Lung Region Segmentation in Chest X-Ray Images using Deep Convolutional Neural Networks

R. D. S. Portela, J. R. G. Pereira, M. G. F. Costa, *Member, IEEE*, C. F. F. Costa Filho, *Member, IEEE*,

***Abstract*— Lung cancer is, by far, the leading cause of cancer death in the world. Tools for automated medical imaging analysis development of a Computer-Aided Diagnosis method comprises several tasks. In general, the first one is the segmentation of region of interest, for example, lung region segmentation from Chest X-ray imaging in the task of detecting lung cancer. Deep Convolutional Neural Networks (DCNN) have shown promising results in the task of segmentation in medical images. In this paper, to implement the lung region segmentation task on chest X-ray images, was evaluated three different DCNN architectures in association with different regularization (Dropout, L2, and Dropout + L2) and optimization methods (SGDM, RMSPROP and ADAM). All networks were applied in the Japanese Society of Radiological Technology (JSRT) database. The best results were obtained using Dropout + L2 as regularization method and ADAM as optimization method. Considering the Jaccard Coefficient obtained (0.97967 ± 0.00232) the proposal outperforms the state of the art.**

***Clinical Relevance*— The presented method reduces the time that a professional takes to perform lung segmentation, improving the effectiveness.**

I. Introduction

According to the World Health Organization (WHO), cancer is the second leading cause of death worldwide, and is responsible for an estimated 9.6 million deaths in 2018. Lung cancer is the most common cause of cancer death when compared to the number of deaths caused by any other cancer type [1]. Medical imaging, such as radiography, assists the physician in decision make about diagnosis and treatment of lung cancer. Computer-Aided Lung Cancer Diagnosis methods may require the segmentation of lung tissue in the radiographic image. Therefore, using tools that apply techniques for automatic segmentation can greatly save effort and help in obtaining more accurate diagnosis. There are several studies published in the literature aiming lung region segmentation in chest X-ray images, using different methods. Chondro et al. [2] used digital imaging techniques, such as histogram equalization and Gaussian filter, to segment the lung region on chest X-ray images. Yassine et al. [3] performed pulmonary segmentation using the Superpixels technique. In recent years, Deep Convolutional Networks, one of the most important Deep Learning Network, have become increasingly prominent in tasks involving medical imaging, such as segmentation, detection, etc., due to the its superiority over other methods used so far in these tasks. Deep Learning is a particular type of machine learning, which uses multiple layers of processing to achieve high levels of abstraction when learning through data representations. Gordienko et al. [4] used U-Net architecture to investigate the ability of network to segment the lung region in radiographic images with and without its bone structures. Islam and Zhang [5] used convolutive networks to segment in the Montgomery and Shenzhen databases. Mayan et al. [6] used the architecture U-Net with an ImageNet Pre-trained encoder.

This study aims to evaluate the ability of Convolutional Networks to perform the task of lung segmentation in chest X-ray images. The article is organized as follows: Section II presents the dataset used, the implemented architectures, as well as their training parameters. Also, section II shows the metrics used for segmentation evaluation. Section III presents the results obtained from simulations and the discussion.

II. Methodology

A. Dataset

The Japanese Society of Radiological Technology (JSRT) database [7] was chosen to evaluate the Deep Convolutional Neural Network (DCNN) models. It is an annotated image database containing 247 chest X-ray images (154 with a single pulmonary nodule and 93 with no pulmonary nodules) with their respective gold standard segmented lung regions. All images in the database are gray level and have 2048x2048 pixels. Fig. 1 shows two examples of database images, one with a pulmonary nodule and another without a pulmonary nodule. To each Chest X-ray image there are two gold standard images: one corresponding to the segmented region of the left lung and the other corresponding to the segmented region of the right lung. To perform convolutional neural network training, it was necessary to fusion the two images. Fig. 2 shows gold standard images of the images shown in Fig. 1, after the fusion operation. Due to hardware memory limitation, it was necessary to resize all images, as well as their respective gold standards, to a resolution of 512x512 pixels.

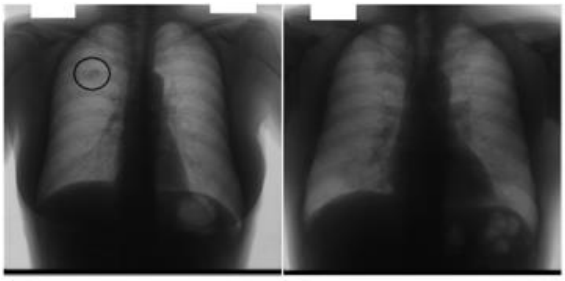
B. Deep Convolutional Neural Network Architectures and Training Parameters

Miyagawa et al. [8] used three architectures to perform lumen segmentation in IVOCT images. Int his paper we use these same architectures for segmentation of the pulmonary region. The first architecture proposed (DCNN-1) is a semantic network. This network consists of 51 layers. Like all semantic networks, it is made up of subsampling layers followed by oversampling layers. The network contains 4

* This research, according for in Article 48 of Decree nº 6.008/2006, was funded by Samsung Electronics of Amazonia Ltda, under the terms of Federal Law nº 8.387/1991, through agreement nº 004, signed with the Center for R&D in Electronics and Information from the Federal University of Amazonas - CETELI/UFAM..

The authors are with Federal University of Amazonas, Amazonas, Brazil (e-mails: portela.ronaldo@hotmail.com, mcosta@ufam.edu.br, ccosta@ufam.edu.br, josergpereira@gmail.com)

subsampling modules. Each module contains two sequence of the following layers: convolution operation, batch normalization operation and ReLu operation followed by a maxpooling operation. In the first subsample layer, 32 feature maps were used, while the remaining layers used 64 feature maps. All convolution filters are 3x3 in size and have zero padding at the edges, so, after the convolution layer, the image dimension does not change. All filters used in subsampling are 2x2 in size with horizontal and vertical steps equal to 2. After the subsampling layer, its size will be reduced by a factor of 2, both horizontally and vertically. In the sequence, follow 4 layers of oversampling, so that the image retrieves the dimension lost in the previous steps and returns to its original size. In all oversampling steps, 4x4 filters were used, with horizontal and vertical steps equal to 2. In the first three oversampling layers, 64 feature maps were used, while, in the fourth oversampling layer, a 32 feature map was used. Subsequently, the image went through a 1x1 convolution layer, where its volume is equal 2 (lung and background). Finally, there is a classification layer. In this layer, a Softmax function classifies each pixel oh the image as belonging to the lung or background. DCNN-1 is shown in Fig. 3(a).



(a) (b)

Figure 1. Example of Chest X-ray images from the database. (a) Image with a nodule (b) Image without a nodule.

(a) (b)

Figure 2. Gold Standard segmented images corresponding to images shown in Figure 1.

The second architecture (DCNN-2) is also a direct network but has Convolution, Batch and ReLu layers between its transposed convolution layers, totaling 75 layers in its structure. ReLu layers added between the transposed convolution steps should give greater nonlinearity and abstraction capability during each oversampling step. DCNN-2 is shown in Fig. 3(b).

The third architecture (DCNN-3) is a network applying Directed Acyclic Graphs (DAG). In this type of architecture, information from the initial subsampling steps is passed on to the final oversampling steps. The information that is passed from the initial layers is linked to the information contained in the final oversampling layers and, as they have the same dimensions, can be connected without any size adjustments being required, as shown in Fig. 3(c).

The database was divided into three sets: training, validation and testing, in the following proportion: 50% - 25% - 25%, respectively. Also, it was maintained the same proportion for images with and without nodules. For example, to compose the training set, 50% of the images without nodule and 50% of the images with nodule were selected.

Dropout, L2 and Dropout + L2 was chosen as regularization methods and SGDM, RMSPROP and ADAM was chosen as optimization methods. In a first step, to choose the best network, the training and validation set was used. After choosing the best network, cross-validation with five folders between the training and test sets was performed.

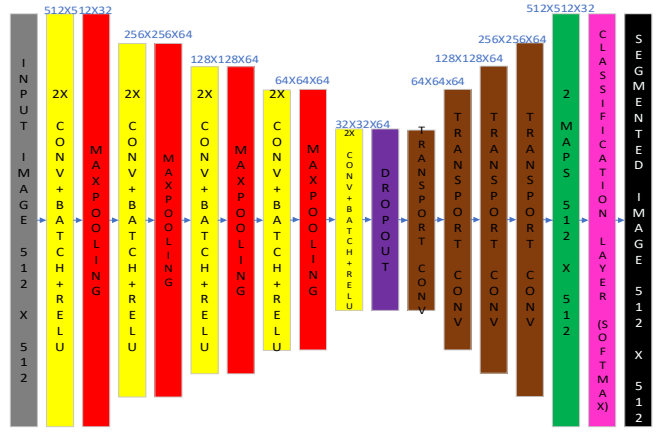
For this work, it was used a computer with Windows 10 operating system, Intel Core I & -8700 CPU @3.20GHz 3.19 GHz processor, 16 GB of RAM and 8GB NVIDIA GeForce GTX 1070 GPU. The environment development was MATLAB R2019a. Parameter values for DCNNs training were adjusted experimentally. **Erro! Fonte de referência não encontrada.** show these values.

C. Evaluation metrics

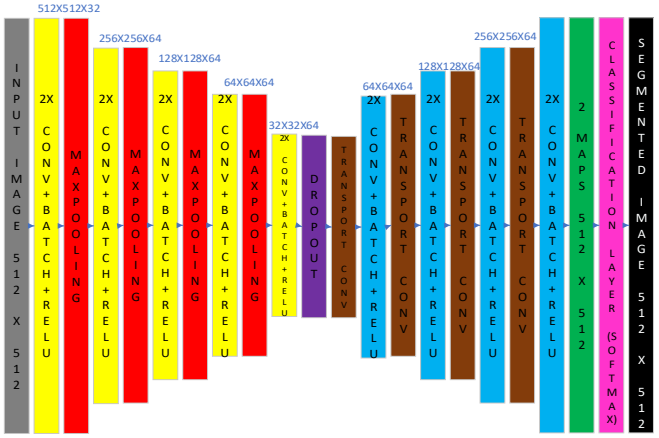
The following metrics were used to evaluate the networks: Global Accuracy, Accuracy, Jaccard Coefficient, Weighted Jaccard Coefficient, Dice Index and Score F1.

III. RESULTS AND DISCUSSION

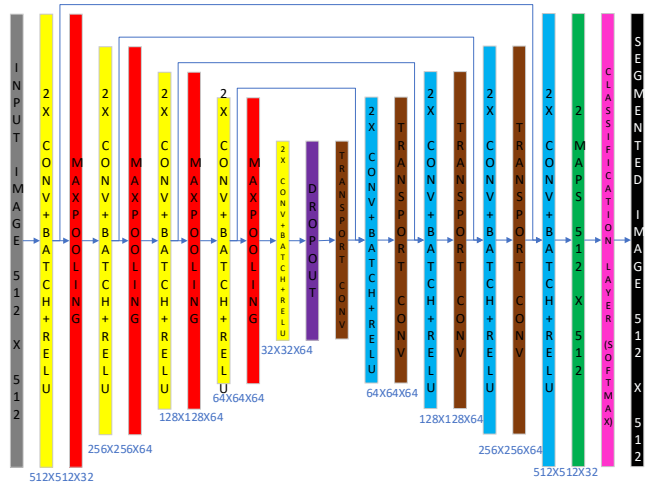
Combining the three architectures, the three regularization methods and the three optimization methods, result in 27 experiments. Table II shows the results of the 27 simulations performed using the training set and the performance obtained using the validation set. The objective was to select the DCNN with best performance in the validation set. As observed, DCNN-1 architecture presents low values for F1 Score, an index that indicates how well the edges of each segmented region align with the edges of the respective gold standard. DCNN-2 architecture outperforms DCNN-1. It is due to the Convolution, Batch and ReLu layers that were inserted between the transposed convolution layers. The best simulation results, however, were obtained with DCNN-3 using Dropout + L_2 regularization in association with the ADAM optimization method. After selecting the best DCNN model, we performed the 5-folder cross-validation using the training and testing sets. The obtained metric values (average ± standard deviation) are shown in Table III. Fig. 4 exemplify the result of a segmented chest X-ray image and presents its respective gold standard. Additionally, a comparison of the results obtained with the existing works in the literature using the same database was made. In [9], a Jaccard Coefficient of 0.903 ± 0.057 was obtained. In [2], a Jaccard Coefficient of 0.963 ± 0.012 and Dice Index of 0.983 ± 0.007 were obtained. In [10], a Jaccard Coefficient of 0.961 ± 0.015 was obtained. As shown in Table 3, in this work, it was obtained a Jaccard Coefficient of 0.97967 ± 0.00232, a Dice Index of 0.98921 ± 0.00163 and F1 Score of 0.97475 ± 0.00357.



(a)



(b)



(c)

Figure 3. (a) First Architecture (DCNN-1) – Direct Network, (b) Second Architecture (DCNN-2) – Direct network and (c) Third Architecture Network (DCNN-3) – Directed Acyclic Graphs

Table I. Parameters used in the network training step

Parameter	Value
Initial Learn Rate (η)	10^{-3}
Learn Rate Drop Factor	0.1
Learn Rate Drop Period	30
Mini Batch Size	1
Max Epochs	50
Squared Gradient Decay Factor (β_2)	0.999
Gradient Decay Factor (β_1)	0.9
Momentum (α)	0.9
Epsilon (ε)	10^{-8}
L_2 Regularization (λ)	10^{-4}
Dropout	50%

When comparing with other works published in the literature, we realize that the results obtained in this work are in accordance with the state of the art until then. This paper helps to highlight the great performance of Deep Convolutional Neural Network in medical images segmentation tasks, and particularly confirms the remarkable performance of DAG networks in segmentation tasks, as previously observed by Miyagawa et al. [8], showing that they can be used to assist the medical professional.

The great advantage of using DCNNs in the lung segmentation, is that there is no need to perform pre or post processing steps, such as threshold application, morphological filters application, histogram equalization etc, because these previous steps are intrinsic to the DCNN.

REFERENCES

[1] WHO. 2018. World Health Organization – WHO, Folha informative – Câncer. Retrieved from https://www.paho.org/bra/index.php?option=com_content&view=article&id=5588:folha-informativa-cancer&Itemid=1094

[2] P. Chondro, C. Y. Yao, S. J. Ruan, and L. C. Chien, "Low order adaptive region growing for lung segmentation on plain chest radiographs". *Neurocomputing,* 275, 1002–1011, September 2017.

[3] B. Yassine, P. Taylor, and A. Story, "Fully automated lung segmentation from chest radiographs using SLICO superpixels". *Analog Integrated Circuits and Signal Processing*, 95(3), 423–428, March 2018.

[4] Y. Gordienko, P. Gang, J. Hui, W. Zeng, Y. Kochura, O. Alienin, S. Stirenko, "Deep learning with lung segmentation and bone shadow exclusion techniques for chest X-ray analysis of lung cancer". *International Conference on Computer Science, Engineering and Education Applications*, 638–647. Springer, Cham. January 2018.

[5] J. Islam, and Y. Zhang. "Towards Robust Lung Segmentation in Chest Radiographs with Deep Learning". *Machine Learning for Health (ML4H) Workshop at NeurIPS*. November 2018. arXiv preprint arXiv: 1811.12638.

[6] M. Frid-Adar, A. Ben-Cohen, R. Amer, and H. Greenspan, "Improving the segmentation of anatomical structures in chest radiographs using U-net with an ImageNet pre-trained encoder". *Image Analysis for Moving Organ, Breast, and Thoracic Images. RAMBO 2018, BIA 2018, TIA 2018 Lectures Notes in Computer Science, vol 11040.* Springer, Cham. September 2018.

[7] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, K. Komatsu, M. Matsui, … K. Doi (1999). "Development of a Digital Image Database for Chest Radiographs with and without a Lung Nodule: Receiver Operating Characteristic Analysis of Radiologists' Detection of Pulmonary Nodules", *American Journal of Roentgenology, 174(1), pp.71-74. Annual meeting of the Radiological Society of North America, Chicago*. November 1999.

[8] M. Miyagawa, M. G. F. Costa, and C. F. F. Costa Filho, "Lumen Segmentation in optical coherence tomography images using

convolutional neural network", *Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Honolulu-USA*. July 2018.

[9] B. Ginneken, M. Stegmann, and M. Loog, "Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database", *Medical Image Analysis*, 10(1), pp.19-40. 2006.

[10] S. Hwang, and S. Park, "Accurate lung segmentation via network-wise training of convolutional networks". *In 3rd Workshop on Deep Learning in Medical Image Analysis*, Quebec. August 2017. arXiv preprint arXiv: 1708.00710

[11] A. A. Novikov, D. Lenis, D. Major, J. Hladûvka, M. Wimmer, and K. Buhler, "Fully convencional architectures for multi-class segmentation in chest radiographs". *IEEE transactions on medical imaging*, 37(8), pp.1865-1876. February 2018.

[12] B. A. Skourt, A. Hassani, A. majda, "Lung CT image segmentation using deep neural networks", *The First Internacional Conference On intelligent Computing in Data Sciences*, pp. 109 – 113. Morocco, 2018.

Table II. Results obtained from the training of architectures using the training set with the performance evaluated using the validation sets

Architecture - Regularization - Optimization	Global Accuracy	Accuracy	Jaccard	Weighted Jaccard	Score F1	Dice
DCNN1 – Dropout – SGDM	0.93935	0.95539	0.86247	0.88047	0.52924	0.96375
DCNN1 – Dropout – RMSPROP	0.98884	0.98704	0.97450	0.97797	0.96889	0.98698
DCNN1 – Dropout – ADAM	0.98763	0.98684	0.95836	0.96477	0.94156	0.98681
DCNN1 – L_2 – SGDM	0.98591	0.98426	0.95520	0.96143	0.91755	0.98419
DCNN1 – L_2 – RMSPROP	0.98900	0.98732	0.97486	0.97828	0.96748	0.98726
DCNN1 – L_2 – ADAM	0.98820	0.98645	0.96000	0.96578	0.94189	0.98638
DCNN1 – Dropout + L_2 – SGDM	0.93560	0.95270	0.85582	0.87428	0.51559	0.95479
DCNN1 – Dropout + L_2 – RMSPROP	0.98931	0.98860	0.97560	0.97891	0.97327	0.98858
DCNN1 – Dropout + L_2 – ADAM	0.98966	0.98818	0.96346	0.96857	0.95279	0.98813
DCNN2 – Dropout - SGDM	0.98048	0.98400	0.94230	0.95172	0.89364	0.98415
DCNN2 – Dropout – RMSPROP	0.98986	0.98970	0.96327	0.96905	0.95366	0.98970
DCNN2 – Dropout – ADAM	0.98929	0.98913	0.96199	0.96795	0.95405	0.98912
DCNN2 – L_2 – SGDM	0.98496	0.98171	0.95321	0.95957	0.92118	0.98155
DCNN2 – L_2 – RMSPROP	0.98878	0.98770	0.96129	0.96692	0.94759	0.98766
DCNN2 – L_2 – ADAM	0.98610	0.98483	0.95567	0.96180	0.93041	0.98477
DCNN2 – Dropout + L_2 – SGDM	0.97637	0.98129	0.93354	0.94430	0.85673	0.98154
DCNN2 – Dropout + L_2 – RMSPROP	0.98876	0.98826	0.96096	0.96693	0.94699	0.98824
DCNN2 – Dropout + L_2 – ADAM	0.98917	0.98871	0.96184	0.96772	0.94960	0.98870
DCNN3 – Dropout - SGDM	0.97900	0.97647	0.95267	0.95903	0.91307	0.97631
DCNN3 – Dropout – RMSPROP	0.99041	0.98849	0.97802	0.98103	0.96777	0.98843
DCNN3 – Dropout – ADAM	0.99168	0.99056	0.98092	0.98352	0.97660	0.99053
DCNN3 – L_2 – SGDM	0.97601	0.96836	0.94567	0.95315	0.89792	0.96770
DCNN3 – L_2 – RMSPROP	0.98769	0.98437	0.97182	0.97569	0.95495	0.98423
DCNN3 – L_2 – ADAM	0.99152	0.98958	0.98052	0.98319	0.97714	0.98953
DCNN3 – Dropout + L_2 – SGDM	0.97918	0.97069	0.95252	0.95914	0.91298	0.97000
DCNN3 – Dropout + L_2 – RMSPROP	0.98930	0.98675	0.97547	0.97884	0.96278	0.98665
DCNN3 – Dropout + L_2 – ADAM	0.99180	0.99027	0.98118	0.98375	0.97873	0.99022

Table III. Result obtained using cross-validation (average ± standard deviation)

Deep Convolutional Network	Global Accuracy	Accuracy	Jaccard	Weighted Jaccard	Score F1	Dice
DCNN-3 – Dropout + L_2 - ADAM	0.99139 ± 0.00098	0.98927 ± 0.00161	0.97967 ± 0.00232	0.98294 ± 0.00191	0.97475 ± 0.00357	0.98921 ± 0.00163

(a) (b) (c)

Fig. 4. Segmentation obtained by Neural Network DCNN-3, using Dropout + L_2 regularization and ADAM optimization (a) left and right lung image, (b) gold standard images and (c) segmented images with gold standard contour superimposed as a black line.