



# HHS Public Access

Author manuscript

*Proc IEEE Int Symp Biomed Imaging*. Author manuscript; available in PMC 2018 November 16.

Published in final edited form as:

*Proc IEEE Int Symp Biomed Imaging*. 2018 April ; 2018: 273–276. doi:10.1109/ISBI.2018.8363572.

## EMPHYSEMA QUANTIFICATION ON SIMULATED X-RAYS THROUGH DEEP LEARNING TECHNIQUES

Mónica Iturrioz Campo<sup>\*†</sup>, Javier Pascau<sup>†</sup>, and Raúl San José Estépar<sup>\*</sup>

<sup>\*</sup>Applied Chest Imaging Laboratory, Brigham and Women's Hospital, Boston, MA, USA

<sup>†</sup>Dept. de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, Spain

### Abstract

Emphysema quantification techniques rely on the use of CT scans, but they are rarely used in the diagnosis and management of patients with COPD; X-ray films are the preferred method to do this. However, this diagnosis method is very controversial, as there are not established guidelines to define the disease, sensitivity is low, and quantification cannot be done. We developed a quantification method based on a CNN, capable of predicting the emphysema percentage of a patient based on an X-ray image. We used real CT scans to simulate X-ray films and to calculate emphysema percentage using the LAA%. The model developed was able to calculate emphysema percentage with an LAA% mean error of 3.96, and it obtained an AUC accuracy of 90.73% for an emphysema definition of 10%, with a mean sensitivity of 85.68%, significantly improving X-ray-based emphysema diagnosis.

### Keywords

X-ray; emphysema quantification; convolutional neural network; regression; COPD

## 1 INTRODUCTION

Emphysema is defined as a condition of the lung characterized by abnormal, permanent enlargement of airspaces distal to the terminal bronchiole, accompanied by the destruction of their walls, and without obvious fibrosis [1]. Emphysema is one of the two conditions that make up COPD (chronic obstructive pulmonary disease), along with chronic bronchitis. COPD is currently the third leading cause of death in the United States [2], and the fourth one worldwide as of 2015 [3]. COPD (and emphysema) cannot be cured; the treatment consists of relieving symptoms and stopping the advance of the disease [4]. Because of this, an early diagnosis is key to ensure the quality of life of patients who suffer from COPD is not decreased.

Current diagnosis techniques include clinical examination, pulmonary functionality tests, chest X-rays and eventually CT scans. The former two are used to confirm or deny the pathological presence of emphysema (relying on the expression of symptoms), whereas imaging techniques also allow for an assessment of the degree of the disease as well as a morphological description (as radiograph represent structure rather than function).

CT scans allow for the quantification of the extent of emphysema using percentage of low-attenuation lung areas (LAA%) below a given threshold [13]. This technique has been shown to correlate well with pathology and can be used to define different degrees of emphysema according to the percentage of total lung volume occupied by emphysema; minimal or no emphysema: <0.5, mild: 5% and <10%, moderate or severe: >10% [5]. Despite the ability of CT scans to diagnose emphysema, CT scanning is not typically used in the initial diagnosis and management of COPD patients.

Although X-ray is routinely used as imaging technique to diagnose chest diseases, emphysema diagnosing on X-rays is a controversial topic. Many studies have been performed to define guidelines that can relate the observable radiological signs of emphysema with emphysema severity [6, 7, 8]. However, a consensus has not been attained as different observers have different success rates when assessing the severity of emphysema in X-ray films, and each observer uses different radiological criteria to reach their conclusions [9, 10]. All studies agree that X-ray-based emphysema diagnosis has a high inter-observer variation when relating radiological signs with emphysema severity with sensitivity ranging from 68% to 77% [11, 12]. Typically, emphysema is diagnosed when 2 or more of the radiological signs of emphysema are present, and severity is a personal assessment of the observer.

Our goal is to develop a method to quantify emphysema on X-rays to obtain a score with diagnostic and clinic relevance. As a proof of concept, we theorize that using a convolutional neural network on a database formed by simulated X-ray films obtained from CT scans with an associated emphysema percentage value; it will be able to correctly associate a numerical score to the images, aiding with early diagnosis. The main purpose of this project is to develop a method capable of diagnosing and quantifying emphysema, in all its degrees, based on X-ray images.

## 2 MATERIAL AND METHODS

### 2.1 Emphysema scoring

Emphysema score was calculated from the CT using the LAA% approach [13]. First, we segmented the lungs using the approach provided in the Chest Imaging Platform ([www.chestimagingplatform.org](http://www.chestimagingplatform.org)). LAA% was computed as the percentage of lung voxels below -950HU.

### 2.2 X-ray generation

We simulated X-ray projects from CT using a parallel projection model, a reasonable approximation for chest examination as the X-ray source is far from the subject. CT units were projected according to the transformation

$$X_P(x_i, y_j) = \frac{1}{N} \sum_{k=1}^N e^{\beta \frac{\max(X_{CT}(x_i, y_j, z_k), -1024) + 1024}{1000}}$$

where  $X_{CT}$  is the CT scan after removal of the CT table and  $X_P$  is the simulated X-ray. The CT table was removed by segmenting the body using an Otsu thresholding and retaining the largest connected component with hole filling.  $\beta$  controls the boosting of X-ray absorption as the tissue density increases. We have chosen  $\beta=0.85$  for our experiments after performing a visual comparison with real chest X-rays. Figure 1 shows three cases with different degrees of emphysema and the corresponding simulated X-ray.

### 2.3 Database

The data used consists of 7,377 subjects with X-ray simulations of the chest generated on both coronal and sagittal planes, and the corresponding emphysema percentage (LAA%) computed on the CT scans. Each image is composed of 512 by 600 grayscale pixels. For this project, we only used the coronal projection as it conveys information from both lungs. We transformed our images by squaring and then normalizing pixel values with the z-score.

### 2.4 Emphysema percentage regression CNN

We construct an 11-layer convolutional neural network with a total number of 1,189,313 parameters. It is composed of 2 sets of 2 convolutional layers separated by a pooling layer, followed by one dropout layer and three fully connected layers, plus the output layer. We use the mean squared error function as the loss function and Adam as the optimizer with a learning rate of 0.001 and a decay of 0.0001. We train the model on 2,705 images (of which 800 were used for validation) for 119 epochs and a batch size of 32. Because of the reduced size of our training data, data augmentation is used (small rotations, vertical flip, and small horizontal and vertical shifts).

## 3 RESULTS

We use our model to predict the emphysema percentage of the remaining 4,671 images not used for training. When comparing the predictions with the actual known emphysema percentage values of the test images, we obtain a mean error of 3.96 and a standard deviation of 4.60. Checking the percentiles, we observe that 76% of the predictions have an error smaller than 5; furthermore, 90% of the predicted values have an error equal to or smaller than 9.62. By plotting the error histogram, we can observe that most errors lie in the range of 0–2.5.

To better assess the overall performance of our test and its validity, we performed a Bland-Altman analysis (Fig. 4), using the known values of emphysema percentage of our test images as our reference method, and the values predicted by our CNN as the alternative. The mean difference is centered around zero (0.2944) which tells us that our model does not have a systematic error. The 95% CI is 11.8803. Our model works remarkably well when predicting emphysema percentages smaller than 10% (the data is caged between the confidence intervals). However, when predicting higher emphysema percentage values, our model tends to underestimate the measurement.

To study how our model performs for the diagnosis of emphysema, we computed ROC curves (Fig. 5) corresponding to different emphysema percentage thresholds, and calculated the AUC as well as the sensitivity and sensibility of all of them to test the prediction

accuracy (Table 1). By obtaining the Youden's J statistic, we can obtain the highest sensitivity-specificity pair for every curve. By increasing the disease threshold (changing the pathological definition of emphysema), our model obtains increasingly better results. At a 5% emphysema threshold, our model has a mean sensitivity of 78.36% and a mean specificity of 76.90%, and it is 85.91% accurate. By increasing the diagnostic criteria for emphysema to 10%, we obtain that our model has a mean sensitivity of 85.68% and a mean specificity of 80.42%, and it is 90.73% accurate.

Table 1 shows the different values obtained at different emphysema thresholds. We can observe that specificity differences across disease threshold levels are statistically insignificant, with an average value of 80%. Sensitivity values increase significantly when the definition of emphysema becomes laxer (higher threshold), as we would expect (more emphysema, easier to diagnose). By increasing the threshold by a 7%, the mean sensitivity increases by 17.72%.

Finally, to get an understanding on how the CNN is analyzing the X-ray simulations, we retrieve the Grad-CAM (Gradient-weighted Class Activation Map) [14] of the output layer of three images with different values of emphysema percentage (1.18%, 10.27%, and 20.73%). Activation maps allow us to see where in the image the net found relevant characteristics, which tells us which parts of the image the CNN uses to predict the emphysema percentage. We can observe that the network pays special attention to the diaphragm and the lungs, ignoring osseous structures.

#### 4. DISCUSSION AND CONCLUSION

The regression convolutional neural network built for this work was able to correctly predict emphysema percentage based on simulated X-ray images. It achieved an error smaller than ten on 90% of our data, yielding a mean error of 3.96 overall. Visual interpretation of X-rays has an accuracy of 77%, and low sensitivity when diagnosing mild to moderate emphysema [12]. Our model was able to make predictions with 85.39% accuracy and reached a maximal mean sensitivity of 78.36% when studying emphysema under the condition that a positive emphysema diagnostic is given at a 5% emphysema percentage threshold, significantly improving the X-ray visual diagnostic technique. By slightly increasing the definition of emphysema to a 10% emphysema percentage, our network improved the obtained results by obtaining an accuracy of 90.73% and a maximal mean sensitivity of 85.65%. However, our results show that even though our model can correctly predict emphysema percentages below 10%, it tends to underestimate measurements above this threshold. Although our model needs to be improved concerning this aspect, it is important to notice that it obtains very accurate scores of small emphysema percentages (mild disease), which are the values that classic image-based diagnosis methods are unable to recognize [12]. When analyzing the activation maps generated by our model, we observed that the main focus of the CNN was the lungs and the diaphragm, and no attention was paid to osseous structures. These findings make sense with the pathological localization of the disease, which affects the lungs [1]. Diaphragm position is very frequently used as a marker for emphysema by doctors, as it tends to flatten and be positioned lower on the chest as the disease advances, making the lungs appear longer [9, 8]. Interestingly, the network appears to learn those features. In cases

with more diffuse emphysema, the prediction is driven by activations within the lung region (Fig. 6 middle). Our results prove to be of significance both at a clinical and diagnostic levels.

While it is true that there exist methods for emphysema quantification on CT scans with similar or even better results than our CNN, the main advantage our model presents is that it relies on X-ray films, which are more frequently used than CT due to its shorter time of image obtainment, lower radiation dose delivered to the patient, and lower monetary cost [10], as well as being part of the follow up of patients with COPD. Furthermore, beyond the assessment of COPD, an emphysema score-based system could be used for lung cancer risk stratification, as it has been identified as a mayor risk factor [15].

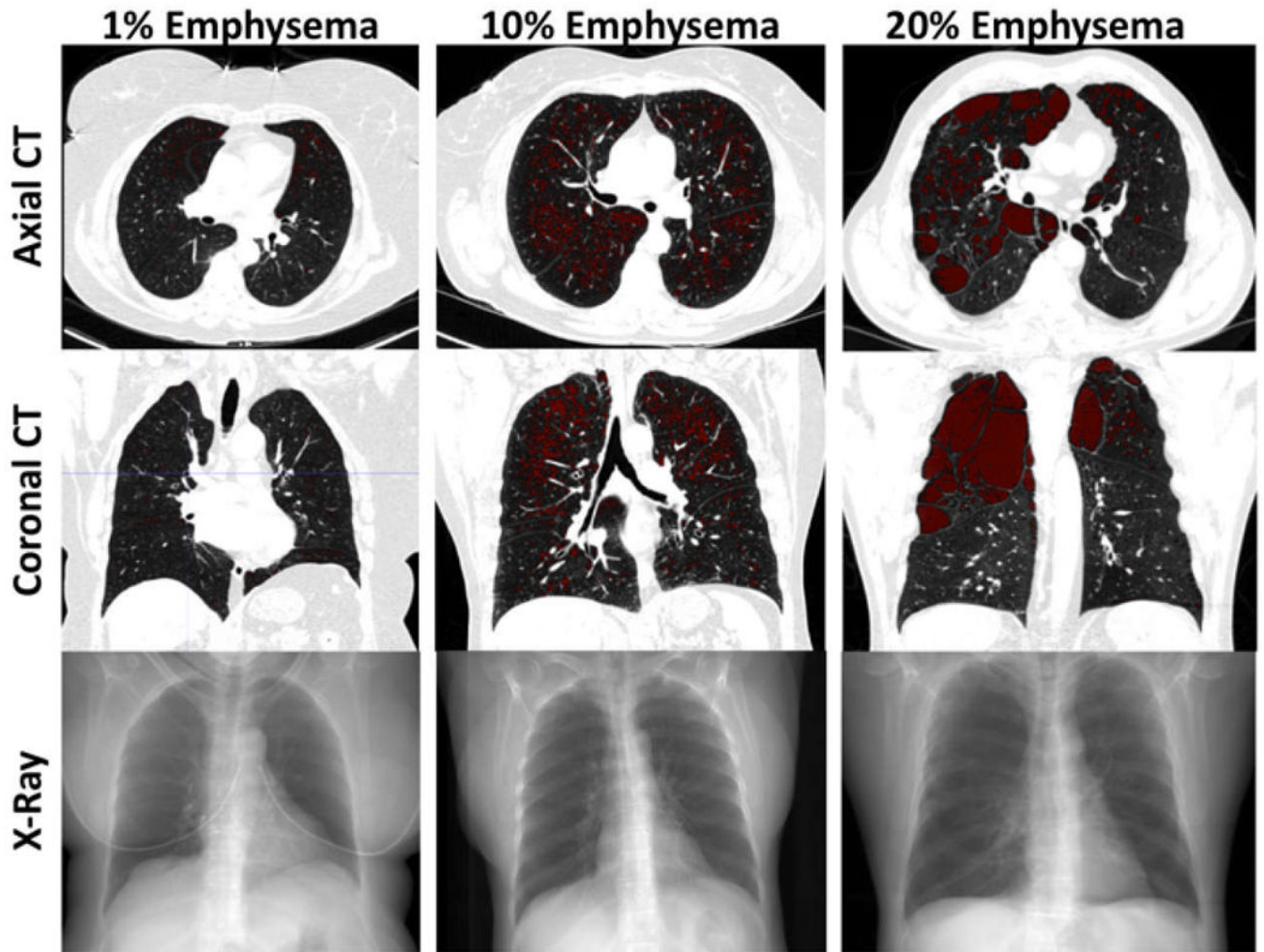
The model developed for this project has room for improvement, and it should not be taken as a functional diagnostic tool. Having a small training dataset (1905 images) seemed a limitation to our problem; however, when we changed the proportions of the training and testing sets so that each contained half of the total data available, our results didn't significantly improve. When studying the activation maps of our CNN, we observed that besides focusing its attention to the lung area, the CNN seemed to pay interest to the corners of the image, where emphysema is never present. This problem could be solved by refining the parameters of the network. Given the good results obtained by our model, which was trained on a database composed of simulations of X-rays obtained from real CT scans, the next step for this study is to be performed with real data. Nevertheless, it serves as a proof of concept –X-ray simulations are a generative model, the ground truth, that can be used to compare with real X-ray films.

## REFERENCES

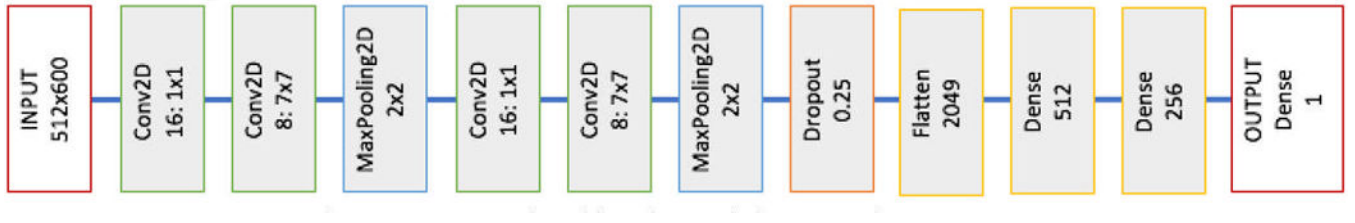
- [1]. Snider GL, Kleinerman J, Thurlbeck WM, & Bengali ZH, "The definition of emphysema: report of a national heart, lung, and blood institute workshop", *AM rev Respoir Dis*, vol. 132, pp. 182–5, 1985
- [2]. Xu J, Murphy SL, Kochanek KD, & Arias E, "Mortality in the United States", *NCHS data brief*, no. 267, 2016.
- [3]. The top 10 causes of death. (n.d.). Retrieved October 13, 2017, from <http://www.who.int/mediacentre/factsheets/fs310/en/>
- [4]. Celli B et al., "Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper", *European Respiratory Journal*, vol. 23, no. 6, pp. 932–946, 2004 [PubMed: 15219010]
- [5]. Kinsella M, Müller NL, Abboud RT, Morrison NJ, & Dybuncio A, "Quantitation of Emphysema by Computed Tomography Using a "Density Mask" Program and Correlation with Pulmonary Function Tests", *Chest*, vol. 97, no. 2, pp. 315–321, 1990 [PubMed: 2298057]
- [6]. Sutinen S, Christoforidis AJ, Klugh GA, & Pratt PC, "Roentgenologic Criteria for the Recognition of Nonsymptomatic Pulmonary Emphysema. Correlation between Roentgenologic Findings and Pulmonary Pathology", *American Review of Respiratory Disease*, vol. 91, no. 1, pp. 69–76, 1965 [PubMed: 14260002]
- [7]. Reid L, & Millard F, "Correlation between radiological diagnosis and structural lung changes in emphysema", *Clinical Radiology*, vol. 15, no. 4, pp. 307–311, 1964 [PubMed: 14213174]
- [8]. Fraser RG, & Thurlbeck WM, "The Accuracy of the Roentgenologic Diagnosis of Chronic Pulmonary Emphysema", *Investigative Radiology*, vol. 2, no. 6, pp. 379–380, 1967
- [9]. Thurlbeck WM, "Emphysema Then and Now", *Canadian Respiratory Journal*, vol. 1, no. 1, pp. 21–39, 1994

- [10]. Miniati M et al., “Value of chest radiography in phenotyping chronic obstructive pulmonary disease”, *European Respiratory Journal*, vol. 31, no. 3, pp. 509–515, 2008 [PubMed: 18057056]
- [11]. Sanders C, Nath PH, & Bailey WC, “Detection of Emphysema with Computed Tomography”, *Investigative Radiology*, vol. 23, no. 4, pp. 262–266, 1988 [PubMed: 3372190]
- [12]. Choroma ska A, & Macura KJ, “Role of computed tomography in quantitative assessment of emphysema”, *Polish Journal of Radiology*, vol. 77, no. 1, pp. 28–36, 2012 [PubMed: 22802863]
- [13]. Müller NL, Staples CA, Miller RR, Abboud RT, ““Density Mask”: An Objective Method to Quantitate Emphysema Using Computed Tomography”, *Chest*, vol. 94, no. 4, pp. 782–787, 1988 [PubMed: 3168574]
- [14]. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, & Batra D, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization”, *arxiv.org* 1610.02391, v. 3, 2016
- [15]. Li Y, Swensen SJ, Karabekmez LG, Marks RS, Stoddard SM, Jiang R, ... Yang P, “Effect of Emphysema on Lung Cancer Risk in Smokers: A Computed Tomography-Based Assessment”, *Cancer Prevention Research*, vol. 4, no. 1, pp. 43–50, 2010 [PubMed: 21119049]



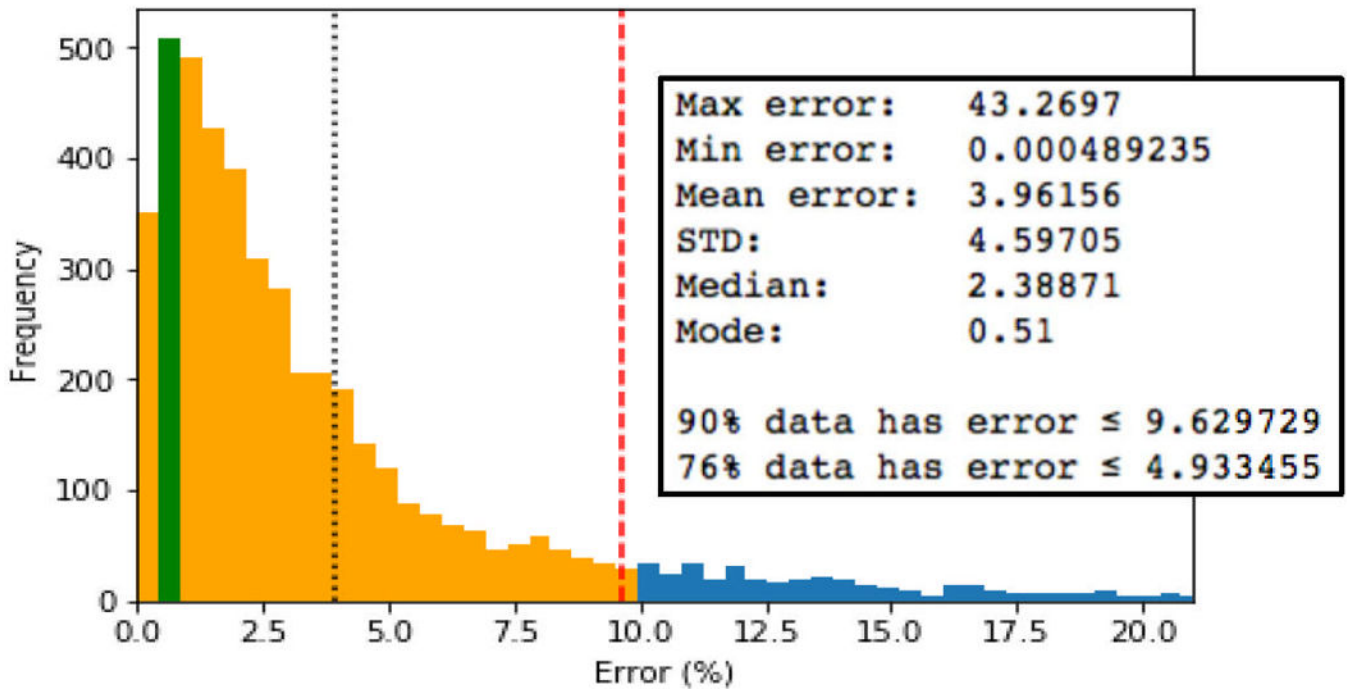


**Fig 1.** X-ray simulations (bottom) obtained from the CT scans (top and middle) of three different patients with significantly different emphysema percentages (mapped in red in the CT scans): 1.18% (left) 10.27% (middle) and 20.73% (right).



**Fig 2.**  
Regression CNN built for this project.





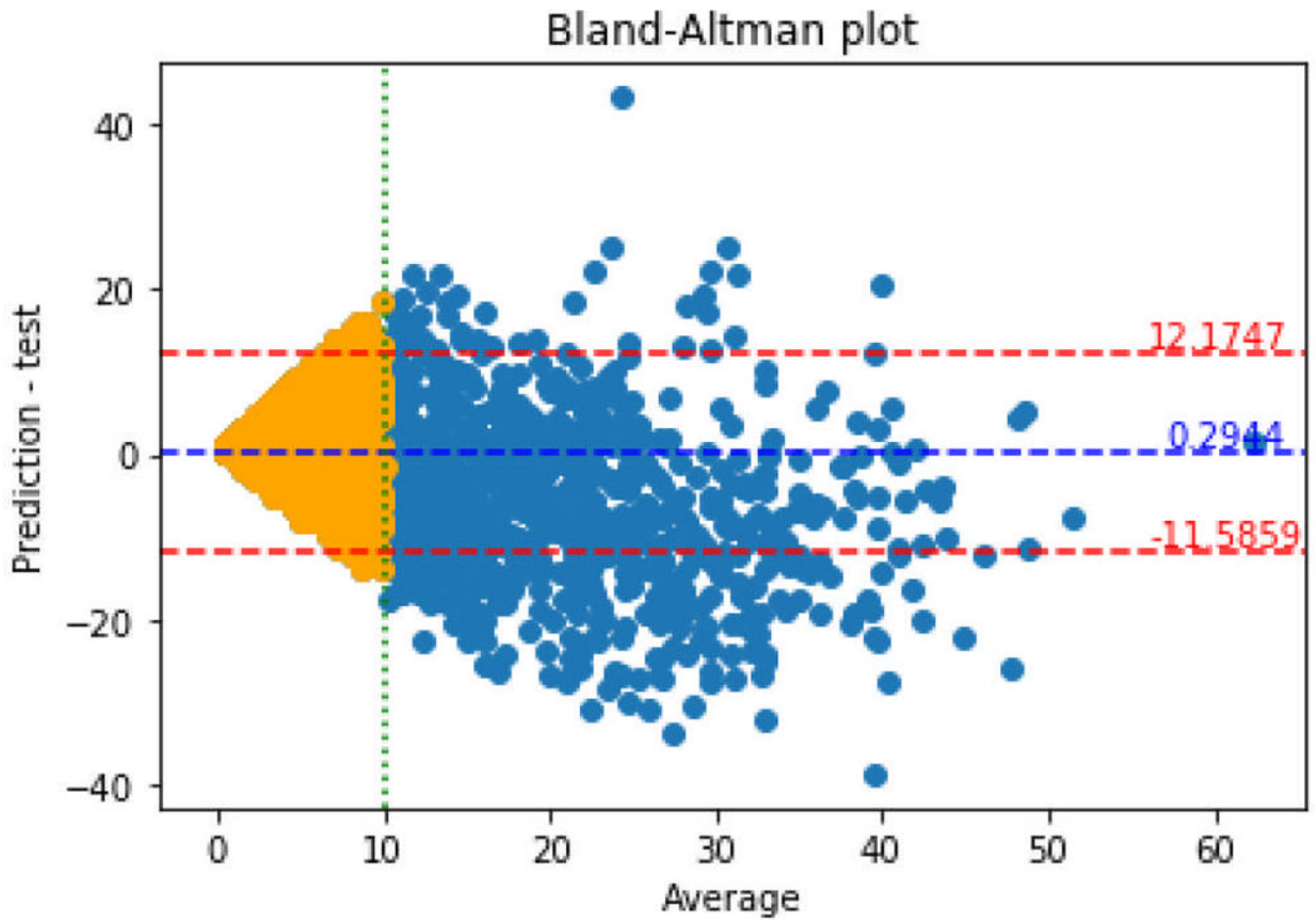
**Fig 3.** Error histogram. The vertical red line marks the 90-percentile and dotted black line the mean. The blue bins correspond to error values above the 90-percentile, and the green bin marks the mode.

Author Manuscript

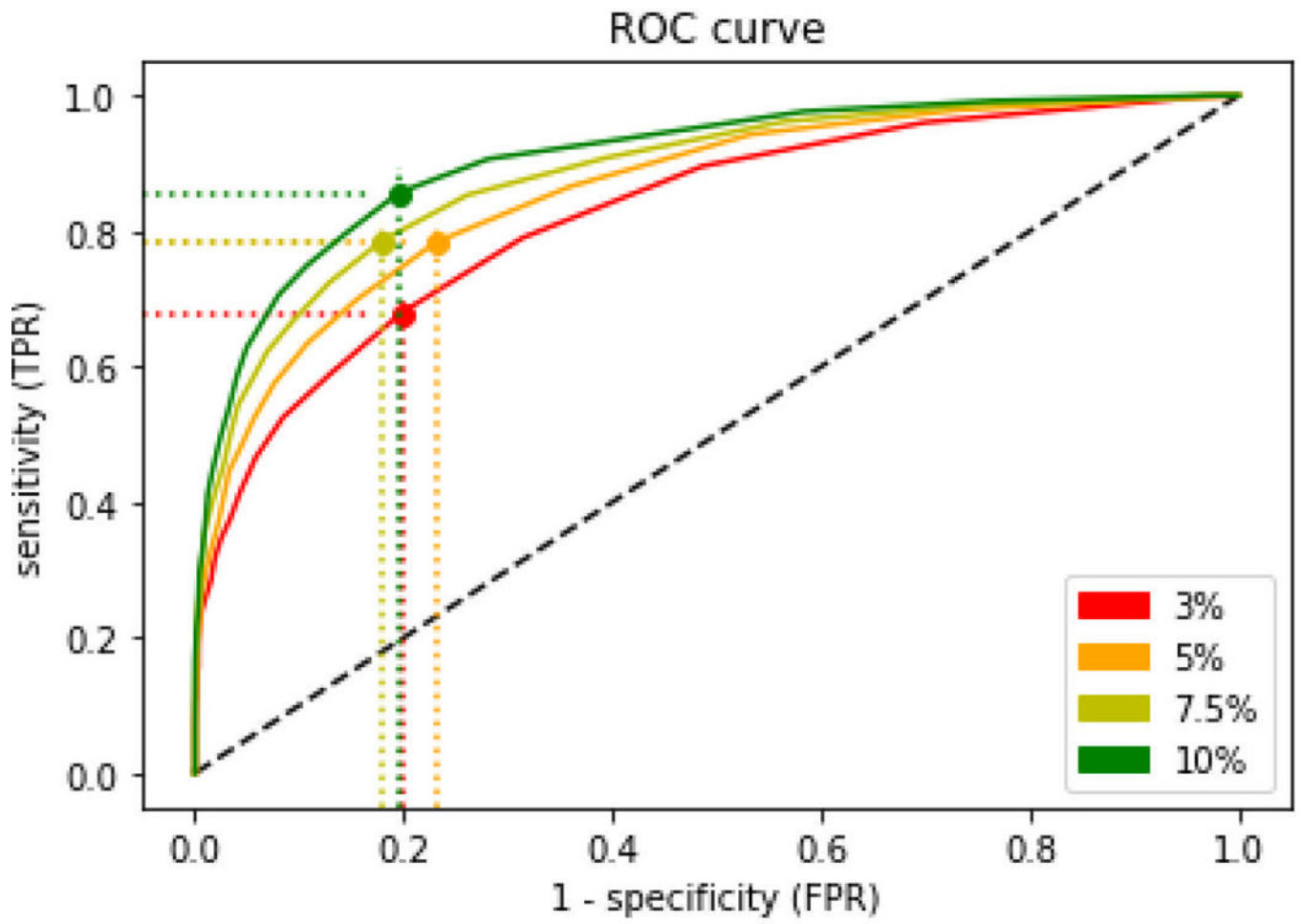
Author Manuscript

Author Manuscript

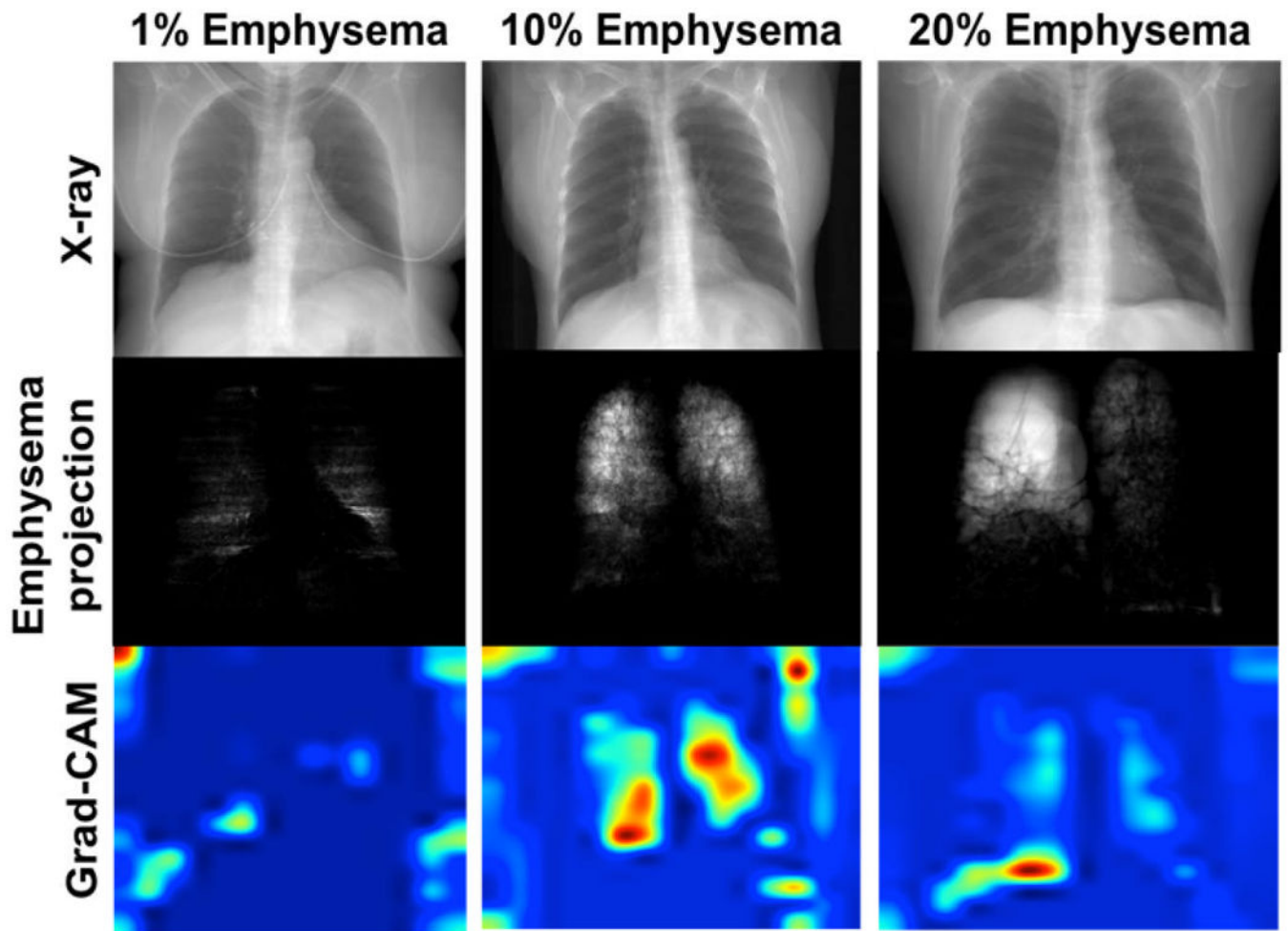
Author Manuscript



**Fig 4.** Bland-Altman plot of our predictions vs. the actual values. The blue horizontal line marks the mean difference (0.2944) and the red horizontal lines the upper and lower limits of the 95%CI (12.1747 and -11.5859). The orange dots show the measurement differences for emphysema percentage values lower than 10%.



**Fig 5.** ROC curves for different diagnostic criteria for emphysema. The dots mark the point in the curve where the Youden's J statistic is the highest.



**Fig 6.** Grad-CAM activation maps (bottom) obtained from X-ray simulations (top), along with the cumulative projection of emphysema voxels detected on CT (middle) of the same patients as Fig. 1 with different emphysema percentages: 1.18% (left) 10.27% (middle) and 20.73% (right).

**Table 1.**

Diagnostic performance of our model for different diagnostic criteria for emphysema.

Disease thresh- old (%emphy- sema)	Sensitivity [95%CI]	Specificity [95%CI]	AUC (%)
3	67.96 [65.89 – 70.03]	80.18 [78.68 – 81.68]	82.48
5	78.36 [76.22 – 80.50]	76.90 [75.45 – 78.35]	85.91
7.5	78.56 [76.13 – 80.99]	82.10 [80.84 – 83.36]	88.39
10	85.68 [83.37 – 87.99]	80.42 [79.16 – 81.68]	90.73

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript