# Coordinate-Unet 3D for segmentation of lung parenchyma

Van-Linh Le[1,2,3]
van-linh.le@u-bordeaux.fr

Olivier Saut[1,2]
olivier.saut@inria.fr

[1] MONC team - INRIA Bordeaux Sud-Ouest, Talence-33400, France
[2] University of Bordeaux (IMB), CNRS and Bordeaux INP, UMR 5251, Talence-33400, France
[3] BRIC Unit, University of Bordeaux, Talence-33400, France

## ABSTRACT

Lung segmentation is an initial step to provide accurate lung parenchyma in many studies on lung diseases based on analyzing the Computed Tomography (CT) scan, especially in Non-Small Cell Lung Cancer (NSCLC) detection. In this work, Coordinate-UNet 3D, a model inspired by UNet, is proposed to improve the accuracy of lung segmentation in the CT scan. Like UNet, the proposed model consists of a contracting/encoder path to extract the high-level information and an expansive/decoder path to recover the features to provide the segmentation. However, we have considered modifying the structure inside each level of the model and using the Coordinate Convolutional layer as the final layer to provide the segmentation. This network was trained end-to-end by using a small set of CT scans of NSCLC patients. The experimental results show the proposed network can provide a highly accurate segmentation for the validation set with a Dice Coefficient index of 0.991, an F1 score of 0.976, and a Jaccard index (IOU) of 0.9535.

## Keywords
Lung segmentation, NSCLC, Unet, Coordinate Convolutional, Deep learning

## 1 INTRODUCTION

Lung cancer is a major disease that accounts for more than one million deaths [33], and Non-Small Cell Lung Cancer (NSCLC) accounts for 85% of all lung cancers[11]. Early detection of lung cancer could reduce the mortality rate and increase the patient's survival rate during treatment operations. In most techniques to capture the image of the cancer patient, Computed Tomography (CT) scan is an effective medical screening that can be used for the diagnosis and detection of lung cancer. As CT acquisition represent a large volume of CT scans with millions of voxels, manual diagnosis and analysis of lung diseases is a difficult task and time-consuming even for experienced radiologists. Thus, automatic computer-aided diagnosis (CAD) for lung CT is a powerful solution to help radiologists. As usual, CAD involves several steps, of which accurate segmentation of the lung is the first step towards the success of the entire CAD system.

In order to response to clinician demands, lung segmentation algorithms have been proposed and have continued to emerge. However, obtaining accurate lung segmentation remains a challenge. In general, lung segmentation algorithms are grouped into two categories: (1) traditional image processing algorithms and (2) artificial intelligence (AI) based algorithms. In traditional image processing, algorithms consider image value, shape, and spatial information to provide lung segmentation. However, most of these methods are computationally expensive, difficult to create representative training features. In recent years, artificial intelligence-based approaches, for example, machine learning and deep learning, have become popular. These methods offer better accuracy under certain ill-defined conditions.

Over the past decade, deep learning methods have been at the forefront of computer vision [12, 7], showing significant improvements over previous methodologies on visual understanding. They are generally becoming a universal solution for image processing applications, with fewer pre-processing steps, but provide better results than other algorithms. Besides classification and object detection tasks [18, 27], segmentation is another success story of deep learning methods in computer vision. These methods (e.g., the FCN model [21] or Unet model [5]) mention labeling the pixels in the image into several categories. Generally, these models consist of two paths: an encoder path to extract and capture the

image's features from a low to a high level and; a decoder path to expand and reconstruct the image features to provide the segmentation map.

In this work, we proposed another 3D architecture that was inspired by UNet model [5], named Coordinate-UNet 3D, for accurate lung segmentation from a 3D volume of CT scan of a NSCLC patient. Our model had the same structure as the original Unet model [5]. However, we have modified the core inside each level of the encoder and the decoder paths. Our model achieved an accurate segmentation (a Dice score of 0.99) based on a small set of 3D images of NSCLC patients. The highlighted points of our method are summarized as follows:

- The proposed model inputs lung CT volume (3D) and outputs the segmentation of lung parenchyma without any post-processing operation.

- The proposed model can provide the results in a short time.

- The proposed model is effective for lung segmentation. Its performance is very stable. It can also be applied to detect nodules in the CT scan of the lung.

This paper is organized as follows: The following section discusses the related works on lung segmentation approaches. Section 3 presents the methodology which details the Coordinate-UNet 3D architecture and the studied datasets. The experimental results will be shown in section 4, and some concluding remarks will be stated in section 5.

## 2 RELATED WORKS

In normal lung CT scans, it is easy to distinguish the lung parenchyma and the non-lung region because the intensities of these regions are very different. However, this task becomes a difficult task when considering abnormal CT images such as the CT scans of NSCLC patients. Therefore, lung segmentation could be a challenge in many studies about lung diseases [8, 35, 26].

As discussed, initial approaches for lung segmentation consider the characteristics of the image or based on the shape knowledge by using the traditional image processing technique such as thresholding based on the grey-level, adaptive thresholding technique, region growing, image registration [4] or the others [34]. However, these methods are computationally expensive, and it is difficult to generalize the learning features.

In medical imaging applications, Convolutional Neural Networks (CNNs) models can be used to analyze lung CT scans [2]. Their applications vary from detecting the lung pathology segmentation [2], or classifying the lung region [2, 37] to segment the lung volume [31].
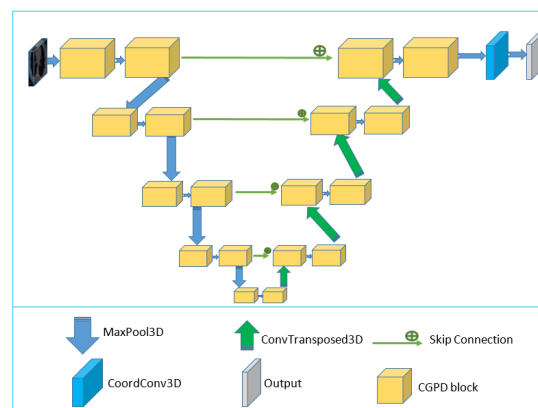


Figure 1: The architecture of the Coordinate-UNet 3D

In the approaches to provide whole volume of lung segmentation, most of applications have chosen to consider independently each slice of CT volume: Ravindra Patil et al. [28] and Brahim A. S. et al. [31] proposed different modifications on UNet model to provide the segmentation of slices in CT volume. Lei Geng et al. [10] presented a combination of VGG-16 [30] and dilated convolution to provide the lung segmentation. Swati P. Pawar et al. [29] have proposed a conditional generative adversarial network to encode the features from the slices of CT image. Then, the encoder and decoder were used to extract the multi-scales features and to give the lung segmentation, respectively. In whole 3D approach, Negahdar et al. [25] have proposed a volumetric segmentation network based on V-net [24] for 3D volumetric medical segmentation.

## 3 METHODOLOGY

In this section, we describe the architecture of the Coordinate-UNet 3D model and the hyper-parameters related to the training process. Then, the studied dataset used to train and validate Coordinate-UNet will be presented in detail.

### 3.1 Network architecture

In recent years, the UNet architecture [5] has been promoted as a good model for medical image segmentation. Generally, the UNet model [5] consists of two paths: a contraction path and an expansion path. The contraction path is like a classical CNN to extract low to high level features from an input representation. On the other hand, the expansive path consists of bottom-up sampling of the feature map followed by bottom-up convolution layers to reconstruct the features at different scales of the input. A concatenation is done between the convolution layer in the contraction path and the up-convolution of the expansion path to obtain more accurate labeling. At the end of the model, a convolution layer is used to map the features to the desired number of classes.

Depending on the requirements of the applications, the depth and layers at each level of the models are different. Figure 1 shows the structure of the Coordinate-UNet 3D model. Our model is inherited from the Unet model [5], but we have modified the UNet structure to adapt to our problem: (1) each level of contracting and expansive path contains two CGPD blocks (described in the next section); (2) the number of channels of the input is doubled before reducing the space to the next level; (3) a coordinate convolutional layer (CoordConv) [20] (instead of the classical convolution layer) is added at the end of the model to provide the segmentation map. Table 1 details the input/output images at each level of Coordinate-Unet 3D model.

Figure 2 illustrates the layers in a CGPD block. Each CGPD block consists of one Convolutional layer, one Group normalization (GN) layer [36], one PreLU activation function [13], and one Dropout layer [32]. Firstly, the GN is used instead of Batch normalization (BN) [14] because BN increases the error rapidly caused by inaccurate batch statistics estimation when the batch size becomes smaller; while GN divides the channels into groups and computes within each group the mean and variance for normalization. GN's computation is independent of batch sizes, and its accuracy is stable in a wide range of batch sizes. Second, the PreLU activation function [13] is used instead of ReLU activation function [23] to make the leakage coefficient a parameter that is learned along with the other network parameters. Finally, the Dropout layer is added to avoid over-fitting.
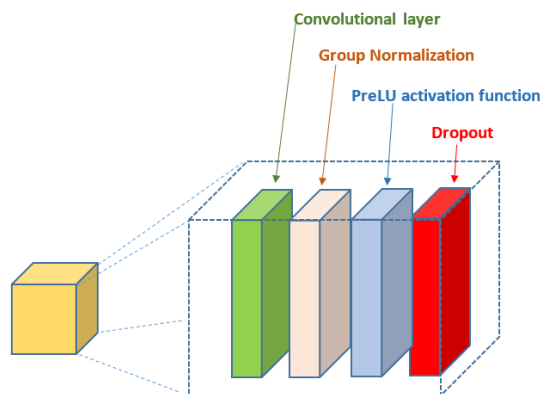


Figure 2: Layers in a CGPD block

As shown in the figure 1, each level of the contracting path includes two CGPD blocks, and the number of input channels is doubled at the second block before applying a max pooling layer to reduce the spatial size of the input and sending it to the next level of the model. At the expansive path, every level consists of an up-sampling layer that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, followed by two CGPD blocks to provide the information context of the feature map. At the end of the expansive path, a Coord-Conv layer (figure 3) is used to map the features to the desired number of classes.
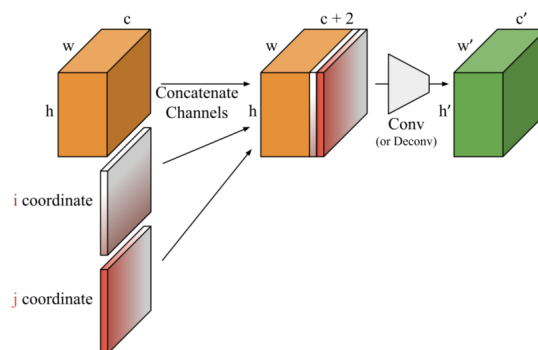


Figure 3: CoordConv layer [20]

As mentioned in [20], CoordConv is an extension of the standard convolution layer. It has the same functional signature as a convolution layer but accomplishes the mapping by first concatenating extra channels to the incoming representation. These channels contain hard-coded coordinates. The CoordConv layer keeps the properties of few parameters and efficient computation from convolutions but allows the network to learn to keep or to discard translation in variance as is needed for the task being learned. This is useful for coordinating transform based tasks where regular convolutions can fail. Because of the presentation of this layer, we give the name Coordinate-UNet 3D for our architecture.

## 3.2 Evaluation metrics

The evaluation of 3D lung parenchyma segmentation is based on the comparison between ground-truth and outputted segmentation from the model. For quantitative analysis of the predicted segmentation, several performance metrics are considered, including Dice score coefficient (DSC), F1-score, Jaccard similarity (IOU) and Matthews correlation coefficient (MCC) [3].

DSC is expressed as in Eq. 1 according to [6]. Here, GT and SP refer to the ground truth and predicted segmentation, respectively.

$$DSC = \frac{2*(|GT| \cap |SP|)}{|GT| + |SP|} \tag{1}$$

The IOU score is represented using Eq. 2 according to [15].

$$IOU = \frac{(|GT \cap SP|)}{|GT \cup SP|} \tag{2}$$

The MCC score is computed according to Eq. 3.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{3}$$

where TP, FP, TN, and FN are True Positive, False Positive, True Negative, and False Negative rates, respectively.

| Level | Input $(c \times w \times h \times d)$ | Output $(c \times w \times h \times d)$ | Level | Input $(c \times w \times h \times d)$ | Output $(c \times w \times h \times d)$ |
|---|---|---|---|---|---|
| Encoder path | | | Decoder path | | |
| Input | - | $1 \times 256 \times 256 \times 48$ | Output | $32 \times 256 \times 256 \times 48$ | $1 \times 256 \times 256 \times 48$ |
| Level 0 | $2 \times 256 \times 256 \times 48$ | $32 \times 256 \times 256 \times 48$ | Level 0 | $96 \times 256 \times 256 \times 48$ | $32 \times 256 \times 256 \times 48$ |
| Level 1 | $32 \times 128 \times 128 \times 24$ | $64 \times 128 \times 128 \times 24$ | Level 1 | $192 \times 256 \times 256 \times 48$ | $64 \times 128 \times 128 \times 24$ |
| Level 2 | $64 \times 64 \times 64 \times 12$ | $128 \times 64 \times 64 \times 12$ | Level 2 | $384 \times 64 \times 64 \times 12$ | $128 \times 64 \times 64 \times 12$ |
| Level 3 | $128 \times 32 \times 32 \times 6$ | $256 \times 32 \times 32 \times 6$ | Level 3 | $768 \times 32 \times 32 \times 6$ | $256 \times 32 \times 32 \times 6$ |
| Level 4 | $256 \times 16 \times 16 \times 3$ | $512 \times 16 \times 16 \times 3$ | Level 4 | $512 \times 16 \times 16 \times 3$ | $512 \times 16 \times 16 \times 3$ |

Table 1: The dimensions of input/output features at each level of proposed model.

### 3.3 Experimental data

Coordinate-UNet 3D was trained and validated on 3D CT images of NSCLC patients. Images were obtained from 2 sources: NSCLC-Radiomics-Interobserver1 dataset [16] consisting of 21 CT scans, and a local dataset consisting of 66 CT scans. Each image has a variable size from $(512 \times 512 \times 60)$ to $(512 \times 512 \times 600)$ pixels. All images were reformatted from standard DICOM to Neuroimaging Informatics Technology Initiative (NIfTI) format created by the National Institutes of Health [19]. The NIfTI file held the 3D image matrix and diverse metadata. Figure 4 shows a 3D image for volumetric measurements of lung parenchyma across three axes.
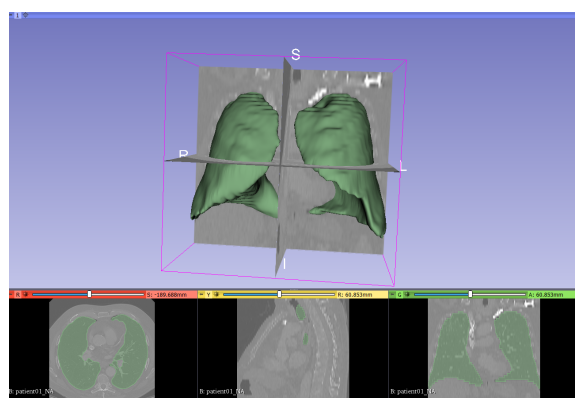


Figure 4: 3D volume of lung with XYZ axes.

*Data splitting and pre-processing:*

The images from 2 datasets were divided into two sets: train/validation and testing set. The train/validation set consists of 21 images from Interobserver1 dataset [16] and 41 images from the local dataset. The testing set consists of remaining 25 images of the local dataset. During the training process, the 62 images were divided into two sets corresponding to the training and validation process with a ratio of 0.8 : 0.2.

According the guideline of CT imaging, a CT scan consists of pixel spacing, axial slice thickness and view in the z axis with various scans. Therefore, the input images are uniformly pre-processed to minimize the vari-

ability within the database. In this work, the input images go through the three following steps for preparing the images:

1. The image intensity of each slice was first truncated in the range of [-1200, 600].

2. Z-normalization was performed on each slice of 3D image.

3. The CT scans were cropped to focus on lung region and converted to $256 \times 256 \times 48$ pixels.

*Data augmentation*

Due to the limitation of the number of samples in the training set, we used online and offline augmentation operations to increase the number of images in the training set. Offline augmentation means that we generate multiple augmented versions from an original image and add them to the dataset, while online augmentation operations are performed during the training process. In our work, we applied warping and flipping operations to generate two new versions from one image. Figure 5 shows an augmented example in our dataset. After the offline augmentation, we obtained 62 x 3 = 186 images for the training and validation processes. In addition to offline augmentation, online augmentation was performed like other segmentation approaches such as spatial flipping and image shearing.



Figure 5: A slice in a 3D studied image with its augmentation. From left to right: original slice, flipped slice, and deformed slice.

### 4 EXPERIMENTS AND RESULTS

Coordinate-UNet 3D was implemented on Pytorch-lightning [9]. The model has been trained in 500 epochs with Dice loss (Dice loss $= 1 - DSC$). The

Adam optimization [17] has been used with an initial learning rate of $10^{-4}$ and reduce to $10^{-6}$ by using cosine annealing schedule [22]. An early stopping has been employed to monitor the validation loss to avoid overfitting.

The proposed model was firstly trained in one fold with 80% of data and validated on 20% of data. The objective was to find the best hyperparameters for our model. After adjusting the hyperparameters, we trained the model on 5-fold cross-validation to ensure the stability of the model. Then, we computed the average of the scores outputted from 5 folds. Table 2 shows the average scores obtained on the validation set by using different kinds of final layer: CoordConv (the second row) and the traditional Convolutional layer (the third row). We see that the model can provide a very accurate lung parenchyma segmentation with both two kinds of Convolutional layers. Moreover, the outputted scores of CoordConv are higher than the traditional Convolutional layer. It provides a high Dice score, the other scores (IOU, F1, MCC) are smaller than Dice but remain in a high accuracy rank.

| Score | Dice index | IOU | F1 | MCC |
|---|---|---|---|---|
| CoordConv | 0.9907 | 0.9535 | 0.9761 | 0.9705 |
| Convolution | 0.9202 | 0.9383 | 0.9643 | 0.9236 |

Table 2: The performance of Coordinate-UNet model on validation set

These scores showed that the Coordinate-UNet 3D model has a good performance on the validation set. Then, the trained model was used to predict the segmentation of the testing images. The outputted segmentation was evaluated by calculating the dice score between prediction and the ground truth. Then, the average Dice score of 25 testing images was considered. We have obtained an average Dice score of 0.776 for all 25 images. Figure 6 and 7 illustrate the two examples in the testing set: one good and one worst predictions. From left to right, each one represents the ground truth and the segmentation generated by our network. The top row shows the segmentation of a slice in the image, the bottom row illustrates the 3D-segmentation.

Figure 8 shows the number of CT scans in several considered ranges of C-index scores. We see that our model worked effectively to provide good segmentation for most of the images in the test set (17 out of 25 CT scans obtain Dice scores greater than 0.75). Even in the image with the worst Dice score (figure 7), we see the segmentation of the lung parenchyma is correct, with the predicted segmentation producing only a minor error outside the lung region. This area really affected the dice score, but it was not that significant. Indeed, we want to see if the model can provide the segmentation of the lung parenchyma.
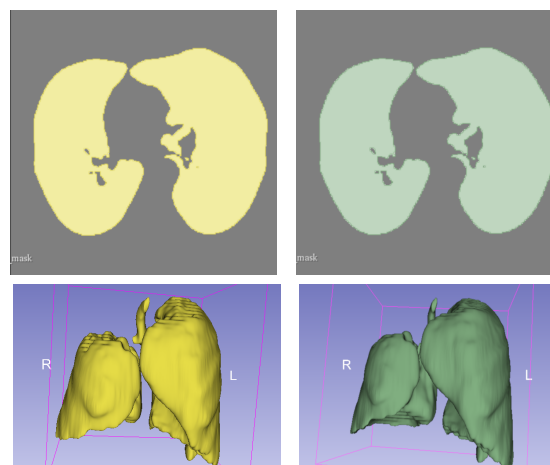


Figure 6: A good prediction in the test set (yellow = ground truth, green = prediction). Top: Segmentation of a slice in 3D image. Bottom: Lung parenchyma segmentation in 3D
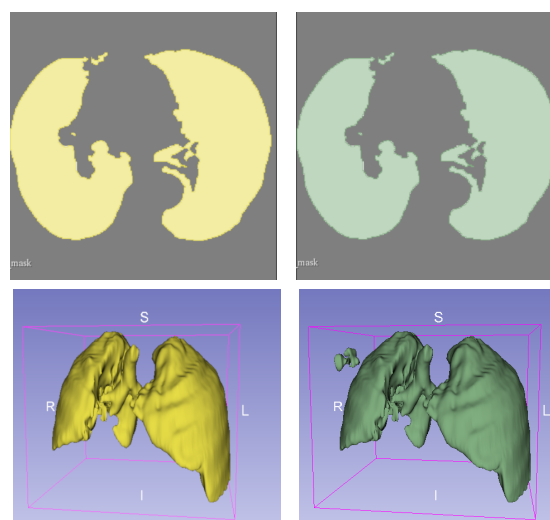


Figure 7: A worst prediction in the test set (yellow = ground truth, green = prediction). Top: Segmentation of a slice in 3D image. Bottom: Lung parenchyma segmentation in 3D

The performance of the model on the testing dataset is good, but this is far from the results on the validation set. The hypothesis is the difference between the training and testing datasets. It is worth noting that the model was trained on a combination of the CT scans from two datasets. After checking the data, we see that 21 images from the Interobserver1 dataset [16] do not have the bronchi, while the bronchi are present in the local dataset.

It is not fair to compare our results with other results because we are on different approaches [28, 31, 10]. However, Coordinate-UNet 3D model can segment lung parenchyma with very satisfactory performance and have the potential to locate and analyze lung lesions. It was employed to segment the lung parenchyma in other
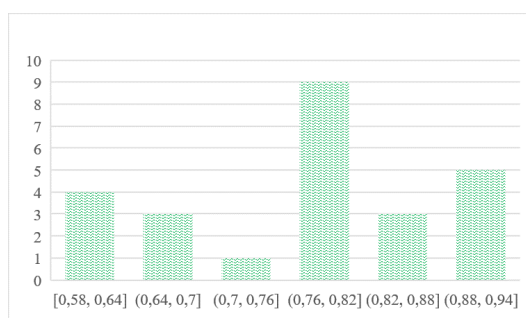
Figure 8: The number of predictions in each range of Dice score

datasets: NSCLC-Radiomics [1], which were used in another work to provide the segmentation of the tumor inside the lung.

## 5 CONCLUSION

In this work, we presented Coordinate-Unet 3D model to provide the lung segmentation. We have obtained an accurate segmentation with 0.99 Dice coefficient index for the validation set. However, the average Dice score on the test set has decreased a little bit. The problem has been found as a difference between the training and testing datasets. Basically, the prediction error is not so high, it can be solved by applying an algorithm to remove the small object in the segmentation map. The advantage of the method is the fact that can work with a whole 3D volume of the CT image and it can be applied to a wide area of different medical image segmentation task. Our objective is to generalize the approach to apply it to another task, for example, to perform lung tumor segmentation. This work has opened some questions that we can address in future work. (1) Concerning the data, we tried to reduce the bias between the images in the datasets, as well as limit the loss of data during the normalization process. (2) Analyze the effect of the presence of bronchi in the CT images by considering two groups of input images with and without the presence of bronchi. (3) This is an intermediate stage in our pipeline which provides the segmentation of NSCLC tumors, it is interesting to compare the results of the whole pipeline with other methods.

## 6 ACKNOWLEDGMENTS

## 7 REFERENCES

[1] Hugo JWL Aerts, Emmanuel Rios Velazquez, Ralph TH Leijenaar, Chintan Parmar, Patrick Grossmann, Sara Carvalho, Johan Bussink, René Monshouwer, Benjamin Haibe-Kains, Derek Rietveld, et al. Decoding tumour phenotype by non-invasive imaging using a quantitative radiomics approach. *Nature communications*, 5(1):4006, 2014.

[2] A Asuntha and Andy Srinivasan. Deep learning for lung cancer detection and classification. *Multimedia Tools and Applications*, 79:7731–7762, 2020.

[3] Pierre Baldi, Søren Brunak, Yves Chauvin, Claus AF Andersen, and Henrik Nielsen. Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*, 16(5):412–424, 2000.

[4] Kenneth R Castleman. *Digital image processing*. Prentice Hall Press, 1996.

[5] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, pages 424–432. Springer, 2016.

[6] Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[7] Andre Esteva, Katherine Chou, Serena Yeung, Nikhil Naik, Ali Madani, Ali Mottaghi, Yun Liu, Eric Topol, Jeff Dean, and Richard Socher. Deep learning-enabled medical computer vision. *NPJ digital medicine*, 4(1):5, 2021.

[8] David S Ettinger, Wallace Akerley, Gerold Bepler, Matthew G Blum, Andrew Chang, Richard T Cheney, Lucian R Chirieac, Thomas A D'Amico, Todd L Demmy, Apar Kishor P Ganti, et al. Non–small cell lung cancer. *Journal of the national comprehensive cancer network*, 8(7):740–801, 2010.

[9] William Falcon and The PyTorch Lightning team. PyTorch Lightning, 3 2019.

[10] Lei Geng, Siqi Zhang, Jun Tong, and Zhitao Xiao. Lung segmentation method with dilated convolution based on vgg-16 network. *Computer Assisted Surgery*, 24(sup2):27–33, 2019.

[11] Peter Goldstraw, David Ball, James R Jett, Thierry Le Chevalier, Eric Lim, Andrew G Nicholson, and Frances A Shepherd. Non-small-cell lung cancer. *The Lancet*, 378(9804):1727–1740, 2011.

[12] Mahmoud Hassaballah and Ali Ismail Awad. *Deep learning in computer vision: principles and applications*. CRC Press, 2020.

---

[1] https://www.plafrim.fr

[13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034, 2015.

[14] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. pmlr, 2015.

[15] Paul Jaccard. The distribution of the flora in the alpine zone. 1. *New phytologist*, 11(2):37–50, 1912.

[16] Petros Kalendralis et al. Fair-compliant clinical, radiomics and dicom metadata of rider, interobserver, lung1 and head-neck1 tcia collections. *Medical Physics*, 47(11):5931–5940, 2020.

[17] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[18] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.

[19] Michele Larobina and Loredana Murino. Medical image file formats. *Journal of digital imaging*, 27:200–206, 2014.

[20] Rosanne Liu, Joel Lehman, Piero Molino, Felipe Petroski Such, Eric Frank, Alex Sergeev, and Jason Yosinski. An intriguing failing of convolutional neural networks and the coordconv solution. *Advances in neural information processing systems*, 31, 2018.

[21] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[22] Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*, 2016.

[23] Andreas Maier, Christopher Syben, Tobias Lasser, and Christian Riess. A gentle introduction to deep learning in medical image processing. *Zeitschrift für Medizinische Physik*, 29(2):86–101, 2019.

[24] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571, 2016.

[25] Mohammadreza Negahdar, David Beymer, and Tanveer Syeda-Mahmood. Automated volumetric lung segmentation of thoracic ct images using fully convolutional neural network. In *Medical Imaging 2018: Computer-Aided Diagnosis*, volume 10575, pages 356–361. SPIE, 2018.

[26] Elise Noel-Savina and Renaud Descourt. Focus on treatment of lung carcinoid tumor. *OncoTargets and therapy*, pages 1533–1537, 2013.

[27] Ajeet Ram Pathak, Manjusha Pandey, and Siddharth Rautaray. Application of deep learning for object detection. *Procedia computer science*, 132:1706–1717, 2018.

[28] Ravindra Patil, Leonard Wee, and Andre Dekker. Auto segmentation of lung in non-small cell lung cancer using deep convolution neural network. In *Advances in Computing and Data Sciences: 4th International Conference, ICACDS 2020, Valletta, Malta, April 24–25, 2020, Revised Selected Papers 4*, pages 340–351. Springer, 2020.

[29] Swati P Pawar and Sanjay N Talbar. Lungseg-net: Lung field segmentation using generative adversarial network. *Biomedical Signal Processing and Control*, 64:102296, 2021.

[30] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[31] Brahim Ait Skourt, Abdelhamid El Hassani, and Aicha Majda. Lung ct image segmentation using deep neural networks. *Procedia Computer Science*, 127:109–113, 2018.

[32] Nitish Srivastava et al. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.

[33] Hyuna Sung et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 71(3):209–249, 2021.

[34] Selin Uzelaltinbulat and Buse Ugur. Lung tumor segmentation algorithm. *Procedia computer science*, 120:140–147, 2017.

[35] Jan P Van Meerbeeck, Dean A Fennell, and Dirk KM De Ruysscher. Small-cell lung cancer. *The Lancet*, 378(9804):1741–1755, 2011.

[36] Yuxin Wu and Kaiming He. Group normalization. In *Proceedings of the European conference on computer vision (ECCV)*, pages 3–19, 2018.

[37] Aleksandr Zotin, Yousif Hamad, Konstantin Simonov, and Mikhail Kurako. Lung boundary detection for chest x-ray images classification based on glcm and probabilistic neural networks. *Procedia Computer Science*, 159:1439–1448, 2019.