Meditadores (repouso)	1,00
Atletas × Meditadores (meditação)	1,04
Meditadores (repouso) × Meditadores (meditação)	0,96
Normais × Meditadores (repouso) × Meditadores (meditação)	0,89
Atletas × Meditadores (repouso) × Meditadores (meditação)	0,79

Tabela 5: Índice Davies–Bouldin para o Teste 1: indivíduos normais, apnéicos e epiléticos.

Grupos	Índice DB
Normais × Apnéicos	1,07
Normais × Epiléticos	2,06
Apnéicos × Epiléticos	1,16
Normais × Apnéicos × Epiléticos	1,35

Os resultados encontrados pela análise usando o Teste 2 são muito parecidos com os do Teste 1, no que tange a distinção entre os grupos: os normais dos meditadores (Figura 11a e Tabela 6), os atletas dos meditadores (Figura 11b e Tabela 6), os atletas e os normais (Figura 11c e Tabela 6) e os normais dos apnéicos e dos epiléticos (Figura 11d e Tabela 7). A maior diferença é com relação à qualidade dessa distinção. O Teste 2 apresentou valores do índice DB que indicam uma melhor distinção entre os grupos analisados.

Tabela 6: Índice Davies–Bouldin para o Teste 2: indivíduos normais, atletas e meditadores.

Grupos	Índice DB
Normais × Atletas	5,79
Normais × Meditadores (repouso)	1,21
Normais × Meditadores (meditação)	1,12
Atletas × Meditadores (repouso)	0,91
Atletas × Meditadores (meditação)	1,03
Meditadores (repouso) × Meditadores (meditação)	0,86
Normais × Meditadores (repouso) × Meditadores (meditação)	0,76
Atletas × Meditadores (repouso) × Meditadores (meditação)	0,69

Tabela 7: Índice Davies–Bouldin para o Teste 2: indivíduos normais, apnéicos e epiléticos.

Grupos	Índice DB
Normais × Apnéicos	1,06
Normais × Epiléticos	1,81
Apnéicos × Epiléticos	1,07
Normais × Apnéicos × Epiléticos	1,18

2. Análise Quantitativa utilizando testes estatísticos

Existem diversos testes estatísticos que permitem comparar parâmetros de populações distintas (como média e variância). Em geral, esses testes irão confirmar ou refutar uma hipótese nula para um determinado grau de precisão.

Alguns desses testes exigem que os grupos de amostras a serem analisados apresentem uma distribuição normal (chamados de testes paramétricos). Portanto, o primeiro procedimento feito foi testar os diferentes grupos quanto à normalidade.

Como os grupos analisados (pessoas normais, atletas, meditadores, apnéicos e epiléticos) eram grupos pequenos, o teste utilizado para analisar a normalidade foi o de Shapiro–Wilk.

O teste de Shapiro–Wilk tem como hipótese nula a de que o grupo analisado não é normal. Se o resultado do teste for "0", a hipótese não é rejeitada; se o resultado for "1", a hipótese é rejeitada. Todos esses resultados são obtidos com uma determinada significância α , que corresponde a um dos parâmetros para a realização do teste. Nos testes realizados neste trabalho, usamos o valor padrão de $\alpha = 0,05$.

Alguns grupos analisados, porém, não apresentaram distribuição normal. Por causa disso, tivemos que optar por testes não-paramétricos, já que, para usar testes paramétricos, todos os grupos analisados teriam que apresentar uma distribuição normal.

O teste não-paramétrico escolhido para a análise dos grupos foi o teste ANOVA de Friedman. Esse teste irá testar a hipótese de que todos os tratamentos analisados possuem a mesma média. Pelas suposições que o teste utiliza, para que a hipótese não seja rejeitada, as variações entre os

tratamentos devem ser maiores do que as variações dentro de um mesmo tratamento.

O parâmetro resultante do teste corresponde ao quociente da variância entre os tratamentos e a variância dentro dos tratamentos (razão F de variâncias). Caso esse quociente seja maior que 1, há diferença entre os grupos analisados. Além disso, a função que realiza o teste no MATLAB (comando `friedman`), retorna o valor da probabilidade p , que indica se a hipótese nula pode ou não ser rejeitada. Quanto mais próximo o valor de p estiver de zero, maior é a probabilidade das médias dos grupos analisados diferirem entre si.

Nas tabelas 8, 9, 10 e 11 abaixo apresentamos os resultados do teste ANOVA de Friedman para o coeficiente α_1 calculado usando o Teste 1 e o Teste 2. O resultado foi exatamente o mesmo para todos os casos, o que nos mostra que, com relação ao coeficiente de correlação a curto prazo, ambos os testes nos dão a mesma precisão de diferenciação dos diferentes grupos analisados. Algo a ser considerado é que, na análise de três grupos distintos, esse não é um teste totalmente conclusivo. Assim, é necessário a realização de testes *post-hoc* para obter resultados mais confiáveis.

Nas tabelas 12, 13, 14 e 15 a seguir, apresentamos os resultados do teste ANOVA de Friedman para o coeficiente α_2 calculado usando os softwares propostos (Teste 1 e Teste 2). Diferentemente dos resultados obtidos na análise feita para o caso anterior, agora encontramos diferenças entre as versões do software. O teste ANOVA de Friedman mostrou melhores resultados (melhor distinção entre os grupos) para o caso em que utilizou-se o software que usa superposição l_k-1 para o cálculo de α_2 . Além disso, os resultados desse teste indicam que as duas implementações encontram valores semelhantes para α_1 .

De forma geral, o resultado do teste indica que ambas as implementações conseguem distinguir os diferentes grupos de forma satisfatória, seja analisando o coeficiente α_1 ou o coeficiente α_2 .

Tabela 8: Resultados do teste de ANOVA de Friedman para α_1 calculado usando o Teste 1: indivíduos normais, atletas e meditadores.

Grupos	Razão F	p
Normais × Atletas	1,33	0,248
Normais × Meditadores (repouso)	1,35	0,238
Normais × Meditadores (meditação)	8,33	0,004
Atletas × Meditadores (repouso)	3,00	0,083
Atletas × Meditadores (meditação)	8,33	0,004
Meditadores (repouso) × Meditadores (meditação)	8,33	0,004
Normais × Meditadores (repouso) × Meditadores (meditação)	13,17	0,001
Atletas × Meditadores (repouso) × Meditadores (meditação)	14,00	0,001

Tabela 9: Resultados do teste de ANOVA de Friedman para α_1 calculado usando o Teste 1: indivíduos normais, apnéicos e epilépticos.

Grupos	Razão F	p
Normais × Apnéicos	8,33	0,004
Normais × Epilépticos	3,00	0,083
Apnéicos × Epilépticos	5,33	0,021
Normais × Apnéicos × Epilépticos	12,17	0,002

Tabela 10: Resultados do teste ANOVA de Friedman para α_1 calculado usando o Teste 2: indivíduos normais, atletas e meditadores.

Grupos	Razão F	p
Normais × Atletas	1,33	0,248
Normais × Meditadores (repouso)	1,35	0,238
Normais × Meditadores (meditação)	8,33	0,004
Atletas × Meditadores (repouso)	3,00	0,083
Atletas × Meditadores (meditação)	8,33	0,004
Meditadores (repouso) × Meditadores (meditação)	8,33	0,004
Normais × Meditadores (repouso) × Meditadores (meditação)	13,17	0,001
Atletas × Meditadores (repouso) × Meditadores (meditação)	14,00	0,001

Tabela 11: Resultados do teste ANOVA de Friedman para α_1 calculado usando o Teste 2: indivíduos normais, apnéicos e epiléticos.

Grupos	Razão F	p
Normais \times Apnéicos	8,33	0,004
Normais \times Epiléticos	3,00	0,083
Apnéicos \times Epiléticos	5,33	0,021
Normais \times Apnéicos \times Epiléticos	12,17	0,002

Tabela 12: Resultados do teste ANOVA de Friedman para α_2 calculado usando o Teste 1: indivíduos normais, atletas e meditadores.

Grupos	Razão F	p
Normais \times Atletas	3,00	0,083
Normais \times Meditadores (repouso)	8,33	0,004
Normais \times Meditadores (meditação)	8,33	0,004
Atletas \times Meditadores (repouso)	12,00	0,001
Atletas \times Meditadores (meditação)	3,00	0,083
Meditadores (repouso) \times Meditadores (meditação)	12,00	0,001
Normais \times Meditadores (repouso) \times Meditadores (meditação)	20,17	$4,2e^{-5}$
Atletas \times Meditadores (repouso) \times Meditadores (meditação)	19,50	$5,8e^{-5}$

Tabela 13: Resultados do teste ANOVA de Friedman para α_2 calculado usando o Teste 1: indivíduos normais, apnéicos e epiléticos.

Grupos	Razão F	p
Normais \times Apnéicos	0,33	0,564
Normais \times Epiléticos	5,33	0,021
Apnéicos \times Epiléticos	12,00	0,001
Normais \times Apnéicos \times Epiléticos	13,17	0,001

Tabela 14: Resultados do teste ANOVA de Friedman para α_2 calculado usando o Teste 2: indivíduos normais, atletas e meditadores.

Grupos	Razão F	p
Normais \times Atletas	1,33	0,248
Normais \times Meditadores (repouso)	8,33	0,004
Normais \times Meditadores (meditação)	8,33	0,004
Atletas \times Meditadores (repouso)	12,00	0,001
Atletas \times Meditadores (meditação)	1,33	0,248
Meditadores (repouso) \times Meditadores (meditação)	12,00	0,001
Normais \times Meditadores (repouso) \times Meditadores (meditação)	20,17	$4,2e^{-5}$
Atletas \times Meditadores (repouso) \times Meditadores (meditação)	18,67	$8,8e^{-5}$

Tabela 15: Resultados do teste ANOVA de Friedman para α_2 calculado usando o Teste 2: indivíduos normais, apnéicos e epiléticos.

Grupos	Razão F	p
Normais \times Apnéicos	0,33	0,564
Normais \times Epiléticos	3,00	0,083
Apnéicos \times Epiléticos	12,00	0,001
Normais \times Apnéicos \times Epiléticos	11,17	0,004

IV. CONCLUSÃO

Ao longo de nosso trabalho, estivemos trabalhando sempre com a análise das flutuações destendenciadas, a fim de aprimorar essa análise e torná-la mais confiável. Os estudos realizados previamente foram essenciais para a compreensão dos conceitos relacionados a ela e para facilitar a manipulação dos códigos que implementam a DFA. Além disso, a familiarização com o ambiente do MATLAB e o conhecimento prévio em programação auxiliaram muito no progresso do projeto. Os principais resultados foram apresentados neste trabalho e mostraram-se muito satisfatórios. Além disso, mostraram que a análise DFA pode tornar-se cada vez mais precisa e mais útil ao auxiliar nos diagnósticos de patologias relacionadas ao sistema nervoso autônomo.

Na continuação deste trabalho, iremos trabalhar mais com a nova abordagem para o cálculo da DFA, utilizando superposição l_k-1 . Iremos analisar de forma mais contundente os efeitos dessa nova abordagem sobre os resultados da DFA. Além disso, procuraremos novas aplicações para o uso da análise destendenciada, para que ela possa ser útil em outras análises além da que foi apresentada nesse trabalho.

V. REFERÊNCIAS

- [1] Peng, C. K., Buldyrev, S. V., Havlin, S., Simons M., Stanley H. E., and Goldberger, A. L., "Mosaic Organization of DNA nucleotides", In: *Physical Review E*, vol. 49, n° 2, p. 1685-1689, 1994.
- [2] Peng, C. K., Havlin, S., Stanley, H. E., and Goldberger, A. L. (1995), "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series", In: *Chaos*, 5(1):82-87.
- [3] Malik, M., Camm, A. J. (1995), "Heart Rate Variability", In: *Futura Publishing Company, Inc.*, Armonk, NY.

- [4] Carvalho, J. L. A., da Rocha, A. F., Nascimento, F. A. O., Souza Neto, J., and Junqueira Jr, L. F. (2002) "Development of a Matlab software for analysis of heart rate variability", In: *ICSP'02, 6th International Conference on Signal Processing*, v. 2, p. 1488-1491.
- [5] Leite, F. S., Carvalho, J. L. A., da Rocha, A. F. (2010), "MatLab Software for Detrended Fluctuation Analysis of Heart Rate Variability", In: *BIOSIGNALS 2010 - International Conference on Bio-inspired Systems and Signal Processing*, p. 225-229.
- [6] Peng, C. K., Mietus, J. E., Liu, Y., Khalsa, G., Douglas, P. S., Bensonb, H., Goldberger, A. L. (1999) "Exaggerated heart rate oscillations during two meditation techniques", In: *International Journal of Cardiology*, 70, p. 101-107.
- [7] Penzel, T., Moody, G. B., Mark, R. G., Goldberger, A. L., Peter, J. H. (2000) "The Apnea-ECG Database", In: *Computers in Cardiology*, 27:255-258.
- [8] Al-Aweel, I. C., Krishnamurthy, K. B., Hausdorff, J. M., Mietus, J. E., Ives, J. R., Blum, A. S., Schomer, D. L., Goldberger, A. L. (1999) "Post-Ictal Heart Rate Oscillations in Partial Epilepsy", In: *Neurology*, 53(7) : 1590-1592.
- [9] Davies, D. L., Bouldin, D. W. (1979) "A Cluster Separation Measure", In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. Pami-1, n°. 2, p. 224-227.
-