

Diagnóstico Automático de Cardiopatia Reumática em Exames Ecocardiográficos

João Francisco B. S Martins¹, Erickson R. Nascimento¹, Gisele L. Pappa¹

¹Departamento de Ciência da Computação – Universidade Federal de Minas Gerais (UFMG)
Av. Antônio Carlos, 6627, Sala 5209, Belo Horizonte, Brasil, 31270-010

Resumo. A cardiopatia reumática (CR) afeta aproximadamente 39 milhões de pessoas no mundo e é a doença cardíaca adquirida mais comum entre crianças e adolescentes. Ecocardiogramas são o padrão-ouro para o diagnóstico de CR, mas uma escassez de profissionais qualificados impede a implementação em larga escala de programas de prevenção e identificação precoce da doença. Nessa direção, esse trabalho propõe um arcabouço baseado em aprendizado de máquina para lidar com os desafios de identificação automática de CR em exames ecocardiográficos, obtendo uma acurácia de 71,18%. O método também gera informação sobre o processo de diagnóstico através de visualizações temporais (estruturas relevantes em um vídeo) e espaciais (estruturas relevantes em um quadro), facilitando a tomada de decisão de diagnóstico clínico.

1. Introdução

A cardiopatia reumática (CR) é causada por válvulas cardíacas danificadas por febre reumática (FR) e afeta cerca de 39 milhões de pessoas em todo o mundo [James et al. 2018]. Ela é a doença cardíaca adquirida mais comum em crianças e jovens adultos [Okello et al. 2013], mas pode ser tratada se detectada em seus estágios iniciais. Em 2013, o Sistema Único de Saúde brasileiro (SUS) registrou 5.169 internações relacionadas à FR e 8.841 relacionadas à CR crônica, totalizando um custo de 33 milhões de dólares, principalmente relacionados a cirurgias cardiovasculares [Ribeiro et al. 2016].

Os ecocardiogramas são hoje o padrão-ouro para o diagnóstico de CR e uma ferramenta eficaz para detecção precoce de CR latente. Seguindo as diretrizes da World Heart Federation (WHF) [Reményi et al. 2012], um ecocardiografista experiente pode realizar um diagnóstico de CR ao identificar anormalidades estruturais (morfológicas) e funcionais na válvula mitral e na válvula aórtica. No entanto, a disponibilidade de profissionais qualificados em regiões subdesenvolvidas é pequena. A Figura 1 mostra alguns exemplos de diferentes visões do coração capturadas por ecocardiogramas.

Aproximadamente 75% das crianças do mundo vivem em regiões de alta prevalência de CR [Vos et al. 2015]. Essa faixa etária é uma das mais acometidas pela doença e, ao mesmo tempo, a menos propensa a manifestar características ecocardiográficas suficientes para levar a um diagnóstico seguro. Por isso, as diretrizes da WHF definem a categoria *borderline* CR para casos subclínicos, sendo CR definitivo a forma mais grave da doença. O critério para diagnóstico subclínico aplica-se apenas a pacientes com idade ≤ 20 anos, considerando que esse grupo se beneficia mais com a detecção precoce e prevenção secundária da doença. Dado o alto número de indivíduos vivendo em área de alta prevalência da doença e a escassez de profissionais, o uso de técnicas automatizadas de diagnóstico apresenta-se como alternativa promissora.

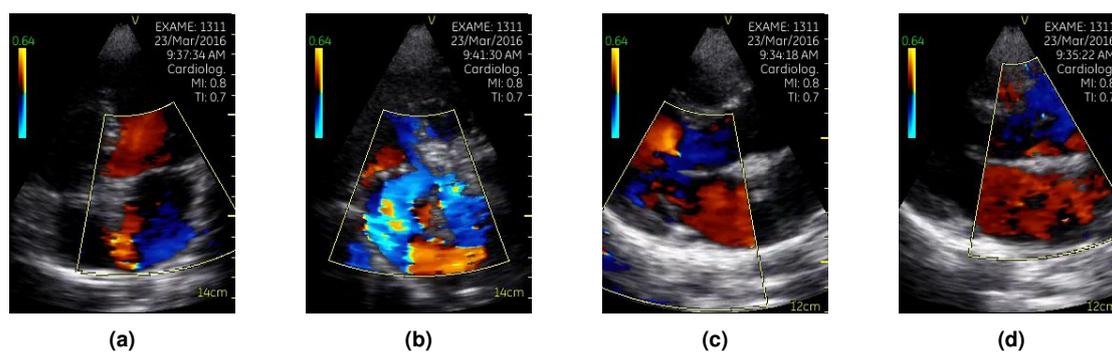


Figura 1. Exemplos de quadros de um ecocardiograma amostrados de quatro visões diferentes de um único exame.

Redes neurais artificiais já vem sendo aplicadas a dados ecocardiográficos 2D desde 1990, mas nos últimos anos o número de publicações na área aumentou consideravelmente devido à popularização de técnicas de Deep Learning (DL) [Martin-Isla et al. 2020]. A literatura de DL em ecocardiografia traz principalmente estudos para identificação das diferentes visões de ecocardiograma (*view*) [Madani et al. 2018], uma vez que para cada exame são capturados vários vídeos com diferentes visões do coração (veja Fig. 1). Outros estudos focam na segmentação das câmaras cardíacas e na classificação de doenças [Madani et al. 2018; Ghorbani et al. 2020], sendo os últimos aplicados principalmente para detecção de anormalidades morfológicas e não funcionais. Porém, nenhuma das pesquisas relacionadas anteriormente aborda diretamente as anormalidades das válvulas do coração - essenciais para o diagnóstico de CR - e praticamente todas elas utilizam uma abordagem 2D para processar imagens, descartando a relação temporal capturada pelos exames.

2. Objetivos e Contribuições

O principal objetivo deste trabalho é propor um novo arcabouço de aprendizado de máquina para o diagnóstico automático de CR. Embora o arcabouço proposto possa ser instanciado para trabalhar com exames ecocardiográficos obtidos por diversos dispositivos, com diferentes recursos, resoluções e formatos de arquivo, ressalta-se que os vídeos utilizados neste trabalho são provenientes de dispositivos portáteis. Esses dispositivos são baratos, fáceis de usar e transportar, geram arquivos pequenos, e por isso são amplamente adotados em programas de prevenção da doença.

Em especial, os dados utilizados neste trabalho foram obtidos em programas de prevenção voltados à detecção precoce da doença em estágio subclínico e prevenção da progressão da CR em áreas subdesenvolvidas do Brasil e Uganda. Devido à escassez de ecocardiografistas experientes, a maioria dos exames não foi coletada por médicos, e a telemedicina foi utilizada para diagnóstico remoto. Esses dois fatores tornam a tarefa de diagnóstico automático muito mais complexa e desafiadora, uma vez que a qualidade das imagens é baixa e nem sempre sua aquisição foi feita da melhor forma possível. Por último, todos os participantes do estudo são crianças, e portanto variações no tamanho do coração dificultam ainda mais esta tarefa. As principais contribuições deste trabalho são:

- Um novo arcabouço para o uso de aprendizado de máquina para o diagnóstico automatizado de doenças cardíacas que dependem de informações temporais e

- da agregação de predições provenientes de várias visões de uma mesma estrutura (e.g., o coração), e adquiridas em um mesmo exame;
- Duas arquiteturas de redes neurais profundas para diagnóstico de CR em vídeos de ecocardiogramas: (i) Uma arquitetura 3D baseada em Redes Neurais Convolucionais (CNN) e (ii) Uma nova arquitetura multitarefa baseada em atenção, que pode ser treinada com pequenos conjuntos de dados e fornecer interpretabilidade em vários níveis;
 - Duas estratégias para agregação de predições feitas por vídeos individuais de um mesmo exame em um diagnóstico final: (i) Uma estratégia de agregação supervisionada baseada em meta-aprendizado das distribuições de predições e (ii) Uma estratégia de agregação não-supervisionada baseada em uma formulação de regularização esparsa que considera amostras fora de distribuição e fornece outra camada de interpretabilidade.

Essas contribuições geraram, até o momento, (i) a publicação de um artigo no *Journal of the American Medical Informatics Association* (JAMIA) [Martins et al. 2021]; (ii) um artigo em avaliação no *Artificial Intelligence in Medicine Journal* (AIMM) [Martins et al. 2022], e (iii) dois resumos estentidos publicados no *Journal of the American College of Cardiology* (JACC) [Nascimento et al. 2020; Nascimento et al. 2021]. O trabalho também foi apresentado, em 2020, no 75º Congresso Brasileiro de Cardiologia, onde ganhou o prêmio de melhor artigo da trilha principal da conferência¹. Em 2021, o trabalho novamente foi eleito o melhor artigo da trilha principal do 30º Congresso Mineiro de Cardiologia.

3. Metodologia

Como mencionado, o arcabouço proposto foi instanciado utilizando dois métodos diferentes baseados em redes neurais profundas, conforme detalhado abaixo.

Redes Neurais Convolucionais 3D + Agregador Supervisionado: A tarefa de identificar anormalidades cardíacas estruturais pode ser realizada de forma eficaz utilizando a classificação de um quadro único do exame. No entanto, o diagnóstico da CR depende de aspectos funcionais do coração, que abrangem vários quadros sequenciais de um ecocardiograma. Assim, a primeira metodologia utiliza a informação temporal dos vídeos para alcançar um diagnóstico mais preciso.

A Figura 2 ilustra a metodologia proposta, que começa alimentando uma CNN 3D [Tran et al. 2015] com vídeos de todas as visões do exame do paciente. Em seguida, as saídas da rede referentes a todos os vídeos de um único paciente são combinadas usando um metaclassificador baseado em Random Forest.

Arquitetura Multitarefa Baseada em Atenção + Agregador Não-supervisionado: Como parte do procedimento de diagnóstico do CR definido pela WHF [Reményi et al. 2012], os médicos anotam, para cada exame, diferentes características diretamente relacionadas a anormalidades morfológicas (estruturais) e funcionais nas válvulas cardíacas do paciente, como a presença e a natureza da regurgitação e da estenose. Esta abordagem tira proveito dessa informação, e propõe uma

¹https://www.sbc2020.com.br/_files/ugd/adad56_754217aea67e41c9bfffcb3fca2044d9f.pdf

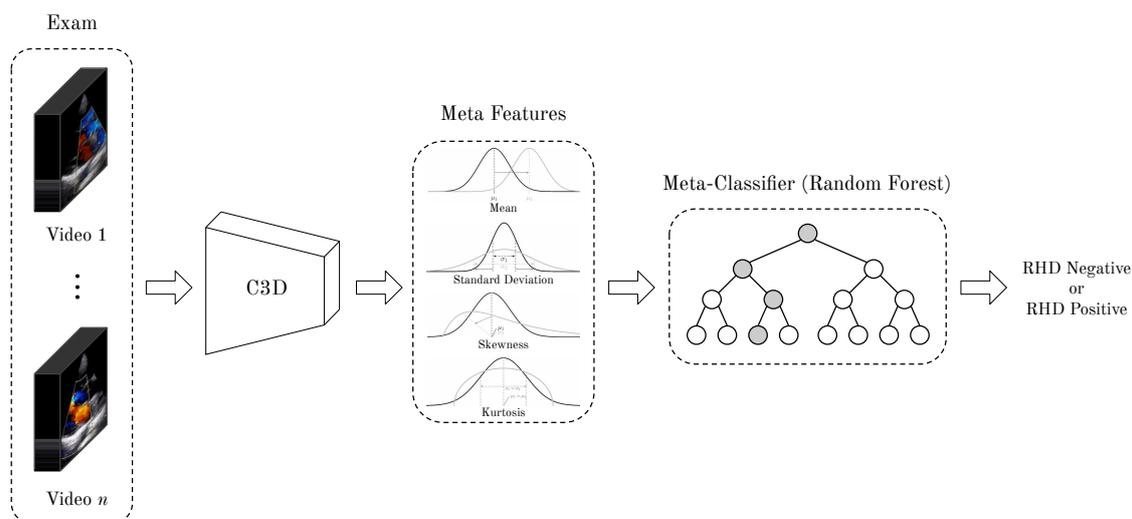


Figura 2. CNN 3D seguida de um metaclassificador supervisionado para agregação de resultados de diferente visões de um exame.

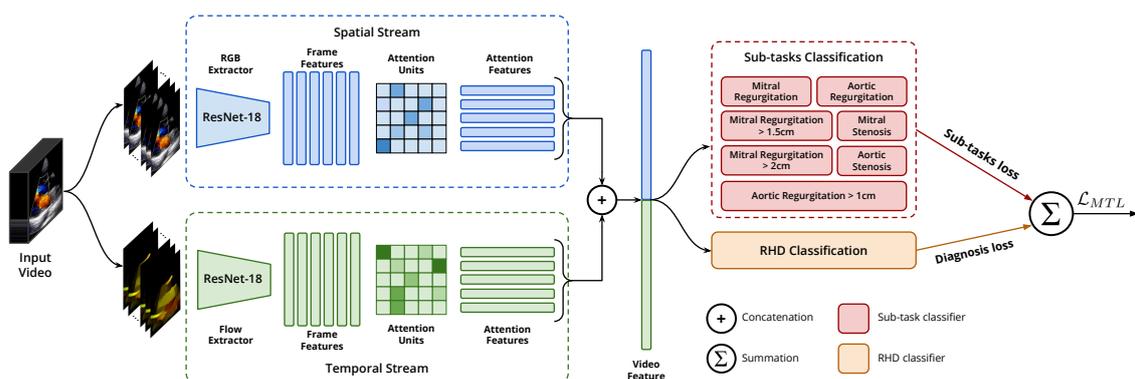


Figura 3. Treinamento de rede multitarefa de dois fluxos.

CNN de dois fluxos em uma configuração de aprendizado multitarefa, que utiliza esses rótulos de anormalidades funcionais do coração como tarefas auxiliares para aumentar sua capacidade de generalização.

O método de classificação consiste em duas etapas principais: i) uma rede multitarefa de dois fluxos treinada para detecção de CR e subtarefas relacionadas e ii) uma estratégia de agregação não-supervisionada que contabiliza previsões das diferentes visões dos vídeos para diagnosticar o exame. A Figura 3 mostra uma visão geral do método. As redes são treinadas em um regime multitarefa processando informações espaciais e temporais de vídeos ecocardiográficos. Unidades de atenção são aplicadas para pesar quadros importantes, onde as características morfológicas e o fluxo sanguíneo estão presentes.

4. Experimentos e Resultados

A base de dados utilizada possui 5,526 vídeos ecocardiográficos com resolução 320×240 pixels, feitos com dispositivos VSCAN[®]. Os vídeos correspondem a 538 exames completos de pacientes únicos (269 CR negativo e 269 CR positivo). Os dados foram adquiridos por técnicos treinados pelo programa de prevenção a CR PRO-

Tabela 1. Comparação entre os métodos propostos e CNN 2D.

Método	Tamanho	Vídeo						Exame			
		Especificidade		Sensibilidade		Acurácia	Especificidade		Sensibilidade		Acurácia
CNN 2D [Zhang et al. 2018]	129M	34,33±	9,02	76,10±	6,92	56,87± _{3,42}	31,29±	10,55	83,30±	7,26	57,26± _{3,93}
(1) Arquitetura CNN 3D	130M	58,57±	10,51	58,00±	10,59	58,20± _{2,70}	64,84±	7,03	65,14±	10,84	64,88± _{5,22}
(2) Método Multitarefa	23M	65,87±	7,81	63,83±	8,46	64,70± _{2,17}	71,71±	10,90	70,70±	9,62	71,18± _{3,10}

VAR [Nascimento et al. 2016]. Devido ao número extremamente baixo de amostras de CR definitivo (apenas 32 exames no total), os diagnósticos *Borderline* e definitivo foram considerados como uma única classe: CR positivo.

A Tabela 1 reporta as medidas de especificidade, sensibilidade e precisão para classificação CR tanto em nível de vídeo quanto de exame utilizando uma estratégia de validação cruzada de 10 partições. Para garantir a significância estatística na comparação dos métodos, utilizou-se um teste de Wilcoxon com intervalo de confiança de 95%. Conforme observado, em relação à precisão, os métodos propostos (referenciados na tabela como 1 e 2), que consideram o tempo, superaram o preditor 2D proposto em Zhang *et al.* em todos os níveis de agregação. Note que o método multitarefa tem um desempenho significativamente melhor do que a CNN 3D, e com um número $5\times$ menor de parâmetros (coluna *Tamanho*). O mesmo padrão pode ser observado nos resultados de especificidade.

5. Conclusões

Este trabalho estabelece as bases para o diagnóstico automático de CR em exames ecocardiográficos. Um arcabouço baseado em técnicas de aprendizado de máquina foi proposto, e instanciado seguindo duas metodologias. A primeira utiliza a 3D (uma CNN 3D) para classificar vídeos dos exames individualmente, e essas predições são posteriormente agregadas por um metaclassificador para emitir um diagnóstico de exame completo.

A segunda estratégia é baseada em uma nova arquitetura de CNN baseada em atenção de dois fluxos, que aproveita anotações de subtarefas criadas naturalmente por especialistas durante o procedimento de diagnóstico para melhorar significativamente a precisão da classificação. Essa abordagem supera significativamente outros métodos e ainda permite que os resultados sejam interpretados, facilitando a adoção do método na tomada de decisões clínicas.

O diagnóstico de CR usando ecocardiogramas é um problema desafiador. No entanto, o diagnóstico automático de CR mostrou-se viável e, com mais pesquisas, tem o potencial de permitir a implementação de programas de prevenção mais amplos, que podem reduzir a carga da doença nos países subdesenvolvidos. Mais do que o diagnóstico da CR, o sistema proposto, se embutido em dispositivos de exame portáteis ou disponibilizado como um aplicativo baseado em nuvem, também tem potencial para permitir a identificação, com baixo custo, de pacientes com maior risco para outras valvopatias e doenças cardiovasculares.

Referências

[Ghorbani et al. 2020] Ghorbani, A. et al. (2020). Deep learning interpretation of echocardiograms. *npj Digital Medicine*, 3(1):1–10.

- [James et al. 2018] James, S. et al. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, 392(10159):1789–1858.
- [Madani et al. 2018] Madani, A., Ong, J. R., Tibrewal, A., and Mofrad, M. R. K. (2018). Deep echocardiography: data-efficient supervised and semi-supervised deep learning towards automated diagnosis of cardiac disease. *NPJ digital medicine*, 1(1):1–11.
- [Martin-Isla et al. 2020] Martin-Isla, C. et al. (2020). Image-based cardiac diagnosis with machine learning: A review. *Frontiers in Cardiovascular Medicine*, 7:1.
- [Martins et al. 2021] Martins, J. et al. (2021). Towards automatic diagnosis of rheumatic heart disease on echocardiographic exams through video-based deep learning. *Journal of the American Medical Informatics Association*. doi:10.1093/jamia/ocab061.
- [Martins et al. 2022] Martins, J. F. et al. (2022). A multi-task learning approach to classify rheumatic heart disease on echocardiographic imaging. *Artificial Intelligence in Medicine (under review)*.
- [Nascimento et al. 2016] Nascimento, B. et al. (2016). Echocardiographic prevalence of rheumatic heart disease in brazilian schoolchildren: Data from the provar study. *Int. journal of cardiology*, 219:439–445.
- [Nascimento et al. 2020] Nascimento, B. et al. (2020). Deep learning for automatic identification of rheumatic heart disease in echocardiographic screening images: data from the atmosphere-provar study. *Journal of the American College of Cardiology*, 75(11_Supplement_1):3577–3577.
- [Nascimento et al. 2021] Nascimento, B. et al. (2021). Spatial-temporal deep-learning for automatic identification of rheumatic heart disease in echocardiographic screening images - data from the provar-atmosphere study. *Journal of the American College of Cardiology*, 77(18_Supplement_1):3243–3243.
- [Okello et al. 2013] Okello, E. et al. (2013). Cardiovascular complications in newly diagnosed rheumatic heart disease patients at mulago hospital, uganda. *Cardiovascular journal of Africa*, 24(3):82.
- [Reményi et al. 2012] Reményi, B. et al. (2012). World heart federation criteria for echocardiographic diagnosis of rheumatic heart disease—an evidence-based guideline. *Nature reviews cardiology*, 9(5):297.
- [Ribeiro et al. 2016] Ribeiro, A., Duncan, B., Brant, L., Lotufo, P., Mill, J., and Barreto, S. (2016). Cardiovascular health in brazil: trends and perspectives. *Circulation*, 133(4):422–433.
- [Tran et al. 2015] Tran, D., Bourdev, L., Fergus, R., Torresani, L., and Paluri, M. (2015). Learning spatiotemporal features with 3D convolutional networks. In *Proc. ICCV*, pages 4489–4497.
- [Vos et al. 2015] Vos, T. et al. (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the global burden of disease study 2013. *The Lancet*, 386(9995):743–800.
- [Zhang et al. 2018] Zhang, J. et al. (2018). Fully automated echocardiogram interpretation in clinical practice: feasibility and diagnostic accuracy. *Circulation*, 138(16):1623–1635.