Recurrent Fully Convolutional DenseNet for Bronchiectasis Detection in CT Imaging

Victoria Mazo [1] **Itamar Tamir** [2] **Eyal Toledano** [1] **Eldad Elnekave** [1]

Abstract

Bronchiectasis, one of several features of Chronic Obstructive Pulmonary Disease (COPD), which is the 4th leading cause of death worldwide, represents an abnormal, often irregular widening of the bronchial airways. It is present in more than 25% of COPD patients and is associated with more severe obstruction and more frequent hospital admissions, therefore accurate bronchiectasis detection is useful in risk stratification and population health management. In this article we propose a Recurrent Fully Convolutional DenseNet architecture for a fully-automated detection of bronchiectasis in CT scans. Convolutional and recurrent layers are combined to learn expressive image representations exploiting the spatial dependencies across axial slices. The network is able to learn not only to locate airways, but also to estimate broncho-arterial ratios in order to decide which bronchi are abnormal. We show that performance of the network improves with increasing numbers of adjacent slices as input.

1. Introduction

In recent years, Deep Learning (DL) has emerged as a powerful approach to learning imaging representations directly from large volumes of data thus dispensing from the need to hand-engineer predictive features (Bengio et al., 2012; Hinton, 2007). Convolutional neural networks (CNNs) demonstarated the ability to learn hierarchically organised low to high-level features directly from raw images (Fukushima, 2007; Lecun et al., 1998) and yield state-of-the-art performance in 2D image classification (Szegedy et al., 2016), object detection (He et al., 2015) and semantic segmentation (Long et al., 2015; Yu & Koltun, 2016) tasks.

[1]Zebra Medical Vision, Shfaim, Israel, [2]Rabin Medical Center, Campus Golda, Israel. Correspondence to: Victoria Mazo <victoria.mazo@zebra-med.com>.

Current state-of-the-art segmentation approaches rely on Fully Convolutional Networks (FCNs), which are more efficient, compared to more traditional sliding-window approaches. They have received increasing interest lately as they unify object localization and segmentation in a single process by extracting both global and local context effectively (Long et al., 2015; Ronneberger et al., 2015). FCNs add upsampling layers to standard CNNs to recover the spatial resolution of the input at the output layer. As a consequence, FCNs can process images of arbitrary size. In order to compensate for the resolution loss induced by pooling layers, FCNs introduce skip connections between their downsampling and upsampling paths (Ronneberger et al., 2015), which help the upsampling path recover fine-grained information from the downsampling layers. For better filters reuse, additional skip connections might be added within blocks of layers (Huang et al., 2016), which improves accuracy of semantic segmentation, as shown for a Fully Convolutional DenseNet (FC-DenseNet) (Jegou et al., 2016).

Segmentation with FCNs has shown promising results on natural images where the objects are typically well-defined and sufficiently large compared to the entire image. Objects in medical images, on the other hand, often have hazy borders, varying size and are ill-defined. Nevertheless we believe DL has excellent potential in this field as well. FCNs have been successfully applied to the challenges of medical imaging segmentation as diverse as the identification of neuronal structures in electron microscopic recordings with U-Net (Ronneberger et al., 2015), multi-slice MRI cardiac segmentation with Recurrent U-Net (Poudel et al., 2016) and Ground Glass Detection with Fully Convolutional DenseNet Network (Author, 2017), but are still missing for many important challenges in Medical Imaging, such as bronchiectasis detection in Chest CT scans (see Fig. 1).

Bronchiectasis is defined to be a localized irreversible dilation of bronchial walls and is diagnosed using computed tomography (CT). It can occur in a solitary segmental bronchus or may be seen diffusely, as is often the case in certain genetic syndromes, such as Cystic Fibrosis. In normal subjects, diameter of the bronchi is smaller or similar

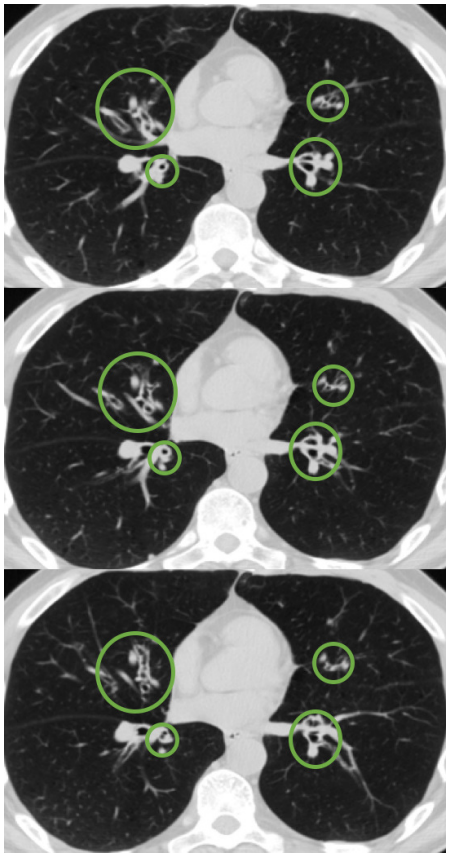

Figure 1. Example of bronchiectasis on three adjacent slices.

to the adjacent artery size, broncho-arterial ratio is around 0.65-0.70, but if the ratio is greater than 1, it is believed to be bronchiectasis.

More than 25% of all patients carrying a diagnosis of Chronic Obstructive Pulmonary Disease (COPD), which is the 4th leading cause of death worldwide (British Lung Foundation, 2007; Gao & Prasad, 2013) and expected to be the 3rd leading cause of death worldwide by 2020, surpassed only by heart disease and stroke (Sin et al., 2006), have CT evidence of bronchiectasis (Obrien et al., 2000). The presence of bronchiectasis in COPD patients is associated with more severe obstruction and more frequent hospital admissions (Martinez-Garcia et al., 2011) and therefore its accurate detection is useful in risk stratification and population health management.

In the last decade two main approaches for automatic bronchiectasis detection have been developed. One technique computes a tree model of the airways and their accompanying arteries followed by automated measurements of broncho-arterial ratios within the lungs (Odry et al., 2008; Kiraly et al., 2008; Yu et al., 2016). Another technique relies upon training a classifier on extracted texture and/or shape features from images to learn discriminative properties of bronchiectasis (Arunkumar, 2012; Elizabeth et al., 2009). These and other conventional computer vision methodologies generally rely upon shape and other basic image qualities as derived from a limited (usually < 200 images) training sample (Linying et al., 2017). This relatively narrow scope of feature identification and limited incorporation of anatomic and physiologic context results in overfitting and hindered applicability of the first generation CAD in clinical practice.

In this work we set out to mimic the reading process carried out by radiologists, who would typically explore a slice together with its neighbor slices from a CT scan stack and draw inferences about the likely presence of abnormal airways by picking up changes in radiological appearance between adjacent slices. To this end we aimed to apply (for the first time to the best of our knowledge) Deep Learning techniques to the task of bronchiectasis detection on CT scans of the chest.

Our hypothesis is that learning imaging features that capture inter-slice spatial correlations can introduce performance gains in bronchiectasis detection. For this purpose we investigate a novel Recurrent Fully Convolutional DenseNet (RFC-DenseNet) architecture, trained end-to-end, that learns to detect and segment abnormal bronchi on axial slices taking into account spatial dependencies across the slices. As an illustrative example of spatial dependencies, Fig. 1 shows three adjacent cross-sectional slices extracted from a Chest CT scan with highlighted areas of abnormal bronchi (Fig. 2 shows examples of true and predicted bronchiectasis segmentation). RFC-DenseNet is a recurrent version of the FC-DenseNet, which has recently showed a superior performance both in semantic segmentation (Jegou et al., 2016) and in medical image segmentation (Author, 2017). It combines the representational power of CNNs with the capability to learn dependencies in sequential data that is typical of recurrent neural networks (Hochreiter & Schmidhuber, 1997).

It is important to emphasize that, since the bronchial tree, similar to all branching anatomical structures, progressively tapers in circumference and as it extends within the lung from the central main bronchi to the peripheral bronchioli, it follows the characterization of an abnormally dilated airway must take into account both the focal appearance of the bronchus as well as its global relationship to the lung in general. The challenge for a network is thus to learn from small positive patches within large lung regions while maintaining adequate levels of sensitivity.

At the slice level, the objective is to binary classify the presence of abnormal bronchi for an input axial CT slice. This is achieved through segmentation, by thresholding minimal

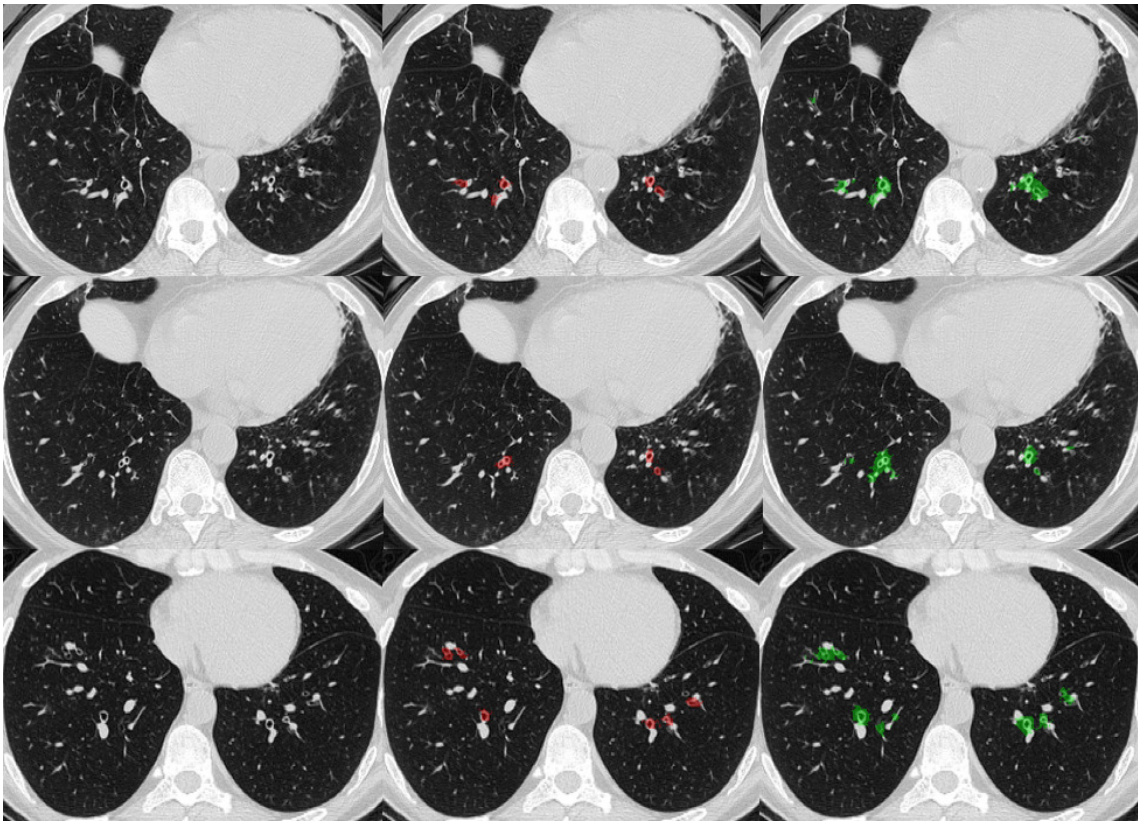

Figure 2. Examples of bronchiectasis segmentation on three adjacent slices. *Left* - axial CT slice, *middle* - segmentation by a radiologist, *right* - predicted segmentation by RFC-DenseNet.

number of positive pixels, in order slice be considered positive. In this way we reduce the amount of training data necessary for achieving good classification performance, and gain the ability to track back network's decisions. By comparing performance of the Recurrent FC-DenseNet to its non-recurrent version, we observed improvement in both segmentation (Dice score) and accuracy per slice while adjacent slices being taken into account.

2. Dataset

We assembled a dataset of 175 individual CT chest studies which had been ordered for a variety of respiratory related clinical indications. Out of them 134 CTs contained bronchiectasis and 46 were of healthy individuals (see Table 1). In CT scans with bronchiectasis every axial slice with abnormal bronchi signs was manually segmented by an expert radiologist (in total 3650 slices were segmented). All healthy studies were verified by a radiologist not to contain bronchiectasis, although they might have had contained small lung nodules ($<$ 3mm). To the best of our knowledge, this represents the largest dataset of Chest CT images to serve as a training dataset for bronchiectasis detection.

The dataset contains only lung slices (on average, 98 slices per patient), which were detected using automatic lungs segmentation (non-lung regions have *not* been masked out). We chose only one series per patient. The dataset includes bronchiectasis of all types: cylindrical, varicose and cystic. It is worth mentioning that we did not exclude studies on the basis of data acquisition protocols, e.g. inspiration/expiration imaging or technical acquisition variations, despite that such variations could alter the CT appearance of the lung tissue. In allowing for data variations we aimed to develop an algorithm geared toward broader clinical application.

3. Recurrent Fully Convolutional DenseNet Architecture

The main idea underlying FCNs is extension of a contracting path, in which a sequence of pooling operators progressively reduces the size of the network, by adding successive

Table 1. Number of CTs scans and slices in training/validation/test subsets.

SUBSET	CT SCANS WITH BRONCHIECTASIS			HEALTHY CT SCANS	
	CTS	SEGMENTED SLICES	ALL SLICES	CTS	ALL SLICES
TRAIN	99	2687	6853	16	3053
VALIDATION	15	490	1087	15	2468
TEST	15	473	1082	15	2666

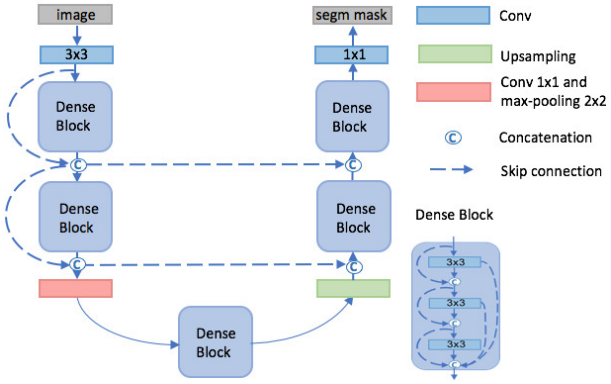

Figure 3. Architecture of a Fully Convolutional DenseNet.

layers where pooling operators are replaced by upsampling operators. In a Fully Convolutional DenseNet (Jegou et al., 2016) the expanding path is characterized by a large number of feature channels, allowing the network to propagate context information to higher resolution layers. In order to localize, high resolution features from the contracting path are combined with the upsampled output. A successive convolution layer can then learn to assemble a more precise output based on this information.

FC-DenseNet consists of $2n + 1$ dense blocks of k_i 3x3 convolution-dropout-batch normalization-ReLU sandwich layers, where $i = 1, ..., 2n + 1$. Within a block, each layer is connected to all the following layers via a concatenation with the skip connections, which is shown in Fig. 3. At the beginning of the contracting path there is an additional 3x3 convolutional layer, in continuation 1x1 convolution-dropout-batch normalization-ReLU sandwich layer and 2x2 max pooling with stride 2 follows every block, and at the end a bottleneck dense block. FC-DenseNet has additional skip connections from the block input to the block output in the contracting path. Every step in the upsampling path consists of a 2x2 transposed convolution with stride $\frac{1}{2}$, a concatenation with the correspondingly cropped feature map from the contracting path and a dense block. At the end of the expanding path, a 1x1 con-

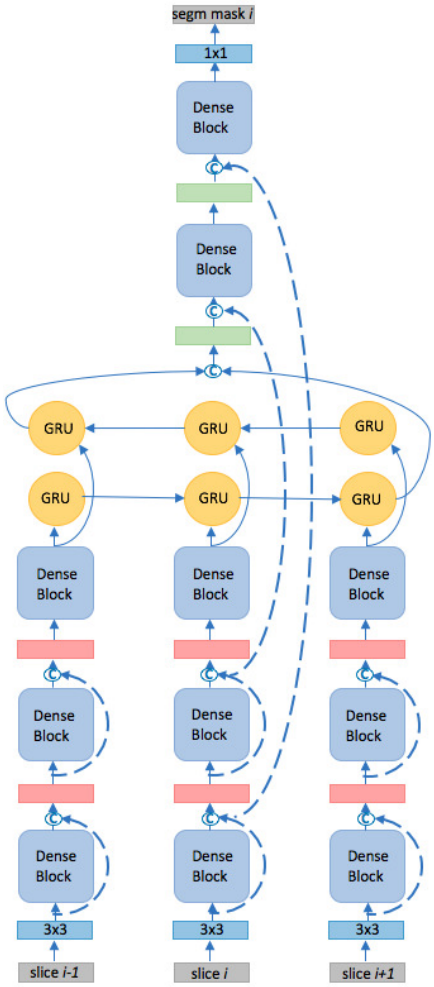

Figure 4. Architecture of a Recurrent Fully Convolutional DenseNet (with 2 adjacent slices).

Table 2. Bronchiectasis segmentation and binary classification results (per slice).

MODEL	# ADJACENT SLICES	ACCURACY	SPECIFICITY	SENSITIVITY	DICE SCORE
FC-DENSENET	0	90.7	94.6	63.6	28.4
RFC-DENSENET	2	91.2	94.8	66.4	37.6
RFC-DENSENET	4	91.5	95.7	62.8	37.8
RFC-DENSENET	6	**91.8**	95.7	64.9	**38.7**

volution is applied followed by a softmax, which results in a probability map of the same dimensions as the input image. Weighted cross-entropy loss function is applied on the probability map. Number of filters in all 3x3 convolutional layers, except the first one (48 filters), is the same and denoted as g (growth rate). The last 1x1 has number of channels equals the number of classes and in all other 1x1 convolutional layers (including transposed convolution) number of filters equals number of input channels.

Recurrent FC-DenseNet receives as an input slice i and its k nearest neighbor slices, and outputs a segmentation mask for the slice i (see Fig. 4). All the $k+1$ slices are feeded into the contracting path of FC-DenseNet (including the bottleneck block). The resulting $k+1$ feature maps are flattened and feeded into a bidirectional GRU (Chung et al., 2014). Dimensions of the GRU layers equal the dimensions of the input. Outputs from the forward and backward directions are concatenated and reshaped into the initial feature map dimensions. The network flow continues with the FC-DenseNet upsampling path and outputs a segmentation map for slice i. Skip connections are only between the middle slice i in the contraction path and the expanding path.

4. Experimental Results

We performed experiments on FC-DenseNet and Recurrent FC-DenseNet with 56 convolutional layers. Block structure of FC-DenseNet is $\{4\ 4\ 4\ 4\ 4\ 4\ 4\ 4\ 4\ 4\ 4\}$ (with $g = 12$), where an element in the parenthesis defines number of layers in a block. Structure of the Recurrent FC-DenseNet is analogous but with a bidirectional GRU in the middle, since all convolutional layers in the contracting path including the bottleneck are reused.

Accuracy, sensitivity and specificity are typical statistical measures to evaluate performance of a binary classification test. They are defined as

$$\begin{aligned} Accuracy &= \frac{TP+TN}{N} \\ Specificity &= \frac{TN}{TN+FP} \\ Sensitivity &= \frac{TP}{TP+FN}, \end{aligned} \tag{1}$$

where TP is the number of true positives, TN of true negatives, FP of false positives, FN of false negatives and N is the number of pixels in a slice.

We measure goodness of segmentation with the help of Dice score (DSC) averaged over all positive slices:

$$DSC = 2\frac{TP}{2TP+FP+FN}. \tag{2}$$

The Dice score not only considers how many positives are found, but also penalizes for false positives. A pixel is considered positive if its probability is higher than 0.5 and a slice is considered positive if number of predicted positive pixels is higher than a threshold based upon the best accuracy and sensitivity received on a validation set.

Results are summarized in Table 2. Both the best accuracy and Dice score were achieved on Recurrent FC-DenseNet with the largest number of adjacent slices (six), 91.8% and 37.8 respectively. Examples of segmentation are shown in Fig. 2. Since most of the abnormal airways go through several slices, looking at adjacent slices above and below helps the network to locate an airway and make a better decision whether it is abnormal or not. We notice that the greatest jump in performance is from zero to two adjacent slices, while increase in accuracy and Dice score slows down between higher number of neighbor slices, meaning that adding a slice above and below the target slice contributes more new information to the network than looking at further slices. RFC-DenseNet shows 4.3% false positive rate (1-specificity), which is relatively low; on the other hand, sensitivity is lower than specificity. This means that the network errs in 4.3% of cases that a slice is healthy and identifies correctly a slice with bronchiectasis in 64.9% of all test cases. We anticipate improvements on sensitivity in future work.

Image intensity was cropped within the window [-1050, 250] in HU. No preprocessing was performed on slices. For augmentation we used random vertical flip and random crop of 55-95% of the original image area, which was resized to 256x256. Cross-entropy loss function was optimized using RMSProp. The initial learning rate was set to 10^{-4} and it is halved every 14 epochs during the training.

We trained the networks for 15-40 epochs (with batch sizes of 4-10) and it took between 12 to 30 hours, depending on a model, using a dual NVIDIA TitanX GPU card. Our code is based on Tensorflow.

5. Conclusions

In this paper we presented the first end-to-end approach for bronchiectasis detection using Deep Learning. We proposed a novel Recurrent Fully Convolutional DenseNet architecture, which is effectively 2.5-dimensional. We explored several variations of the network with varying number of adjacent slices included in input and conclude that performance of the network improves with the number of adjacent slices. The best accuracy, 91.8%, was achieved with the maximal number of adjacent slices - six, which is encouraging, considering the complexity of bronchiectasis detection. The network learned not only to locate airways, but also to estimate broncho-arterial ratios in order to decide which bronchi are abnormal. Since neural networks are capable to learn well and generalize, we have not excluded portions of the images that are not expected to present abnormal airways, and, indeed, we did not see a significant number of false posivites in these regions.

References

Arunkumar, R. Quantitative Analysis of Bronchiectasis Using Local Binary Pattern and Fuzzy Based Spatial Proximity. *International Conference on Recent Trends in Information Technology*, pp. 72–76, 2012.

Author, N. N. Suppressed for anonymity, 2017.

Bengio, Y., Courville, A. C., and Vincent, P. Unsupervised Feature Learning and Deep Learning: A Review and New Perspectives. *arXiv:1206.5538 [cs.CV]*, 2012.

British Lung Foundation. Invisible lives: Chronic Obstructive Pulmonary Disease (COPD) finding the missing millions, 2007. URL http://www.thehealthwell.info/node/22715.

Chung, J., Gülçehre, Ç., Cho, K. H., and Bengio, Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv:1412.3555 [cs.CV]*, 2014.

Elizabeth, D. Shiloah, Kannan, A., and Nehemiah, H. Khanna. Computer–aided Diagnosis System for the Detection of Bronchiectasis in Chest Computed Tomography Images. *International Journal of Imaging Systems and Technology*, 19(4):290–298, 12 2009.

Fukushima, K. Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. *Biological Cybernetics*, 36(4):193–202, 2007.

Gao, J. and Prasad, N. Chronic obstructive pulmonary disease in China: the potential role of indacaterol. *Journal of Thoracic Disease*, 5(4):549–558, 2013.

He, K., Zhang, X., Ren, S., and Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *arXiv:1502.01852 [cs.CV]*, 2015.

Hinton, G. E. Learning Multiple Layers of Representation. *Trends in Cognitive Sciences*, 11(10):428–434, 2007.

Hochreiter, S. and Schmidhuber, J. Long Short-Term Memory. *Neural Comput.*, 9(8):1735–1780, 1997.

Huang, G., Liu, Z., and Weinberger, K. Q. Densely Connected Convolutional Networks. *arXiv:1608.06993 [cs.CV]*, 2016.

Jegou, S., Drozdzal, M., Vazquez, D., Romero, A., and Bengio, Y. The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation. *arXiv:1611.09326 [cs.CV]*, 2016.

Kiraly, P. A., Odry, B., Godoy, M. C. B., Geiger, B., Novak, C. L., and Naidich, D. Computer-aided Diagnosis of the Airways: Beyond Nodule Detection. *Journal of Thoracic Imaging*, 23(2):105–113, 2008.

Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. Gradient-based Learning Applied to Document Recognition. In *Proceedings of the IEEE*, pp. 2278–2324, 1998.

Linying, L., X., L., Chunwu, Z., Xinming, Z., and Yanfeng, Z. A Review of Ground Glass Opacity Detection Methods in Lung CT Images. *Current Medical Imaging Reviews*, 13(1):20–31, 2017.

Long, J., Shelhamer, E., and Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *In: CVPR*, 2015.

Martinez-Garcia, M. A., Soler-Cataluna, J. J., Sanz, Y. D., Serra, P. C., Lerma, M. A., Vicente, J. B., and Perpia-Tordera, M. Factors Associated with Bronchiectasis in Batients with COPD. *Chest*, 140(5):1130–1137, 2011.

Obrien, C., Guest, P. J., Hill, S. L., and Stockley, R. A. Physiological and Radiological Characterisation of Patients Diagnosed with Chronic Obstructive Pulmonary Disease in Primary Care. *Thorax*, 55(8):635–642, 2000.

Odry, B. L., Kiraly, A. P., Novak, C. L., Naidich, D. P., and Lerallut, J. F. An evaluation of Automated Broncho-arterial Ratios for Reliable Assessment of Bronchiectasis. In *Medical Imaging 2008: Computer-Aided Diagnosis*, volume 6915, pp. 69152M, 2008.

Poudel, R. P. K., Lamata, P., and Montana, G. Recurrent Fully Convolutional Neural Networks for Multi-slice MRI Cardiac Segmentation. *arXiv:1608.03974 [cs.CV]*, 2016.

Ronneberger, O., Fischer, P., and Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *MICCAI*, 9351:234–241, 2015.

Sin, D. D., Anthonisen, N. R., Soriano, J. B., and Agusti, A. G. Mortality in COPD: Role of Comorbidities. *Eur. Respir. Journal*, 28(6):1245–1257, 2006.

Szegedy, C., Ioffe, S., and Vanhoucke, V. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *arXiv:1602.07261 [cs.CV]*, 2016.

Yu, F. and Koltun, V. Multi-Scale Context Aggregation by Dilated Convolutions. *arXiv:1511.07122 [cs.CV]*, 2016.

Yu, N., Li, H., Wu, B., Li, Y., Deng, L., and Guo, Y. Computerized Identification of Bronchiectasis Using a 3D Quantitative CT Protocol. *Journal of Medical Imaging and Health Informatics*, 6(5), 2016.