Detecção e segmentação de sons cardíacos

TAAD - 2022

Beatriz Negromonte - 2017109753
Francisco Relvão - 2018285965

Motivação

Doenças cardiovasculares são umas das causas líderes mundiais de morte, especialmente em zonas rurais.

Motivação

A detecção de sons cardíacos depende:

- Experiência de auscultação do profissional de saúde (em declínio nos últimos anos);
- Equipamento médico e hospitalar

É necessário automatizar o diagnóstico de doenças cardiovasculares.

Objetivos

- Prever S1 e S2 ao longo de um ficheiro de áudio;
- Obter a frequência cardíaca calculada a partir de S1 e S2.

Para isso iremos:

- Treinar vários modelos
- Escolher o que melhor se adequa aos dados
- E fazer previsões com o melhor modelo

1

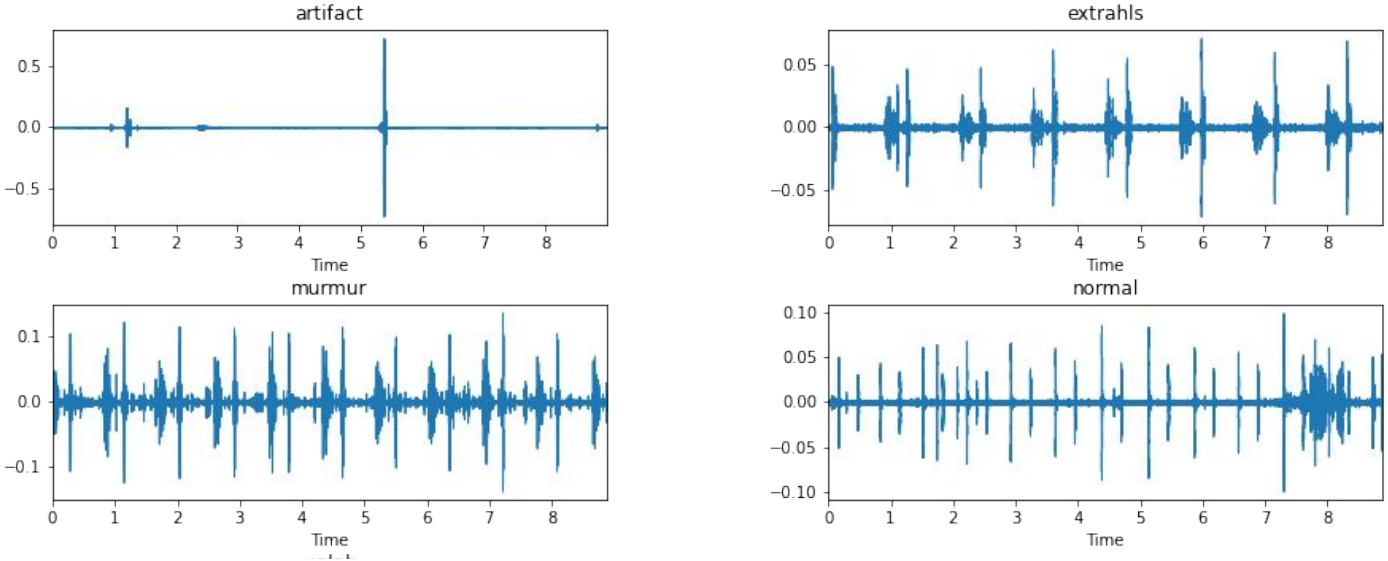
Sons cardíacos e Transformadas

Sons cardíacos: S1 e S2

Fig. (1). Relationship between the heart sounds and the ECG wave-form.

Amine, Debbal & Reguig, Fethi. (2010). Features for Heartbeat Sound Signal Normal and Pathological. Recent Patents on Computer Science. 1. 1-8. 10.2174/1874479610801010001.

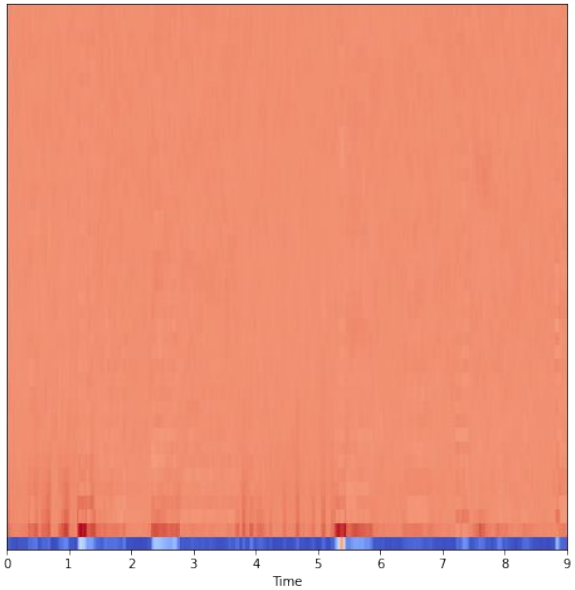
Sons cardíacos: S1 e S2



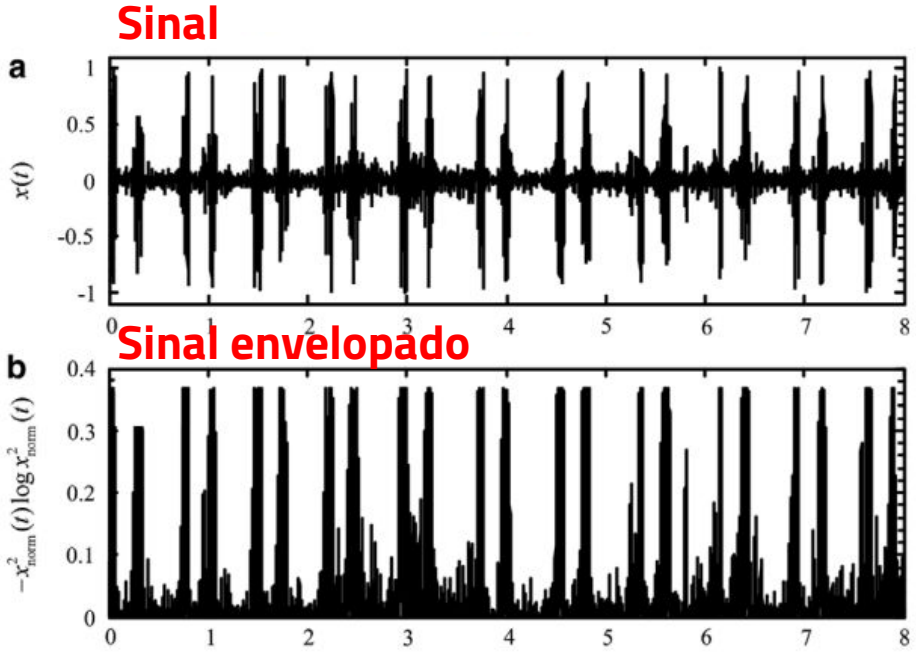
Coeficientes espectrais das frequências de Mel - MFCCs

Objetivo: simplificar e filtrar o sinal consoante a percepção do ouvido humano

Analisar o comportamento da frequência perceptual ao longo do tempo num ficheiro de áudio.



Envelopamento de Shannon



Choi, Samjin and Zhongwei Jiang. "Comparison of envelope extraction algorithms for cardiac sound signal segmentation." *Expert Syst. Appl.* 34 (2008): 1056-1069.

- Técnica popular de extração de envelope de sinais cardíacos.
- Dá ênfase local a sinais de intensidade média e atenua sinais de baixa e alta intensidade, sendo a atenuação maior para sinais de baixa intensidade.

$$\text{Shannon energy} = -x^2_{\text{norm}}(t)\log x^2_{\text{norm}}(t).$$

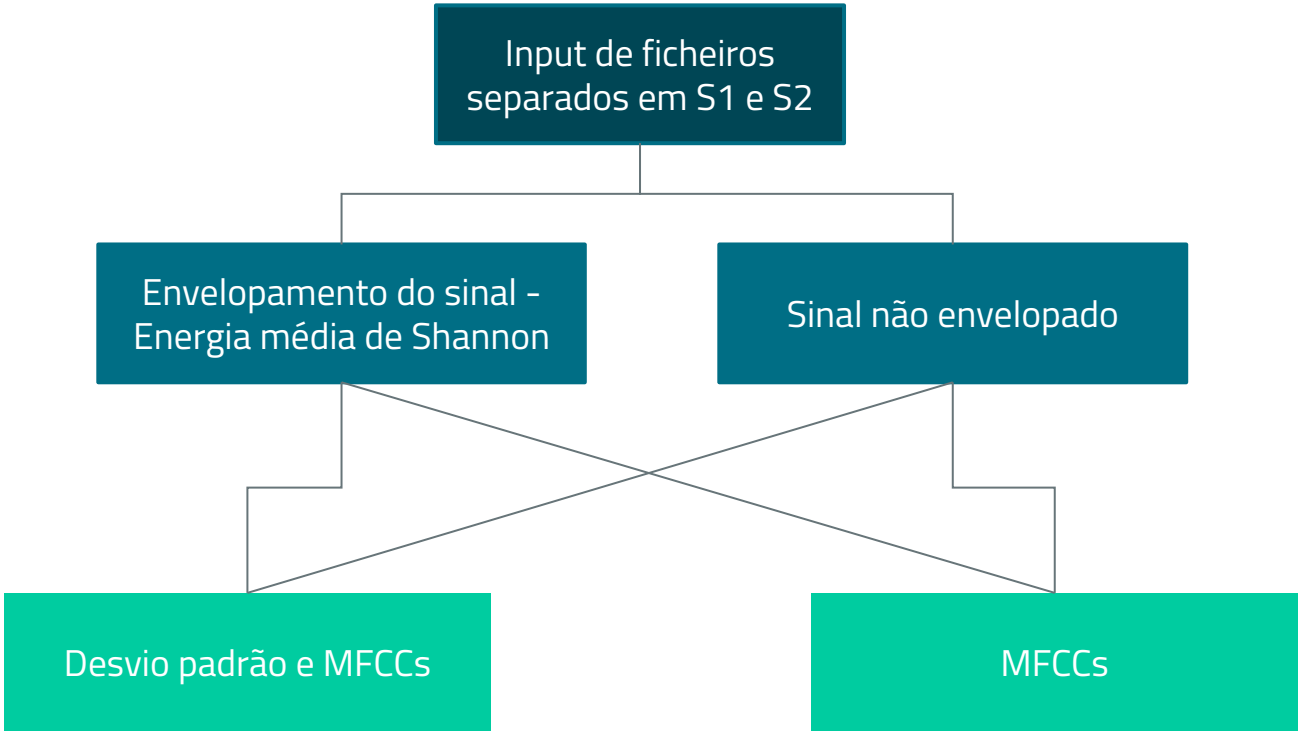
2

Treino de redes neurais

Base de dados

- Base de dados com 176 áudios, retirada do set A do Classifying Heart Sounds challenge
- Os áudios foram obtidos através da aplicação iStethoscope Pro, sendo gravados pelos usuários no dia a dia com uso do telemóvel.
- Classificações: Normal, Murmur, Extra Heart Sound, Artifact e Unlabelled
- Dos 176 áudios, foram usados 21 com classificação “Normal” para separar os trechos de S1, S2 e Noise e treinar a rede neural.
- Para os áudios usados no treino, tinham-se dados com informações acerca da localização exata do S1 e do S2

Metodologias: Características de audio



Rede Neural: Arquitetura Linear

Input: Vetores de 40 a 80 MFCCs ou MFCCs + Desvio-padrão com ou sem envelope

Camadas: 3 camadas com 512 neurónios, ativação Relu e dropout de 50%

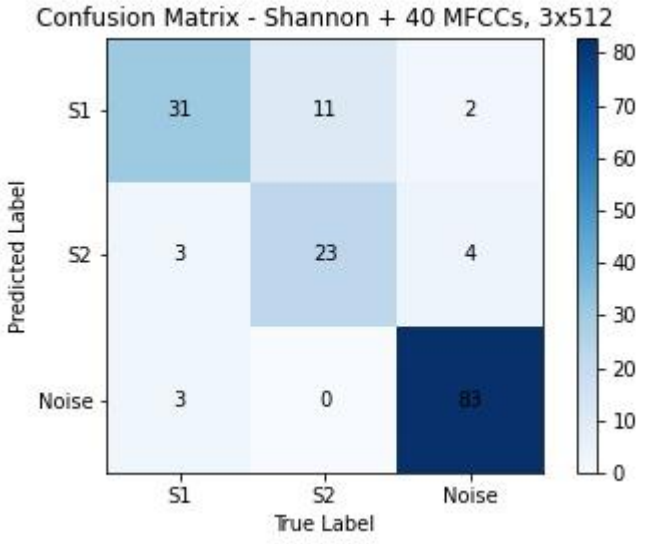
Output: Ativação Softmax com três categorias (S1, S2, Noise). A saída é também um vetor

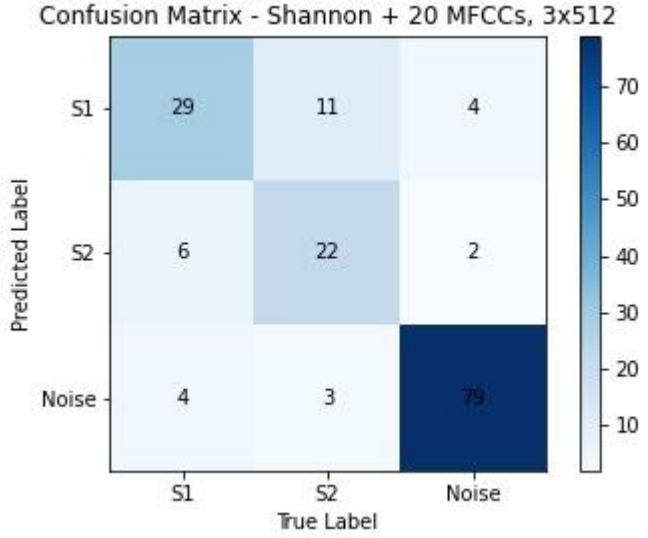
Resultados: Arquitetura Linear

Com envelope e MFCCs

Previsão
S1: 23%
S2: 21%
Ruído: 56%

Accuracy
S1: 69%
S2: 75%
Ruído: 97%





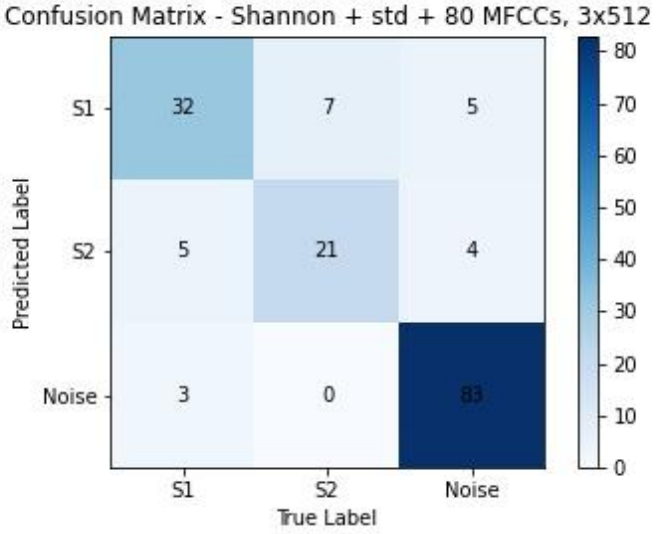
Previsão
S1: 24%
S2: 22%
Ruído: 53%

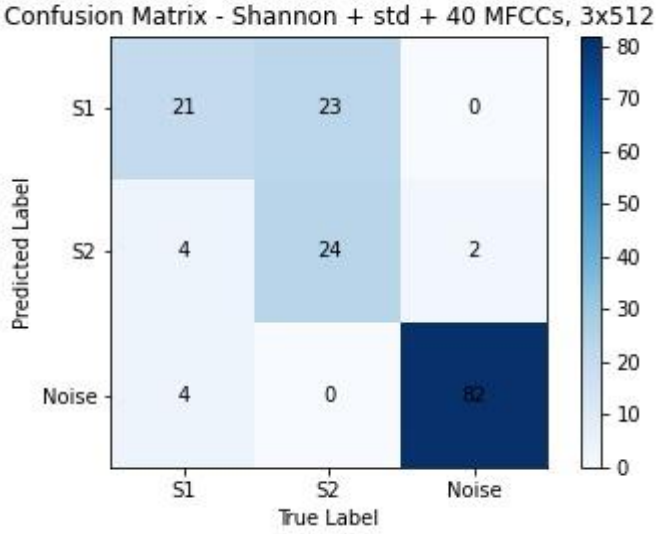
Accuracy
S1: 64%
S2: 72%
Ruído: 93%

Em todas as matrizes: S1 (28%) S2 (19%) Ruído(53%)

Resultados: Arquitetura Linear

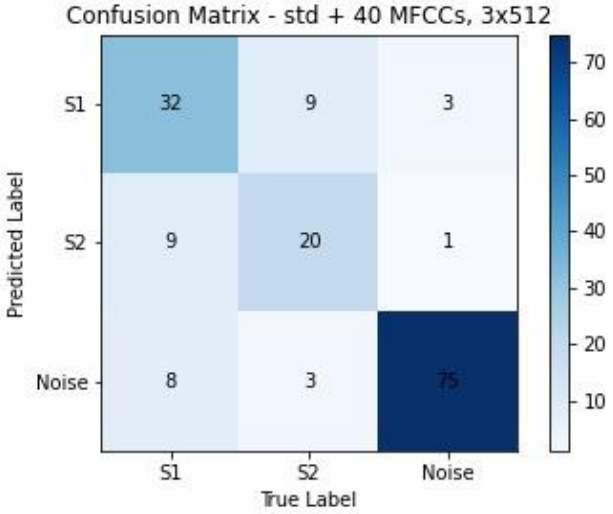
Com envelope, MFCCs e Desvio padrão

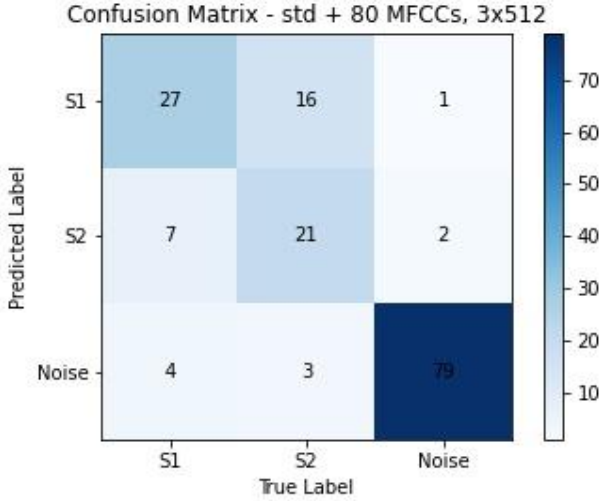




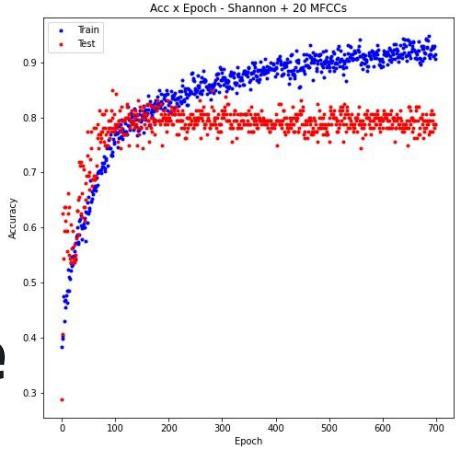
Resultados: Arquitetura Linear

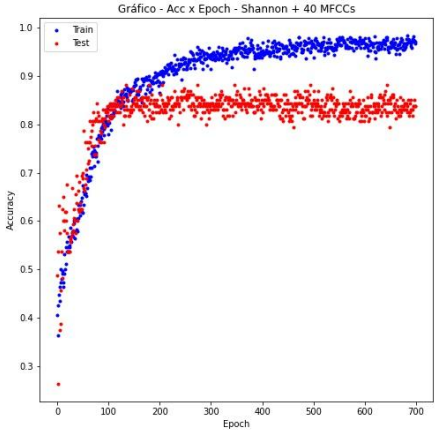
Sem envelope, com MFCCs e Desvio padrão

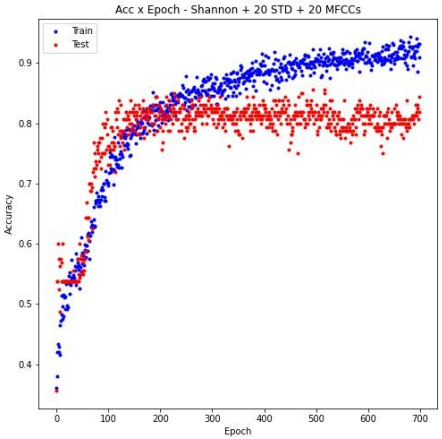


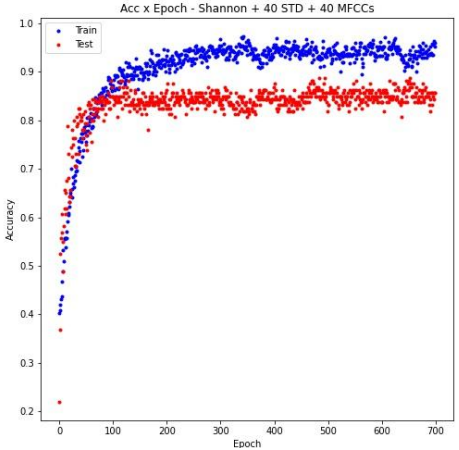


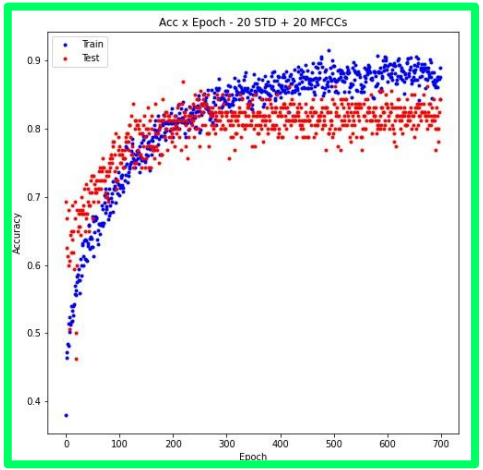
Gráficos de Precisão: Teste e Treino

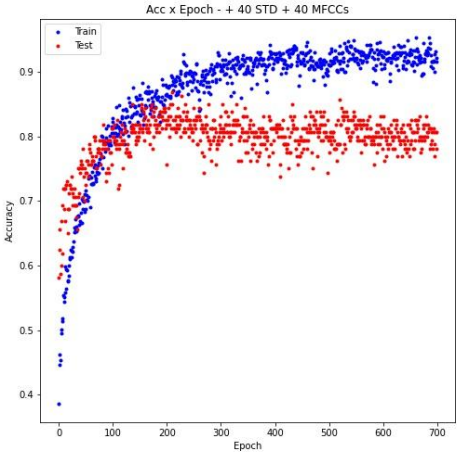




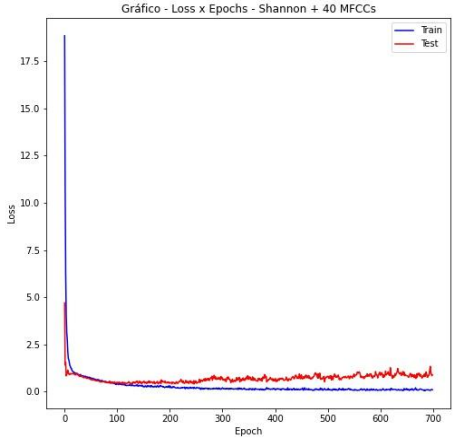


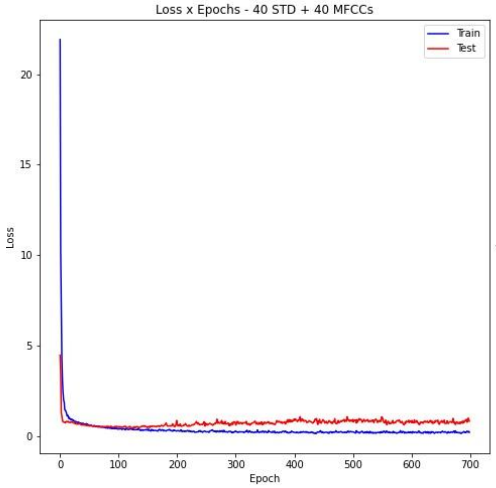


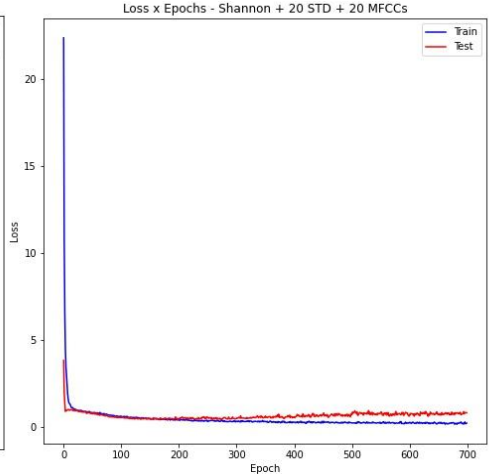


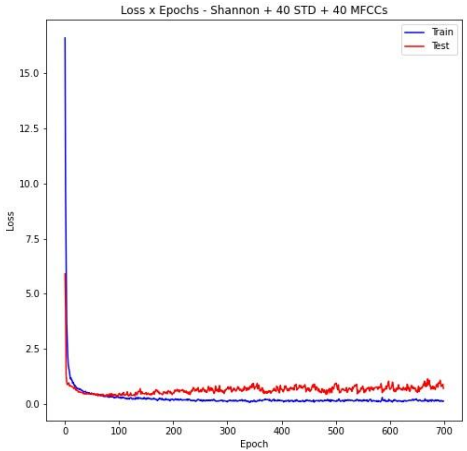


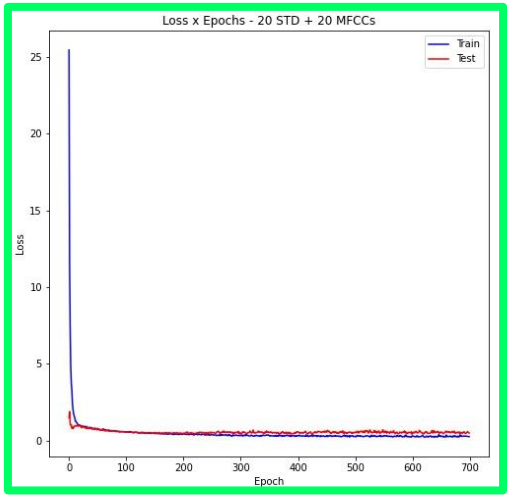
Gráficos de Perda: Teste e Treino

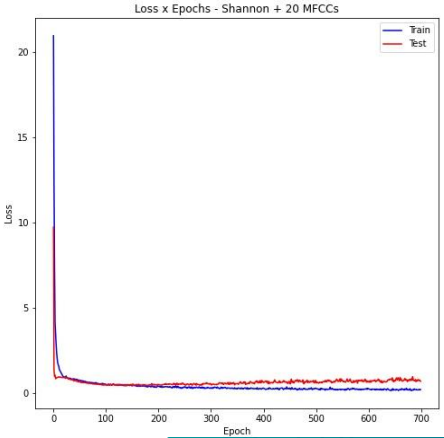












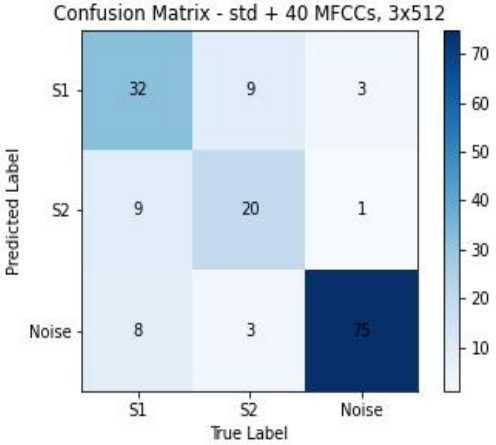
Rede Neural: Arquitetura Funil

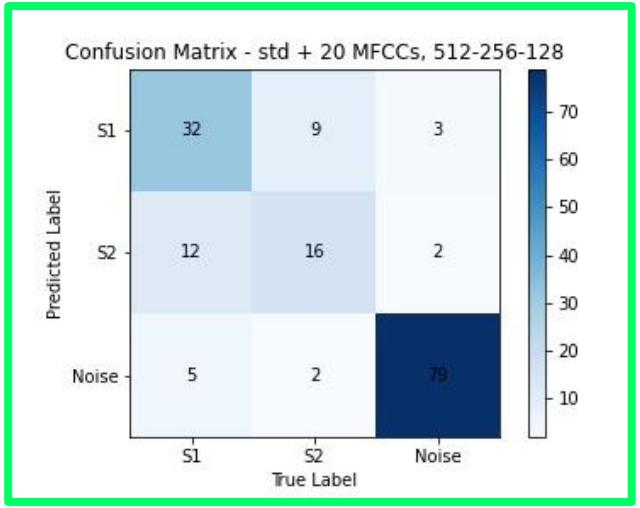
Input: Vetores 20 MFCCs + 20 Desvio-padrão

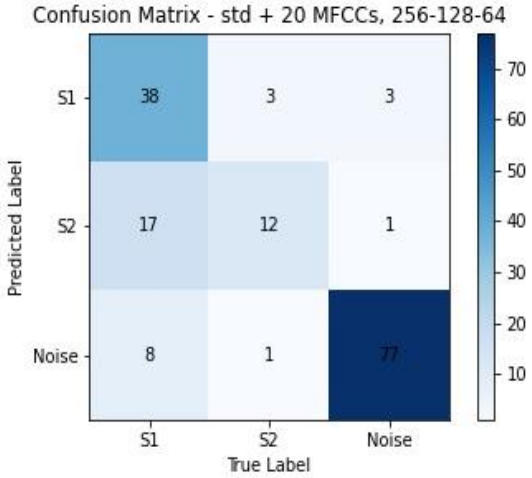
Camadas: 3 camadas com ativação Relu e dropout de 50%. O número de neurônios é reduzido na direção do output: input-512-256-128-output

Output: Ativação Softmax com três categorias (S1, S2, Noise). A saída é também um vetor

Melhor modelo: 20 MFCCs e 20 D.Ps.



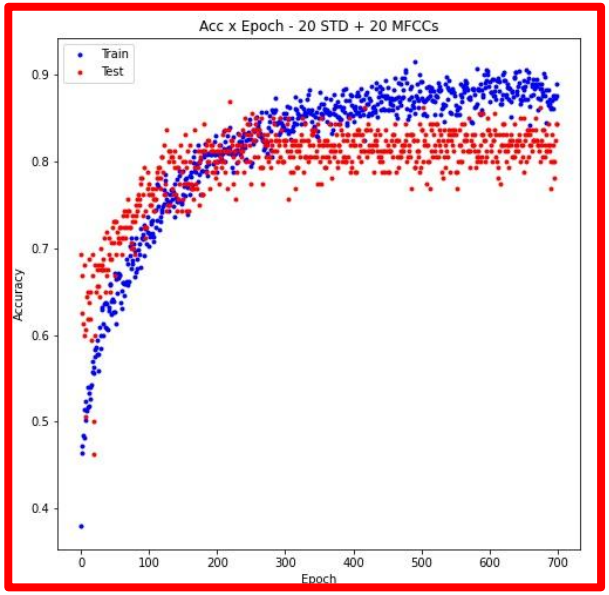


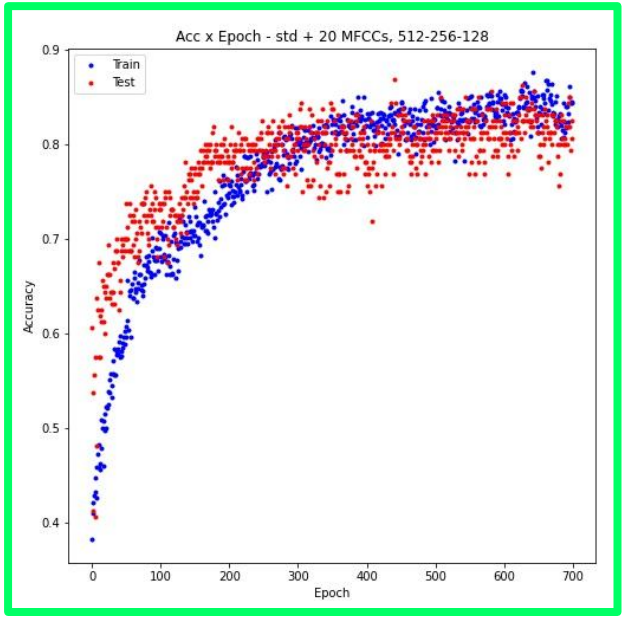


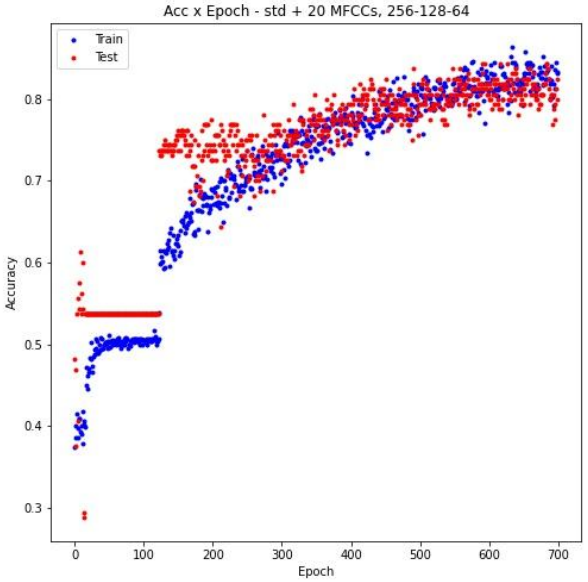
Em todas as matrizes: S1 (28%) S2 (19%) Ruído(53%)

Accuracy	Previsão
S1: 71%	S1: 30%
S2: 52%	S2: 17%
Ruído: 93%	Ruído: 49%

Gráficos: Precisão

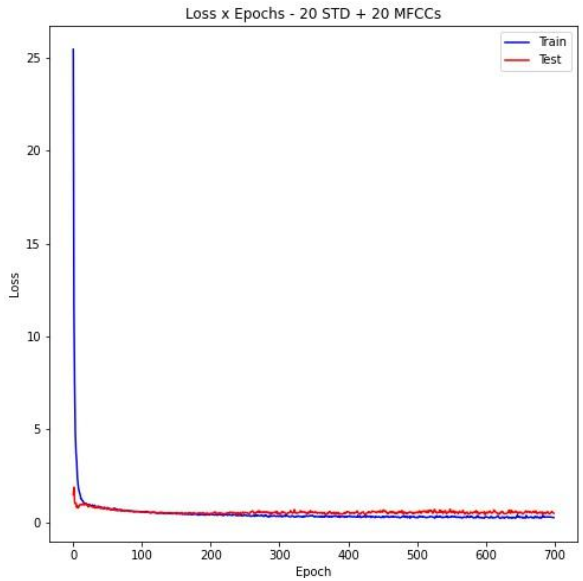


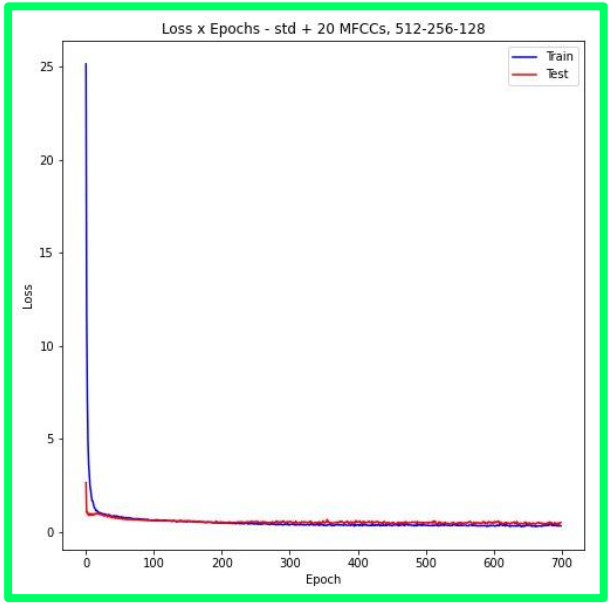


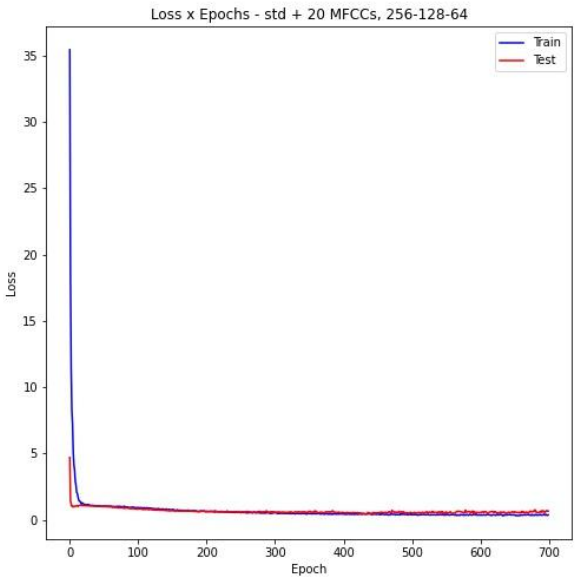


Menor Complexidade

Gráficos: Perdas





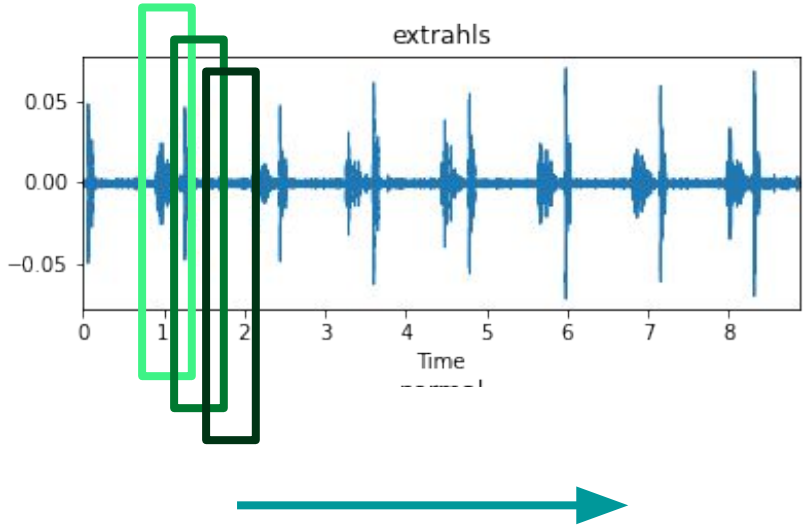


3

Previsão das redes neurais treinadas

Metodologia

Analisar o áudio com uma janela deslizante e obter as características desejadas consoante a rede treinada escolhida.



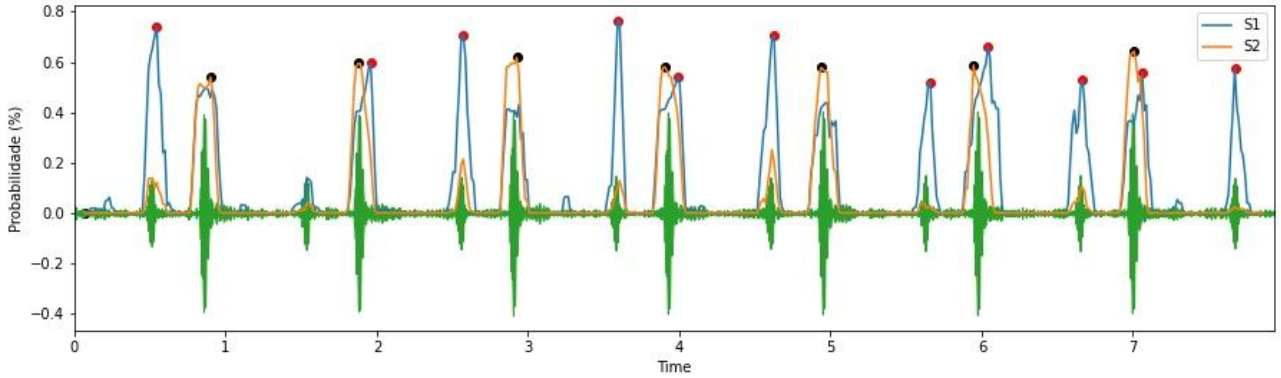
Resultados

Para a melhor rede: 20 MFCCs e Desvios-Padrões (Modelo 1)

No entanto, este modelo não origina bons resultados em previsões. Será também utilizado o Modelo 1 com envelope (Modelo 2).

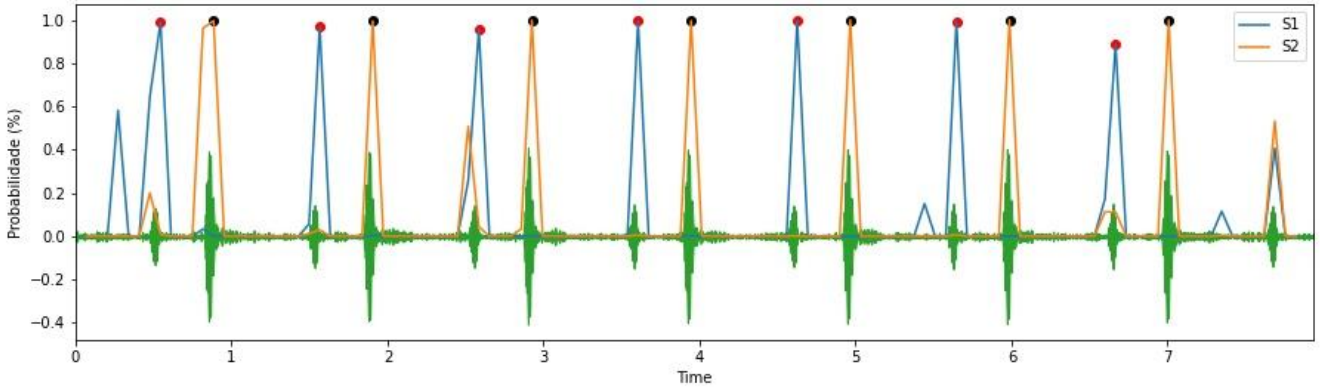
Foram escolhidos 4 áudios sem sons anormais, pois o treino não continha essa amostra.

Resultados: Áudio 1



Modelo 1

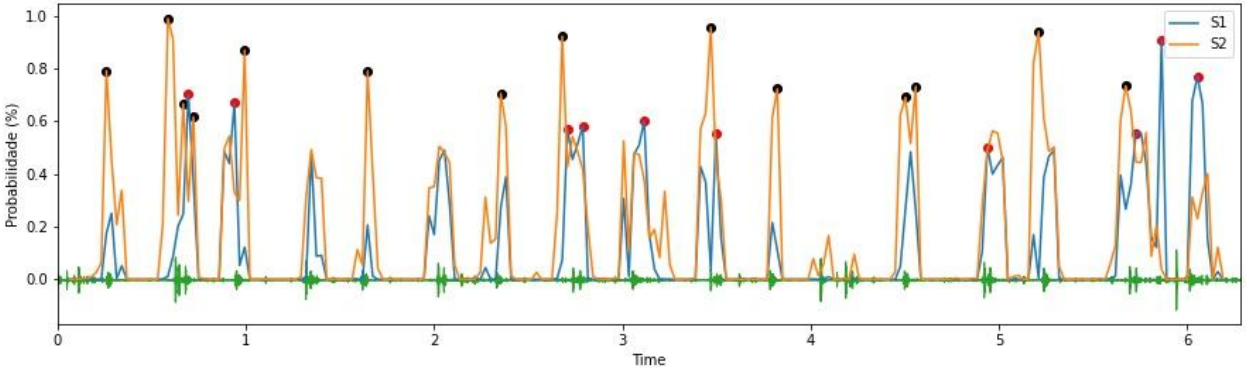
Ritmo cardiaco:
105 bpm a partir de S1.
87.5 bpm a partir de S2.

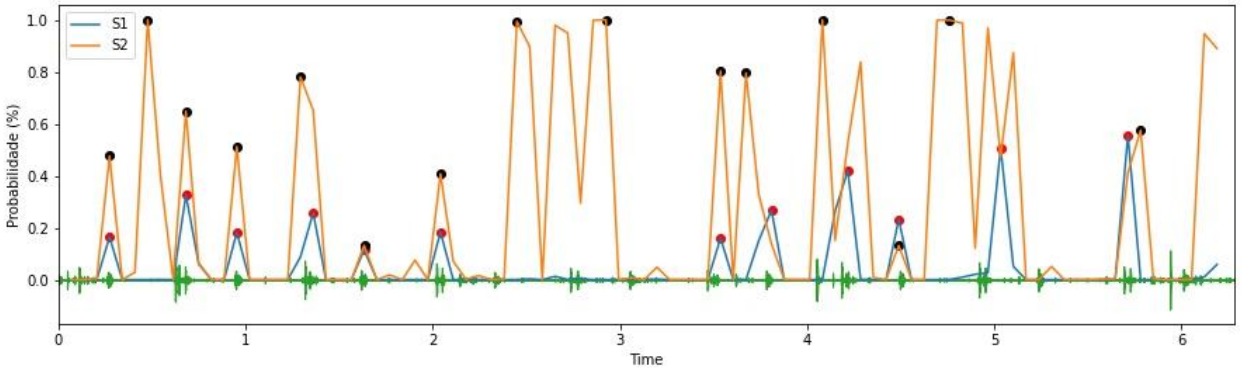


Modelo 2

Ritmo cardiaco:
58.8 bpm a partir de S1.
58.8 bpm a partir de S2.

Resultados: Áudio 2





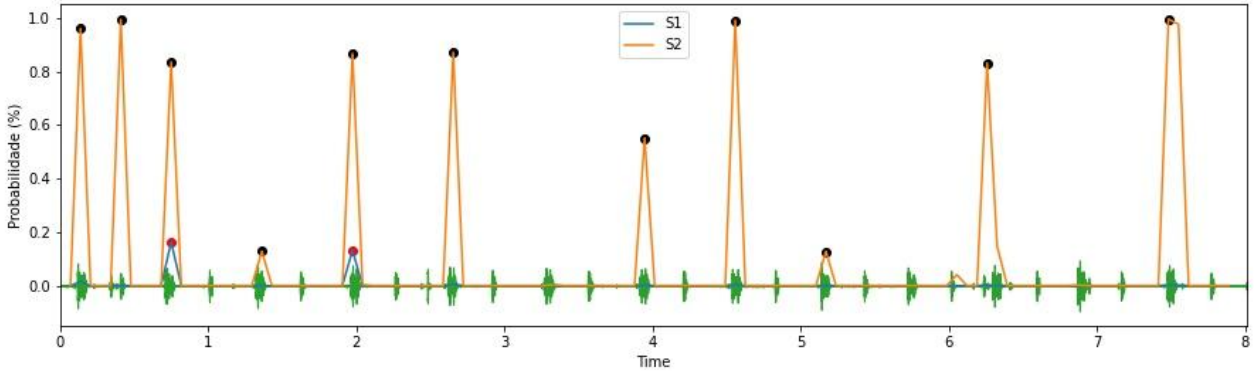
Modelo 1

Ritmo cardiaco:
93.8 bpm a partir de S1.
141.5 bpm a partir de S2.

Modelo 2

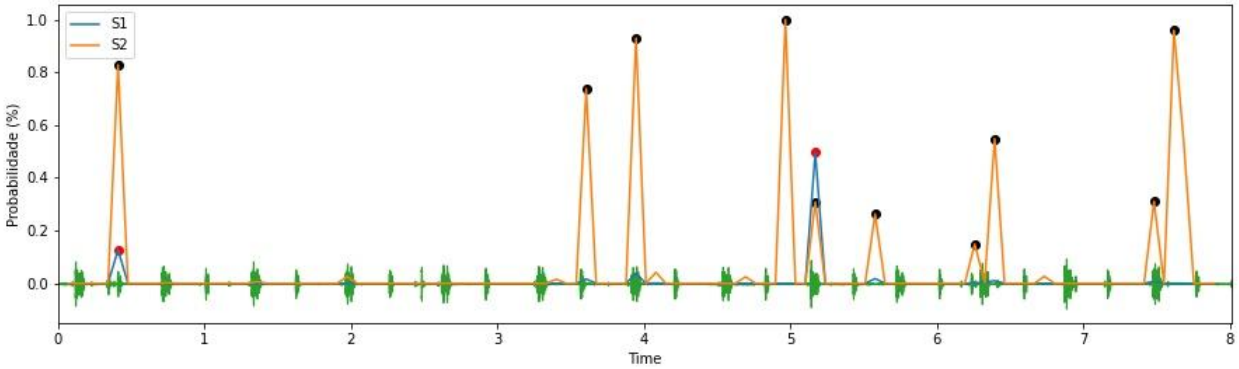
Ritmo cardiaco:
119 bpm a partir de S1.
147 bpm a partir de S2.

Resultados: Áudio 3



Modelo 1

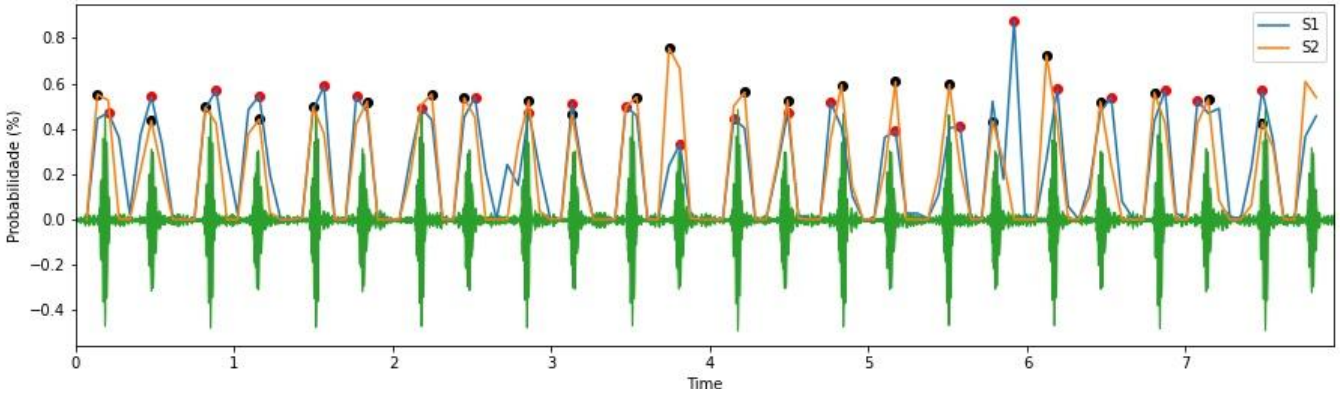
Ritmo cardiaco:
--- bpm a partir de S1.
76 bpm a partir de S2.

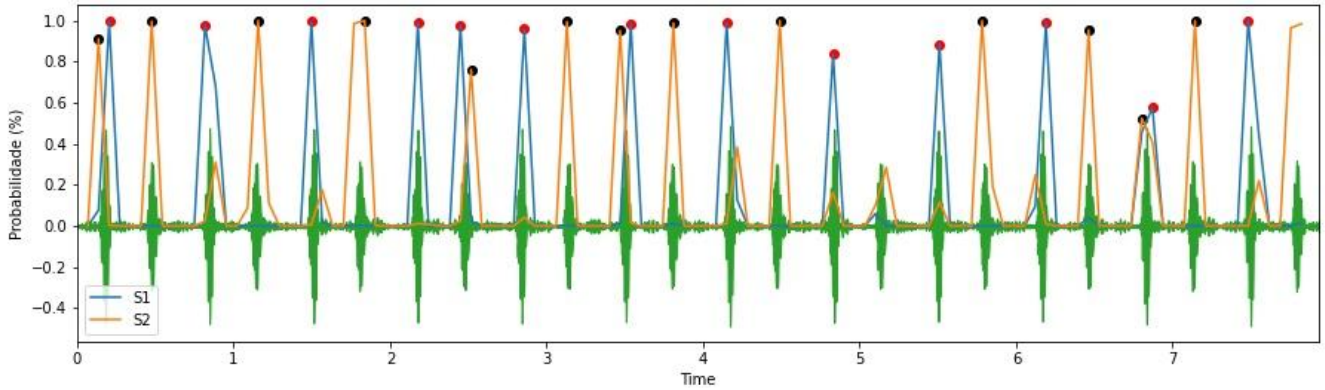


Modelo 2

Ritmo cardiaco:
--- bpm a partir de S1.
72 bpm a partir de S2.

Resultados: Áudio 4





Modelo 1

Ritmo cardiaco:
179 bpm a partir de S1.
179 bpm a partir de S2.

Modelo 2

Ritmo cardiaco:
99 bpm a partir de S1.
99 bpm a partir de S2.

4

Conclusões

Sobre a qualidade dos resultados

- No treino:
  - Podemos observar overfitting para todos os modelos treinados, na curva de perda;
  - O modelo escolhido para previsão, apesar de não ocorrer overfitting, tem uma precisão de cerca de 80% após 700 épocas, o que não é desejável;
  - Nas matrizes de confusão, os nossos modelos têm mais dificuldade em detectar S1, no entanto, o melhor modelo tem mais dificuldade em detectar S2
  - Não foi possível encontrar um balanço entre overfitting e precisão.

Sobre a qualidade dos resultados

- A partir das conclusões anteriores, a amostra de sons usada:
  - Tem demasiados audios com ruído, não sendo uma amostra equilibrada com os 3 tipos de som, em validação cruzada.

Sobre a qualidade dos resultados

- Na Previsão:
  - Os resultados são insatisfatórios para previsão e cálculo do batimento cardíaco;
  - O método de janela deslizante afeta o cáculo dos MFCCs pelo príncipio da incerteza;
  - Mesmo no modelo com menor overfitting, este continua com overfitting presente para áudios com grande amplitude e os sons bem demarcados.

Melhorias:

- No pré-processamento de dados, existir uma amplificação dos sinais S1 e S2 para uma melhor segmentação destes;
- Filtragem de ruído de alta e baixas frequências - filtro passa banda;
- Para o treino, considerar a possibilidade de se utilizar CNNs