

# Localized energy-based normalization of medical images: application to chest radiography

R.H.H.M. Philipsen\*, P. Maduskar, L. Hogeweg, J. Melendez, C.I. Sánchez and B. van Ginneken, *Member, IEEE*

**Abstract**—Automated quantitative analysis systems for medical images often lack the capability to successfully process images from multiple sources. Normalization of such images prior to further analysis is a possible solution to this limitation. This work presents a general method to normalize medical images and thoroughly investigates its effectiveness for chest radiography (CXR). The method starts with an energy decomposition of the image in different bands. Next, each band's localized energy is scaled to a reference value and the image is reconstructed. We investigate iterative and local application of this technique. The normalization is applied iteratively to the lung fields on six datasets from different sources, each comprising 50 normal CXRs and 50 abnormal CXRs. The method is evaluated in three supervised computer-aided detection tasks related to CXR analysis and compared to two reference normalization methods. In the first task, automatic lung segmentation, the average Jaccard overlap significantly increased from  $0.72 \pm 0.30$  and  $0.87 \pm 0.11$  for both reference methods to  $0.89 \pm 0.09$  ( $p < 0.01$ ) with normalization. The second experiment was aimed at segmentation of the clavicles. The reference methods had an average Jaccard index of  $0.57 \pm 0.26$  and  $0.53 \pm 0.26$ ; with normalization this significantly increased to  $0.68 \pm 0.23$  ( $p < 0.01$ ). The third experiment was detection of tuberculosis related abnormalities in the lung fields. The average area under the Receiver Operating Curve increased significantly from  $0.72 \pm 0.14$  and  $0.79 \pm 0.06$  using the reference methods to  $0.85 \pm 0.05$  ( $p < 0.01$ ) with normalization. We conclude that the normalization can be successfully applied in chest radiography and makes supervised systems more generally applicable to data from different sources.

**Index Terms**—Normalization, chest radiography, energy, CAD

## I. INTRODUCTION

THE automated interpretation and quantitative analysis of medical images has gained a lot of attention the last decades. Many techniques depend on the image's absolute intensity values to perform image interpretation; for example for the measurement of emphysema scores from CT scans [1], or the segmentation of anatomical structures from chest X-rays (CXR) [2], [3]. However, it is well-known that the appearance and statistical properties of medical images are influenced by acquisition settings, scanner technology and often proprietary—

This work was supported by the European and Developing Countries Clinical Trials Partnership (EDCTP), the Evaluation of multiple novel and emerging technologies for TB diagnosis, in smear-negative and HIV-infected persons, in high burden countries (TB-NEAT) project. Asterisks indicates corresponding author.

\*R.H.H.M. Philipsen is with the Diagnostic Image Analysis Group, Department of Radiology and Nuclear Medicine, Radboud University Medical Center, 6500 HB Nijmegen, The Netherlands (e-mail: rick.philipsen@radboudumc.nl).

All authors are with the Diagnostic Image Analysis Group, Department of Radiology and Nuclear Medicine, Radboud University Medical Center, 6500 HB Nijmegen, The Netherlands

post-processing techniques. These changes can have a huge impact on the performance of common image processing techniques, such as image retrieval, image segmentation or computer-aided detection (CAD). Many methods only work well on test data that is similar to the training data (and this is typically the case in experimental validation in the scientific literature); a decrease in performance is often noticed when the method is applied to data not included in or not closely resembling the training data. Different approaches have been proposed to make algorithms robust to those changes, such as the use of large training sets with large variation, the use of invariant image features [4]–[6], the transformation of feature vectors [7], or the use of an image normalization technique. All solutions have their own drawbacks. Firstly, different training sets for each new scanner successfully solves the problem but is not feasible for a larger number of machines. It implies that for each scanner a new annotated training database has to be created which is time-consuming and laborious. A related solution, with the same drawbacks, would be a sufficiently large training set that contains images from multiple sources representing a large variability in CXR appearances. For example, [8] used two different data sources each with a different training set. Secondly, some image features are invariant to certain changes in images. For example, features derived from gradient orientations or isophote curvature are invariant to any monotonic gray level transformation. In [4], a set of differential gray-value invariant features is proposed and showed some applications on artificial data. [5] and [6] proposed an algorithm invariant to gray-scale transformations by pre-processing with a histogram equalization gray-scale transformation. In [9] a set of local binary pattern features is described that is gray-scale invariant. A more recent paper [10], proposes a differential Radon transform based on hidden Markov models. All in all, the number of proposed invariant features is limited and these features are only invariant to some classes of transformation. This limits the applicability of such an approach. A final alternative would be to apply a suitable feature vector transformation [7] such as feature selection, feature extraction by PCA or whitening, or non-linear transformations. Considering these drawbacks, we propose an image normalization based approach in the current work.

Image normalization is a technique that changes the intensity values in an image in order to resemble the characteristics of a pre-selected set of reference data. Common approaches are linearly re-scaling of the image intensities to zero mean and unit variance, histogram matching or histogram equalization. However, for complex images like medical images, more elaborate image normalization techniques are needed.

In literature, several papers in different medical areas have proposed normalization solutions. For example, in mammography, Kallenberg et al. [11] successfully applied a normalization technique in a CAD system trained with digitized films and applied it to full-field digital mammography data. The normalized images were modeled using a non-linear characteristic curve describing the relation between optical density and x-ray exposure. In pathology, Macenko et al. [12] proposed a method to normalize histology slides using singular value decomposition and histogram scaling. Other intensity standardization examples can be found in the field of magnetic resonance imaging [13]–[15]. These studies calculate a set of landmarks in the image's histogram which are subsequently used to calculate a mapping to the landmarks of a reference histogram. However, these techniques are bound to the specific characteristics of the target modalities.

In this work we focus on digital chest radiography, the most commonly used imaging modality to find pulmonary abnormalities. It is widely used for the detection and screening of tuberculosis (TB). TB remains a major global health problem in spite of the availability of cheap treatment. In 2012, approximately 8.6 million people developed TB and another 1.3 million people died from it, and most of these deaths were entirely preventable. A key factor in reducing TB mortality is early detection of TB cases [16]. Digital chest radiography is an important tool in many of the screening algorithms recommended by the World Health Organization [17]. It is cheap, fast and previous studies suggest that it has high levels of sensitivity if the examination is standardised and a coding system is used [18]–[20]. In this light, CAD software is being developed for the automatic detection of TB [21]–[24]. Our supervised software is able to do a fully automated, objective and reproducible analysis of a digital CXR within one minute [21], [22]. However, the system does not perform well if applied to data from scanners which the system is not trained with.

Although radiographic enhancements have already received some attention [25]–[27], to the best of our knowledge, CXR normalization has been addressed only once before [28]. However, this technique standardizes pixel intensities based on the intensities of the mediastinum and does not seem ideal. Furthermore, current supervised systems in literature usually avoid the problem by adhering to one or two image sources and use a separate training set for each source [2], [8], [29]–[35]. In [36] histogram equalization is used for image standardization.

In this paper, we propose a general framework for medical image normalization based on localized energy standardization. The technique is based on the decomposition of images in energy bands and the standardization of each energy band with respect to a reference. Dividing the image in energy bands allows the technique to isolate different structures in different bands and apply a different normalization to each of them. This makes our normalization not a function that maps each intensity value to a fixed corresponding different value. Additionally, the proposed technique can be restricted to a localized region of interest in the image. This prevents that intensity values in irrelevant areas of the images, for example the projection of regions of air outside the body, affect the

normalization within a region of interest. The effectiveness of this technique is evaluated on three different automated tasks in CXR interpretation: lung segmentation, clavicle segmentation and the detection of textural abnormalities. An initial experiment, using simple energy normalization tested on a small dataset evaluated using only the lung segmentation system, was presented in [37].

The different data sets used in this work are described in Section II. Section III and IV provide a detailed explanation of the method and the evaluation procedure, respectively. In Section V, the experiments and results are shown which are discussed in Section VI. Section VII concludes the proposed work.

## II. DATA

Six datasets with 100 posterior-anterior (PA) CXRs each were used. The CXRs were acquired from five different sources: four sources with 100 CXRs and one source with 200 CXRs. From the latter, two different datasets were distinguished: 100 CXRs with high kVp settings, and 100 with low kVp settings to simulate the effect of acquisition parameter variation. Taking into account other settings that affect dose such as tube current, collimation and exposure time, the latter uses, on average, a higher dose and contains less noise. Details can be found in Table I. Each dataset contained 50 normal CXRs and 50 CXRs containing one or more abnormalities. Four datasets (OdelcaDR, Atomex, Philips, OdelcaDR-kV) contained abnormalities consistent with active TB. The Siemens set contained abnormalities encountered in a Western hospital in daily clinical practice and the abnormal CXRs in the JSRT dataset contained a single pulmonary nodule. Before any further processing was done, all images were rescaled to a width of 1024 pixels. Fig. 1 shows one normal and one abnormal example image from each source, depicted with the original DICOM window center/width settings.

## III. METHODS

In this section, the proposed localized energy-based image normalization algorithm is described.

### A. Localized energy-based image normalization

Let  $I(\mathbf{x})$  be the intensity value of image  $I$  at position  $\mathbf{x}$  where  $\mathbf{x}$  can correspond to a pixel in a 2D image, i.e.  $\mathbf{x} = (x, y)$ , or a voxel in a 3D image, i.e.  $\mathbf{x} = (x, y, z)$ . The image  $I$  can be decomposed in  $B$  energy bands  $\{I_i, i = 1, \dots, B\}$  as follows:

$$I_i(\mathbf{x}) = I_{i-1}(\mathbf{x}) - L_{i-1}(\mathbf{x}; \sigma_{i-1}) \quad (1)$$

where  $I_0(\mathbf{x}) = I(\mathbf{x})$ ,  $L_i(\mathbf{x}; \sigma)$  is defined as

$$L_i(\mathbf{x}; \sigma) = I_i(\mathbf{x}) * G(\mathbf{x}; \sigma) \quad (2)$$

and  $G(\mathbf{x}; \sigma)$  is a Gaussian kernel with scale parameter  $\sigma$  and  $*$  denotes the convolution operation. The scale parameter is set as  $\sigma^i = 2^i$  so each energy band contains structures at different scales [40]. This decomposition method is simple,



Fig. 1: Two images from each of the six datasets are displayed. The top row contains normal CXRs and the bottom row contains abnormal CXRs. From left to right the images come from: OdelcaDR, Atomed, Philips, Siemens, OdelcaDR-kV and JSRT. Except for the JSRT images, all images are shown using original DICOM window center/width settings.

TABLE I: Characteristics of the six different datasets.

Dataset	Name	Scanner	Detector	kVp	Source
A.	OdelcaDR	OdelcaDR, Delft Imaging Systems, The Netherlands	Slotscan	130	Symptomatic patients from a TB clinic in Zambia
B.	Atomed	Atomed mobile X-ray, Delft Imaging Systems, The Netherlands	Canon CXDI	120	TB prevalence survey in the Gambia
C.	Philips	DigitalDiagnost, Philips Medical Systems, The Netherlands	Pixium 4600	90	Find&Treat screening program in the UK [38]
D.	Siemens	THORAX/MULTIX FD, Siemens, Germany	Siemens FD-X	125	Clinical practice patients in an academic hospital in The Netherlands
E.	OdelcaDR-kV	OdelcaDR, Delft Imaging Systems, The Netherlands	Slotscan	100	Symptomatic patients from a TB clinic in Zambia
F.	JSRT	Unknown	Unknown	-	Digitized films from the JSRT database [39]

and as the multidimensional Gaussian filter is separable, also very fast. When larger image data is used, one could decide to further optimize the algorithm by using a decimated frequency decomposition [41].

Let  $\Omega \subseteq I$  be a region of interest defined inside  $I$ . This region can correspond to the whole image, an anatomical structure, such as lung fields in CXRs, or a specific geometric region. We define the energy value  $e_i(I, \Omega)$  of the energy band  $I_i$  of  $I$  inside  $\Omega$  as

$$e_i(I, \Omega) = [\frac{1}{N-1} \sum_{\mathbf{x} \in \Omega} (I_i(\mathbf{x}) - m_i(I, \Omega))^2]^{\frac{1}{2}} \quad (3)$$

with  $m_i(I, \Omega) = \frac{1}{N-1} \sum_{\mathbf{x} \in \Omega} I_i(\mathbf{x})$  and  $N = \sum_{\mathbf{x} \in \Omega} 1$ . This energy measure corresponds to the standard deviation of the intensity values inside the region of interest and summarizes the underlying brightness dispersion. A different measure of energy, such as power spectrum density, can be used without loss of generality at the expense of additional computational complexity.

Using  $R$  reference images  $I_1^{\text{ref}}, \dots, I_R^{\text{ref}}$ , a global reference energy value for each energy band is calculated. The reference

energy  $e_i^{\text{ref}}(\Omega)$  for the  $i$ -th energy band inside  $\Omega$  is defined as:

$$e_i^{\text{ref}}(\Omega) = \frac{1}{R} \sum_{j=1}^R e_i(I_j^{\text{ref}}, \Omega) \quad (4)$$

Using these reference values, the normalized image  $I_{\text{norm}}$  of  $I$  with respect to  $\Omega$  is calculated as:

$$I_{\text{norm}}(\mathbf{x}, \Omega) = \sum_{i=1}^B \lambda_i(\Omega) I_i(\mathbf{x}) = \sum_{i=1}^B \frac{e_i^{\text{ref}}(\Omega)}{e_i(\Omega)} I_i(\mathbf{x}) \quad (5)$$

The value  $\lambda_i(\Omega)$  measures how close the intensity values of the  $i$ -th energy bands are to the reference values. In order to have a complete correspondence of energies between the image  $I$  and the reference images  $I_1^{\text{ref}}, \dots, I_R^{\text{ref}}$ , the described normalization procedure is repeated iteratively until  $\lambda_i(\Omega) \approx 1$ . Algorithm 1 shows the iterative procedure for the proposed localized energy-band normalization algorithm.

### B. CXR normalization

To specifically apply the described method to CXRs, the region of interest needs to be defined. For the region of interest,

**Algorithm 1** Localized energy-based normalization algorithm

```

1: input Original image  $I_{original}$ 
2: input Region of interest  $\Omega$ 
3: input  $R$  reference images  $\{I_j^{ref}\}$ 
4: input Number of energy bands  $B$ 
5: Initialize  $I_{norm} = I_{original}$ 
6: for all Reference images  $I_j^{ref}$  do
7:   Decompose  $I_j^{ref}$  in  $B$  energy bands  $I_{j,i}^{ref}$ 
8: end for
9: for  $i \leftarrow 1$  to  $B$  do
10:   Calculate reference energy value  $e_i^{ref}(\Omega)$ 
11: end for
12: repeat
13:   Set  $I = I_{norm}$ 
14:   Decompose  $I$  in energy bands  $I_i$ 
15:   for  $i \leftarrow 1$  to  $B$  do
16:     Calculate energy value  $e_i(\Omega)$ 
17:     Set  $\lambda_i(\Omega) = \frac{e_i^{ref}(\Omega)}{e_i(\Omega)}$ 
18:   end for
19:   Set  $I_{norm} = \sum_{i=1}^B \lambda_i(\Omega) I_i(x)$ 
20: until  $\{\lambda_i(\Omega) \approx 1\}$ 
21: output  $I_{norm}$ 
```

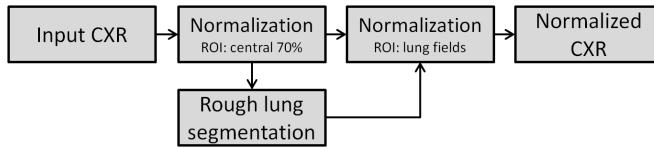


Fig. 2: Flowchart of the CXR normalization process.

the lung fields are selected as pulmonary abnormalities occur in this area. As the location of the lung fields is unknown at the time of analysis and supervised lung segmentation algorithms cannot be applied successfully without normalization, the normalization technique is applied in two stages. In the first stage  $\Omega$  is chosen as the central 70% of the image. With this choice of  $\Omega$ , primarily statistics from the lung fields are gathered. Only one iteration is performed. After this first stage a rough lung segmentation can be obtained from a supervised method [2], and the resulting lung region was set as the region of interest in the second stage to give optimal lung standardization. The process is summarized in the flowchart in Fig. 2.

#### IV. EVALUATION DESIGN

To evaluate the effectiveness of the normalization method, three tasks typically carried out for automatic CXR interpretation and analysis were investigated. First, a method to segment the lungs, which is a common first step in computer-aided detection systems as most systems focus on lung abnormalities only. Second, the clavicles were segmented using a clavicle segmentation system, which can be beneficial for the suppression of overprojecting bony structures. Third, a system aimed at detecting TB related textural abnormalities in the lung fields.

A set of  $R = 50$  images from the same scanner as dataset B were selected as reference images as this was the main

source of data. All images in this set were normal images. An initial set of pilot experiments, with varying  $B$  in the range of 3 to 9, was conducted to determine the optimal number of frequency bands and we chose  $B = 6$  as the optimal number of bands. Fig. 3 shows an example of the frequency decomposition. As described above, the first stage of the normalization only uses one iteration. For the second stage however, two different settings were applied: 1 and 10 iterations. In this specific application, one iteration of the normalization gives the desired  $\lambda_i(\Omega)$  values of approximately 1. However, to equalize the images to a maximum extent, a maximum number of 10 iterations was applied. Fig. 4 plots the average values of  $\lambda_i(\Omega)$  (averaged over all images) as a function of the number of iterations. The standard deviation of  $\lambda_i(\Omega)$  over all 600 images is shown in Fig. 5.

The performance of the three tasks is analyzed on the available datasets (see Section II) after applying the normalization method to the images and compared with the performance obtained when a baseline normalization method, and when histogram equalization was used. For this baseline method, the intensity values were linearly re-scaled to zero mean and unit standard deviation according to:

$$I_{baseline}(\mathbf{x}) = \frac{I(\mathbf{x}) - \mu}{\sigma}. \quad (6)$$

To calculate the histogram equalized image, the image's values  $I(\mathbf{x})$  are mapped according to:

$$I_{hist}(\mathbf{x}) = round\left(\frac{cdf(I(\mathbf{x})) - cdf_{min}}{MN - cdf_{min}}(L - 1)\right), \quad (7)$$

where  $M$  and  $N$  represent the image's width and height respectively,  $cdf$  represent the image's cumulative distribution function,  $cdf_{min}$  is the minimum non-zero value of the  $cdf$  and  $L$  is the number of grey levels used in the image.

The performance of the three tasks is analyzed on the available datasets (see Section II) after applying the normalization methods to the images and compared with the performance obtained when the baseline and histogram equalization method were used.

The performance  $P^k$  of each supervised task  $k$  was evaluated while training with one dataset  $d_i$  and testing on the others:  $P^k(d_i, d_j)$   $i \neq j$ . If  $i = j$ ,  $P^k(d_i, d_i)$  was obtained in leave-one-out cross-validation.

#### V. EXPERIMENTS AND RESULTS

Visual results of the normalization are shown in the second column of Fig. 8. The effectiveness of the normalization was evaluated with three in-house developed, previously published, supervised systems that are briefly described below.

##### A. Lung segmentation

Segmenting the lung fields is an important first step in many CAD algorithms because it restricts further analysis to this area only. In this paper, a pixel classification technique was used [2]. It was shown to outperform a range of other methods that employ statistical shape and appearance modeling.

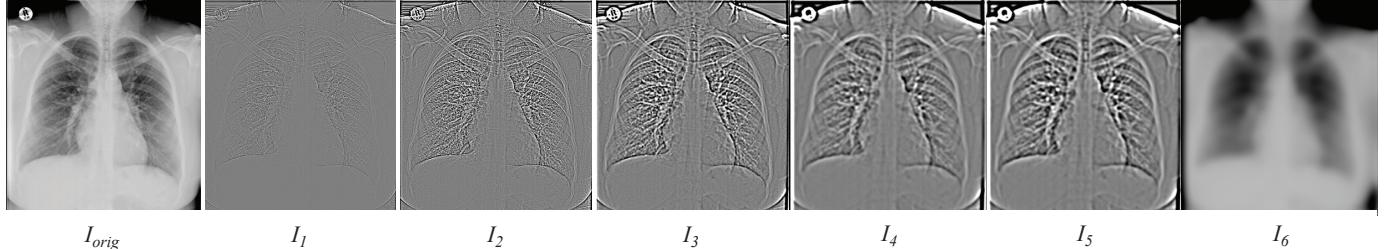


Fig. 3: The original image  $I_{\text{orig}}$  is decomposed into six frequency bands. From left to right, frequency band  $I_1$  contains highest frequency information and  $I_6$  contains lowest frequency information.

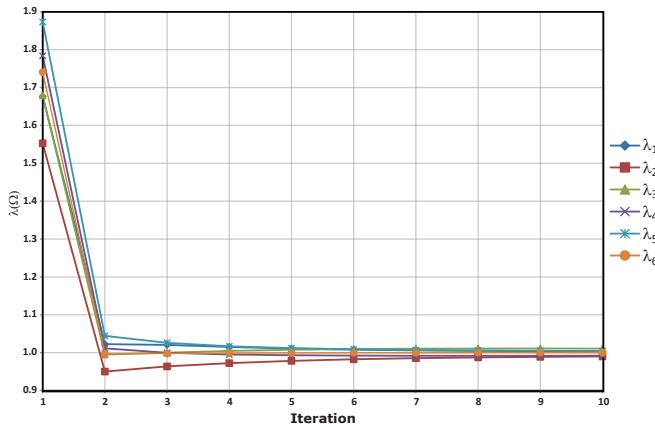


Fig. 4: Values of  $\lambda_i(\Omega)$  plotted as a function of the number of iterations for each frequency band. Reported values are averaged over all 600 images.

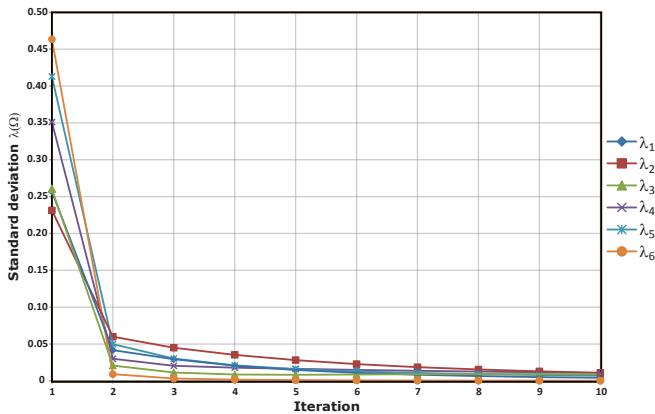


Fig. 5: Standard deviation of  $\lambda_i(\Omega)$  plotted as a function of the number of iterations for each frequency band. Reported values are calculated using all 600 images.

**1) Method:** Images were first re-scaled to a width of 256 pixels. Next, a kNN pixel classifier was used with features based on Gaussian derivatives (up through second order calculated at five exponentially spaced scales: 1-16 pixels), the original pixel value and the  $x, y$  position of each pixel. Resulting posterior probability images were post-processed using smoothing and a sequence of mathematical morphological operations to obtain a binary lung mask image.

**2) Evaluation:** The reference unobscured lung field annotations were manually outlined by a CRSS certified "B" grade reader [42], except for the JSRT dataset which came from the SCR database (Segmentation in Chest Radiographs) [2].

The performance measure used in this experiment is the Jaccard overlap index  $J$ . This is an accepted measure that works well with larger objects whose shapes are not too complex. For each image, the overlap  $J$  between the automatic segmentation  $A$  and the manual segmentation  $B$  was calculated as:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{TP}{FP + FN + TP}, \quad (8)$$

where TP is the number of true positive pixels (pixels correctly labeled as object), FP is the number of false positive pixels (pixels labeled as object, which are background) and FN is the number of false negative pixels (pixels labeled as background which are object). Significance testing was done using Wilcoxon signed-rank tests.  $p$ -values less than 0.05 are considered significant.

**3) Results:** Example results of the lung segmentations are shown in the third column of Fig. 8. The performance of the baseline method is clearly significantly worse than the other methods. For each method and for each training set, the overlap results of all images are shown in Fig. 6. The figure is divided in six columns with four boxplots: one column for each source of training. The blue, red, magenta and green boxes represent the results of the baseline method, histogram equalization, normalization with one iteration and the iterated normalization, respectively. Table II shows all values of  $P^k(d_i, d_j)$  and its variation. Less brightly colored cells in the columns implies that the segmentation performance is less dependent on the utilized training set. From the table and boxplots it is seen that the segmentation performance has improved compared to the baseline and histogram equalization when the proposed normalization method is used. The variability of the results has clearly reduced, demonstrating that the segmentation method is less affected by changes in the image characteristics. The average Jaccard index after applying the baseline method was  $0.72 \pm 0.30$ , which significantly improved to  $0.87 \pm 0.11$  ( $p < 0.01$ ) using histogram equalization and in turn, the proposed normalization, irrespective of the number of iterations, showed a significantly higher performance with an average Jaccard index of  $0.89 \pm 0.09$  ( $p < 0.01$ ). Although the average values for the normalization and iterated method are the same, the latter performed significantly better ( $p < 0.01$ ).

TABLE II:  $P^k(d_i, d_j)$  values for the lung segmentation with different methods. Reported values are the mean and standard deviation of the Jaccard indices ( $J$ ). Color coding shows column-wise differences with the diagonal value: green shows improvement, red shows decline and intensity reflects the difference's magnitude.

Training set	Test set										Baseline					Histogram equalization					Normalization					Iterated normalization				
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F						
A	$\mu$	.89	.56	.48	.68	.92	.76	.87	.85	.86	.87	.90	.86	.89	.91	.90	.89	.91	.92	.89	.91	.90	.88	.91	.92					
	$\sigma$	.08	.26	.20	.13	.06	.16	.12	.12	.10	.08	.09	.07	.09	.07	.05	.09	.07	.03	.09	.08	.05	.09	.07	.03					
B	$\mu$	.04	.89	.90	.88	.29	.85	.71	.90	.88	.85	.85	.87	.78	.90	.90	.90	.86	.91	.79	.90	.90	.90	.86	.91					
	$\sigma$	.09	.08	.06	.07	.21	.07	.18	.08	.06	.09	.10	.05	.15	.07	.05	.06	.08	.04	.14	.08	.05	.06	.08	.04					
C	$\mu$	.08	.89	.92	.90	.36	.86	.72	.89	.92	.89	.89	.91	.78	.90	.92	.91	.88	.91	.79	.90	.92	.91	.87	.91					
	$\sigma$	.11	.08	.06	.07	.21	.07	.22	.08	.05	.07	.09	.04	.14	.07	.05	.05	.08	.04	.14	.07	.05	.06	.08	.04					
D	$\mu$	.06	.88	.91	.92	.35	.88	.81	.89	.90	.91	.89	.92	.79	.89	.91	.92	.86	.91	.80	.89	.91	.92	.87	.91					
	$\sigma$	.11	.11	.06	.06	.22	.06	.16	.09	.07	.09	.03	.15	.08	.06	.06	.08	.04	.15	.08	.05	.06	.08	.03						
E	$\mu$	.83	.84	.87	.86	.92	.90	.83	.90	.90	.90	.88	.92	.87	.86	.91	.90	.90	.93	.86	.91	.90	.91	.93	.93					
	$\sigma$	.12	.10	.06	.06	.07	.05	.14	.07	.07	.07	.08	.08	.12	.06	.05	.07	.06	.03	.12	.06	.05	.06	.06	.03					
F	$\mu$	.32	.90	.92	.91	.66	.94	.81	.89	.90	.90	.88	.93	.82	.91	.92	.91	.89	.94	.82	.91	.92	.91	.89	.94					
	$\sigma$	.19	.08	.05	.05	.19	.03	.18	.09	.06	.07	.09	.03	.14	.07	.05	.06	.08	.03	.14	.07	.05	.06	.08	.03					

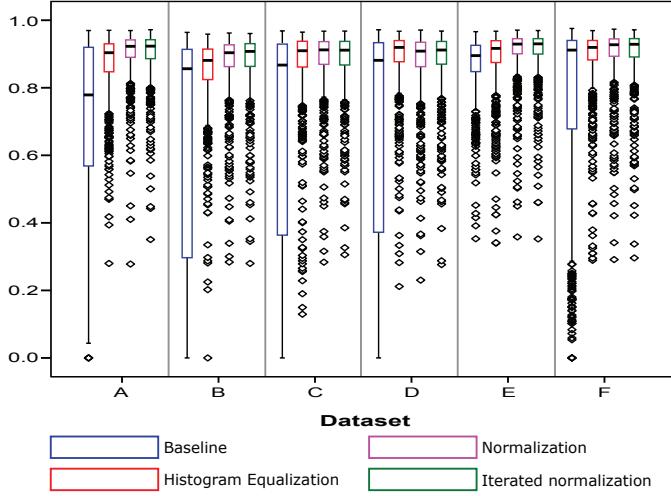


Fig. 6: Box plots of the Jaccard index calculated for the lung segmentation task for all images. The horizontal axis indicates the training set. Blue, red, magenta and green boxes show the results for the baseline method, histogram equalization, normalization, and iterated normalization, respectively. The central line of the box represents the median, the box edges the 2nd and 3rd quartile and the whiskers the extremes of the data excluding outliers. Outlier points (black dots) are further than 1.5 inter-quartile range away from the box.

### B. Clavicle segmentation

For TB it is known that abnormalities most frequently occur in the upper lobes of the lungs [43], [44]. However, analysis of this region, due to superimposition of the clavicles, is a challenging task for CAD systems. Therefore, segmentation of the clavicles can be advantageous to suppress them or to reduce false positive detections in this area. The system used in this paper was previously described by Hogeweg et al. [45].

1) *Method:* The segmentation process is performed in a number of stages. First, three different zones of the clavicle are classified with a pixel classifier using 59 features: Gaussian filtered images (up through second order at 7 exponentially spaced scales: 1-64 pixels), the original image, two Hessian

matrix derived features at 5 different exponentially spaced scales (1-16 pixels) and 6 position features. The output of the pixel classification is used as input for active shape modeling. The resulting segmentation is subsequently used in a dynamic programming step that optimizes a cost function combining the appearance information from the different clavicle zones.

2) *Evaluation:* The reference clavicle annotations came from the SCR database [2] for the JSRT dataset, whereas the other cases were outlined by a CRRS certified "B" grade reader [42]. For each clavicle, the part superimposed on the lungs until the superior border of the scapula was outlined. As we are only interested in the part of the clavicle inside the lung field, we do not consider the part of the clavicles that is outside of the lung regions. As for the lung segmentation, we calculate the Jaccard index (Eq. 8), and only take the intersection of the manual lung segmentations and the automatically segmented clavicles into account.

3) *Results:* Example results for the clavicle segmentation task are shown in the fourth column of Fig. 8. The overlap results are shown in Fig. 7. The blue, red, magenta and green boxes represent the baseline, histogram equalization, normalization and iterated normalization method, respectively. Table III shows all values of  $P^k(d_i, d_j)$  and its variation for clavicle segmentation. Less brightly colored cells in the columns implies that the segmentation performance is less dependent on the utilized training set. The overlap averaged for all different combination of train and test sets, increased significantly from  $0.57 \pm 0.26$  for the baseline method and  $0.53 \pm 0.26$  for the histogram equalization method to  $0.69 \pm 0.23$  ( $p < 0.01$ ) for the normalization method and increased to  $0.68 \pm 0.23$  ( $p < 0.01$ ) for the iterated method. It can be seen from the table and the boxplots that the normalization procedure is successful. Although the increase is less substantial than what was observed for the lung segmentation task, the effect is significant and the variability has decreased after normalization. There was no significant difference between the normalization and iterated normalization methods ( $p = 0.19$ ).

TABLE III:  $P^k(d_i, d_j)$  values for the clavicle segmentation task with different methods. Reported values are the mean and standard deviation of the Jaccard indices ( $J$ ). Color coding shows column-wise differences with the diagonal value: green shows improvement, red shows decline and intensity reflects the difference's magnitude.

Training set	Test set																								
	Baseline						Histogram equalization						Normalization						Iterated normalization						
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F	
A	$\mu$	.80	.47	.37	.39	.65	.53	.77	.56	.54	.54	.57	.52	.82	.75	.68	.69	.78	.62	.80	.72	.64	.67	.64	.63
	$\sigma$	.10	.24	.24	.24	.25	.24	.10	.25	.25	.23	.29	.25	.09	.07	.05	.09	.07	.03	.09	.08	.05	.09	.07	.03
B	$\mu$	.43	.80	.73	.65	.45	.63	.59	.64	.50	.52	.65	.50	.75	.81	.68	.71	.81	.74	.78	.80	.68	.73	.81	.74
	$\sigma$	.13	.12	.19	.17	.17	.17	.18	.24	.28	.23	.25	.21	.15	.07	.05	.06	.08	.04	.14	.08	.05	.06	.08	.04
C	$\mu$	.34	.67	.78	.67	.25	.64	.55	.55	.62	.55	.35	.49	.70	.68	.77	.70	.49	.73	.70	.67	.77	.70	.49	.73
	$\sigma$	.13	.22	.12	.13	.18	.15	.18	.27	.21	.18	.31	.21	.14	.07	.05	.05	.08	.04	.14	.07	.05	.06	.08	.04
D	$\mu$	.30	.67	.78	.76	.33	.77	.56	.50	.62	.62	.30	.58	.72	.67	.75	.77	.46	.79	.74	.68	.78	.78	.49	.80
	$\sigma$	.16	.20	.11	.13	.24	.12	.18	.26	.19	.18	.29	.21	.15	.08	.06	.06	.08	.04	.15	.08	.05	.06	.08	.03
E	$\mu$	.70	.58	.37	.42	.82	.43	.47	.46	.33	.35	.53	.29	.74	.74	.52	.60	.82	.45	.75	.74	.48	.58	.81	.45
	$\sigma$	.15	.23	.28	.23	.12	.28	.18	.22	.24	.22	.22	.21	.12	.06	.05	.07	.06	.03	.12	.06	.05	.06	.06	.03
F	$\mu$	.45	.51	.75	.74	.18	.80	.66	.44	.68	.71	.19	.79	.69	.54	.76	.74	.22	.82	.69	.52	.75	.75	.22	.82
	$\sigma$	.16	.32	.13	.14	.24	.11	.19	.30	.18	.16	.28	.10	.14	.07	.05	.06	.08	.03	.14	.07	.05	.06	.08	.03

TABLE IV:  $P^k(d_i, d_j)$  values for the texture abnormality detection system with different methods. Reported values are dataset  $A_z$  values. Color coding shows column-wise differences with the diagonal value: green shows improvement, red shows decline and intensity reflects the difference's magnitude.

Training set	Test set																							
	Baseline						Histogram equalization						Normalization						Iterated normalization					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
A	.83	.69	.76	.82	.73	-	.82	.77	.83	.82	.69	-	.86	.88	.84	.78	.81	-	.83	.86	.78	.85	.81	-
	.68	.89	.90	.84	.46	-	.81	.87	.87	.77	.72	-	.87	.92	.89	.88	.77	-	.87	.91	.92	.86	.82	-
B	.81	.73	.91	.83	.44	-	.80	.86	.90	.78	.72	-	.78	.91	.95	.85	.74	-	.83	.92	.95	.87	.76	-
	.63	.58	.80	.83	.62	-	.85	.74	.80	.75	.72	-	.83	.93	.91	.86	.78	-	.81	.87	.90	.86	.78	-
C	.81	.50	.48	.64	.71	-	.83	.80	.82	.76	.70	-	.87	.88	.87	.83	.83	-	.89	.89	.86	.84	.83	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

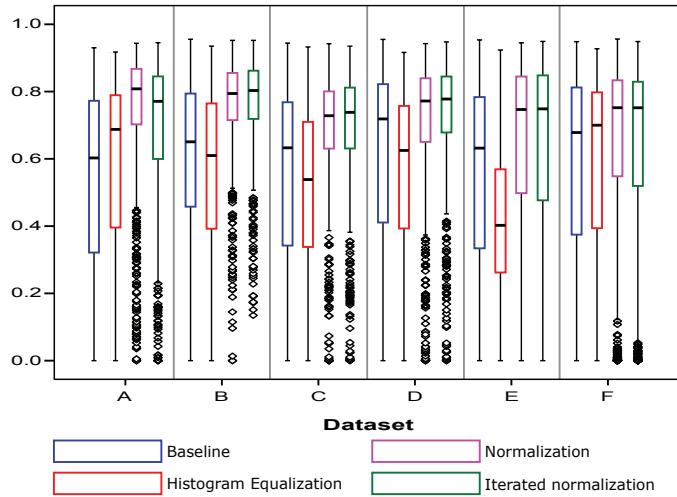


Fig. 7: Box plots of the Jaccard index calculated for the clavicle segmentation task for all images. The horizontal axis indicates the training set. Blue, red, magenta and green boxes show the results for the reference method, histogram equalization, normalization, and iterated normalization, respectively. The central line of the box represents the median, the box edges the 2nd and 3rd quartile and the whiskers the extremes of the data excluding outliers. Outlier points (black dots) are further than 1.5 inter-quartile range away from the box.

### C. Texture abnormality detection

Many abnormalities in chest radiographs are textural in nature. The system used in this paper was previously described by Hogeweg et al. [21] and aims at the detection of TB related abnormalities such as consolidations and small and large opacities.

1) *Method:* The system detects textural abnormalities within automatically segmented lung fields. To avoid potential performance bias by the quality of the segmented lung fields, the manual lung segmentations were used. The calculation of features was done on Gaussian derivative filtered images (up through second order at 4 exponentially spaced scales: 1-8 pixels) and the original image. For each image, the first four moments of small circular image patches (32 pixel radius), subsampled every 8 pixels were calculated, leading to a total of 100 features. In addition, the relative  $x$  and  $y$  position in the image, the relative  $x$  and  $y$  position within the bounding box of the lung fields, the distance to the lung wall and the distance to the hilum were added. Thus, a total of 106 features was used in a kNN pixel classifier to give each pixel a probability score. After morphological post-processing, the resulting probability map (heatmap) was converted to an abnormality score using a 95% quantile rule.

2) *Evaluation:* Contrary to the other two tasks, this task is only evaluated with five datasets: the system is not designed to detect focal lesions such as nodules and the JSRT dataset only contains nodular abnormalities. All abnormalities in the abnormal images were outlined by a CRRS certified "B" grade

reader [42]. Each of the abnormal images contained either a small or large opacity or consolidation.

The performance of the texture system was calculated using receiver operating characteristic (ROC) analysis and calculation of the area under the ROC curve  $A_z$  for each set.

3) *Results:* Example results of the texture abnormality detection system are shown as a heat map in the last column of Fig. 8. Performance results are summarized in Table IV. Each number is the  $A_z$  value for a combination of training and test set. It can be seen that in general the reported diagonal values are the highest within the columns, which is expected as the same source is used for training and testing. The  $A_z$  values for the normalization and iterated normalization method are significantly higher ( $0.85 \pm 0.05$ ,  $p < 0.01$ ) than that of the baseline method ( $0.72 \pm 0.14$ ) and histogram equalization ( $0.79 \pm 0.06$ ), which indicates that the normalization methods are effective. The difference between the normalization and iterated normalization methods are small and not significant ( $p = 0.68$ ).

## VI. DISCUSSION

In this study, an iterative algorithm to standardize medical images has been presented. The algorithm can successfully normalize structures of interest using a localized energy-based normalization technique. The effectiveness of the normalization was evaluated by performing three different tasks, common in automated chest radiography image processing, on 600 CXRs from different sources. In this section, we first discuss the proposed method, and then we will discuss the experiments after which we conclude with the limitations of the system.

Two key elements in the proposed method are the applied energy band scaling and the addition of a region of interest. The decomposition of the images was done using a 2D Gaussian filter as it is simple and fast. The filter is separable, and thus can be applied as two 1D filters, making the operational complexity of this approach only  $\mathcal{O}(n)$  while other approaches such as discrete Fourier transform [46], discrete cosine transform [47] and power spectrum analysis have complexity  $\mathcal{O}(n \log n)$  if the optimized fast Fourier transform is used. While the complexity of the popular discrete wavelet transform [48] is similar, this decomposition method results in both frequency and spatial information, which is not needed in this application. The decomposition into frequency bands provides a separation of structures of different sizes, which is a useful property, as it allows for applying specific scaling factors to each band, which can enhance or suppress specific structures. For example, a combination of high  $\lambda_1(\Omega)$ ,  $\lambda_2(\Omega)$  values with low  $\lambda_3(\Omega) \dots \lambda_B(\Omega)$  leads to unsharp masking, a known technique in medical image analysis for image enhancement [49]–[51]. By taking all  $\lambda_i(\Omega)$  values equal to a reference value, the image's frequency information is standardized among images, which gives them similar appearance and intensity characteristics. In our approach we have chosen images from the same scanner used for dataset B as the reference. This could potentially bias the performance of set B. However, as we are comparing differences between normalization methods,

and not per dataset within one method of normalization, the potential bias does not influence the results and the main outcome of the paper. The addition of a region of interest allows for a better normalization of specific structures which is beneficial in applications focusing on specific regions. In our application, after an initial normalization stage with the central 70% of the image set as region of interest, a second stage of normalization was performed utilizing the lung fields as region of interest. This assured that the lungs, compared to other regions in the image, were optimally standardized.

The iterative approach of the normalization has the advantage that it can standardize structures of interest better, and thus match the reference energy values more closely. In our evaluation, we have distinguished two cases: normalization with one and with ten iterations in the second stage. The latter case ensures all  $\lambda_i(\Omega)$  to be roughly equal to 1. In our application, one iteration of the method was sufficient for two out of three evaluation methods, as the performance of the different tasks did not improve after more iterations. Investigation of  $\lambda_i(\Omega)$  as a function of the number of iterations, showed that the factors were close to 1 after just one iteration. However, they approximated 1 more closely and showed less variation for a higher number of iterations. Additionally, other applications might benefit more from the iterated approach.

The proposed pre-processing normalization technique has been shown to be effective in lung and clavicle segmentation and texture abnormality detection. Ideally, a normalization technique makes subsequent supervised analysis completely independent of the training and test set. With our results, we have shown that the proposed method can be successfully applied on data from different sources and that it performs significantly better than a simple baseline method and histogram equalization.

In the next paragraphs, each of the evaluation tasks will be discussed. In the first task, a supervised method to automatically segment the lung fields was presented and it was shown that the normalization procedure improved the results and showed less variability compared to both reference methods. With an average overlap of  $0.89 \pm 0.09$  the system shows good performance. Although Fig. 6 shows a considerable amount of outliers, the normalization method decreases the variation substantially. Investigation of the outliers showed that these were mostly seen in CXRs containing large abnormalities, which are more difficult to segment correctly. Comparison of the results to literature is difficult as most datasets used are not publicly available. However, performance reported in literature on the JSRT dataset is in the range of 0.71 to 0.95 [2], [52]–[58], as is our result:  $0.90 \pm 0.09$ . However, it must be noted that current results are calculated on a smaller subset of the full dataset, the reported value is the average for all the different training sets and only a small dataset of 100 CXRs was used for training. Assessing the performance of other state of the art lung segmentation algorithms ([52], [54]) using the proposed normalization algorithm as a preprocessing step, is not possible as these are not publicly available. However, our pixel classification based lung segmentation technique is still among the top ranked algorithms [52] in the literature.

The second task aimed at the automatic segmentation of

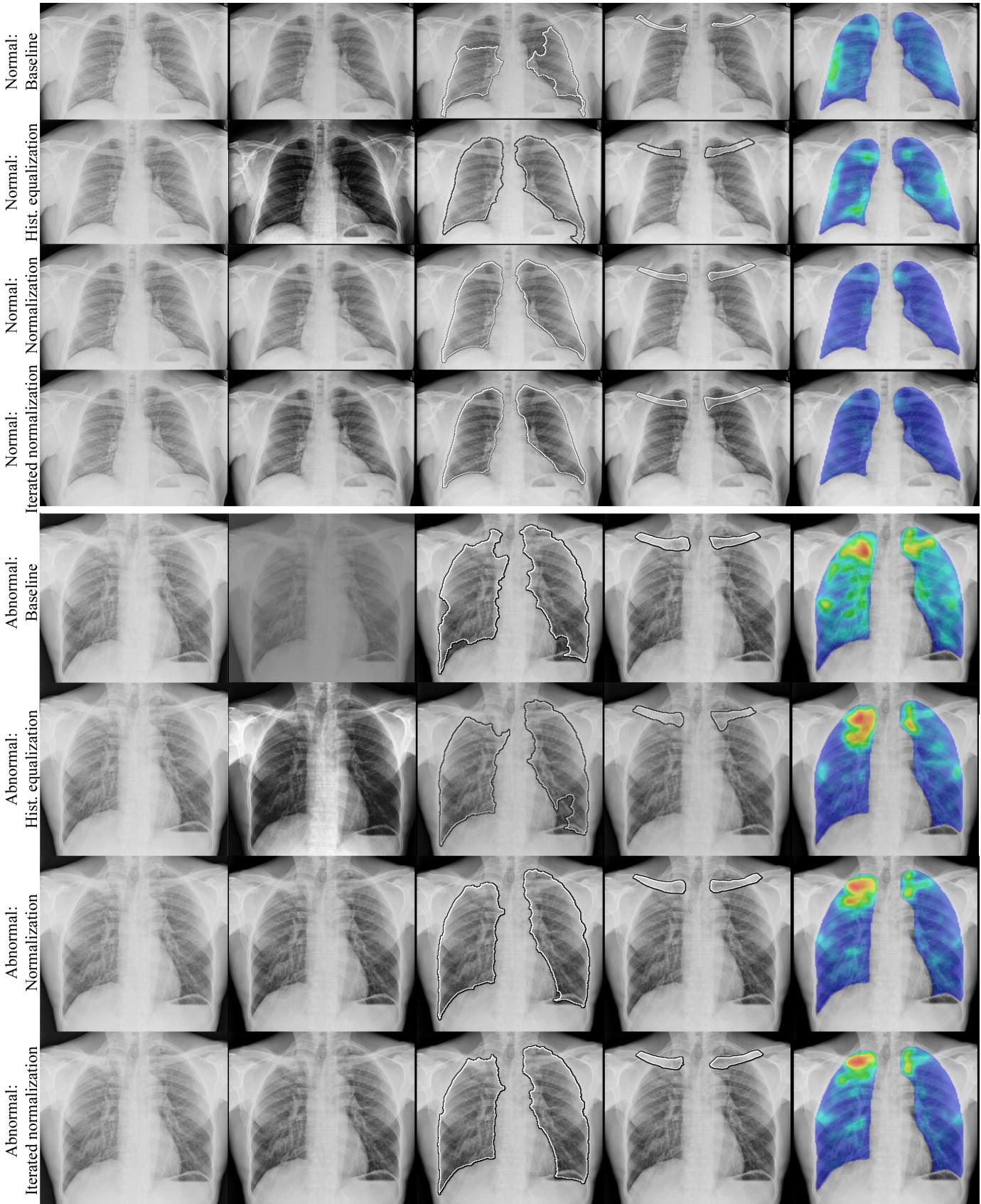


Fig. 8: Example results for the different tasks using different normalization methods. Column 1: original DICOM image; column 2: normalized image; column 3: lung segmentation; column 4: clavicle segmentation; column 5: abnormality detection heatmap. Row 1 and 5: baseline method; row 2 and 6: histogram equalization; row 3 and 7: normalization method; row 4 and 8: iterated normalization method results. The first four rows show a normal example image from dataset B and the last four rows show an abnormal example image from dataset A. Results are obtained using dataset A for training.

clavicles. With a significant increase and less variability in Jaccard index, the normalization method is successful again. The decrease in performance compared with lung segmentation results reflect the complexity of the task: due to superimposed structures in the top of the lungs, automated clavicle segmentation is more difficult than lung segmentation.

The third task focused on the detection of abnormalities related to TB. For this task, the JSRT data was omitted as this contained only focal lesions and the manually segmented lung masks were used as input to prevent any bias due to the quality of the lung segmentations from the first task. The results, calculated in terms of area under the ROC curve, in Table IV show that the normalization procedure markedly improves the detection of TB related abnormalities compared to the baseline and histogram equalization method.

Although the proposed technique works well for the performed tasks, there are some limitations. First, as the results indicate, the proposed technique does not lead to systems that are completely independent of the training data used. Although the algorithm can cope with (quantum) noise present in the high frequency bands, as they will be also normalized, it can not completely overcome anatomical noise. Especially the posterior ribs may vary in appearance, as their appearance depends on the bone mineral density, beam energy and proprietary postprocessing. The last one often includes highly non-linear processing which cannot be undone. Applying the proposed technique to raw unprocessed CXR images may overcome this problem but such images are scarcely available in practice. Alternatively, one could consider using rib suppression/enhancement software [59]–[61]. Another reason for suboptimal normalization is under- and overexposure of CXRs. Although this has been greatly reduced with digital chest radiography, some images suffer from a very dense (white) appearing heart and ribs, which makes it impossible to assess the lung tissue behind it. As this intensity information is not contained in the image, our normalization method cannot retrieve it. Focusing the normalization to the lungs introduces another limitation. It is advantageous for the normalization of the lungs, but comes at the expense of the appearance of the overall image. Visual inspection of the images showed that there were differences in gray values appearance outside the lung regions. This does not affect the automated analysis, but can hamper interpretation by radiologists.

The presented technique is not restricted in its application to CXRs: it can be used in 2D as well as 3D images. A preliminary study outside the scope of this paper has shown the normalization to be effective in Computed Tomography scans. Further research should show if it can be used in combination with other normalization techniques and its relevance in other medical imaging techniques.

## VII. CONCLUSION

A generic method to normalize medical images was presented and its application in chest radiography was evaluated. The method uses an energy decomposition of the image, after which the energy of each band is scaled separately to a reference energy to acquire a normalized image. Applying this

technique iteratively and locally gives maximum standardization of structures of interest in the image. In our application, the normalization was first applied to the central 70% of the CXRs and then locally to the lung fields. The method was evaluated on six datasets from different sources for three automated tasks in chest radiography interpretation: lung segmentation, clavicle segmentation and texture abnormality detection. A marked increase in performance was seen with the normalization method compared to a baseline and histogram equalization method in each of the three tasks. Future work would focus on other application areas.

## ACKNOWLEDGMENT

This work was supported by the European and Developing Countries Clinical Trials Partnership (EDCTP), the Evaluation of multiple novel and emerging technologies for TB diagnosis, in smear-negative and HIV-infected persons, in high burden countries (TB-NEAT) project.

## REFERENCES

- [1] N. L. Müller, C. A. Staples, R. R. Miller, and R. T. Abboud, "Density mask. an objective method to quantitate emphysema using computed tomography," *Chest*, vol. 94, pp. 782–787, 1988.
- [2] B. van Ginneken, M. B. Stegmann, and M. Loog, "Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database," *Med Image Anal*, vol. 10, pp. 19–40, 2006.
- [3] S. Juhász, . Horváth, L. Nikházy, G. Horváth, and A. Horváth, "Segmentation of anatomical structures on chest radiographs," in *XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010*, ser. IFMBE Proceedings, P. Bamidis and N. Pallikarakis, Eds. Springer Berlin Heidelberg, 2010, vol. 29, pp. 359–362.
- [4] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever, "General intensity transformations and differential invariants," *J Math Imaging Vis*, vol. 4, pp. 171–187, 1994.
- [5] J.-L. Chen and A. Kundu, "Rotation and gray scale transform invariant texture identification using wavelet decomposition and hidden markov model," *IEEE Trans Pattern Anal Mach Intell*, vol. 16, pp. 208–214, 1994.
- [6] W.-R. Wu and S.-C. Wei, "Rotation and gray-scale transform-invariant texture classification using spiral resampling, subband decomposition, and hidden markov model," *IEEE Trans Image Process*, vol. 5, pp. 1423–1434, 1996.
- [7] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, pp. 1345–1359, 2010.
- [8] Y. Arzhaeva, D. M. J. Tax, and B. van Ginneken, "Dissimilarity-based classification in the absence of local ground truth: application to the diagnostic interpretation of chest radiographs," *Pattern Recognit*, vol. 42, pp. 1768–1776, 2009.
- [9] T. Ojala, M. Pietikäinen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans Pattern Anal Mach Intell*, vol. 24, pp. 971–987, 2002.
- [10] A. Jalil, T. Cheema, A. Manzar, and I. Qureshi, "Rotation and gray-scale-invariant texture analysis using radon and differential radon transforms based hidden markov models," *Image Processing, IET*, vol. 4, pp. 42–48, 2010.
- [11] M. Kallenberg and N. Karssemeijer, "Computer-aided detection of masses in full-field digital mammography using screen-film mammograms for training," *Phys Med Biol*, vol. 53, pp. 6879–6891, 2008.
- [12] M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, X. Guan, C. Schmitt, and N. E. Thomas, "A method for normalizing histology slides for quantitative analysis," *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1107–1110, 2009.
- [13] L. G. Nyúl and J. K. Udupa, "On standardizing the MR image intensity scale," *Magn Reson Med*, vol. 42, pp. 1072–1081, 1999.

- [14] L. G. Nyúl, J. K. Udupa, and X. Zhang, "New variants of a method of MRI scale standardization," *IEEE Trans Med Imaging*, vol. 19, pp. 143–150, 2000.
- [15] A. Madabhushi and J. K. Udupa, "New methods of MR image intensity standardization via generalized scale," *Med Phys*, vol. 33, pp. 3426–3434, 2006.
- [16] World Health Organization, "Global tuberculosis report 2013," World Health Organization, 2013.
- [17] World Health Organization, "Systematic screening for active tuberculosis: principles and recommendations," World Health Organization, 2013a.
- [18] S. Den Boon, E. D. Bateman, D. A. Enarson, M. Borgdorff, S. Verver, C. J. Lombard, E. Irusen, N. Beyers, and N. W. White, "Development and evaluation of a new chest radiograph reading and recording system for epidemiological surveys of tuberculosis and lung disease," *Int J Tuberc Lung Dis*, vol. 9, pp. 1088–1096, 2005.
- [19] S. Graham, G. K. Das, R. Hidvegi, R. Hanson, J. Kosiuk, Z. Al, D. Menzies *et al.*, "Chest radiograph abnormalities associated with tuberculosis: reproducibility and yield of active cases," *Int J Tuberc Lung Dis*, vol. 6, p. 137, 2002.
- [20] J. P. Zellweger, R. Heinzer, M. Touray, B. Vidondo, and E. Altpeter, "Intra-observer and overall agreement in the radiological assessment of tuberculosis," *Int J Tuberc Lung Dis*, vol. 10, pp. 1123–1126, 2006.
- [21] L. Hogeweg, C. Mol, P. A. de Jong, R. Dawson, H. Ayles, and B. van Ginneken, "Fusion of local and global detection systems to detect tuberculosis in chest radiographs," in *Med Image Comput Comput Assist Interv*, ser. Lect Notes Comput Sci, vol. 6363, 2010, pp. 650–657.
- [22] L. E. Hogeweg, "Automatic detection of tuberculosis in chest radiographs," Ph.D. dissertation, Radboud Universiteit Nijmegen, 2013.
- [23] S. Jaeger, A. Karargyris, S. Candemir, J. Siegelman, L. Folio, S. Antani, and G. Thoma, "Automatic screening for tuberculosis in chest radiographs: a survey," *Quant Imaging Med Surg*, vol. 3, p. 89, 2013.
- [24] S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z. Xue, K. Palaniappan, R. Singh, and S. Antani, "Automatic tuberculosis screening using chest radiographs," *IEEE Trans Med Imaging*, 2013.
- [25] S. Dippel, M. Stahl, R. Wiemker, and T. Blaffert, "Multiscale contrast enhancement for radiographies: Laplacian pyramid versus fast wavelet transform," *IEEE Trans Med Imaging*, vol. 21, pp. 343–353, 2002.
- [26] C. Fivez, P. Vuylsteke, P. Wambacq, P. Suetens, and E. Schoeters, "Multi-resolution contrast amplification in digital radiography with compensation for scattered radiation," in *Proc Int Conf Image Process*, vol. 1, 1996, pp. 339–342.
- [27] M. Stahl, T. Aach, and S. Dippel, "Digital radiography enhancement by nonlinear multiscale processing," *Med Phys*, vol. 27, pp. 56–65, 2000.
- [28] S. Tservas and D. Iakovidis, "Patient specific normalization of chest radiographs and hierarchical classification of bacterial infection patterns," in *Imaging Systems and Techniques (IST), 2010 IEEE International Conference on*, July 2010, pp. 156–160.
- [29] B. van Ginneken, "Computer-aided diagnosis in chest radiography," Ph.D. dissertation, Utrecht University, The Netherlands, 2001.
- [30] Y. Shi, F. Qi, Z. Xue, L. Chen, K. Ito, H. Matsuo, and D. Shen, "Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics," *IEEE Trans Med Imaging*, vol. 27, pp. 481–494, 2008.
- [31] S. Jaeger, A. Karargyris, S. Antani, and G. Thoma, "Detecting tuberculosis in radiographs using combined lung masks," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 4978–4981.
- [32] T. Xu, I. Cheng, R. Long, and M. Mandal, "Novel coarse-to-fine dual scale technique for tuberculosis cavity detection in chest radiographs," *EURASIP Journal on Image and Video Processing*, vol. 3, pp. 1–18, 2013.
- [33] P. Maduskar, L. Hogeweg, R. Philipsen, S. Schalekamp, and B. van Ginneken, "Improved texture analysis for automatic detection of tuberculosis (TB) on chest radiographs with bone suppression images," in *Medical Imaging*, ser. Proceedings of the SPIE, vol. 8670, 2013, p. 86700H.
- [34] O. Mohd Rijal, H. Ebrahimian, and N. Noor, "Determining features for discriminating ptb and normal lungs using phase congruency model," in *Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on*, 2012, pp. 341–344.
- [35] J. H. Tan, U. R. Acharya, C. Tan, K. T. Abraham, and C. M. Lim, "Computer-assisted diagnosis of tuberculosis: A first order statistical approach to chest radiograph," *J Med Syst*, vol. 36, pp. 2751–2759, 2012.
- [36] P. Annangi, S. Thiruvenkadam, A. Raja, H. Xu, X. Sun, and L. Mao, "A region based active contour method for x-ray lung segmentation using prior shape and low level features," in *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, April 2010, pp. 892–895.
- [37] R. Philipsen, P. Maduskar, L. Hogeweg, and B. van Ginneken, "Normalization of chest radiographs," in *Medical Imaging*, ser. Proceedings of the SPIE, vol. 8670, 2013, p. 86700G.
- [38] A. Story, R. W. Aldridge, I. Abubakar, H. R. Stagg, M. Lipman, J. M. Watson, and A. C. Hayward, "Active case finding for pulmonary tuberculosis using mobile digital chest radiography: an observational study," *Int J Tuberc Lung Dis*, vol. 16, pp. 1461–1467, 2012.
- [39] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, "Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules," *AJR Am J Roentgenol*, vol. 174, pp. 71–74, 2000.
- [40] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans Commun*, vol. COM-31,4, pp. 532–540, 1983.
- [41] T. Stathaki, *Image Fusion: Algorithms and Applications*. Academic Press, 2008.
- [42] R. Dawson, P. Masuka, D. J. Edwards, E. D. Bateman, L.-G. Bekker, R. Wood, and S. D. Lawn, "Chest radiograph reading and recording system: evaluation for tuberculosis screening in patients with advanced HIV," *Int J Tuberc Lung Dis*, vol. 14, pp. 52–58, 2010.
- [43] P. Eini, H. Owaysee Osquee, M. Sajjadi Nasab, F. Nasiroghli Khiyabani, and A. H. Rahighi, "Chest radiological features among patients with smear positive pulmonary tuberculosis," *Caspian J Intern Med*, vol. 4, pp. 777–780, 2013.
- [44] V. Kumar, A. K. Abbas, and J. C. Aster, *Robbins Basic Pathology*, 9th ed. Saunders Elsevier, 2012.
- [45] L. Hogeweg, C. I. Sánchez, P. A. de Jong, P. Maduskar, and B. van Ginneken, "Clavicle segmentation in chest radiographs," *Med Image Anal*, vol. 16, pp. 1490 – 1502, 2012.
- [46] Z. Wang, "Fast algorithms for the discrete w transform and for the discrete fourier transform," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 32, no. 4, pp. 803–816, Aug 1984.
- [47] A. B. Watson, "Image compression using the discrete cosine transform," *Mathematica journal*, vol. 4, no. 1, p. 81, 1994.
- [48] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," vol. 11, no. 7, pp. 674–693, Jul 1989.
- [49] B. H. Brinkmann, A. Manduca, and R. A. Robb, "Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction," *IEEE Trans Med Imaging*, vol. 17, pp. 161–171, 1998.
- [50] A. Polesel, G. Ramponi, and V. J. Mathews, "Image enhancement via adaptive unsharp masking," *IEEE Trans Image Process*, vol. 9, pp. 505–510, 2000.
- [51] G. Ramponi, N. Strobel, S. K. Mitra, and T. H. Yu, "Nonlinear unsharp masking methods for image contrast enhancement," *J Electron Imaging*, vol. 5, pp. 353–366, 1996.
- [52] S. Candemir, S. Jaeger, K. Palaniappan, J. Musco, R. Singh, Z. Xue, A. Karargyris, S. Antani, G. Thoma, and C. McDonald, "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," *IEEE Trans Med Imaging*, vol. 33, pp. 577–590, 2014.
- [53] G. Coppini, M. Miniati, S. Monti, M. Paterni, R. Favilla, and E. M. Ferdegiani, "A computer-aided diagnosis approach for emphysema recognition in chest radiography," *Med Eng Phys*, vol. 35, pp. 63–73, 2013.
- [54] A. Dawoud, "Lung segmentation in chest radiographs by fusing shape information in iterative thresholding," *Computer Vision, IET*, vol. 5, pp. 185–190, 2011.
- [55] B. van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny, and M. A. Viergever, "Active shape model segmentation with optimal features," *IEEE Trans Med Imaging*, vol. 21, pp. 924–933, 2002.
- [56] D. Seghers, D. Loeckx, F. Maes, D. Vandermeulen, and P. Suetens, "Minimal shape and intensity cost path segmentation," *IEEE Trans Med Imaging*, vol. 26, pp. 1115–1129, 2007.
- [57] Y. Shi and D. Shen, "Hierarchical shape statistical model for segmentation of lung fields in chest radiographs," in *Medical Image Computing and Computer-Assisted Intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 11, 2008, pp. 417–24, PMID: 18979774.
- [58] T. Yu, J. Luo, and N. Ahuja, "Shape regularized active contour using iterative global search and local optimization," in *Computer Vision and Pattern Recognition*, vol. 2, 2005, pp. 655–662.
- [59] A. Karargyris, S. Antani, and G. Thoma, "Segmenting anatomy in chest x-rays for tuberculosis screening," vol. 2011, pp. 7779–7782, 2011.

- [60] K. Suzuki, H. Abe, H. MacMahon, and K. Doi, "Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN)," *IEEE Trans Med Imaging*, vol. 25, pp. 406–416, 2006.
- [61] L. Hogeweg, C. Mol, P. A. de Jong, and B. van Ginneken, "Rib suppression in chest radiographs to improve classification of textural abnormalities," in *Medical Imaging*, ser. Proceedings of the SPIE, vol. 7624, 2010, pp. 76240Y1–76240Y6.