

Classificação de doenças cardíacas através de eletrocardiogramas

Mikaela Baldasso IME-UFRGS Marcio Valk IME-UFRGS

Introdução

Doenças cardiovasculares (DCV) continuam sendo a principal causa de morbidade e mortalidade no mundo todo, representando 31% de todas as mortes globais, [2].

(ECG's) e fonocardiogramas (PCG's) são séries temporais.

Objetivo: Encontrar medidas de distâncias que sejam capazes de capturar diferenças entre dois ECG's (ou PCG's), quando elas existem.

Abordagens comuns na análise de sinais como a filtragem e a autocorrelação podem ser usadas na classificação dos sinais.

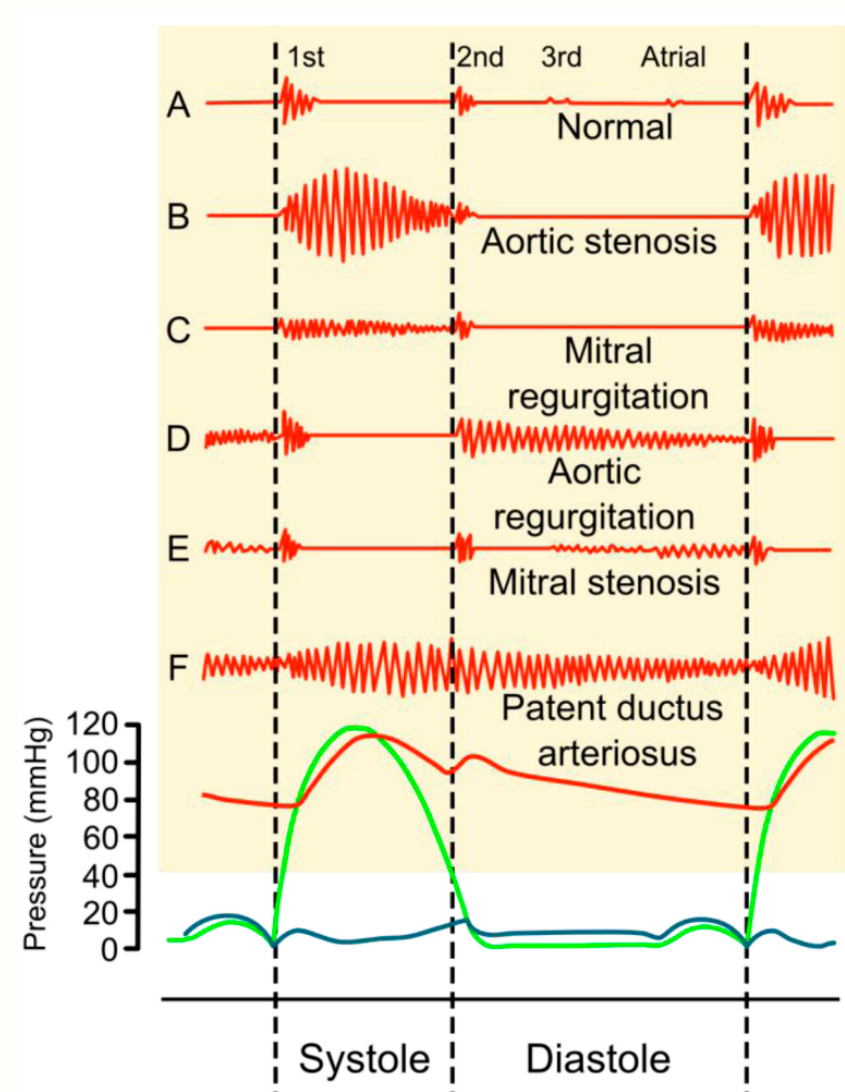


Figure 1: Fonocardiograma de sons cardíacos. Vermelho indica pressão aórtica, verde indica pressão ventricular e azul a pressão atrial. Fonte: [2].

Descrição dos dados

A Figura 2 apresenta 4 imagens de gravações de sinais disponibilizadas pelo "National Institute of General Medical Sciences": dois de pacientes saudáveis e dois de pacientes com algum problema cardíaco.

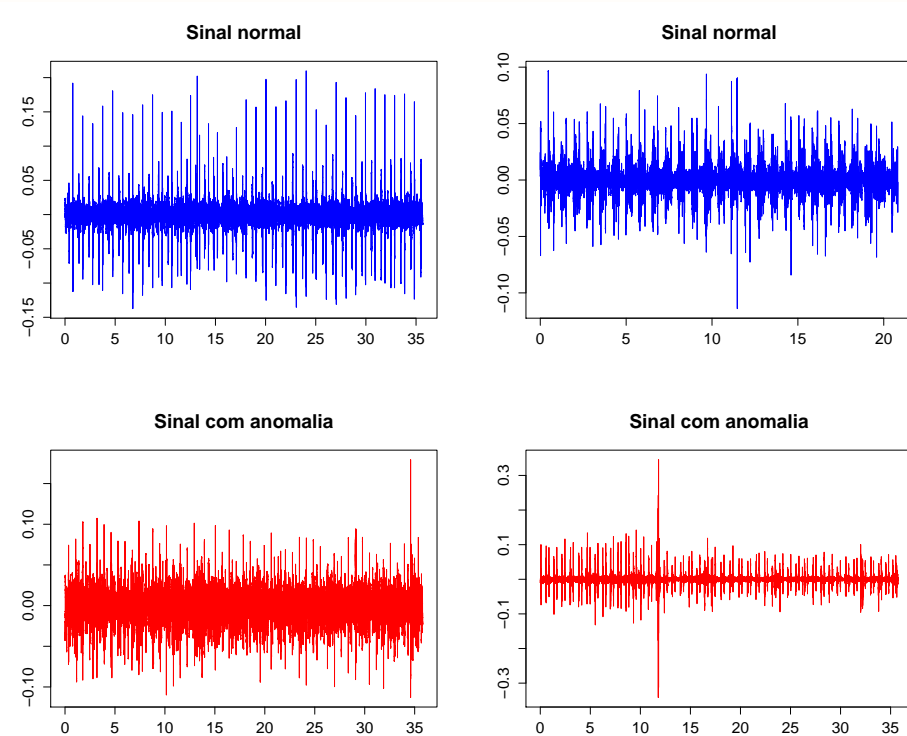


Figure 2: Exemplos de sinais cardíacos com arritmia e sem arritmia.

Metodologia

A metodologia adotada nesse trabalho segue basicamente aquela descrita em [1]. Uma decomposição wavelet é realizada e então é calculada a energia de Shannon (ASE).

$$e(n) = -\frac{1}{N_w} \sum_{j=1}^{N_w} d^2(j) \log(d^2(j)),$$

onde N_w é o comprimento da janela e $d(j)_{j=1}^{N_w}$ são os coeficientes wavelets de uma parte específica da decomposição. Por fim, ainda é calculada a função de autocorrelação da ASE, chamada de ACF.

$$r(m) = \frac{1}{r(0)} \sum_{n=0}^{N-m-1} e(n)e(n+m), \quad m > 0$$

onde $e(n+m)$ é a versão deslocada no tempo do sinal de $e(n)$ da ASE com um intervalo de tempo m para $m = 0, 1, \dots, M$.

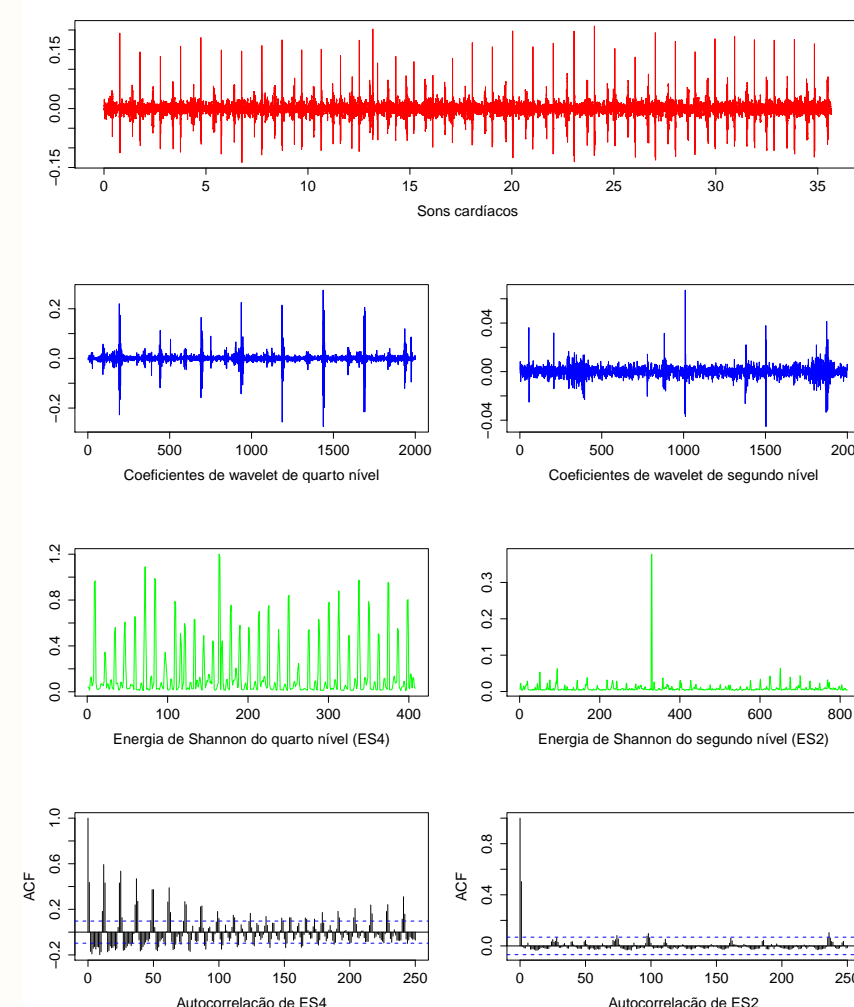


Figure 3: Autocorrelações dos coeficientes do som cardíaco com ruídos.

Função kernel gaussiano como medida de dissimilaridade

Em classificação e agrupamento de séries temporais é fundamental encontrar uma medida de dissimilaridade, como, por exemplo, a função kernel gaussiano encontrada em Deng2016.

Se n o número de sinais e \mathbf{r}_i o adACFC do sinal $i \in \{1, \dots, n\}$, que é um só vetor com as componentes dos coeficientes de aproximação e detalhes da TDW, $\mathbf{r}^{ad} = [\mathbf{r}^a, \mathbf{r}^d] \in \mathcal{R}^{2M}$, define-se a medida de similaridade baseada na função kernel gaussiano entre os sinais i e j por

$$\omega(i, j) = \exp \left\{ -\frac{\|\mathbf{r}_i - \mathbf{r}_j\|^2}{\delta^2} \right\},$$

em que δ^2 é a largura do kernel.

Avaliação do Desempenho dos Métodos de Classificação

Considerando os indivíduos que tem a doença (VP), que não tem a doença (FP), indivíduo que não tem a doença mas é doente (FN) ou não doente (VN). As comparações que levam em conta esses aspectos são as mais precisas e, no entanto, as mais usuais são

$$\text{Sensitividade: } Se = \frac{VP}{VP + FN}$$

$$\text{Especificidade: } Es = \frac{VN}{FP + VN}$$

$$\text{Precisão/Acurácia: } Pr = \frac{VP + VN}{N}$$

Resultados

A Figura 4 apresenta um dendrograma com uma possível separação, em que podemos observar 3 grupos.

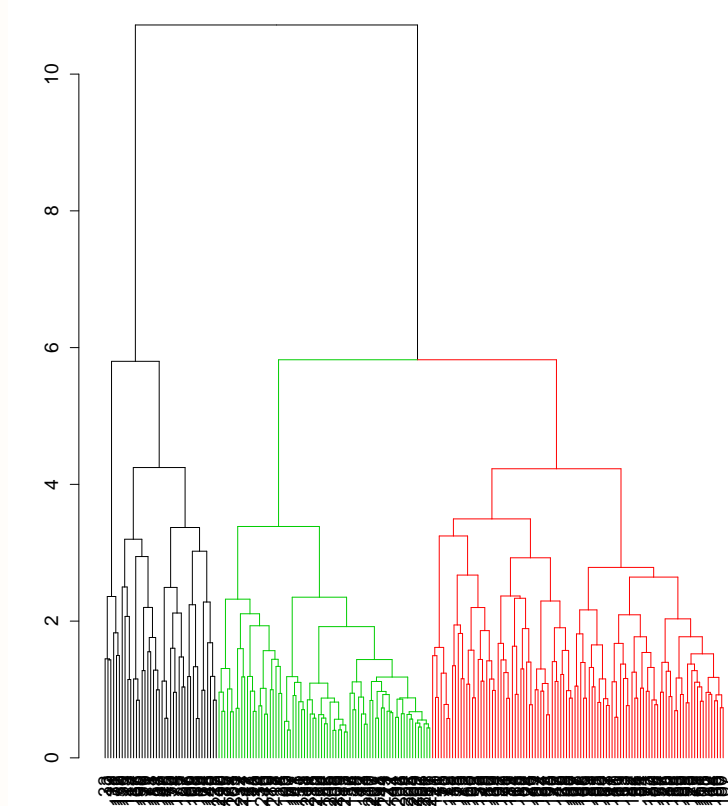


Figure 4: Dendrograma com os grupos separados por cores. Preto (não classifica), vermelho (anormal) e verde (normal).

Os resultados obtidos nesse trabalho podem ser comparados, ou são comparáveis, com os resultados de [3] o qual apresenta uma precisão de 78%. Os resultados preliminares apresentam uma sensibilidade de 57%, especificidade de 76.92% e uma precisão de 65.73%.

Referências

- [1] Shi-Wen Deng and Ji-Qing Han. Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps. *Future Generation Computer Systems*, 60:13–21, 2016.
- [2] Chengyu Liu et.al. An open access database for the evaluation of heart sound algorithms. *Physiological Measurement*, 37(12):2181, 2016.
- [3] Philip Langley and Alan Murray. Abnormal heart sounds detected from short duration unsegmented phonocardiograms by wavelet entropy. In *Computing in Cardiology Conference (CinC)*, 2016, pages 545–548. IEEE, 2016.