LUNG COMPUTED TOMOGRAPHY IMAGE SEGMENTATION USING U-NET CONVOLUTIONAL NEURAL NETWORK

Yana dos Santos Pereira[1] – yana.pereira@discente.ufopa.edu.br
Davi Guimarães da Silva[1,2,3] – davi.guimaraes@ifpa.edu.br
Regina Cely Barroso[4] – cely_barroso@hotmail.com
Anderson Alvarenga de Moura Meneses[1,2] – anderson.meneses@ufopa.edu.br
[1] Laboratory of Computacional Intelligence, Federal University of Western Pará – PA, Brazil
[2] Graduate Program in Society, Nature and Development, Federal University of Western Pará – PA, Brazil
[3] Federal Institute of Education, Science and Technology of Pará – PA, Brazil
[4] Laboratory of Applied Physics to Biomedical and Environmental Sciences, Physics Institute, State University of Rio de Janeiro – RJ, Brazil

Abstract. *Image segmentation is one of the main resources in computer vision. Nowadays, this procedure can be made with high precision using Deep Learning, and this fact is important to applications of several research areas including medical image analysis. Image segmentation is currently applied to find tumors, bone defects and other elements that are crucial to achieve accurate diagnoses. The objective of the present work is to perform segmentation of lung computed tomography from a dataset of the website Kaggle (www.kaggle.com) using U-Net, a Deep Convolutional Neural Network with Deep Learning, used in biomedical image segmentation. The dataset contains 267 volumes, which are composed of the 2D images and the masks that corresponds to the manual segmented lungs, the reference image called ground truth. In the present work, the dataset was subdivided in such a way that 80% of the volumes were dedicated for training and 20% were used for testing. The results were evaluated using the Dice Similarity Coefficient as metric and the value 0.83 was the mean obtained segmenting images of test data.*

Keywords: Deep Learning, Biomedical Image Segmentation, Fully Convolutional Networks, U-Net

1. INTRODUCTION

Deep Learning (DL) is a subfield of machine learning which is based on learning representations from data with emphasis on learning successive layers of increasingly meaningful representations, almost always using models called neural networks (Chollet, 2018). Nowadays the application of DL has presented promising results in image segmentation. Zhang et al. (2015) designed Convolutional Neural Networks (CNNs, Lecun et al., 1998) architectures to segment infant brain tissues in Magnetic Resonance (MR) images which is a process even more difficult than for adults due to low tissue contrast, increased noise and

Anais do XXV ENMC – Encontro Nacional de Modelagem Computacional, XIII ECTM – Encontro de Ciências e Tecnologia de Materiais,
9º MCSul – Conferência Sul em Modelagem Computacional e IX SEMENGO – Seminário e Workshop em Engenharia Oceânica
19 a 21 Outubro 2022

ongoing white matter myelination. The CNNs segmentation of isointense-phase brain image outperformed competing methods on a set of manual process.Oktay et al. (2018) extended U-Net model to an attention U-Net model for pancreas segmentation, which presented results 2% to 3% higher than other models second the Dice Similarity Coefficient (DSC). This performance improvement in pancreas segmentation is important in many clinical applications of liver segmentation in 3D images (Shen et al., 2017). Conversely, qualitative or even quantitative analysis of lesions and other regions of interest can be performed through the application of DL (Liu et al., 2021), a factor that represents important advances for the entire healthcare sector. The use of DL generates results that do not rely on the subjectivity of the observer and imply a decrease of errors involving diagnostics.

Although several Artificial Neural Networks (ANNs, Haykin, 1998) have been used in biomedical image segmentation, in this work, the process is performed using U-Net, a Fully Convolutional Neural Network (FCN or Fully Convnet) that has been consolidating as one of the most prominent ANNs to this type of images.

The objective of the present work is the segmentation of lung computed tomography using the FCN, delimit the lungs in the tomography, the region of interest that is focus of analysis for many purposes, for example, the detection of pulmonary nodules for early diagnosis of lung cancer.

The content of the article is structured as follows. The first section gives the introduction and provides the purpose of the study. Second section contains the theoretical framework. Third section presents the adopted methodology. Fourth section is dedicated to the results and discussions. Finally, the fifth section contains the conclusion of the work.

2. THEORETICAL FRAMEWORK

2.1 Biomedical Image Segmentation

Image segmentation is considered the most important medical imaging process and corresponds to the extraction of the Region of Interest (ROI), subdividing an image into areas based on a specified description, such as segmenting body organs and certain tissues to detect tumors and other elements and this division can be applied for both 2D and 3D data (Ashour, et al., 2021). The binary segmentation, applied in the present work, subdivide the image in a white part that corresponds to the mask containing the region of interest, and a black part as background. Once obtained, in addition to detecting tumors or other abnormalities, the masks can measure tissue growth analyzing the growth of possible tumors and help in treatment planning.

2.2 Fully Convolutional Neural Networks

FCNs are classified as a specific type of CNNs, that contains a convolutional path connected to a Full Connected (FC) layer (Goodfellow et al., 2016). The FCs has Multilayer Perceptron (MLP) as main representative and they are used to classification. The main difference between the designs of CNNs and FNCs is that the last one has a deconvolutional path, also called expansion path, instead of the dense layer. Therefore, Fully Convnets reduce the number of parameters once there are no FC layers, speeding up learning and inference. As output, the FNCs generate a pixel vector which size corresponds to the input data (Nie et al., 2016).

Anais do XXV ENMC – Encontro Nacional de Modelagem Computacional, XIII ECTM – Encontro de Ciências e Tecnologia de Materiais,
9º MCSul – Conferência Sul em Modelagem Computacional e IX SEMENGO – Seminário e Workshop em Engenharia Oceânica
19 a 21 Outubro 2022

2.3 U-Net

U-Net is an FCN created by Ronneberger, Fischer and Brox (2015) in order to cope with biomedical images segmentation. The architecture of the ANN consists of a contracting path to capture context and expanding path that enables precise localization (Figure 1). Precisely because of the fact that the database of biomedical images uses not to be very large, the authors created a network that is able to be trained end-to-end and presents accurate results.
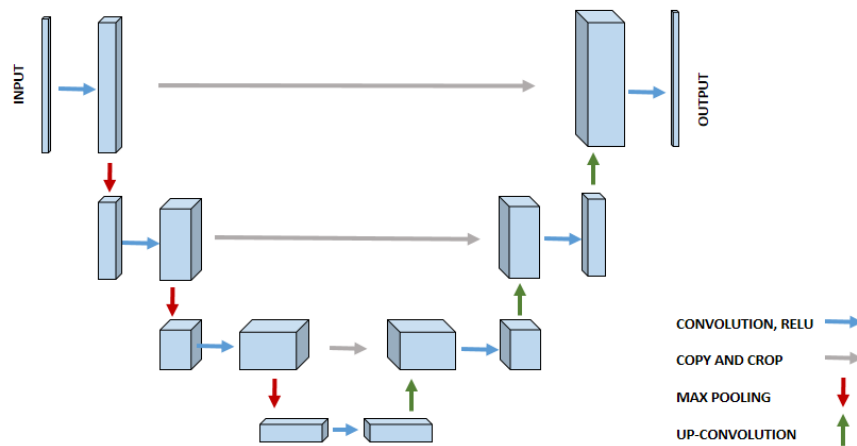

Figure 1 – Architecture of U-Net

Each blue box corresponds to a multi-channel feature map. The path indicated by the red arrows on the left side of Figure 1 corresponds to convolutional path, also known as downsampling path. On the right side, indicated by green arrows, there is the deconvolutional path or upsampling path.

2.4 Related Works

Paiva et al. (2022) used U-Net to segment microcomputed tomography images which the Region of Interest (ROI) corresponds to lenses of an imago specimen of the frog *Thoropa miliaris* and compared the performance to methods of semiautomatic segmentation. The research concluded that the automatic segmentation using Fully Convnet was much faster than the semiautomatic processes and it also showed high accuracy.

Moura and Meneses (2021) applied U-Net in the segmentation of heart computed tomography images and compared the performance of different models. The authors concluded that there were not significant difference between the architectures, therefore the chosen model could be the one with the best performance during the tests.

Saood and Hatem (2021) segmented lung Computed Tomography (CT) using U-Net and SegNet in order to detect and label infected tissues in the lungs and contribute to verify the diagnosis of COVID-19 patients. The results of the work presented that U-Net showed better performance as a multi-class segmentor.

Anais do XXV ENMC – Encontro Nacional de Modelagem Computacional, XIII ECTM – Encontro de Ciências e Tecnologia de Materiais,
9º MCSul – Conferência Sul em Modelagem Computacional e IX SEMENGO – Seminário e Workshop em Engenharia Oceânica
19 a 21 Outubro 2022

3. MATERIAL AND METHODS

3.1. Programming Language and Packages

The programming language used for the implementation of the U-Net was Python 3.9.12. The libraries applied were Keras (www.keras.io) using TensorFlow 2.6 (www.tensorflow.org) as backend for building the model, Scikit-learn (www.scikit-learn.org) for splitting the data into training and testing, Numpy (www.numpy.org) to mathematical operations on the arrays composed by the pixels of the images and for verifying the statistical information of the performance of the model on test predictions, Opencv (www.opencv.org) and Matplotlib (www.matplotlib.org) for image visualization and for building the performance graphics and Pandas (www.pandas.pydata.org) for storing the Dice coefficients of the predictions in dataframes.

3.2. Hardware

The hardware used for training and testing was a processor Intel Core i7-6700 CPU with 15,5 GB of memory. Furthermore, a graphics card NVIDIA GK208B [GeForce GT 730] was employed in the processing.
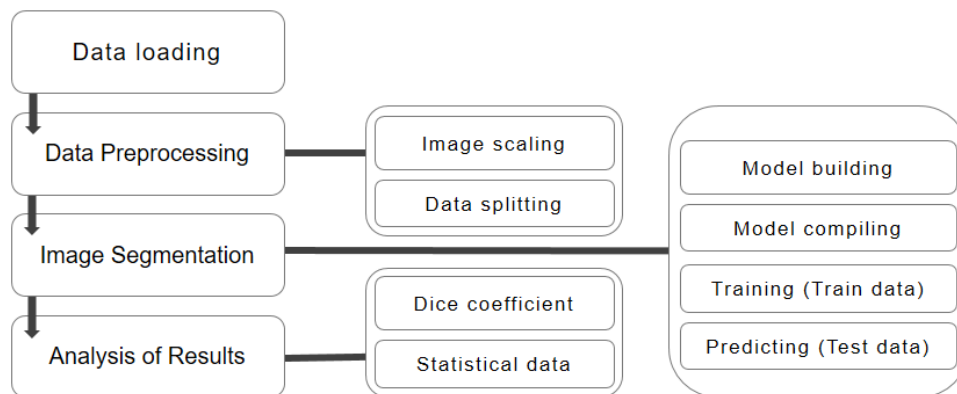
3.3. Methodological Stages

Figure 2 – Methodological flowchart of the work

The dataset was downloaded from the website Kaggle which contains several open source datasets. The dataset used is composed of 267 lungs CT scans and each volume presents a 2D image and a corresponding mask representing the ROI highlighted in white and the background in black.

Before starting the segmentation there were three stages of preprocessing that had to be performed. First step was subdividing the volumes in training and testing. Therefore, 80% of the initial data became training and validation data, corresponding to 214 volumes, while 20% of the initial data became testing data, equivalent to 53 volumes. Secondarily, the images were resizing for ensuring the correct reading of the data by the model and avoid possible differences in file dimensions. The initial data contained 512 by 512 pixels on its dimensions and a normalization from minimum and maximum values as represented below was utilized in order to concentrate the pixels scale on values from 0 to 1.

Anais do XXV ENMC – Encontro Nacional de Modelagem Computacional, XIII ECTM – Encontro de Ciências e Tecnologia de Materiais,
9º MCSul – Conferência Sul em Modelagem Computacional e IX SEMENGO – Seminário e Workshop em Engenharia Oceânica
19 a 21 Outubro 2022

$$Normalization = image\ pixels - \frac{min}{max - min} \qquad (1)$$

Once the preprocessing was completed, a process of Data Augmentation (Shorten and Khoshgoftaar, 2019) was realized to enable an increase in the number of training volumes based on the original ones. In order to add more variability in the dataset, certain geometric transformations were applied, such as horizontal and vertical displacement of the lungs in the images, rotation and zoom. With these changes, from each original image 8 new images were generated, improving quality of the training data and avoiding overfitting.

The training of the model was performed through 10 epochs with 200 steps per epoch and two callbacks were applied to save the best weight configuration in a .h5 format file and to update the value of the optimization rate between the epochs. Furthermore, of the 214 volumes reserved for training, only 80% (171 volumes) were actually used for training while the remaining 20% (43 volumes) were used for model validation.

The metric used for both training and testing was the DSC. The DSC corresponds to a comparison between the result showed by the model and the mask equivalent to the image segmented that is the Ground Truth (GT), the reference image. The metric is a reason between the double of the intersection of the compared images and the sum of pixels of the both images and can be calculated by the equation below.

$$Dice\ coef = \frac{2[n(prediction) \cap n(GT)]}{n(prediction) + n(GT)} \qquad (2)$$

Finally, the testing data was loaded in the model and the Dice coefficient from each of the 53 volumes was saved on a data frame. Then, the statistical information of the table generated by Numpy functions were mean, median, standard deviation and maximum and minimum values.

4. RESULTS AND DISCUSSION

Training and validation runtime was about 272 seconds with an average of 27.2 seconds per epoch and 136 milliseconds per step. Runtime model for testing data, for its part, was about 0.1 second and approximately 2 milliseconds per volume. The distribution of the DSCs of training data and testing data is presented as follows.
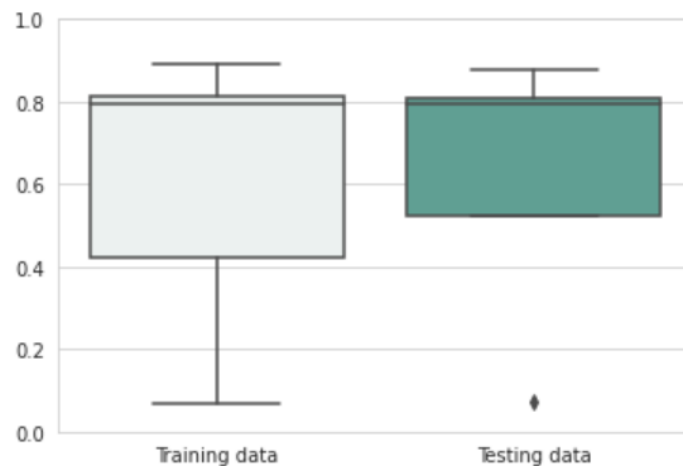

Figure 3 – DSCs distribution boxplot

Anais do XXV ENMC – Encontro Nacional de Modelagem Computacional, XIII ECTM – Encontro de Ciências e Tecnologia de Materiais,
9º MCSul – Conferência Sul em Modelagem Computacional e IX SEMENGO – Seminário e Workshop em Engenharia Oceânica
19 a 21 Outubro 2022

Table 1 – Statistical results of training and testing segmentation

	Training Data	Testing Data
Mean	0.796	0.793
Median	0.814	0.811
Standard Deviation	0.067	0.071
Maximum Value	0.892	0.867
Minimum Value	0.423	0.521

The best results occurred to images that had low noise image, the lung borders are not scattered and there was a significant contrast between the lung and background. This correlation is evident when the predictions of maximum (Figure 4) and minimum Dice (Figure 5) are compared to the images they are equivalent.
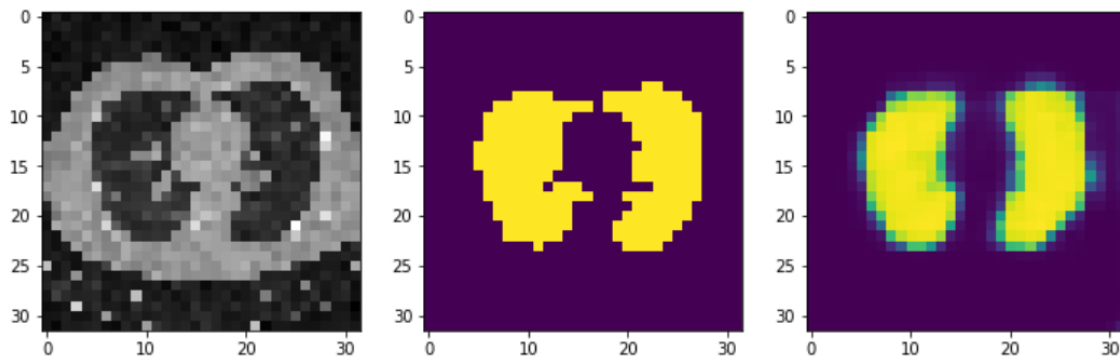

Figure 4 – Computed tomography, equivalent Ground Truth and prediction by U-Net, respectively, of the volume with maximum DSC of testing data.
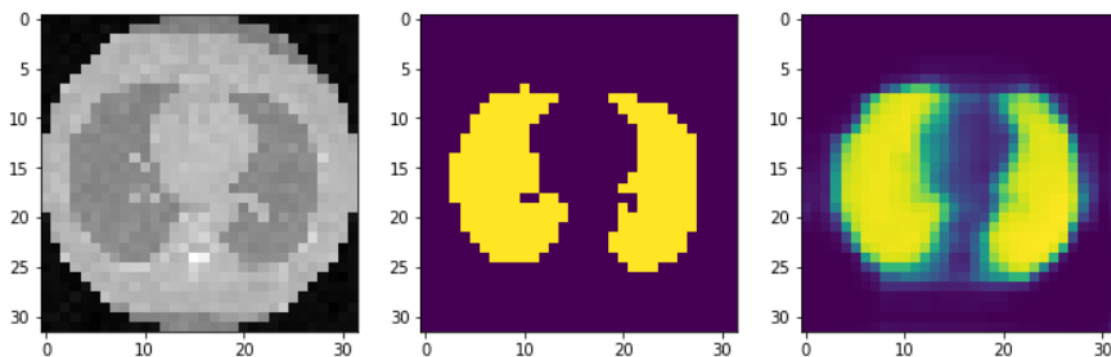

Figure 5 - Computed tomography, equivalent Ground Truth and prediction by U-Net, respectively, of the volume with minimum DSC of testing data.

Therefore, it is possible to observe that the training data had the maximum overall DSC (0.892) and also the minimum overall DSC (0.423), over an interval longer than the DSC of testing data, even though the training data presented the lowest standard deviation (0.067).

Anais do XXV ENMC – Encontro Nacional de Modelagem Computacional, XIII ECTM – Encontro de Ciências e Tecnologia de Materiais,
9º MCSul – Conferência Sul em Modelagem Computacional e IX SEMENGO – Seminário e Workshop em Engenharia Oceânica
19 a 21 Outubro 2022

5. CONCLUSION

In this work, U-Net was used to segment lungs in CT data. The dataset was subdivided in training data and testing data and the model was trained using the masks (GTs) also available in the dataset. The predictions performed by the model to the testing data were compared to the GT, presenting satisfactory results in terms of the metrics used as basis throughout the process (DSC) even though the low number of volumes in the dataset.

Therefore, the present research confirms the efficiency of the U-Net architecture in order to segment biomedical images, factor which enables its implementation to future works of interest to both DL and health care applications.

REFERENCES

Ashour, A., Guo, Y. and Mohamed W. (2021) Medical Image Segmentation. Computer Aided Oral Maxillofacial Surgery. Available at:https://doi.org/10.1016/B978-0-12-823299-6.00001-8.

Chollet, F. (2018) *Deep learning with Python*.

Haykin, S. (1998*) Neural Networks: A Comprehensive Foundation*.

Indolia, S. *et al.* (2018) "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," in *Procedia Computer Science*. Elsevier B.V., pp. 679–688. Available at: https://doi.org/10.1016/j.procs.2018.05.069.

Lecun, Y. et al. (1998) Gradient-Based Learning Applied to Document Recognition.

Liu, X. *et al.* (2021) "A review of deep-learning-based medical image segmentation methods," *Sustainability (Switzerland)*, 13(3), pp. 1–29. Available at: https://doi.org/10.3390/su13031224.

Nie, D. et al. (2016) "Fully convolutional networks for multi-modality isointense infant brain image segmentation," in Proceedings - International Symposium on Biomedical Imaging. IEEE Computer Society, pp. 1342–1345. Available at: https://doi.org/10.1109/ISBI.2016.7493515.

Oktay, O. et al. (2018) *"Attention U-Net: Learning Where to Look for the Pancreas."* Available at: http://arxiv.org/abs/1804.03999.

Paiva, K. *et al.* (2022) "Performance evaluation of segmentation methods for assessing the lens of the frog Thoropa miliaris from synchrotron-based phase-contrast micro-CT images," *Physica Medica*, 94, pp. 43–52. Available at: https://doi.org/10.1016/j.ejmp.2021.12.013.

Ronneberger, O., Fischer, P. and Brox, T. (2015) "U-net: Convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, pp. 234–241. Available at: https://doi.org/10.1007/978-3-319-24574-4_28.

Saood, A. and Hatem, I. (2021) "COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet," *BMC Medical Imaging*, 21(1). Available at: https://doi.org/10.1186/s12880-020-00529-5.

Moura, M. and Meneses, A.A.M. (2021) "*Evaluation Of Unet Convolutional Neural Network Parameters For Segmentation Of Heart CT Images*". Available in the annals of XXIV National Meeting of Computacional Modeling (ENMC).

Shen, D., Wu, G. and Suk, H.-I. (2017) *"Deep Learning in Medical Image Analysis."* Available at: https://doi.org/10.1146/annurev-bioeng-071516.

Shorten, C. and Khoshgoftaar, T.M. (2019) "A survey on Image Data Augmentation for Deep Learning," Journal of Big Data, 6(1). Available at: https://doi.org/10.1186/s40537-019-0197-0.

Zhang, W. et al. (2015) "*Deep convolutional neural networks for multi-modality isointense infant brain image segmentation*," NeuroImage, 108, pp. 214–224. Available at: https://doi.org/10.1016/j.neuroimage.2014.12.061.

Anais do XXV ENMC – Encontro Nacional de Modelagem Computacional, XIII ECTM – Encontro de Ciências e Tecnologia de Materiais,
9º MCSul – Conferência Sul em Modelagem Computacional e IX SEMENGO – Seminário e Workshop em Engenharia Oceânica
19 a 21 Outubro 2022