

Automatic Segmentation and ROI detection in cardiac MRI of Cardiomyopathy using q -Sigmoid as preprocessing step

Eduardo Coltri, Gabriel S. Figueredo Costa, Kelvin Lins Silva, Pedro Zigante Martim, Leila Cristina C. Bergamasco
Computer Science Department, Centro Universitário FEI, São Bernardo do Campo, SP, Brasil
[unifecoltri, unifgacosta, unifksilva, unifpmartim, leila.cristina]@fei.edu.br

Abstract—The growth of data volume is a reality in all segments of our society. Despite of the benefits such as personalized experiences, accurate and fast information, new challenges had arisen. For healthcare industry, for example, it was noted an increase of radiologists workload which may cause visual fatigue and, consequently, errors during diagnosis. Artificial intelligence was pointed as an option to support physicians analysis and reduce the visual fatigue. Thus, this paper focus on the proposal of a novel strategy to enhance cardiac magnetic resonance images (MRI) and automatically detect their region-of-interest (ROI) using a convolutional neural network (CNN). Our object of study is the disease of Cardiomyopathy and the desirable ROI is the left ventricle from axial slices. We evaluated q -Sigmoid performance using it as a preprocessing step and validate the results through modified CNNs: U-Net and ResNet.

Index Terms—Cardiac MRI, cardiomyopathy, ROI detection, CNN

I. INTRODUÇÃO

Big Data é um conceito da área da Computação, que estuda como tratar e analisar informações a partir de um grande volume de dados com o objetivo de proporcionar melhores tomadas de decisões. Nos últimos anos esse termo vem se tornando cada vez mais relevante, uma vez que a sociedade vem gerando uma quantidade cada vez maior de dados. Até o ano de 1986, os dados que a humanidade havia produzido de todas as formas, até então, possíveis – como livros, imagens e filmes –, ocupavam a ordem de grandeza do *petabyte* (10^{15} bytes). Em 2007, já havia sido produzido aproximadamente 300 *exabytes* (10^{18} bytes) de dados. Nos últimos anos, a humanidade ultrapassou 8 *zettabytes* (10^{21} bytes) de dados produzidos [3].

Na medicina, o *Big Data* representa um grande potencial para auxiliar em estudos epidemiológicos, para intervenções em pandemias e para a avaliação de perfis de saúde em hospitais. Tornando assim possível a adoção de propostas coerentes com as demandas e necessidades, gerando estratégias e decisões efetivas. O uso de imagens na medicina é considerado um dos recursos mais importantes na definição dos diagnósticos médicos e para auxiliar esses diagnósticos, técnicas de processamento de imagens e *Machine Learning* têm sido utilizadas com o objetivo de oferecer melhores parâmetros para a elaboração de um diagnóstico mais preciso,

indicando áreas suspeitas, bem como anormalidades pouco aparentes.

Em paralelo, o barateamento e aperfeiçoamento dos *hardwares* favoreceram um aumento no volume de dados disponíveis, uma vez que a captação de tais imagens se torna cada vez mais precisa, com altas taxas de resolução. Por exemplo, o volume de dados provenientes de imagens cardíacas em 2014, se encontrava no mesmo patamar do volume de dados provenientes do Twitter - uma das redes sociais mais utilizadas no mundo [17].

Com o objetivo de auxiliar na análise e disponibilização de diagnósticos, médicos se apoiam em *softwares* semi-automáticos para analisar tais imagens. Uma etapa necessária desse processo, dentro da Cardiologia, consiste na segmentação das imagens para então quantificar as estruturas. Previamente à segmentação, é também realizada a seleção da região de interesse (ROI) nas imagens na qual, como o próprio nome sugere, áreas das imagens são destacadas para futura análise.

Tais *softwares* colaboram com a velocidade nas análises médicas, entretanto, nota-se que alguns de seus processos não são completamente automáticos sendo necessária intervenção manual de um especialista durante a etapa de segmentação. Duas situações se destacam: a necessidade de identificação manual do ROI em cada exame e em situações na qual determinada anomalia causa deformações agressivas na estrutura cardíaca ocasionando uma segmentação incorreta pela ferramenta.

Dentro do processo de segmentação, uma das etapas que podem colaborar para o sucesso da estratégia é a normalização da imagem, a qual permite alterações da intensidade dos *pixels* da imagem visando a diminuição de ruídos na imagem oriundos de intensidades extremamente altas e/ou baixas.

Uma das técnicas de normalização possíveis é a q -Sigmoid que apresentou bons resultados em [18], na remoção ou redução de ruídos presentes em imagens de ultrassom, de forma a indicar potencial de utilização em outros contextos de imagens médicas, como imagens de ressonância magnética cardíaca (RMC), por exemplo.

Considerando o contexto inserido, o presente trabalho tem por objetivo apresentar os resultados preliminares da identificação do ROI e consequentemente a segmentação do

ventrículo esquerdo de maneira automática, com o auxílio de técnicas de pré-processamento como a q -Sigmoid. Para avaliar tal estratégia, será utilizada a base de Cardiomiopatia do desafio *Automatic Cardiac Diagnosis Challenge* (ACDC) e a combinação de duas redes neurais modificadas: a U-Net e a ResNet. O restante do artigo está dividido da seguinte forma: a Seção 2 apresenta os principais conceitos utilizados no trabalho. A Seção 3 apresenta a metodologia que foi aplicada e desenvolvida. Já na Seção 4 são apresentados e discutidos os resultados encontrados e por fim, a conclusão e os trabalhos futuros estão descritos na Seção 5.

II. CONCEITOS

A. Rede Neural Convolutacional

O termo *Deep Neural Network*, se refere às Redes Neurais Artificiais com múltiplas camadas, e, nas últimas décadas, essa ferramenta vem sendo considerada uma das ferramentas mais populares devido a sua capacidade de processar e manipular uma grande quantidade de dados, superando desempenho de métodos clássicos em diferentes campos, especialmente no reconhecimento de padrões. Uma das redes neurais profundas mais populares é a Rede Neural Convolutacional (*Convolutional Neural Network* - CNN) existe algum tipo de processamento com dados de imagens. [20]

A CNN é um algoritmo que pode captar uma imagem de entrada e atribuir parâmetros, como pesos e tendências a vários aspectos dela, sendo capaz de diferenciá-la de outras imagens. O pré-processamento que uma CNN exige é muito menor em comparação com outros algoritmos de classificação, que em geral, possuem filtros feitos manualmente. Ao utilizar uma CNN, ela mesma aprende quais filtros são mais adequados. Além disso, a arquitetura desse tipo de rede, executa um melhor ajuste ao conjunto de dados da imagem, devido à capacidade de reutilização dos pesos e à redução no número de parâmetros envolvidos, através da convolução – operação matemática linear entre matrizes.

B. U-NET

A crescente utilização de CNN's em tarefas de segmentação gerou a necessidade da criação de arquiteturas eficientes que realizassem esta ação com eficiência e bons resultados.

Dessa forma foi proposta uma arquitetura em um formato de U, como apresentada na Figura 1, dividida em dois caminhos: o de contração, responsável pela extração das características da imagem, e o de expansão, responsável pela construção da imagem segmentada [19]. A grande vantagem desta arquitetura é a propagação do contexto da imagem de entrada durante a construção da imagem segmentada.

Suas aplicações ocorrem principalmente em segmentação de imagens médicas, devido excelentes resultados entre diversas tarefas dessa modalidade [4] [13] [19]

C. ResNet

A arquitetura Resnet é um modelo de CNN densa com alimentações da entrada juntamente com a saída para a próxima camada entre cada grupo de filtros convolucionais

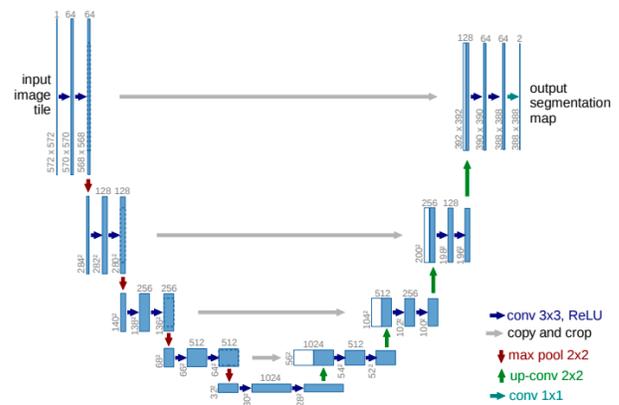


Fig. 1: Arquitetura da U-Net 2D

Fonte: [19]

3x3 [12], como apresentado na Figura 2. Estas alimentações permitem que esta modalidade de rede possa ser densa sem ocasionar problemas de aprendizado. Tais operações acontecem por meio da aplicação dos valores de entrada juntamente com os de saída ao final de cada bloco de pares de filtros convolucionais, surgindo a denominação de blocos residuais.

D. q -Sigmoid

A tarefa de normalização permite alterações da intensidade dos *pixels* da imagem visando a diminuição de ruídos na imagem oriundos de intensidades extremamente altas e/ou baixas obtendo melhores resultados durante o processamento das fases seguintes.

Tais métodos incluem funções matemáticas como a função Sigmoid (Equação 1), na qual o termo x se refere ao valor do *pixel* sendo lido enquanto e se refere a constante de Euler.

$$S(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

O método q -Sigmoid é uma adaptação da função Sigmoid, sendo considerado mais generalista, uma vez que o parâmetro q possibilita a quantificação do nível da não extensividade de um sistema físico, possuindo assim uma representação mais próxima ao mundo real [18] [22].

As equações foram propostas da seguinte maneira:

- q -Sigmoid para $q < 1$:

$$\tilde{I}_1(I; \beta, \alpha, \lambda, q) = \frac{2}{1 + \left[1 + \lambda(1 - q) \left(\frac{|I - \beta|}{\alpha} \right)^{\frac{1}{1-q}} \right]}, \quad (2)$$

- q -Sigmoid para $q > 1$:

$$\tilde{I}_2(I; \beta, \alpha, \lambda, q) = \begin{cases} \frac{1}{1 + [1 + \lambda(1 - q) * F(I)]^{\frac{1}{1-q}}}, & \text{se } I \neq \beta \\ 1 & \text{senão} \end{cases}, \quad (3)$$

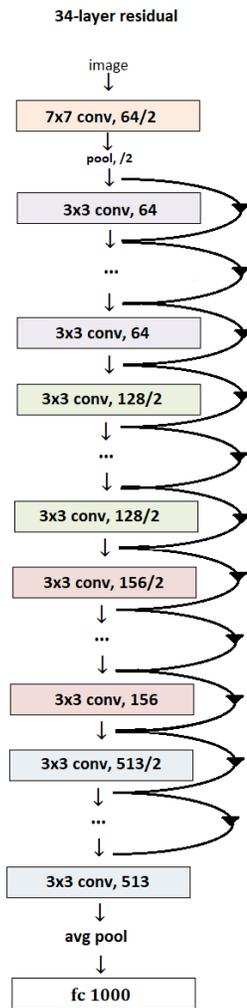


Fig. 2: Arquitetura da ResNet-34

Fonte: Adaptado de [12]

tal que:

$$F(I) = -\frac{1}{\left(\frac{|I-\beta|}{\alpha}\right)^q}$$

Seus parâmetros α e β definem respectivamente o valor médio da luminância e o desvio padrão da região alvo, q representa o parâmetro não-extensivo da Entropia de Tsallis, λ é o fator de amplificação da intensidade no valor do *pixel* analisado e finalmente I determina o valor do pixel em análise.

E. Cardiomiopatia

Segundo [2], a Cardiomiopatia é um grupo de doenças que afetam o músculo cardíaco, na qual no início da doença os sintomas podem ser difíceis de perceber sendo a Cardiomiopatia dilatada (CMD) e a Cardiomiopatia hipertrófica (CMH) os subtipos da doença que mais ocorrem.

A CMH confere uma taxa anual de mortalidade de cerca de 1%, e a maioria dos pacientes possuem pouca ou nenhuma

incapacidade, tendo uma expectativa de vida normal. A visibilidade ligada à CMH relaciona-se em grande parte ao seu reconhecimento como a causa mais comum de morte súbita em adolescentes e adultos jovens, especialmente em atletas. Ela é caracterizada por um aumento da largura da parede ventricular [15].

A CMD, por sua vez, é a doença do miocárdio mais comum no mundo, sendo responsável por 25% dos casos de insuficiência cardíaca, ela é uma doença do músculo cardíaco com dilatação e alterações associadas à massa e ao volume. Essa doença pode ser genética ou idiopática – se manifesta de forma natural –, e nela, predomina-se a disfunção sistólica, podendo evoluir para a insuficiência cardíaca ou não. Essa doença é caracterizada por um aumento da cavidade do ventrículo esquerdo [2]. As Figuras 3a e 3b, apresentam exemplos de fatias de exames de RMC de casos com CMD e CMH, respectivamente, e a Figura 3c apresenta um exemplo de fatia de exame de RMC sem anomalias.

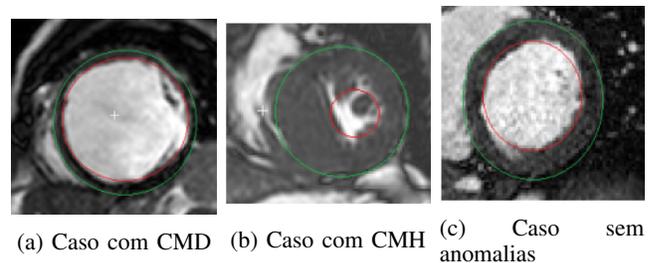


Fig. 3: Diferentes tipos de Cardiomiopatia

Fonte: Adaptado de [5]

III. METODOLOGIA

A. Pré-processamento

O pré processamento das imagens deu-se pela aplicação de uma normalização de seus *pixels* para mantê-los no domínio entre 0 e 1, enquanto as imagens foram redimensionadas para 128x128 *pixels*.

Adicionalmente, foi aplicado o método de q -Sigmoid com $\alpha=0.1$, $\beta=0.3$, $\lambda=0.5$ e $q=2$ para serem realçadas as regiões alvo do ventrículo esquerdo. Os valores de α e β foram definidos após testes iniciais em uma amostra de 10% das imagens, enquanto q e λ tiveram seus valores baseado nos estudos de [18].

Ressalta-se que nesta primeira fase de implementação não foi implementada uma variação nos parâmetros da q -Sigmoid, dado que nesta fase inicial do projeto o objetivo foi compreender quais as consequências de se aplicar tal técnica no pré-processamento em imagens de RMC.

B. Configuração das redes neurais

Para implementação da CNN foi utilizada uma arquitetura modificada de uma U-NET, na qual o caminho de contração foi substituído por uma ResNet-34, buscando extrair a maior quantidade de características possíveis com essa modalidade densa de CNN. Além disso, foi aplicada a função de perda

baseada no índice de Jaccard para penalizar os erros nas amostras de teste e auxiliar durante o aprendizado da CNN.

Uma vez criado o modelo, o mesmo foi treinado por 50 épocas em lotes de 32 imagens. A quantidade definida de épocas se baseou no trabalho de [14], enquanto o tamanho do lote de imagens foi definido de acordo com as limitações do hardware utilizado e o tamanho das imagens.

C. Base de Dados

Para o treinamento de uma rede neural, é fundamental dispor de bases de dados completas e consistentes, para que os parâmetros da rede sejam ajustados apropriadamente. Por conseguinte será usada uma base, amplamente utilizada em literaturas relacionadas a esse contexto, com exames de RMC, no qual existem pacientes com CMD, CMH e sem anomalias. [4] [13]

A base de dados do ACDC é composta por 2000 imagens de exames de RMC criados a partir de diagnósticos clínicos reais, tais exames foram adquiridos no Hospital Universitário de Dijon, na França e suas coletas ocorreram em um período de 6 anos utilizando dois equipamentos de Ressonância [6].

O conjunto de imagens disponíveis foi obtido a partir de 100 pacientes alvos que geraram gravações de RMC com imagens 3D, nos quais após adquirir os cortes do eixo longo, uma série de cortes do eixo curto foram propiciados. Os pacientes foram referenciados e classificados por dois especialistas médicos, gerando um conjunto de imagens dividido em cinco tipos de anomalias diferentes: sem anomalia, CMD, CMH, infarto do miocárdio e ventrículo direito anormal.

Dado o escopo da presente pesquisa apenas os exames que se referem a Cardiomiopatias e pacientes sem anomalias foram utilizados, que correspondem a 60 casos divididos igualmente entre CMH, CMD e sem anomalia.

D. Software e Hardware

Para o desenvolvimento do projeto será utilizada a linguagem Python 3.6 por meio das bibliotecas NumPy [11], Tensorflow [1], Keras [9], Opencv [7], NiBabel [8] e Segmentation Models [23]. Os recursos computacionais consistem em uma máquina com processador AMD Ryzen 5 com 4 Núcleos 3.7Ghz, possuindo memória RAM de 16GB e placa de vídeo NVIDIA GeForce GTX1080 TI - 12GB, com sistemas operacional Windows 10.

E. Treinamento e Validação

O conjunto de imagens foi dividido na proporção 70-30, sendo que 70% das imagens foram utilizadas para a fase de treinamento e 30% para validação.

Com a finalidade de avaliar os resultados obtidos durante o processamento de imagens feito pela rede neural foi computado o coeficiente Sørensen–Dice, no qual compara a similaridade entre duas regiões segmentadas, sendo muito utilizada para avaliação de métodos de segmentação de imagens [21].

A fórmula do coeficiente de Sørensen–Dice é dada pela Equação 4 em que X representa o *ground truth* que consiste

na região segmentada manualmente por um especialista, e Y que representa a região segmentada pelo algoritmo. Ambos os valores são vetores binários, sendo que, se o valor da sobreposição dos intervalos resultar em 1, é caracterizado uma sobreposição perfeita e quando o resultado é igual a 0, é caracterizado uma não sobreposição entre as duas regiões.

$$d_c = \frac{2|X \cap Y|}{|X| + |Y|} \quad (4)$$

Além disso foi utilizado como base da função de perda o Índice de Jaccard, também conhecido como *Intersection Over Union* (IOU), descrito em [16]. Definido por meio da Equação 5, é uma medida conhecida pela similaridade entre dois conjuntos X e Y, definida como o tamanho da interseção dividido pelo tamanho da união dos dois conjuntos, de forma que J seja igual a 1 caso a intersecção entre S e T ($S \cap T$) seja igual a 0 [10].

$$J = \frac{|X \cap Y|}{|X \cup Y|} \quad (5)$$

IV. RESULTADOS

As Tabelas I e II apresentam os resultados obtidos pelas redes neurais que foram treinadas utilizando a *q*-Sigmoid na etapa de pré-processamento ou apenas a normalização dos pixels, respectivamente. Ambos obtiveram valores de Dice superiores a 0.80 indicando uma performance robusta, obtendo valores próximos ao de 0.886 apresentado em [14].

Nota-se que a melhor performance do *q*-Sigmoid ocorreu para os casos de CMD (Tabela I). Entretanto tais resultados não se mostraram superiores quando comparados à técnicas mais simples como normalização de *pixels*.

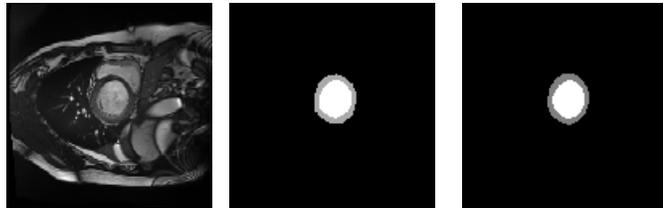
Uma das possíveis razões para o resultado inferior para casos de CMH e sem anomalia foram os parâmetros utilizados no *q*-Sigmoid para esse experimento, que podem ter limitado o desempenho da rede para essas duas situações. Nas Figuras 4 e 5 são apresentados casos de segmentações realizadas com sucesso e sem sucesso pela *q*-Sigmoid. É possível verificar que a maiorias dos casos com segmentações incorretas, nota-se que as bordas do ventrículo esquerdo foram minimamente encontradas ou nenhuma borda foi encontrada. Em específico para os casos de CMH uma vez que o ventrículo esquerdo possui um tamanho reduzido, notou-se que a *q*-Sigmoid destacou regiões diferentes além da cavidade ventricular, dificultando o aprendizado da segmentação de pacientes com este tipo de anomalia. Para trabalhos futuros espera-se aplicar técnicas que possibilitem a variação dos parâmetros da *q*-Sigmoid assim como a expansão da base de imagens objetivando aumentar o aprendizado da arquitetura construída.

TABLE I: Arquitetura treinada com pré-processamento *q*-Sigmoid

Tipo	DICE
CMD	0.852877
Sem Anomalia	0.822208
CMH	0.805707

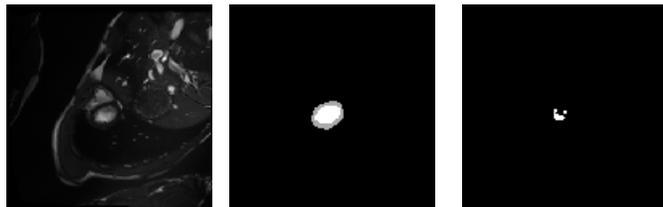
TABLE II: Arquitetura treinada sem pré-processamento q -Sigmoid

Tipo	DICE
CMD	0.913155
Sem Anomalia	0.888433
CMH	0.889447



(a) Imagem Original Paciente 1, CMD (b) Máscara de segmentação esperada (c) Máscara de Segmentação Gerada

Fig. 4: Exemplo de segmentação realizada com sucesso, índice de *Dice* de 0.95



(a) Imagem Original Paciente 67, Sem Anomalia (b) Máscara de segmentação esperada (c) Máscara de Segmentação Gerada

Fig. 5: Exemplo de segmentação realizada erroneamente, índice de *Dice* de 0.33

V. CONCLUSÃO E TRABALHOS FUTUROS

No presente trabalho foi proposta a utilização da técnica de normalização q -Sigmoid e a combinação de uma arquitetura modificada de uma U-NET com uma ResNet-34. O modelo treinado com o pré-processamento da q -Sigmoid obteve resultados iniciais promissores, em especial para casos de CMD. Para trabalhos futuros, espera-se empregar métodos que variem os parâmetros da q -Sigmoid buscando encontrar a melhor configuração para etapa de pré-processamento assim como a expansão da base de imagens objetivando aprimorar os resultados do modelo.

AGRADECIMENTOS

Os autores agradecem a infraestrutura disponibilizada pelo Laboratório *Immersive Lab*, do Centro Universitário FEI para a realização dos testes e ao Instituto Nacional de Ciência e Tecnologia- Medicina Assistida por Computação Científica (INCT-MACC)

REFERENCES

- [1] M. Abadi, A. Agarwal, and P. B. and. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [2] F. M. Albanesi Filho. Cardiomiopatias. *Arquivos Brasileiros de Cardiologia*, 71(2), 1998.
- [3] S. Bach. Afinal, o que é big data? <https://administradores.com.br/artigos/bigdata>, 05 2019.
- [4] C. F. Baumgartner, L. M. Koch, M. Pollefeys, and E. Konukoglu. An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation. *ArXiv*, abs/1709.04496, 2017.
- [5] L. C. C. Bergamasco. RecuperaÇÃO de objetos mÉdicos 3d utilizando harmÔnicos esfÉricos e redes de fluxo. 2015.
- [6] O. Bernard, A. Lalande, and C. Zotti. Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: Is the problem solved? *IEEE Transactions on Medical Imaging*, 37:2514–2525, 2018.
- [7] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- [8] M. Brett, C. J. Markiewicz, and M. Hanke. Nibabel. <https://github.com/nipy/nibabel>.
- [9] F. Chollet et al. Keras. <https://github.com/fchollet/keras>, 2015.
- [10] S. Fletcher and M. Z. Islam. Comparing sets of patterns with the jaccard index. *Australasian Journal of Information Systems*, 22, 2018.
- [11] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, and et al. Array programming with numpy. *Nature*, 585(7825):357–362, Sep 2020.
- [12] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.
- [13] F. Isensee, P. F. Jaeger, and P. M. Full. Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features. *ArXiv*, abs/1707.00587, 2017.
- [14] Z. Liu, P. Li, J. Li, Q. Xie, and X. Wang. Left ventricular full segmentation from cardiac magnetic resonance imaging via multi-task learning. *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, pages 71–75, 2021.
- [15] B. J. Maron. Hypertrophic cardiomyopathy. *Jama*, 287(10), 2002.
- [16] I. e. a. Martire. A novel probabilistic jaccard distance measure for classification of sparse and uncertain data. 2017.
- [17] A. Pah, L. Rasmussen-Torvik, S. Goel, P. Greenland, and A. Kho. Big data: What is it and what does it mean for cardiovascular research and prevention policy. *Current Cardiovascular Risk Reports*, 9, 12 2014.
- [18] P. Rodrigues, G. A. Wachs-Lopes, R. M. Santos, E. Coltri, and G. Giraldi. A q -extension of sigmoid functions and the application for enhancement of ultrasound images. *Entropy*, 21, 2019.
- [19] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. *ArXiv*, abs/1505.04597, 2015.
- [20] V. Sekar, M. Zhang, C. Shu, and B. C. Khoo. Inverse design of airfoil using a deep convolutional neural network. *AIAA Journal*, 57(3):993–1003, 2019.
- [21] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2015.
- [22] C. Tsallis. Nonextensive statistics: theoretical, experimental and computational evidences and connections. *Brazilian Journal of Physics*, 29:1 – 35, 03 1999.
- [23] P. Yakubovskiy. Segmentation models. https://github.com/qubvel/segmentation_models, 2019.