Utilização de Redes Neurais Completamente Convolucionais para identificação e medição de crânios fetais

Everton Leonardo Skeika*, Mauren Louise Sguario Coelho de Andrade*, Hugo Valadares Siqueira◇
*Departamento Acadêmico de Informática
◇Departamento Acadêmico de Engenharia Elétrica
Universidade Tecnológica Federal do Paraná (UTFPR)
Ponta Grossa, Brasil
everton_skeika@hotmail.com; mlsguario@utfpr.edu.br; hugosiqueira@utfpr.edu.br

***Abstract*—Ultrasonography is a non-invasive and non-radioactive imaging technique frequently used in the monitoring of fetal development during pregnancy. The precise assessment is important for the analysis of fetal growth, diagnosis of malformations and possible congenital diseases, thus ensuring the well-being of the mother and the fetus during pregnancy. However, for accurate measurement of anatomical structures of the fetus, specialized knowledge of the obstetrician is necessary. In addition to being a tedious and time-consuming process, contour extraction is influenced by his experience. In this sense, this work proposes the adaptation of the fully convolutional neural network V-Net for segmentation of fetal skulls present in two-dimensional ultrasound images, in addition to the construction of an algorithm for calculating the circumference of them. Preliminary results suggest that the methodology proposed can be automatically used as an auxiliary tool in the monitoring of fetal development.**

***Resumo*—A ultrassonografia é uma técnica de diagnóstico por imagem não invasiva e não radioativa frequentemente utilizada no acompanhamento do desenvolvimento fetal durante a gestação. A avaliação precisa é importante para a análise do crescimento do feto, diagnóstico de malformações e de possíveis doenças congênitas, garantindo assim o bem-estar da mãe e do feto durante a gravidez. Porém, para uma medição precisa de estruturas anatômicas do feto é necessário conhecimento especializado do médico obstetra. Além de ser um processo tedioso e demorado, a extração do contorno é influenciada pela sua experiência. Neste sentido, o objetivo do trabalho é propor a adaptação da rede neural completamente convolucional V-Net para segmentação de crânios fetais presentes em imagens de ultrassonografia bidimensional, além da construção de um algoritmo para o cálculo da circunferência dos mesmos. Resultados preliminares sugerem que a metodologia pode ser utilizada de forma automática como ferramenta auxiliar no acompanhamento do desenvolvimento fetal.**

I. Introdução

A ultrassonografia é um método de diagnóstico de patologias que utiliza ondas ultrassônicas para geração de imagens em tempo real. Devido a sua natureza não invasiva e não radioativa é a modalidade de escolha em muitas aplicações clínicas devido ao seu custo reduzido em comparação a outras formas de captura de imagem, como a Tomografia Computadorizada ou a Ressonância Magnética [1] [2].

Tal técnica auxilia no diagnóstico médico em diversas áreas tais como obstetrícia, ginecologia, oftalmologia, neurologia e cardiologia, além de sua utilização como ferramenta comum em procedimentos terapêuticos [3].

Na obstetrícia, a ultrassonografia é amplamente abordada para a avaliação do desenvolvimento fetal durante a gravidez, em que, ao utilizar a imagem gerada pelo equipamento de ultrassom, o especialista obtém as medidas do comprimento da cabeça do feto, as medidas relativas ao corpo, bem como a análise de seus movimentos, a fim de identificar e prevenir o surgimento de doenças congênitas [4].

Algumas das principais medidas para avaliação do desenvolvimento fetal são: circunferência da cabeça (*head circumference* - HC), diâmetro biparietal (*biparietal diameter*), circunferência do abdômen (*abdominal circumference*), largura do fêmur (*femur length*), largura do humerus (*humerus length*), dentre outras [5].

Na prática, a delimitação da área a ser mensurada é feita manualmente por um médico obstetra. Este processo requer conhecimento especializado sendo um processo maçante e que consome tempo [6]. Além disso, a extração do contorno dos ossos e órgãos em formação é influenciada pela experiência do avaliador [7].

Neste sentido, a fim de facilitar tal processo e auxiliar nas análises dos resultados, técnicas de segmentação de imagens e medição automáticas são candidatas [7]. No entanto, a segmentação automática de fetos em imagens de ultrassom continua a ser uma tarefa desafiadora pelo fato de que a imagem gerada pelo ultrassom pode apresentar várias distribuições de intensidade devido a diferentes condições de aquisição. Além disso, uma série de ruídos como sombras acústicas, ruído *speckle* e baixo contraste podem dificultar o reconhecimento das bordas [7] [2].

Dada a importância do problema recentemente foi construído um desafio público chamado HC18[1], cujo objetivo é desafiar pesquisadores da área a projetar um algoritmo que possa medir automaticamente a circunferência da cabeça do

[1]https://hc18.grand-challenge.org/

feto mediante imagens de ultrassonografia bidimensional (2D). O desafio possibilita a submissão de resultados, tornando possível a comparação com o de outros pesquisadores.

Dentre as diversas técnicas de segmentação de imagens existentes na literatura, uma que vem obtendo grande eficácia são as baseadas em Redes Neurais Convolucionais (*Convolutional Neural Networks*- CNN) [8]. Esta é uma arquitetura baseada na estrutura do córtex visual de mamíferos, a qual tornou-se popular e bem-sucedida em diversas tarefas como reconhecimento visual e detecção de objetos *Faster-RCNN* [9], classificação de imagens *GoogLeNet* [10] e segmentação de objetos em imagens por meio da Rede Neural Completamente Convolucional (*Fully Convolutional Neural Network*- FCN) [11].

A. Trabalhos relacionados

Devido à relevância do problema, trabalhos recentes que visam o aprimoramento da FCN para a segmentação de órgãos e ossos fetais em imagens de ultrassonografia bidimensional vem sendo observados. Em Li e colaboradores [12] os autores propuseram a utilização de FCN para a segmentação automática do líquido amniótico e tecidos fetais em imagens de ultrassom 2D. O trabalho de Sundaresan e colaboradores [13] utilizou FCN para localização automática do coração fetal em *frames* de vídeo de ultrassonografia (ecocardiografia). O trabalho de Sinclair e colaboradores [14] utilizam FCN e *Ramanujan Approximation* II para mediação automática da circunferência da cabeça e o diâmetro biparietal em imagens de ultrassonografia 2D.

B. Contribuição

Nesse sentido, este trabalho propõe a segmentação automática de crânios fetais em imagens de ultrassonografia bidimensional (2D), utilizando para tanto a rede Neural Completamente Convolucional V-Net[2], com o objetivo de reconhecer e medir a circunferência do crânio fetal HC. Pretende-se uma padronização das medidas e possivelmente melhorando a confiabilidade da idade gestacional do feto. Além disso, este trabalho propõe uma nova solução para o desafio HC18, a fim de comparar os resultados com os demais trabalhos já submetidos.

II. Rede Neural Completamente Convolucional V-Net

Criada por Milletari, Navab e Ahmadi em 2016, a Rede Neural Completamente Convolucional V-Net foi projetada para a segmentação de imagens de ressonância magnética tridimensionais (3D) da próstata [15].

Sua arquitetura realiza convoluções com o objetivo de extrair características dos dados, a qual pode ser descrita em duas partes, parte esquerda (compressão) e direita (descompressão), tal como sumariza a Figura 1.

- **Detalhes resumidos da arquitetura:** Como mencionado, a parte esquerda consiste em um caminho de compressão.

[2]https://arxiv.org/pdf/1606.04797.pdf

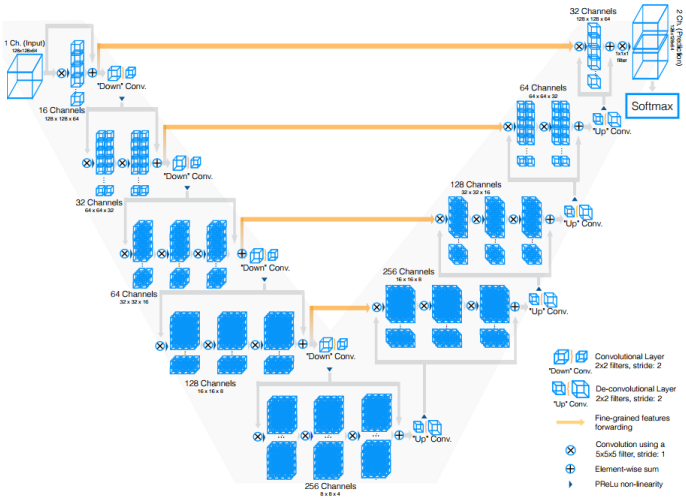

Figura 1. Representação esquemática da arquitetura da rede V-Net proposta por Milletari; Navab; Ahmadi (2016). **Melhor visto em formato digital.**

Este lado é dividido em diferentes estágios que operam em diferentes resoluções possuindo de uma a três camadas convolucionais. À medida que os dados prosseguem por diferentes estágios ao longo do caminho de compactação, sua resolução é reduzida. Este processo reduz o custo computacional da rede. Já o lado direito extrai as características e expande os mapas de características com baixa resolução com o objetivo de juntar e montar as informações necessárias para realizar a segmentação. Os dois mapas de características computados pela última camada convolucional produzirão saídas do mesmo tamanho que o volume de entrada. Estas serão processadas pela última camada da rede que gerará a probabilidade de cada *pixel* da imagem pertencer ao fundo (*background*) ou ao objeto alvo (*foreground*).

III. Metodologia

Para a execução do presente trabalho seis etapas foram propostas: aquisição do *dataset*, pré-processamento do *dataset*, adaptação da rede V-Net, pós-processamento das segmentações imprecisas, implementação do algoritmo para o cálculo do HC, submissão ao desafio HC18 e avaliação dos resultados.

A. Aquisição do Dataset

Para a realização do trabalho foi utilizado o *dataset* do desafio HC18, o qual foi disponibilizado por Heuvel e colaboradores [16]. O conjunto de imagens é composto por 1.334 imagens de ultrassons 2D em escala de cinza com resolução de 800 por 540 *pixels*. Destas, 999 estão separadas para treinamento do algoritmo e 335 para o teste. Além disso, para cada uma das imagens de treinamento existe uma *ground truth*, estas que são anotações manuais da circunferência da cabeça realizadas por um ultrassonografista treinado. A Figura 2 ilustra algumas das imagens presentes no *dataset* de treinamento, em que (a)-(c)

representam três imagens de ultrassom 2D e (d)-(f) as suas respectivas *ground truth*.

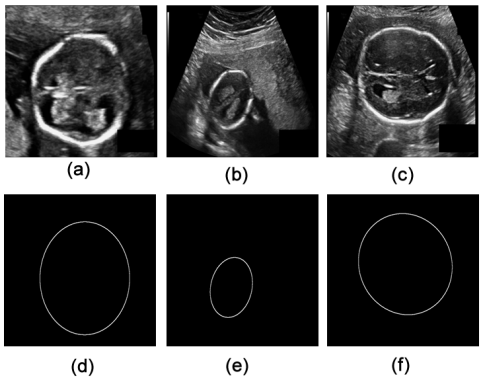

Figura 2. Representação das imagens de ultrassonografia bidimensional e as *ground truth* contidas no *dataset*

B. Pré-Processamento

Nesta etapa foram feitas duas alterações no *dataset*: primeiramente foi realizada a redução da dimensão das imagens para a resolução de 512 por 512 *pixels* a fim de diminuir o custo computacional no treinamento da rede. A segunda foi realizada nas imagens de *ground truth*, onde foi criado um algoritmo em linguagem de programação MATLAB® para preencher as 999 imagens, fazendo-as parecer como elipses sólidas (Figura 3).

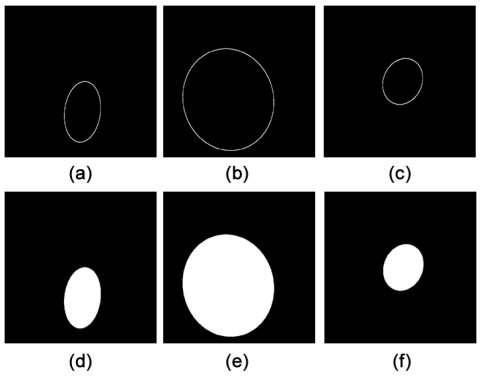

Figura 3. Resultado do preenchimento das *ground truth*

Esta alteração foi necessária pois nos resultados iniciais obtidos as elipses resultantes da segmentação ficaram com suas bordas incompletas. Isto tornaria impraticável o cálculo da elipse de uma forma precisa, pois não haveria como delimitar a área faltante com exatidão.

C. Adaptação da rede V-Net

A rede utilizada para a realização das segmentações é uma adaptação da rede original V-Net 3D projetada por Milletari et al. [15]. A rede utilizada como base foi adaptada para a segmentação de imagens 2D e está disponível no sitio[3] sob a licença GPL.

[3]https://github.com/FENGShuanglang/2D-Vnet-Keras

Inicialmente foram realizadas alterações básicas para a execução do *dataset*. Com as imagens em resolução 512x512, a rede foi treinada com diversas épocas e passos. A Figura 4 ilustra três segmentações obtidas pela rede em 300 épocas de treinamento.

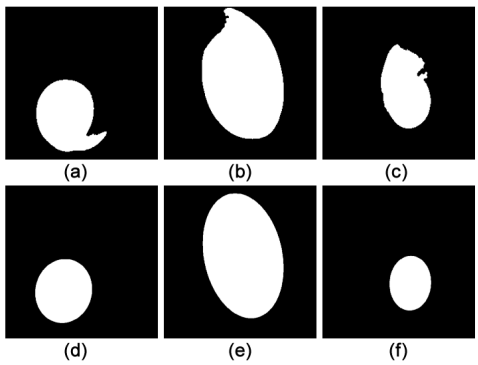

Figura 4. Resultados iniciais obtidos pela rede

Observa-se pela Figura 4 que os resultados das segmentações em (a)-(c) são ineficientes comparadas as suas *ground truth* (d)-(f). Esta ineficiência está relacionado ao *overfitting*[4].

Visando uma diminuição do sobreajuste foram aplicadas alterações na arquitetura da rede, como a reestruturação da técnica de *batch normalization*, alteração no uso de *dropout*, alteração na função de ativação de PReLU para ELU em todas as etapas convolução, edição nos valores de *data augmentation* e aumento no *batch* de geração dos dados sintéticos de 2 para 4, alteração na função de ativação da última camada da rede de *sigmoid* para *hard_sigmoid*. Resultados iniciais demonstram uma melhoria no desempenho da rede. Ainda destacamos que é possível que os resultados possam ser avaliados pelas métricas de sensibilidade, especifidade, precisão, acurácia, Dice e Jaccard.

Outro ponto que está em desenvolvimento é a adaptação de outras técnicas presentes na literatura na arquitetura da rede como a aplicação de convoluções dilatadas (*dilated convolutions*) e conexões residuais agregadas (*aggregated residual connections*) propostas por Vesal et al. [17].

D. Pós-Processamento

A fim de aprimorar ainda mais os resultados das segmentações foi criado um algorítimo em linguagem Python OpenCV, para pós-processamento, em que dada uma imagem segmentada defeituosa seja gerada uma nova elipse encaixando-a aos seus contornos.

Tal processo pode ser observado na Figura 5, na qual: (a) ilustra o resultado falho da segmentação obtida pela rede V-Net adaptada; (b) representa o encaixe de uma nova elipse aos contornos da segmentação ineficiente realizada pelo algoritmo

[4]overfitting (em português, sobreajuste) é um termo estatístico utilizado para indicar que um modelo se adequa bem ao conjunto de dados previamente observado, mas que não é eficaz para prever resultados fora do seu conjunto de treinamento.

proposto; (c) o resultado do preenchimento da elipse; (d) ilustra a *ground truth* utilizada como métrica avaliativa.

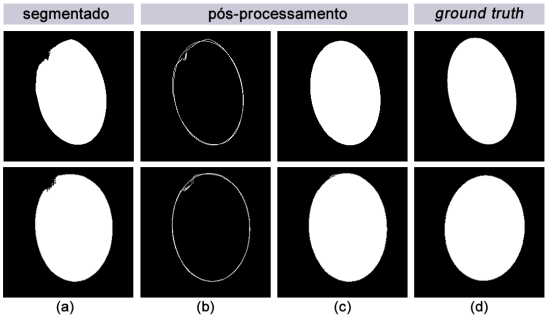

Figura 5. Figura ilustrativa dos resultados obtidos após o algoritmo de pós-processamento

E. Implementação do Algoritmo para Cálculo do HC

Para a submissão dos resultados ao desafio será necessário a construção de um algoritmo para encontrar cinco valores para cada uma das elipses geradas pela segmentação sendo: centro x, centro y, semieixo a, semieixo b e ângulo de rotação.

Para a extração dos dados foi desenvolvido um algoritmo em Python com a utilização da biblioteca OpenCV, em que os valores resultantes serão dispostos em um arquivo CSV para a submissão dos resultados ao desafio.

F. Submissão ao desafio HC18

Um problema real e de grande impacto na medicina é a medição automática do raio elíptico do crânio fetal. Neste sentido, este trabalho mostrou as etapas para o desenvolvimento de um método baseado em redes neurais convolucionais para a solução do problema.

Para tal, etapas de aquisição de dados, pré-processamento e ajuste da rede são necessários, como discutidos. Os resultados preliminares têm se mostrado promissores nas etapas anteriores à medição final.

A continuidade do trabalho se dará obtendo-se os resultados práticos com a rede V-net reestruturada. Além disso, será feita uma análise comparativa entre o desempenho desta com outras abordagens existentes.

IV. Conclusão

Este artigo apresentou o desenvolvimento de uma abordagem para realização de segmentação de crânios fetais em imagens de ultrassom bidimensional utilizando como base a rede Neural Completamente Convolucional V-Net. Como passos futuros propõe-se a agregação novas técnicas presentes em outras arquiteturas de rede presentes na literatura, além da análise comparativa com trabalhos de outros autores.

Referências

[1] S. Rueda, S. Fathima, C. L. Knight, M. Yaqub, A. T. Papageorghiou, B. Rahmatullah, A. Foi, M. Maggioni, A. Pepe, J. Tohka *et al.*, "Evaluation and comparison of current fetal ultrasound image segmentation methods for biometric measurements: a grand challenge," *IEEE Transactions on medical imaging*, vol. 33, no. 4, pp. 797–813, 2014. [Online]. Available: https://doi.org/10.1109/TMI.2013.2276943

[2] L. Wu, J.-Z. Cheng, S. Li, B. Lei, T. Wang, and D. Ni, "Fuiqa: Fetal ultrasound image quality assessment with deep convolutional networks," *IEEE transactions on cybernetics*, vol. 47, no. 5, pp. 1336–1349, 2017. [Online]. Available: https://doi.org/10.1109/TCYB.2017.2671898

[3] A. M. Al-Karmi, M. A. Dinno, D. A. Stoltz, L. A. Crum, and J. C. Matthews, "Calcium and the effects of ultrasound on frog skin," *Ultrasound in medicine & biology*, vol. 20, no. 1, pp. 73–81, 1994. [Online]. Available: https://doi.org/10.1016/0301-5629(94)90019-1

[4] N. M. Zayed, A. M. Badwi, A. Elsayad, M. S. Elsherif, and A.-B. Youssef, "Wavelet segmentation for fetal ultrasound images," in *Proceedings of the 44th IEEE 2001 Midwest Symposium on Circuits and Systems. MWSCAS 2001 (Cat. No. 01CH37257)*, vol. 1. IEEE, 2001, pp. 501–504. [Online]. Available: https://doi.org/10.1109/MWSCAS.2001.986220

[5] G. Carneiro, B. Georgescu, S. Good, and D. Comaniciu, "Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree," *IEEE transactions on medical imaging*, vol. 27, no. 9, pp. 1342–1355, 2008. [Online]. Available: https://doi.org/10.1109/TMI.2008.928917

[6] S. M. Jardim and M. A. Figueiredo, "Segmentation of fetal ultrasound images," *Ultrasound in medicine & biology*, vol. 31, no. 2, pp. 243–250, 2005. [Online]. Available: https://doi.org/10.1016/j.ultrasmedbio.2004.11.003

[7] W. Lu, J. Tan, and R. Floyd, "Automated fetal head detection and measurement in ultrasound images by iterative randomized hough transform," *Ultrasound in medicine & biology*, vol. 31, no. 7, pp. 929–936, 2005. [Online]. Available: https://doi.org/10.1016/j.ultrasmedbio.2005.04.002

[8] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587. [Online]. Available: https://doi.org/10.1109/CVPR.2014.81

[9] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99. [Online]. Available: https://doi.org/10.1109/TPAMI.2016.2577031

[10] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9. [Online]. Available: https://doi.org/10.1109/CVPR.2015.7298594

[11] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440. [Online]. Available: https://doi.org/10.1109/CVPR.2015.7298965

[12] Y. Li, R. Xu, J. Ohya, and H. Iwata, "Automatic fetal body and amniotic fluid segmentation from fetal ultrasound images by encoder-decoder network with inner layers," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2017, pp. 1485–1488. [Online]. Available: https://doi.org/10.1109/CVPR.2015.7298965

[13] V. Sundaresan, C. P. Bridge, C. Ioannou, and J. A. Noble, "Automated characterization of the fetal heart in ultrasound images using fully convolutional neural networks," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 2017, pp. 671–674. [Online]. Available: https://doi.org/10.1109/ISBI.2017.7950609

[14] M. Sinclair, C. F. Baumgartner, J. Matthew, W. Bai, J. C. Martinez, Y. Li, S. Smith, C. L. Knight, B. Kainz, J. Hajnal *et al.*, "Human-level performance on automatic head biometrics in fetal ultrasound using fully convolutional neural networks," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 714–717. [Online]. Available: https://doi.org/10.1109/EMBC.2018.8512278

[15] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, 2016, pp. 565–571. [Online]. Available: https://doi.org/10.1109/3DV.2016.79

[16] T. van den Heuvel, D. de Bruijn, C. L. de Korte, and B. van Ginneken, "Automated measurement of fetal head circumference," Jul. 2018. [Online]. Available: https://doi.org/10.5281/zenodo.1322001

[17] S. Vesal, N. Ravikumar, and A. Maier, "A 2d dilated residual u-net for multi-organ segmentation in thoracic ct," *arXiv preprint arXiv:1905.07710*, 2019.