

Identificação Automática de Anomalias em Sons de Auscultação Cardíaca

Thalles A. Silva¹, Bruno N. de Carvalho¹

¹Universidade Federal do Maranhão (UFMA)
Caixa Postal 65.085–580 - São Luís – MA - Brasil

thalles.al.slv@gmail.com, bruno.rockout@gmail.com

Resumo. *Estima-se que 15 milhões de pessoas morreram de doenças cardíacas em 2015, sendo assim a maior causador de mortes mundialmente [WHO 2017]. A auscultação é um dos exames médicos mais populares para diagnosticar vários de doenças cardíacas e respiratórias. Entretanto para detectar anomalias através do som é necessário muita experiência e conhecimento do examinador. Esse problema atrai interesse para classificação dos sons cardíacos normais e anormais afim de melhor auxiliar o medico na tomada de decisão e, também, na prevenção de doenças.*

1. Introdução

A ausculta cardíaca é importante para a identificação dos focos, bulhas, ritmo, frequência e outras as possíveis alterações que ocasionam doenças. É um dos mais simples e acessíveis métodos para monitoração da saúde e detecção de anomalias. Porém esse exame é dependente da experiência e habilidade do examinador.

O objetivo desse trabalho é apresentar uma ferramenta de classificação de anomalias através dos som dos batimentos cardíacos, utilizando máquina de vetores de suporte (SVM, *Support Vector Machine*), um método de aprendizado supervisionado de máquina, e o algoritmo KNN(*K-Nearest Neighbors*) para classificação e reconhecimento de padrões das características segmentadas do som. Tais ferramentas podem ter o potencial de auxiliar o examinador a obter melhores e mais precisos resultados em exames de auscultação, sem perder a acessibilidade e simplicidade.

Em 2016, *PhysioNet*, um centro de dados online que oferece acesso à grandes quantidades de gravações de sinais fisiológicos, promoveu um desafio que consistiu da classificação de sons cardíacos entre "normais", "anormais" e "muito ruidosos" Aqui, no decorrer deste trabalho, será desenvolvido um meio de classificação de sons de auscultação cardíaca com o intuito de servir como uma opinião inicial para que um paciente possa ser ou não direcionado a um médico especializado.

2. Fundamentação Teórica

Nesta seção estaremos abordando algumas formalidades referentes às técnicas utilizadas no decorrer deste trabalho, brevemente introduzindo alguns conceitos das características extraídas das faixas de áudio e das técnicas de classificação aqui utilizadas.

2.1. Características Acústicas

Neste trabalho foram utilizadas 193 características acústicas obtidas a partir de 5 descritores de artefatos em áudio – *mel-frequency ceptral coefficient* (MFCC), *mel spec-*

togram, *chroma features*, *spectral contrast* e *tonnetz* – que serão brevemente abordados em seguida.

Utilizado em trabalhos como [Kurniawan et al. 2015], [Logan et al. 2000] e [Hermansky 1990], o MFCC já foi amplamente explorado, tendo já sido comprovado eficiente na identificação de fala por ser capaz de expressar a amplitude espectral de forma compacta e fiel. Já o *Mel Spectrogram* consiste em um artefato obtido durante uma das etapas da obtenção do MFCC. Mas informações sobre ambos métodos podem ser obtidos em [Rabiner and Juang 1993].

As características croma (*chroma features*) são características de audio obtidas através da projeção de todo o espectro em 12 bins referentes aos semitons musicais, sendo assim uma ferramenta poderosa na descrição de faixas de audio [Ellis 2007]. Este nos diz a intensidade de cada um dos semitons em cada espaço de tempo (frame).

Contraste espectral (*spectral contrast*), melhor abordado em [Jiang et al. 2002], descreve as frequências obtidas a partir de uma análise baseada em oitavas musicais da frequência do espectrograma dos sinais de áudio.

Proposto por [Harte et al. 2006], o método *tonnetz* detecta variações harmônicas em sinais de audio, através de análise temporal de alterações de tom em um momento n a partir da distância euclidiana entre os frames $n - 1$ e $n + 1$ de forma que um pico na função de detecção demarca uma transição de uma região harmonicamente estável à outra.

2.2. Classificadores

As ferramentas de classificação utilizadas neste trabalho foram a Support Vector Machine (SVM) [Vapnik 1998], já amplamente conhecida e utilizada dentro das mais diversas áreas da inteligência computacional, e o algoritmo de KNN, um dos – se não o mais – comumente conhecido e básico dos algoritmos de classificação.

O classificador SVM tem o trabalho de segmentar as duas classes através de um hiperplano que maximiza uma margem de separação denotada por δ . Esse hiperplano está localizado entre as semi-retas $H1$ e $H2$, que separam as classes $+1$ e -1 respectivamente, atingindo seu ótimo quando a distância para as duas linhas for máxima. Dois pontos da classe $+1$ definem a reta $H1$ e três, da classe -1 , definem a reta $H2$. Esses pontos são chamados de vetores de suporte (SV - Support Vector) [Soares 2008].

Para conjuntos de dados com padrões não-linearmente separáveis, o classificador SVM mapeia os padrões de entrada em um vetor de características com alta dimensão, para que a separação ocorra no novo espaço. Porém, a tarefa de encontrar uma função de transformação de espaço não é trivial. Para resolver o problema, a transformação ocorre com o auxílio de funções não-lineares que são denominadas de Kernel, que torna possível a construção de um hiperplano de separação ótimo para o espaço de características sem considerar explicitamente o espaço [Soares 2008].

KNN é um dos algoritmos não-paramétricos mais importantes em na área de reconhecimento de padrões e é um algoritmo de aprendizado supervisionado e de classificação. As regras de classificação do KNN são geradas pelas próprias amostras de treinamento sem quaisquer dados adicionais. Algoritmos de classificação de KNN predizem as amostras de teste categoricamente de acordo com as K amostras de treina-

mento que são os vizinhos mais próximos à instância de teste, e a julgam de acordo com a categoria que tem maiores chances deste pertencer [Soucy and Mineau 2001].

3. Metodologia

A metodologia aqui proposta consiste na extração das características explicitadas na seção 2, normalização da base de dados, dentro de um intervalo de 0 a 1, e treinamento dos modelos por meio de validação cruzada *v-fold* com 10 partições, tanto para o algoritmo de SVM quanto para KNN. Uma visão superficial do método pode ser vista na imagem 1.

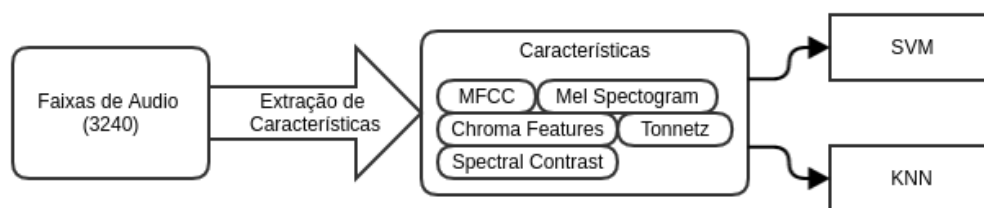


Figure 1. Sequência de atividades

A base utilizada é advinda de [Clifford et al. 2016], um desafio online que encoraja o desenvolvimento de algoritmos para classificação de sons cardíacos nos mais diversos ambientes. A etapa de extração de características foi sumariamente efetuada com ajuda da biblioteca *librosa*, um pacote em python para análise de faixas de áudio que permite obtenção de dados sobre estes sinais. As etapas de normalização e classificação foram efetuadas no WEKA [Eibe Frank and Witten 2016].

Ambos modelos foram treinados sob de validação cruzada *v-fold*, com $v = 10$, onde os parâmetros ótimos foram obtidos por meio de *grid-search*, buscando uma acurácia ótima. A base de treinamento utilizada no treinamento é composta por 3240 instâncias – 2575 normais e 675 anormais – e os resultados obtidos serão exibidos a seguir.

4. Resultados

Demonstrados agora os resultados obtidos com a metodologia proposta neste trabalho. Apresentaremos taxa de acerto obtida com ambos os métodos (tabela 1), bem como as matrizes de confusão (tab elas 2 e 3), de forma a permitir interpretar de maneira mais sólida os resultados obtidos com a base de treinamento disponível.

Trabalhamos com duas classes: anormais (A), referente à classe de sons cardíacos que apresentam alguma anomalia, e normais (B), referente à classe dos sons cardíacos considerados saudáveis.

Table 1. Comparativo dos resultados obtidos com SVM e KNN

	Correct (%)	Incorreto (%)
SVM	91.9753	8.0247
KNN	91.6049	8.3951

Table 2. Matriz de Confusão da SVM

	A	B
A	506	159
B	101	2474

Table 3. Matriz de Confusão do KNN (K=7)

	A	B
A	477	188
B	84	2491

De forma geral pode-se notar que os resultados obtidos com ambos métodos são consideravelmente similares, com uma variação de apenas 0,3704% entre as taxas de acerto obtidas pela SVM com parâmetros ótimos e o KNN com $k = 7$.

Entretanto quando observamos as matrizes de confusão pode-se notar que SVM teve uma taxa de acerto maior quanto aos casos de anormais, o que é mais desejável no caso de um sistema de auxílio à opinião médica – é preferível que um paciente seja erroneamente diagnosticado doente que diagnosticado saudável quando este de fato sofre de alguma patologia, tendo em vista que diagnóstico, quanto mais próximo mais provável de salvar uma vida.

5. Conclusão

Neste trabalho foi desenvolvido uma abordagem para classificação de gravações de auscultação cardíaca entre normais e anormais, com base em descritores de sinais de áudio (*mel-frequency ceptral coefficient*, *mel spectrogram*, *chroma features*, *spectral contrast* e *tonnetz*) e algoritmos de classificação supervisionados (KNN e SVM).

Os resultados obtidos apresentam margens de acerto satisfatórias, apesar de ainda não ótimas. Parte das inconsistências obtidas nos resultados podendo ser decorrente do desequilíbrio nos números entre instâncias de áudio de pacientes saudáveis (normal) e não saudáveis (anormal), a primeira apresentando quase 4 vezes mais casos para análise.

De forma geral é possível considerar que o método desenvolvido neste trabalho é capaz de auxiliar opiniões de um médico especialista mas sem jamais o substituir. A margem de erro, de aproximadamente 8% jamais pode ser desconsiderada tendo em vista que isso poria a vida de pessoas em risco.

Assim o trabalho aqui desenvolvido ainda requer uma especialização maior no seu desenvolvimento. É preciso explorar novas formas de representar as gravações de auscultação, fazendo com que estas venham a ser melhor representadas com maior fidelidade e aumentando a confiabilidade de suas características. Precisa-se também buscar uma organização dos dados apresentados aos classificadores de forma que este seja capaz de identificar com maior confiança, dado que uma base desequilibrada influencia consideravelmente durante o treinamento.

References

- Clifford, G. D., Liu, C., Moody, B., Springer, D., Silva, I., Li, Q., and Mark, R. G. (2016). Classification of normal/abnormal heart sound recordings: the physionet/computing

- in cardiology challenge 2016. In *Computing in Cardiology Conference (CinC), 2016*, pages 609–612. IEEE.
- Eibe Frank, M. A. H. and Witten, I. H. (2016). Data mining: Practical machine learning tools and techniques.
- Ellis, D. (2007). Chroma feature analysis and synthesis. <https://labrosa.ee.columbia.edu/matlab/chroma-ansyn/>. (Accessed on 06/30/2017).
- Harte, C., Sandler, M., and Gasser, M. (2006). Detecting harmonic change in musical audio. In *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*, pages 21–26. ACM.
- Hermansky, H. (1990). Perceptual linear predictive (plp) analysis of speech. *the Journal of the Acoustical Society of America*, 87(4):1738–1752.
- Jiang, D.-N., Lu, L., Zhang, H.-J., Tao, J.-H., and Cai, L.-H. (2002). Music type classification by spectral contrast feature. In *Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on*, volume 1, pages 113–116. IEEE.
- Kurniawan, F., Khalil, M. S., and Malik, H. (2015). Robust tampered detection method for digital audio using gabor filterbank. *Proc. ICIPCS*, pages 75–82.
- Logan, B. et al. (2000). Mel frequency cepstral coefficients for music modeling. In *ISMIR*.
- Rabiner, L. R. and Juang, B.-H. (1993). Fundamentals of speech recognition.
- Soares, H. B. (2008). Análise e classificação de imagens de lesões da pele por atributos de cor, forma e textura utilizando máquina de vetor de suporte.
- Soucy, P. and Mineau, G. W. (2001). A simple knn algorithm for text categorization. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pages 647–648. IEEE.
- Vapnik, V. (1998). *Statistical learning theory. 1998*, volume 3. Wiley, New York.
- WHO, W. H. O. (2017). The top 10 causes of death. <http://www.who.int/mediacentre/factsheets/fs310/en/>. (Accessed on 04/06/2017).