

# Segmentação automática dos lobos pulmonares em imagens de CT utilizando U-Net: comparando abordagens 2D e 2D-estendida

Jean Antonio Ribeiro, Leticia Rittner  
{j265739@dac.unicamp.br, lrittner@unicamp.br}

Departamento de Engenharia de Computação e Automação (DCA)  
Faculdade de Engenharia Elétrica e de Computação (FEEC)  
Universidade Estadual de Campinas (Unicamp)  
Campinas, SP, Brasil

**Resumo** – O desenvolvimento de algoritmos eficientes e robustos em ferramentas automatizadas para segmentação pulmonar e seus lobos é fundamental para o diagnóstico e acompanhamento de doenças pulmonares. Este trabalho apresenta uma abordagem para segmentação automática de lobos pulmonares utilizando redes neurais profundas em imagens de CT. Para isso, foram utilizadas duas abordagens em conjunto com a rede U-Net: 2D e 2D-estendida. Na abordagem 2D, a rede é treinada fatia por fatia utilizando convolução 2D e, na predição, os resultados são empilhados para formar a segmentação volumétrica. Na abordagem 2D-estendida, utiliza-se também convolução 2D, mas são colocadas na entrada da rede 3 fatias (como canais), para reter parte da informação volumétrica. Testes utilizando um dataset público, que inclui 50 imagens com anotações, para cada um dos lobos pulmonares, mostraram que a abordagem 2D-estendida pode alcançar uma média de *Dice* superior a 91%, sendo ligeiramente superior à abordagem 2D.

**Palavras-chave** – segmentação automática, lobos pulmonares, fissuras, rede neural convolucional, U-Net, tomografia computadorizada

## 1. Introdução

O pulmão humano é dividido em cinco lobos que são separados pelas fissuras lobares. O pulmão direito tem três lobos: inferior, superior e médio. O pulmão esquerdo tem somente dois lobos: superior e inferior. Os sistemas brônquico e vascular são isolados com conexões mínimas entre os lobos, e os lobos são considerados unidades com funções relativamente independentes. Assim, as doenças pulmonares podem ser limitadas a um lobo individual.

Geralmente, imagens de tomografia computadorizada (CT) são utilizadas para identificar anormalidades ou pequenas lesões pulmonares, dificilmente vistas em imagens radiográficas. No entanto, a análise dessas imagens não é simples, principalmente quando o radiologista precisa de uma solução rápida. A quantidade de fatias geradas durante o exame por paciente é grande, podendo conter centenas de fatias, cada uma das quais precisa ser analisada pelo radiologista. Este processo é demorado, difícil de ser realizado e envolve algum grau de variabilidade entre observadores.

Além disso, os limites lobares, definidos pelas fissuras pulmonares, muitas vezes são parcialmente invisíveis nas imagens de CT [3]. Em pulmões saudáveis, as bordas dos lobos são definidas por fissuras visíveis, que muitas das vezes podem estar incompletas, dificultando a identificação correta dos limites lobares [13].

Alterações morfológicas em lobos específicos também podem ocorrer durante o progresso patológico de

doenças pulmonares [3]. Doenças como fibrose ou enfisema podem obscurecer a forma e a aparência das fissuras lobares. Em pacientes com COVID-19, as imagens de CT apresentam sombras em vidro fosco e os lobos são preenchidos com líquido pleural, tornando invisíveis partes das fissuras lobares [16].

Na segmentação de lobos pulmonares, George et al. [4] desenvolveram um algoritmo utilizando redes neurais profundas, para identificar os limites lobares, e *random walker* (RW), para gerar as segmentações finais. O treinamento e a inferência foram realizados em fatias axiais 2D. Em relação a U-Net 3D, a U-Net 2D possui a vantagem que não utilizar, nas imagens de entradas, a profundidade existente nos volumes de CT. Por sua vez, Imran et al. (2020) [8] combinaram características da rede V-Net [5] e redes progressivas aninhadas holisticamente [6]. Para evitar o *overfitting*, foram utilizadas somente fatias axiais, em que pelo menos um lobo do pulmão está presente.

Visto que a localização e distribuição da doença pulmonar é um fator significativo na determinação de um tratamento adequado, muitos trabalhos tem sido desenvolvidos com o objetivo de alcançar uma segmentação adequada dos lobos pulmonares. Do ponto de vista técnico, a segmentação precisa do lobo pulmonar pode melhorar as tarefas clínicas subsequentes, incluindo a previsão de malignidade do nódulo, avaliação e quantificação de doenças pulmonares, reduzindo o espaço de busca para lobos mais propensos a serem afetados [3].

## 2. Proposta

A seguir, são descritos os datasets que foram utilizados nos experimentos. Também é descrita a implementação das etapas de pré-processamento, treinamento e predição nas abordagens 2D e 2D-estendida.

### 2.1. Dataset

Foram utilizados quatro datasets: Medical Segmentation [10] (9 volumes), VESSEL12 [14] (23 volumes), MOSMED [11] (172 volumes) e Tang, Zhang and Xie [15] (50 volumes). Nos 3 primeiros datasets, as anotações foram geradas de forma automática pelo *framework lung-mask* [7], consideradas neste trabalho como padrão prata (*silver standard*). O único dataset com padrão ouro (*gold standard*) disponível é o de Tang, Zhang and Xie [15], que segundo os autores, foram geradas por radiologistas.

### 2.2. Método

Primeiramente, todas as imagens de CT foram normalizadas para valores da Unidade Hounsfield (HU) no intervalo de  $[-1024, 600]$  e depois normalizadas para um intervalo entre 0 e 1. O otimizador utilizado foi Adam [9], com *learning rate* de  $5 \times 10^{-4}$  e batches de 6 imagens. O modelo foi implementado utilizando a biblioteca de código aberto PyTorch Lightning. Todos os experimentos foram executados com 60 épocas. O modelo foi treinado e avaliado nas vistas axiais, cada uma com tamanho de  $512 \times 512$ . Com o objetivo de reduzir a quantidade de memória utilizada na GPU e ajudar na generalização do modelo, um *RandomCrop* de  $64 \times 256 \times 256$  foi realizado em cada uma das fatias, em ambas as abordagens 2D e 2D-estendida. Para otimizar os pesos do modelo, durante o treinamento, foi utilizada a função de custo *Dice*.

Foram desenvolvidas duas abordagens para a segmentação dos lobos pulmonares em imagens de CT: 2D e 2D-estendida. Para isso, foi utilizada a rede U-Net, que é uma rede amplamente utilizada na segmentação de imagens médicas [3]. A rede é treinada fatia por fatia utilizando convolução 2D, contudo, na abordagem 2D-estendida, são colocadas na entrada da rede 3 fatias, como se fossem três canais, obedecendo a uma determinada regra de espaçamento. A ideia de utilizar 3 fatias consecutivas, como entrada da rede, foi proposta inicialmente por Pereira et al. [12], com o intuito de trazer a informação volumétrica para a U-Net 2D.

Na abordagem 2D-estendida, três variações de espaçamento foram implementadas: sem espaçamento (*consec.*), a cada 2 fatias (*salta 1*) e a cada 3 fatias (*salta 2*). Na variação do tipo *consec.*, as imagens são formadas pelas fatias  $n - 1$ ,  $n$  e  $n + 1$ ; na variação do tipo *salta 1*, as imagens são formadas pelas fatias  $n - 2$ ,  $n$  e  $n + 2$ ; e,

finalmente, na variação *salta 2*, as imagens são formadas pelas fatias  $n - 3$ ,  $n$  e  $n + 3$ . Na abordagem 2D, cada imagem é gerada por cada uma das fatias que formam o volume 3D (Fig. 1).

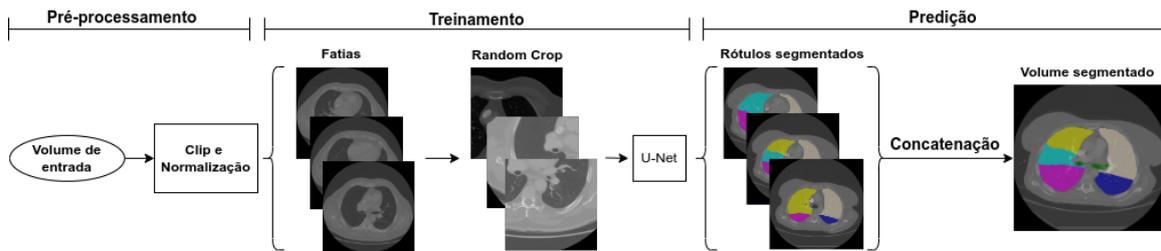
Na etapa de predição, as fatias são empilhadas para formar a segmentação volumétrica. Em seguida, o desempenho do volume é por meio do Coeficiente *Dice* [2]. A saída da rede é composta de seis canais: fundo, lobo direito superior (RUL), lobo direito do meio (RML), lobo direito inferior (RLL), lobo esquerdo superior (LUL) e lobo esquerdo inferior (LLL). O conjunto de treinamento é composto por todos os volumes dos datasets com padrão-prata (Medical Segmentation, VESSEL12 e MOSMED). O conjunto de teste é composto de 40 volumes do dataset de Tang, Zhang and Xie [15]. Os 10 volumes restantes, deste último dataset, foram separados para utilização nos experimentos com *finetuning*.

## 3. Experimentos e Resultados

A segmentação do lobo pulmonar foi avaliada por meio da métrica *Dice Similarity Coefficient (Dice)*. O *Dice* é amplamente utilizado para avaliar a similaridade entre duas amostras [2].

Na avaliação dos resultados, foram realizados 2 conjuntos de experimentos. No primeiro conjunto, o modelo foi treinado somente com máscaras padrão-prata. A predição do modelo foi comparada às máscaras padrão-ouro, no conjunto de testes. No segundo conjunto, o modelo também foi treinado com as máscaras padrão-prata e foi realizado o *finetuning* dos pesos, utilizando-se os 10 volumes com padrão-ouro, separados para este fim. O objetivo aqui foi verificar se, na falta de um conjunto grande com anotações manuais, o *finetuning* de um modelo pré-treinado com máscaras geradas automaticamente é capaz de melhorar a performance do modelo. Dentro de cada conjunto de experimentos (com e sem *finetuning*), foram comparadas as abordagens 2D e 2D-estendida, esta última variando-se o espaçamento das fatias de entrada.

Na abordagem utilizando a rede U-Net 2D-estendida sem *finetuning*, os valores médio de *Dice* foram superiores a 90%, sendo 95,3% o maior deles, correspondente a abordagem *consec.* Na abordagem *consec.* com *finetuning*, os resultados foram melhores do que nas abordagens com espaçamento *salta 1* e *salta 2*. Em geral, nas abordagens utilizando 2D-estendida, os resultados foram ligeiramente melhores do que na abordagem 2D. Por causa da anatomia do pulmão e da pouca visibilidade apresentada pela fissura horizontal direita em imagens de CT, segmentar o lobo superior direito é mais difícil do que os outros lobos [1]. Consequentemente os valores de *Dice*, desse lobo, foram inferiores aos demais (Tabela 1).



**Figura 1.** Fases do algoritmo implementado utilizando a rede U-Net 2D. No caso da abordagem 2D-estendida, cada fatia representa uma imagem com três fatias.

Finetuning	Espaçamento	Dice	LUL	LLL	RUL	RML	RLL
Não	<i>consec.</i>	0.909±0,045	0.958±0,031	0.955±0,026	0.903±0,071	0.778±0,128	0.950±0,024
	<i>salta 1</i>	0.920±0,042	0.961±0,032	0.955±0,026	0.914±0,068	0.817±0,125	0.951±0,023
	<i>salta 2</i>	0.915±0,044	0.961±0,028	0.955±0,023	0.908±0,074	0.804±0,124	0.945±0,025
	2D	0.908±0,043	0.957±0,028	0.954±0,024	0.899±0,072	0.794±0,120	0.935±0,027
Sim	<i>consec.</i>	0.896±0,051	0.955±0,038	0.953±0,031	0.895±0,085	0.739±0,136	0.938±0,037
	<i>salta 1</i>	0.873±0,068	0.946±0,042	0.940±0,035	0.863±0,097	0.695±0,197	0.922±0,049
	<i>salta 2</i>	0.863±0,065	0.943±0,044	0.938±0,036	0.844±0,098	0.676±0,171	0.915±0,055
	2D	0.891±0,061	0.949±0,038	0.945±0,036	0.884±0,083	0.741±0,176	0.937±0,036

**Tabela 1.** Média dos valores de *Dice* obtidos no dataset de teste de Tang, Zhang and Xie [15]. O treinamento do modelo foi realizado com e sem *finetuning*. Para cada lobo, é mostrado o valor médio de *Dice* correspondente a variação de espaçamento entre fatias que é utilizado na abordagem 2D-espandida.

Na Figura 2, é mostrada uma imagem de CT, a anotação *Ground Truth* (GT), correspondente para cada um dos lobos pulmonares, e uma imagem de predição. Na imagem de predição, é possível observar que alguns voxels foram segmentados de forma incorreta, quando comparados com a anotação GT. Cada tom de cor, na anotação GT e na imagem de predição, representa um lobo do pulmão. Geralmente, muitas das segmentações incorretas são causadas por fissuras incompletas que estão presentes nas imagens de CT, por exemplo, originadas na ocorrência de lesões patológicas próximas às bordas dos lobos.

#### 4. Conclusões

Este trabalho apresentou uma abordagem para segmentação automática de lobos pulmonares utilizando redes neurais profundas em imagens de CT. Para isso, foram utilizadas duas abordagens em conjunto com a rede U-Net: 2D e 2D-estendida. A qualidade das segmentações produzidas foi avaliada por meio do Coeficiente *Dice*. Os valores de *Dice* alcançados pelo modelo ainda são baixos, quando comparados com o estado da arte. Especificamente, o baixo valor de *Dice*, apresentado pela segmentação do lobo superior direito, indica que a presente abordagem precisa ser melhorada. Nas segmentações produzidas pela abordagem 2D-estendida *consec.*, o valor médio de *Dice* foi superior a 91% sem *finetuning* e 89% utilizando *finetuning*. Na abordagem 2D, o valor médio de *Dice* foi de 90,8% e 89,1%, nas abordagens com e sem

*finetuning*, respectivamente. Neste caso, a utilização de *finetuning* ajudou na generalização da rede. Em comparação com a abordagem 2D, a abordagem 2D-estendida foi ligeiramente superior.

Baseado nos resultados experimentais, percebe-se que esta abordagem é uma boa base para projetos futuros, envolvendo a segmentação dos lobos pulmonares. Entretanto, as segmentações finais apresentam ruídos e precisam de ajustes. Uma solução, para este problema, seria a utilização de outras redes, juntamente com técnicas de regularização e argumentação. Trabalhos futuros incluem treinar o modelo em um dataset maior e avaliar outros recursos, tanto computacionais quanto metodológicos, com o objetivo de melhorar a precisão da segmentação em diferentes datasets. Esse dataset deverá conter imagens sintéticas e de CT que representem uma ampla variedade de doenças pulmonares, como COVID-19 e câncer de pulmão. Devido à diferença nos tamanhos dos lobos, uma função de custo que utilize uma média ponderada por lobos deverá ser desenvolvida, com o objetivo de melhorar a qualidade das segmentações produzidas.

#### Agradecimentos

Este projeto foi parcialmente apoiado pelas bolsas da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) 88887.513444/2020-00.

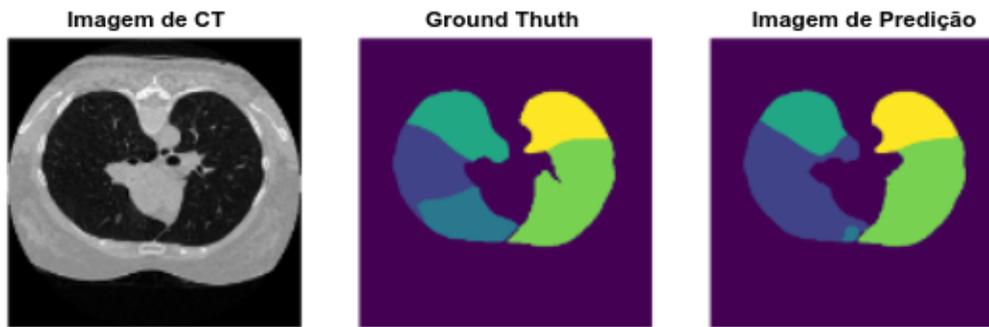


Figura 2. Demonstração de uma imagem de CT, anotação GT e imagem de predição.

## Referências

- [1] P. Cronin, B. Gross, A. Kelly, S. Patel, E. Kazerooni, and R. Carlos. Normal and accessory fissures of the lung: Evaluation with contiguous volumetric thin-section multidetector CT. *Eur. J. Radiol.*, 75:e1–8, 11 2009.
- [2] L. R. Dice. Measures of the Amount of Ecologic Association Between Species. *Ecology*, 26(3):297–302, 1945.
- [3] T. Doel, D. J. Gavaghan, and V. Grau. Review of automatic pulmonary lobe segmentation methods from CT. *Comput. Med. Imaging Graphics*, 40:13–29, 2015.
- [4] K. George, A. Harrison, D. Jin, Z. Xu, and D. Mollura. Pathological Pulmonary Lobe Segmentation from CT Images Using Progressive Holistically Nested Neural Networks and Random Walker. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 195–203, Cham, 09 2017. Springer.
- [5] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. Davidson, S. P. Pereira, M. J. Clarkson, and Dean C. Barratt. Automatic Multi-Organ Segmentation on Abdominal CT With Dense V-Networks. *IEEE Trans. Med. Imaging.*, 37(8):1822–1834, 2018.
- [6] A. P. Harrison, Z. Xu, K. George, Le Lu, R. M. Summers, and D. J. Mollura. Progressive and Multi-path Holistically Nested Neural Networks for Pathological Lung Segmentation from CT Images. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2017*, pages 621–629, Cham, 2017. Springer.
- [7] J. Hofmanninger, F. Prayer, J. Pan, S. Röhrich, H. Prosch, and G. Langs. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. *Eur. Radiol. Exp.*, 4:50, 08 2020.
- [8] A. Imran, A. Hatamizadeh, S. P. Ananth, X. Ding, N. Tajbakhsh, and D. Terzopoulos. Fast and automatic segmentation of pulmonary lobes from chest CT using a progressive dense V-network. *Comput. Methods Biomech. Biomed. Eng.: Imaging Visualization*, 8(5):509–518, 2020.
- [9] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations, ICLR 2015*, San Diego, CA, USA, 05 2015.
- [10] MedSeg. COVID-19 CT segmentation dataset. [urlhttp://medicalsegmentation.com/covid19](http://medicalsegmentation.com/covid19), 2020. acessado em 10/08/2022.
- [11] S. P. Morozov, A. E. Andreychenko, N. A. Pavlov, A. V. Vladzomyrskyy, N. V. Ledikhova, V. A. Gombolevskiy, I. A. Blokhin, P. B. Gelezhe, A. V. Gonchar, and V. Yu. Chernina. MosMedData: Chest CT Scans With COVID-19 Related Findings Dataset, 2020.
- [12] M. Pereira, I. Fantini, R. Lotufo, and L. Rittner. An extended-2d cnn for multiclass alzheimer’s disease diagnosis through structural mri. In *Medical Imaging 2020: Computer-Aided Diagnosis*, volume 11314, pages 438–444. SPIE, 2020.
- [13] B. N. Raasch, E. W. Carsky, E. J. Lane, J. P. O’Callaghan, and E. R. Heitzman. Radiographic anatomy of the interlobar fissures: a study of 100 specimens. *AJR.*, 138:1043–1049, 06 1982.
- [14] R. D. Rudyanto, S. Kerkstra, and E. M. [van Rikxoort] et al. Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: the VESSEL12 study. *Med. Image Anal.*, 18(7):1217–1232, 2014.
- [15] H. Tang, C. Zhang, and X. Xie. Automatic Pulmonary Lobe Segmentation Using Deep Learning. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 1225–1228, Venice, Italy, 03 2019. IEEE.
- [16] S. Zheng, W. Nie, L. Pan, B. Zheng, Z. Shen, L. Huang, C. Pei, Y. She, and L. Chen. A dual-attention V-network for pulmonary lobe segmentation in CT scans. *IET Image Proc.*, 15(8):1644–1654, 2021.