

Lung Segmentation of MR Images: A Review

Tatyana Ivanovska, Katrin Hegenscheid, René Laqua, Sven Gläser,
Ralf Ewert, and Henry Völzke

Abstract Magnetic resonance imaging (MRI) is a non-radiation based examination method, which gains an increasing popularity in research and clinical settings. Manual analysis of large data volumes is a very time-consuming and tedious process. Therefore, automatic analysis methods are required. This paper reviews different methods that have been recently proposed for automatic and semi-automatic lung segmentation from magnetic resonance imaging data. These techniques include thresholding, region growing, morphological operations, active contours, level sets, and neural networks. We also discuss the methodologies that have been utilized for performance and accuracy evaluation of each method.

1 Introduction

Magnetic resonance imaging (MRI) has non-invasive, non-ionizing nature and provides a superior soft tissue contrast. This technique is rapidly developing over the last years. For example, continuous technological advances such as parallel imaging [16, 54] and integrated parallel acquisition techniques (iPAT) [30] significantly reduce the acquisition time and enable standardized imaging acquisition at a very high spatial and temporal resolution providing morphological and functional information of the highest detail. This makes MRI attractive in clinical practice and research settings.

T. Ivanovska (✉) • H. Völzke
Institute of Community Medicine, Ernst-Moritz-Arndt University, Greifswald, Germany
e-mail: tetyana.ivanovska@uni-greifswald.de; voelzke@uni-greifswald.de

K. Hegenscheid • R. Laqua
Institute of Diagnostic Radiology and Neuroradiology, Ernst-Moritz-Arndt University,
Greifswald, Germany
e-mail: katrin.hegenscheid@uni-greifswald.de; laquar@uni-greifswald.de

S. Gläser • R. Ewert
Department of Internal Medicine B, Division of Pulmonary Medicine, Ernst-Moritz-Arndt
University, Greifswald, Germany

Lung imaging is one of the most challenging topics in MRI due to motion artifacts and low signal. Nevertheless, in recent years lung MRI is becoming increasingly popular in comparison to the pulmo-imaging “gold standard”, computed tomography (CT). High quality MRI assists in detection of numerous pulmonary diseases. Moreover, MRI of the lung is important for specific clinical applications. For instance, it can contribute to decision making the case of such diseases as lung cancer [67], malignant pleural mesothelioma, and acute pulmonary embolism. The advantages of MR over CT are not limited to the lack of ionizing radiation, which is of particular interest for the assessment of lung disease in children, pregnant women, or in patients, who require frequent follow-up examinations (e.g., immunocompromised patients with fever of unknown origin). Chest wall invasion by a tumor and mediastinal masses are accepted indications benefiting from MRI superior soft tissue contrast. Dynamic examinations to study respiratory mechanics and contrast enhanced first pass perfusion imaging reach far beyond the scope of CT [25].

MRI of the lung has multiple application areas. For example, conventional (so-called proton or 1H) MRI provides anatomical details and is standardly used for assessing lung volumes, boundaries, detection of nodules, infiltrates, and masses. There is a significant number of protocols designed for these purposes [5, 6]. In Fig. 1, two example slices of anatomical MR images from two different sequences, namely, T1-weighted VIBE (volume-interpolated breath hold examination) and T2-weighted HASTE (half-Fourier single shot turbo spin-echo), are shown. The VIBE sequence has a higher spatial resolution and a smaller slice thickness than the HASTE sequence (for instance, $1.8 \times 1.8 \times 3$ mm vs. $2.3 \times 1.8 \times 5$ mm on 1.5 Tesla Magnetom Avanto Siemens device) and is preferable for lung volumetry, whereas lung infiltrates serve as key pathologies for applying the HASTE sequence.

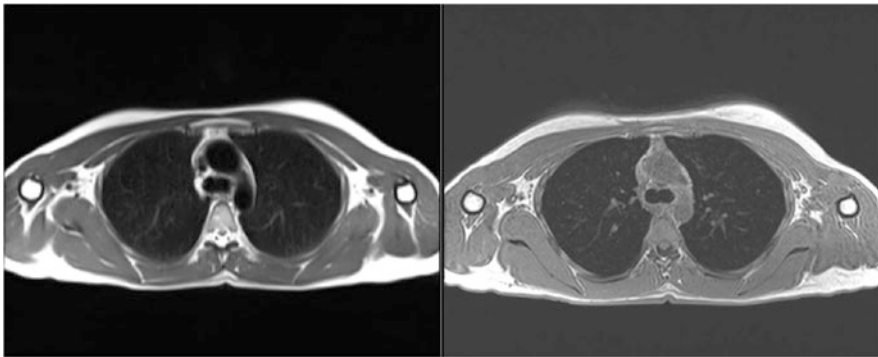


Fig. 1 An example of two slices from two anatomical MR thoracic scans from the same participant. *Left*: T2-weighted Half-Fourier Acquisition Single-Shot Turbo Spin-Echo sequence (HASTE), $256 \times 206 \times 44$; *Right*: T1-weighted volume-interpolated breath hold sequence (VIBE), axial orientation, $512 \times 416 \times 88$. Image courtesy of Study of Health in Pomerania, Germany [18]

MRI with non-radioactive noble gases, such as hyperpolarized ^3He (helium-3) or ^{129}Xe (xenon-129), is applied for detection and evaluation of functional, e.g., ventilation, abnormalities [40, 42, 62]. MR angiography (MRA) gains its popularity for assessment of pulmonary vascular disease [25].

(Semi-)automatic segmentation is an essential step in medical image analysis, medical visualization, and computer-aided diagnosis. Since manual segmentation of large volumes of MR data is a very laborious, observer-dependent, and time-consuming process, which is prone to inter- and intra-observer variability, automatized segmentation methods for extraction of different organs and structures are actively developed.

This paper aims to provide an overview of recently published literature on automatic and semi-automatic segmentation techniques for MR images of human lungs. Our main source of references was the Internet: we have searched for the terms *lung*, *segmentation*, *volumetry*, *pulmonary*, *magnetic resonance imaging (MRI)* on Google Scholar, PubMed, and IEEE-Xplore.

The paper is organized as follows. First, a short introduction to human lung anatomy is given in Sect. 2. In Sect. 3, we introduce the methods and the data, which they are applied to. The performance of the methods is discussed in Sect. 4. Section 5 concludes the paper.

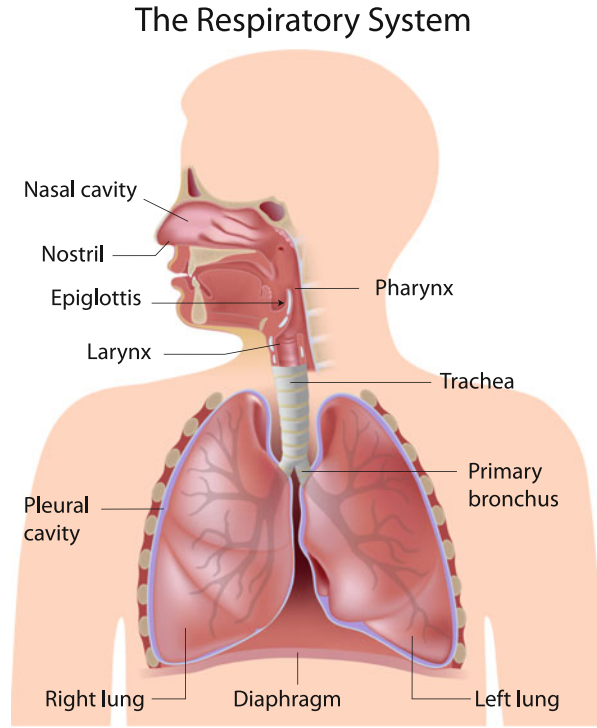
2 Pulmonary Anatomy

In this section, we briefly present the human pulmonary system. For a more detailed description, we refer to specialized medical literature, for example [35].

Human lungs are located in the chest wall, consisting of the rib cage, diaphragm, and mediastinum, the area between the lungs, which includes all of the organs in the chest, such as heart, vessels, trachea, except the lungs. The right lung is composed of three lobes, while the left lung has only two lobes. The lobes are further divided into sub-lobar segments, which are defined by the branching pattern of the airway tree. The lobar fissures, space between the surface of the lobes, allow the lobes to rotate relative to one another to accommodate body posture-related changes in chest wall geometry.

Air travels into the lungs through the nose or mouth and through the pharynx, larynx, and into the tracheobronchial (airway) tree (cf. Fig. 2). The pulmonary blood circulation is provided by pulmonary arteries and veins.

Fig. 2 A schematic description of the human respiratory system



3 Lung Segmentation

Segmentation is the first and essential step in computer analysis of medical imaging data. Segmentation results are further used, for instance, for volumetric measurements, visualization, and biomechanical modeling.

In thoracic MR images, various anatomical entities are presented, such as lungs (including lobes and fissures), bronchial tree, and pulmonary vessels. Whereas in high-resolution CT images all these structures are clearly visible and can be distinguished, lung imaging is a challenge in MRI. Here, usually only organ borders can be identified. Moreover, such fine structures as the complete bronchial tree or the fissures are hardly visible in anatomical MR images. Therefore, the task of lung segmentation from MR images usually consists of lung region detection, separation of the lungs from the large airways, and exclusion or inclusion of the pulmonary vessel regions.

Numerous approaches have been proposed for CT data [13, 38, 53], whereas there are only a few methods for MRI data presented in the literature. Here, we classify and discuss these methods for automatic and semi-automatic lung segmentation from MRI data. The literature can be roughly divided into the following categories: intensity-based, model-based, and active contour-based approaches. A summary of published works is presented in Table 1. The table rows are ordered as the techniques are discussed in the paper body. We leave table cells dashed, if the correspondent information is missing in the original papers.

Table 1 A summary of published works on MR lung segmentation

Authors	Main points	Data		
		Resolution	Orientation	Sequence
Sensakovic et al. 2006 [52]	<i>Intensity-based</i> 2D, slice based, Histogram thresholding, Morphological operations	256 × 256	Axial	T1 SPGR
Ivanovska et al. 2012 [22]	<i>Intensity-based</i> 3D, Clustering, CC Analysis, Trachea removal, Lung separation, Morphological operations	2.3 × 1.8 × 5 mm	Axial	T1 VIBE
Kullberg et al. 2009 [27]	<i>Intensity-based</i> 3D, Thresholding, Distance transform, Region growing, CC Analysis	2.1 × 2.1 × 8 mm	Coronal	Whole body multigradient
Lui et al. 2013 [33]	<i>Intensity-based</i> 2D, slice based, Semi-automatic Histogram thresholding, Clustering, Trachea removal, Morphological operations	256 × 256, 8–14 slices	Coronal	³ He Fast GRE

(continued)

Table 1 (continued)

Authors	Main points	Data		
		Resolution	Orientation	Sequence
Plathow et al. 2005 [45]	<i>Intensity-based</i> 2D, slice based, Semi-automatic Interactive region growing Morphological operations	$3.75 \times 3.75 \times$ 3.8 mm	Coronal	3D FLASH
		$2.5 \times 1.6 \times$ 10 mm	Coronal	2D trueFISP
Heydarian et al. 2012 [19]	<i>Intensity-based</i> 2D and 3D K-Means Region growing	$256 \times 128,$ 14 slices	Coronal	T1 SPGR
Kirby et al. 2012 [26]		$128 \times 128,$ 14 slices	Coronal	^3He
Tokuda et al. 2009 [58]	<i>Intensity-based</i> 2D, Semi-automatic Confidence connected- ness Fuzzy connectedness Region growing (ITK)	$3.125 \times$ 3.125×5 mm	Coronal	3D FLASH + GRAPPA
Woodhouse et al. 2005 [66]	<i>Intensity-based</i> 2D, Semi-automatic Thresholding	256×256	Coronal	^1H SSFSE
		112×128	Coronal	^3He
Virgincar et al. 2012 [64]	<i>Intensity-based</i> 2D, Semi-automatic Region Growing Morphological operations	192×192	Coronal	SSFP
		192×192	Coronal	^3He SPGRE
Lelieveldt et al. 2000 [28, 29]	<i>Model-based</i> 3D, Multiorgan Energy potential Minimization	—	Thoracic scout	—

(continued)

Table 1 (continued)

Authors	Main points	Data		
		Resolution	Orientation	Sequence
Tustison et al. 2011 [60]	<i>Model-based</i> 3D, Automatic Multiorgan Shape prior, PCA Ventilation-based defects N4 correction	80 × 128, 19–28 slices	Axial	³ He
Ray et al. 2003 [48]	<i>Active contours</i> 2D, Automatic Contour merging Snakes Slice based	—	Coronal	¹ H
Middleton and Damber, 2004 [39]	<i>Active contours</i> <i>Neural network</i> Binary classification Refinement with Snakes 3D	256 × 256 × 35	Axial	T1-weighted
Tetzlaff et al. 2010 [57]	<i>Semi-automatic</i> 2D Graph Cuts	128 × 128	Coronal	2D FLASH
Böttger et al. 2007 [9]	<i>Semi-automatic</i> 2D and 3D Parametric active contours Simplex mesh	—	Coronal	HASTE

The table rows are ordered as the techniques are discussed in the paper body

3.1 Intensity-Based Segmentation

Intensity-based segmentation steps are similar in MR and CT data. Namely, intensity-based approaches usually consist of the following steps. First, lungs are extracted and separated from the large airways. Second, if needed, anterior and posterior lung junctions are corrected. Third, the lung contours are smoothed and the cavities are filled.

Sensakovic et al. proposed a multi-step approach based on thresholding, shape descriptors, and morphological operations [50–52]. The method works on a slice-by-slice basis. In each slice, the thorax is separated from the background using a histogram-based thresholding and a boundary smoothing [17]. Once the thorax segmented image is computed, a histogram-based thresholding [17] is applied to it to create a lung-segmented images. Thereafter, the connected components

are analyzed, and all regions that fail to satisfy the descriptor thresholds (area, compactness, center of mass, and perimeter) are classified as non-lung. If no lung region is detected, the threshold is decreased and the procedure is repeated. After this procedure, which is referred as the “core stage”, a refinement step (“the correction stage”) starts. This step is important especially in the presence of disease and acquisition artifacts. It aims to capture valid lung regions that may have been obscured by disease or acquisition artifacts. First, a gray scale erosion operator [55] is applied to the thorax-segmented image. It lowers the intensity values of the lung regions masked by acquisition artifacts or disease. Second, a thresholding procedure, similar to the one in the core stage, is applied. A series of rolling ball filters [2] is applied to the internal boundaries the lungs. Candidate lung regions from the correction stage lung-segmented image that satisfy the circularity criterion are combined with lung regions from the core stage, and the final lung-segmented image is constructed.

Ivanovska et al. implemented a 3D method for lung segmentation [22]. The approach consists of four steps: the main extraction and three refinement steps. First, the lung and trachea are segmented. Any automatic clustering technique (for example, K-Means [34]) is applicable here. Second, the trachea and main bronchi are found using a 3D connected component analysis [17]. The separation of the large airways from the lungs is done slice-wise using 2D Watershed [63] on the inverted distance map [55]. Third, the lungs are separated from each other using a 3D Watershed procedure on a subimage, which represents an anterior or posterior junction. The subimage is selected such that the computational costs of the watershed procedure stay relatively low. Fourth, the lung cavities are corrected with a morphological closing [55].

Kullberg et al. presented an automatic approach for lung segmentation as a part of the adipose tissue depots segmentation from the whole-body MRI datasets [27]. Segmentation of lungs is performed with thresholding and morphological operations [17]. To increase robustness to erroneous inclusion of intestinal gas, the lungs are segmented one by one. First, the right lung is segmented, since the position of the liver increases the distance between the right lung and the intestines. The segmentation is performed as follows. Initial thresholding is followed by a distance transform [55]. Then, a region growing [17] is applied to the distance transform image. The condition for the growing process is continuous decrease of the distance transform values. The result is dilated to ensure inclusion of the whole lung. An intensity threshold is applied in the dilated lung to determine the final result. The threshold value is computed from the fitted Gaussian function to the histogram from the dilated lung region.

Lui et al. proposed a semi-automatic segmentation algorithm to isolate areas of ventilation from hyperpolarized ^3He MRI [33]. First, the background is removed by determining an optimal threshold from a sampled background noise distribution located outside of the lung field. Second, the lung mask is refined with a four-class fuzzy C-Means [56] clustering. Third, the trachea is semi-automatically removed

with a seeded region growing method [56]. Finally, morphological operators [55] are used to refine the result.

Plathow et al. applied a semi-automatic method based on an interactive region growing technique [56] to evaluate lung volumes during the breathing cycles using dynamic MRI [45]. First, displacement of the chest wall and the diaphragm as a surrogate of the lung volume and lung surface was measured in dynamic 2D MRI images during the breathing cycle. Thereafter, vital capacity (VC) was calculated using a model, consisting of a half ellipse, an ellipsoid cone, and a prisma [44]. An automatic segmentation in 3D dynamic MRI failed due to intensity inhomogeneities and motion artifacts. Therefore, the 3D data was segmented in a slice-by-slice manner using interactive region growing. For each dataset, the 2 segmentation results were combined into a voxel-based 3D description of the current breathing state. The voxel-based segmentation results were transformed into triangular meshes using the marching cubes algorithm [31]. The isosurfaces were then smoothed in a post-processing step. Total surface lung areas were computed by summing up the area of each individual triangle.

Heydarian et al. proposed a semi-automated method for generating ^3He measurements of individual slice (2D) or whole lung (3D) gas distribution [19]. The same technique was also used by Kirby et al. [26]. The authors applied hierarchical K-Means [34] clustering for the ^3He and a seeded region-growing algorithm [56] for anatomical MR images. Thereafter, the images were registered semi-automatically using landmark-based registration [41]. Finally, the segmented areas were postprocessed with a morphological closing algorithm with a structuring element disk [55]. The vertical central region was excluded from the processing to avoid connecting the right and left lung areas.

Tokuda et al. measured lung volumes in dynamic 3D lung images [58]. The segmentation was performed in a slice-by-slice manner. In each slice, the lung area was segmented using a combination of confidence-connectedness and fuzzy-connectedness region growing algorithms, implemented in the Insight Segmentation and Registration Toolkit [21]. First, a rough segmentation of the lung area was obtained with the confidence-connectedness region growing method. The algorithm extracted a connected set of pixels whose pixel intensities are consistent with the pixel statistics (the mean and variance across a neighborhood) of predefined seed points. The pixels, connected to the seed points, whose values are within the confidence intervals are grouped together. The width of the confidence interval is controlled by a multiplier parameter. Second, the mean and variance of intensities in the lung area were then used for the fuzzy connectedness algorithm [21] to compute an affinity map, which represents degrees of adjacency and the similarity of pairs of nearby voxels. Finally, the refined lung area was extracted by thresholding the affinity maps. The multiplier and threshold parameters were tuned such the segmented lung area included the blood vessels.

Woodhouse et al. used a combination of anatomical and ^3He MR images to compare the ventilated and thoracic lung volumes in groups of smokers and never-smokers [66]. The processing was done manually in a slice-by-slice manner. ^3He images were also segmented semi-automatically with adaptive thresholding [56].

The threshold value was derived from the signal-to-noise (SNR) [56] value of each image and the mean signal in a manually selected region of interest in each image.

Virgincar et al. analysed ^{129}Xe and SSFP (steady state free precession) ^1H images in a semi-automatic manner [64]. The anatomical MRI was segmented by a region growing method [56]. The seeds were placed manually in the lowest intensity areas of the right and left lung in the central slice of the thoracic cavity. Thereafter, a threshold range was computed from the intensities of the seeds. Then, the morphological closing [55] was applied to the extracted lung. The ^1H and ^{129}Xe were acquired over different breath-holds, they required registration. It was performed using either affine transform [3] or similarity transform [20]. After the registration, ventilation images were also segmented with the region growing method.

3.2 *Model-Based Segmentation*

Lelieveldt et al. developed a model-based method to simultaneously segment lungs as well as some other nearby organs [28, 29]. An anatomical model of thorax is built by modeling of individual organs with implicit surfaces from manually delineated training images and subsequent grouping of single organ models into a tree structure. The whole hierarchical scene is described by a boundary model, which characterizes the scene volume as a boundary potential (or energy) function. Thereafter, a model is matched to image data. First, an initial parameter set for the pose and scale parameters is selected. Second, an automatic thresholding based on empirical histogram evaluation discriminates air from tissue. A set of target boundary points is obtained. Third, a model is placed and the energy function is minimized with Levenberg-Marquardt nonlinear by fitting the minimization [32] by fitting the target boundary points to the model. The authors reported that their method was designed for a coarse pre-segmentation and a more accurate local segmentation process is further required.

Tustison et al. proposed an automated segmentation method for differentiation of the ventilated lung volumes on ^3He MRI [60]. The method consists of two main steps: template and statistical model construction and individual subject processing. All images in the database are registered to a normalized space. A normalized unbiased template is built from seven representative subjects with a symmetric diffeomorphic registration algorithm [4]. Thereafter, a principal component analysis (PCA) model [15] from an image database is built. Each image is transformed to the template using an affine transformation, so that the presence of any global shape differences in the statistical model is avoided. Then the processing of individual subjects starts. Each dataset is mapped to the template, the bias field is corrected [59], and then a shape-based level set procedure [43] starts to extract the lungs, where the PCA model is used.

3.3 *Active Contours and Neural Network Segmentation*

Ray et al. implemented a slice-based active contour approach for automatic segmentation of lungs from 1H MRI [48]. Initial snakes are placed automatically within the lung cavities that are to be segmented. Then all snakes are evolved independently of each other. Finally, the union of the regions covered by all snakes is considered as the final results. For the snake evolution, the authors proposed to apply a modified gradient vector flow (GVF) with Dirichlet boundary conditions [47].

Middleton and Damber proposed a combined method, consisting of neural networks and active contours (snakes) to segment lungs in MRI [39]. First, a neural network was trained for binary classification of each pixel as a “boundary” or a “non-boundary”. The inputs to the neural network are 7×7 image patches of the pixel to be classified. The weights were determined by training on MR sections with lung boundary pixels segmented by an expert observer. The resulting edge-point image was used as the external energy for the snake evolution.

Tetzlaff et al. analysed lung areas on 2D dynamic MR images [57]. The segmentation was performed semi-automatically, using Graph Cut algorithm [10]. The algorithm was initialized by adapting a bounding box to the size of the region of interest, and then the areas inside and outside of the lungs were roughly marked. If the segmentation leaked into the thoracic wall, an additional scribble had to be drawn at the point of leakage.

Böttger et al. presented a new segmentation approach [8, 9], based on parametric active contours [24]. Discrete surface meshes [12] were used here. The authors proposed to use an additional speed term, which was derived from the magnitude of the Gaussian gradient image. New attractor forces were introduced into the simplex mesh deformation scheme [7]. These forces were obtained from the user defined attractor points. To cope with complex surface geometries and avoid oscillating behavior during the evolution, the mesh was refined during the deformation automatically after every 50 iterations.

4 Discussion

4.1 *Methodical Analysis*

There are several aspects that define the relationships between the described algorithms. As it is shown above, most of the presented techniques are pipelines consisting of well-known algorithms, such as region growing, intensity clustering or thresholding, morphological operations [55, 56], combined together for a specific purpose. The novelty of the presented pipelines is in the application areas. Therefore, the algorithms are usually designed for a specific data type, i.e., MR sequence, and would often require substantial changes to be adapted to other data. Most of the presented techniques are developed for coronal or axial anatomical

MRI. Unfortunately, often no exact data protocols are given in the papers (cf. Table 1). Some techniques are designed for detection of ventilation defects in MR with noble gases (^3He and ^{129}Xe). Since ^3He images are often acquired concurrent with the traditional anatomical MR images, some techniques of ^3He segmentation are based on the prior segmentation of ^1H images and subsequent image registration, for instance, [19]. Moreover, some methods were designed only for 2D (slice-wise) processing [48, 51] and do not consider any 3D object consistency and connectivity. More complicated approaches, such as presented in [22, 60], were developed for a 3D organ extraction. Most of the presented techniques consist of one or two steps and include the user interaction. Usually, the more users are involved in the processing, the simpler the algorithmic pipeline is, since the complicated parts are accomplished by the users. For example, for many algorithms [19, 33] the mediastinum area is manually pre-excluded from the processing, and the user initiates the computations by placing the initial seed points [33, 58, 64] or adapting threshold values [66]. Of course, semi-automatic algorithms are less tedious for the users than the completely manual segmentation, but still can be inappropriate for processing of thousands of datasets. The fully automated pipelines, such as [22, 60], have a significant number of algorithmic steps and pre-selected parameters.

Therefore, the practitioners should take into account the following facts, while selecting, which technique to apply:

- What type of data is available?
- How many datasets need to be processed?
- Is 2D or 3D processing required?
- What level of user interaction is accepted?

For example, if the user interaction and 2D processing are acceptable, then slice-by-slice algorithmic solutions for ^1H or ^3He MR data, similar to the ones presented in [19, 26, 66], can be considered.

4.2 Performance Evaluation

All automatic and semi-automatic methods require some form of quantitative measure of the accuracy. Usually, ground truth masks are obtained manually and the results produced by the algorithm are compared to it. Here, region based metrics, such as DICE coefficient [14] or sensitivity and specificity, are often applied. Moreover, for certain applications the correlation to the clinical examinations, such as spirometry, is important. Hence, the correlation coefficients between these measurements are assessed.

However, there is no unique system of quality measures used in the community, which complicates the evaluation of the algorithms' efficiency. This section reviews the approaches, presented above, from the quantitative and qualitative perspectives.

A summary of the test sets and details of validation for each method is given in Table 2. The order of the rows in kept the same, as in Table 1.

Table 2 A summary of performance measurements of the segmentation methods on MR lung segmentation

Authors	Test set	Evaluation
Sensakovic et al. 2006 [52]	101 slices	Area of overlap (regions) $AOM \geq 82\%$
Ivanovska et al. 2012 [22]	10 datasets	Volume fractions (regions) $TPVF \geq 97\%$ Error fractions $\leq 5\%$
Kullberg et al. 2009 [27]	24 datasets	No evaluation
Lui et al. 2013 [33]	10 datasets	DICE's coefficient (regions) $DICE = 0.96 \pm 0.1$
Plathow et al. 2005 [45]	20 datasets	Correlation with spirometry Spearman's correlation $r > 0.83$ Significance level $P < 0.005$
Heydariian et al. 2012 [19]	4 datasets	Correlation with manual Pearson's correlation $r > 0.98$ Significance level $P < 0.0001$ $DICE > 0.9$
Kirby et al. 2012 [26]	15 datasets 3 pathologies	Ventilation Defect Volume (VDV) Ventilation Volume (VV) Correlation with manual Pearson's correlation $r \geq 0.84$ Significance level $P < 0.0001$ $DICE \geq 0.88$
Tokuda et al. 2009 [58]	2 datasets	No evaluation
Woodhouse et al. 2005 [66]	18 datasets Smokers and never-smokers	Ventilation Volume (VV) Correlation with manual $ICC \geq 0.98$
Virgincar et al. 2012 [64]	44 datasets	Ventilation Defect Percentage (VDP) Correlation with manual Pearson's $r = 0.97, P < 0.0001$
Lelieveldt et al. 2000 [28, 29]	15 Datasets	No evaluation

(continued)

Table 2 (continued)

Authors	Test set	Evaluation
Tustison et al. 2011 [60]	43 datasets for defect evaluation 18 datasets for ventilation based segmentation	Defect evaluation: $ICC = 0.85$ Segmentation: STAPLE method from 4 observers $Sens = 0.898$, $Spec = 0.905$ (regions, no ground truth)
Ray et al. 2003 [48]	10 datasets	$FOM = 69\%$ (edges) Error = 6% (regions)
Middleton and Damper 2004 [39]	13 datasets	Segmentation performance $F \geq 0.84$ (regions)
Tetzlaff et al. 2010 [57]	10 datasets	Correlation to spirometry $r \geq 0.97$
Böttger et al. 2007 [8, 9]	10 datasets	$DICE \geq 0.88$ Surface distance ≤ 2 mm Hausdorff distance ≤ 20 mm

4.2.1 Region-Based and Edge-Based Metrics

Sensakovic et al. [51, 52] and Sensakovic and Armato [50] applied their technique to a random sample of 101 thoracic MR sections. True lung regions were manually delineated by two radiologists. 90% of the patient data included abnormalities, such as mesothelioma, scarring, enlarged lymph node and others. To measure the segmentation quality, the authors used an area-of-overlap (AOM) measure, also known as the Jaccard coefficient [23]. The AOM measure of two regions, A and B, is defined as the number of pixels contained within the intersection of the regions divided by the number of the pixels contained within the union of the regions:

$$AOM = \frac{A \cap B}{A \cup B}, AOM \in [0, 1], \quad (1)$$

where 0 corresponds to disjoint regions and 1 is the complete overlap. The AOM is similar to the Dice coefficient [14], but the normalizing part (the denominator) is different. The reported AOM values were equal to 0.82 ± 0.16 and 0.83 ± 0.13 , when compared to the first and second observers, respectively.

Ivanovska et al. [22] tested the proposed method on ten randomly selected participants with normal lungs. The accuracy evaluation was done by the methodology proposed by Udupa et al. [61]. Namely, the following volume measures (VF: volume fractions) were computed: TPVF (true positives is the fraction of voxels

in the intersection of the automatic and manual segmentation results), FPVF (false positives is the fraction of voxels falsely identified by the automatic segmentation), FNVF (false negatives is the fraction of voxels defined in manual segmentation, but missed by the automatic method). Additionally, TPVE (true positive volume error) was calculated. Let C_{exp} and C_{auto} denote the binary masks produced the expert and the automatic pipeline, correspondingly. Then, the measures are defined as

$$TPVF = \frac{|C_{auto} \cap C_{exp}|}{|C_{exp}|}, \quad (2)$$

$$FNVF = \frac{|C_{exp} - C_{auto}|}{|C_{exp}|}, \quad (3)$$

$$FPVF = \frac{|C_{auto} - C_{exp}|}{|C_{exp}|}. \quad (4)$$

Let V_{auto} and V_{exp} denote the volumes, computed by the algorithm and the expert, correspondingly. Then $TPVE$ is defined as

$$TPVE = \frac{|V_{auto} - V_{exp}|}{|V_{exp}|}. \quad (5)$$

The accuracy measures have been computed for left and right lungs separately. For the left lung: $TPVF = 97.06 \pm 1.36 \%$, $FPVF = 2.97 \pm 1.08 \%$, $FNVF = 2.94 \pm 1.36 \%$, $TPVE = 1.58 \pm 1.16 \%$. For the right lung: $TPVF = 97.33 \pm 1.57 \%$, $FPVF = 3.96 \pm 1.69 \%$, $FNVF = 2.66 \pm 1.57 \%$, $TPVE = 2.37 \pm 1.74 \%$. Moreover, the authors acquired the ground truth from two independent experts and observed that agreement about 93.5%, which shows that the automatic result accuracy lies with the variation interval between the experts.

Lui et al. [33] validated their method on four healthy and six asthmatic subjects, calculating the ventilated lung volume (LVL). To assess the accuracy of the method, the authors calculated a Dice coefficient [14], comparing their results to the manually obtained ground truth. The Dice coefficient measures the overlap between the regions A and B :

$$DICE = \frac{2|A \cap B|}{|A| + |B|}, \quad DICE \in [0, 1], \quad (6)$$

when 0 there is no overlap, 1 is the perfect match. The reported Dice value is 0.96 ± 0.01 . Bland-Altman analysis [36] was used to determine the 95% limits of agreement calculated from the mean and standard deviation of the volume difference between the segmentation results and the ground truth. The authors processed a total of 109 coronal slice for ten subjects. It showed that the means LVLs of the semi-automatic approach were 3.88 ± 0.75 L and 3.83 ± 1.11 L for healthy and asthmatics subject, respectively. The differences to the manual and standard spirometric measurements were statistically insignificant.

Ray et al. [48] applied their approach to ten different datasets. The authors measured Pratt's figure of merit (FOM) [1]. $FOM \in [0, 1]$. The FOM quantifies the comparison between ideal edges (ground truth) and detected edges of the image.

$$FOM = \frac{1}{\max(I_A, I_I)} \sum_{i=1}^{I_A} \frac{1}{1 + d_i \alpha^2}, \quad (7)$$

where I_A and I_I are the detected and ideal edge images, respectively. d_i is the distance between the actual and ideal edge i , and α is a penalty factor for displaced edges. The average obtained FOM is 0.69. The authors also utilized a region-based measure, namely, the percentage error, computed as

$$Error = \frac{\sum_i |Segm(i) - I_g(i)|}{\sum_i I_g(i)} 100\%, \quad (8)$$

where $Segm$ and I_g are the segmentation result and the ground truth, correspondingly. This metric is a combination of FP (false positives) and FN (false negatives) measures. The mean percentage error for ten datasets was about 6%.

Tustison et al. [60] used data from seven random subjects to build the unbiased template. Experimentally, it was seen that such number of datasets provided a satisfactory compromise between quality of results and required computational time. For the PCA statistical model 156 images from normal subjects were used. For the evaluation two comparative analyses were performed. First, the number of defects scored by two human readers in 43 subjects. Here, the intraclass correlation coefficient (ICC) [37]. This metric is a general measurement of agreement or consensus, where the measurements used are assumed to be parametric (continuous and has a Normal distribution). There was a high agreement between the algorithm and the readers ($ICC = 0.85$ and $ICC = 0.86$ for the first and second observers, correspondingly). Second, the simultaneous truth and performance estimation (STAPLE) [65] was performed on 18 subjects in which the ventilation defects were manually segmented by four human readers. Here, the sensitivity and specificity are defined as

$$Sens = \frac{TP}{TP + FN}, \quad (9)$$

$$Spec = \frac{TN}{TN + FP}, \quad (10)$$

where TP , FN , FP , TN denote the true positives, the false negatives, false positives, and true negatives, respectively. The STAPLE results yielded the best sensitivity and specificity combination for the algorithm ($Sens = 0.898$, $Spec = 0.905$).

Heydarian et al. [19] tested their method on two healthy subjects and two subjects with chronic obstructive lung disease (COPD). The results of hierarchical K-means 2D and 3D segmentation were compared to an expert observer's manual

segmentation results using linear regression, Pearson correlations [11] and the Dice similarity coefficient. 2D hierarchical K-means segmentation of ventilation volume (VV) and ventilation defect volume (VDV) was strongly and significantly correlated with manual measurements (VV: $r = 0.98, P < 0.0001$; VDV: $r = 0.97, P < 0.0001$) and mean Dice coefficients were greater than 0.92 for all subjects. 3D hierarchical K-means segmentation of VV and VDV was also strongly and significantly correlated with manual measurements (VV: $r = 0.98, P < 0.0001$; VDV: $r = 0.64, P < 0.0001$) and the mean Dice coefficients were greater than 0.91 for all subjects.

Kirby et al. [26] applied the same approach for segmentation of ventilation defect volume (VDV) and ventilation volume (VV) on five patients with asthma, five patients with chronic obstructive pulmonary disease (COPD), and five patients with cystic fibrosis (CF). They compared semi-automatic and manual measurements and observed strong significant correlations between the VDV values generated by each method (asthma: $r = 0.89, P < 0.0001$; COPD: $r = 0.84, P < 0.0001$; CF: $r = 0.89, P < 0.0001$). The spatial agreement for VV values was measured with the Dice coefficient (asthma: 0.95; COPD: 0.88; CF: 0.9).

Middleton and Damper [39] applied their technique to 13 datasets. To evaluate the performance, they utilized the following region-based measures: precision (P), recall (R), effectiveness E [49], and segmentation performance $F = 1 - E$. The measures are defined as:

$$P = \frac{TP}{TP + FP}, \quad (11)$$

$$R = \frac{TP}{TP + FN}, \quad (12)$$

$$E = 1 - \frac{PR}{(1 - \alpha)P + \alpha R}, \quad (13)$$

where $\alpha = 0.5$, i.e., the precision and recall are weighted equally. The mean of F was 0.866 and 0.844 for the left and right lungs, respectively.

Böttger et al. [9] applied their tool to 10 datasets and compared the semi-automatically obtained results to two expert readings. For evaluation, the average surface distance, Hausdorff distance, and the Dice coefficient of two compared segmentation were computed. For all segmentation, the average surface distance was significantly lower than 2 mm, and the Hausdorff distance was lower than 20 mm, apart from one outlier. The Dice coefficient was higher than 0.88.

4.2.2 Other Metrics

Plathow et al. [45] measured on a dynamic 2D MRI displacement of the chest wall and the diaphragm as a surrogate of the lung volume and lung surface in 20 healthy subjects. Moreover, a 3D volumetric evaluation of the breathing cycle was done,

using a dynamic 3D MRI. The results were correlated to spirometry and the vital capacity (VC) [46] was measured. VC using spirometry was $4.3 \pm 1.0\text{L}$; using the 2D model, VC was $4.9 \pm 1.2\text{L}$. Using the 3D MRI VC was $4.65 \pm 0.9\text{L}$. Correlation between spirometry and the 2D and 3D models was highly significant: the Spearman's correlation coefficient $r > 0.83$ and significance level $P < 0.005$ [11]. The differences of absolute VC values between spirometry, 2D and 3D MRI were insignificant.

Tokuda et al. [58] assessed lung motion sequences from two volunteers, the quality of segmentation was determined by visual observation of the original images.

Woodhouse et al. [66] measured the ventilated lung volumes from a combination of ^3He and proton single-shot fast spin echo (SSFSE) coronal MR images in groups of "healthy" smokers (five subjects), smokers with moderate COPD (five subjects), and never-smokers (eight subjects). The results (volume values) from slice-wise semi-automatic segmentation and manual segmentations from two observers were compared to each other. It showed high agreement (Pearson correlation coefficient $r \geq 0.94, P < 0.05$). The main source of disagreement occurred in the central five slices of each dataset, with the middle slice centered on the bifurcation of the trachea.

Virgincar et al. [64] developed their method for ^{129}Xe MR ventilation and anatomical MR images and tested in on a group of forty four participants: 24 healthy subjects, 10 subjects with COPD, 9 subjects with Global Initiative for Chronic Obstructive Lung Disease (GOLD), and 10 age-matched control (AMC) subjects. Ventilation images were quantified by two methods: the ventilation defect percentage (VDP), as the ratio between the thoracic cavity volume and ventilated volume, was computed from the semi-automatically segmented images; an expert computed the ventilation score percentage (VDS%). For the ventilation images, the intensity histograms from the thoracic cavity volume were analysed, and the coefficient of variation (CV) was computed there. The study showed that there was a correlation between VDS% and VDP ($r = 0.97, P < 0.0001$), and between VDS% and CV ($r = 0.82, P < 0.0001$).

Lelieveldt et al. [28, 29] tested their pipeline on 15 MR scans and assessed the total model matching. However, no separate evaluation of the lung segmentation results was done.

Kullberg et al. [27] used the segmented lungs only as a marker for further detection of adipose tissue in the body. No additional evaluation of the lung segmentation accuracy was done.

Tetzlaff et al. [57] tested their method on 10 datasets from healthy patients and compared the semi-automatically obtained volumes with the spirometric volumes. The comparison showed high agreement (mean Pearson correlation coefficient is ≥ 0.97).

5 Conclusion

The paper provides a review of existing automatic and semi-automatic methods for human lung segmentation from MR data. The categorization of the methods is done according to the main segmentation strategy. The approaches are related and compared to each other from the application point of view. Namely, such aspects as data types, which the techniques are designed for, the level of user involvement and the methods' complexity are discussed. Accuracy of the segmentation methods is crucial according to the nature of the work. The review of performance evaluation approaches is also done. Unfortunately, there is no standardized evaluation system, which would make the methods really comparable to each other. Moreover, for certain application areas there is no existing ground truth, which additionally complicates the evaluation.

References

1. Abdou, I.E., Pratt, W.K.: Quantitative design and evaluation of enhancement/thresholding edge detectors. *Proc. IEEE* **67**(5), 753–763 (1979)
2. Armato, S.G., MacMahon, H.: Automated lung segmentation and computer-aided diagnosis for thoracic ct scans. *Int. Congr. Ser.* **1256**, 977–982 (2003)
3. Avants, B.B., Tustison, N., Song, G.: Advanced normalization tools (ANTs). *Insight J.* **2**, 1–35 (2009)
4. Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C.: A reproducible evaluation of ants similarity metric performance in brain image registration. *Neuroimage* **54**(3), 2033–2044 (2011)
5. Biederer, J., Hintze, C., Fabel, M., Jakob, P., Horger, W., Graessner, J., Bolster, B., Heller, M.: MRI of the lung—ready . . . get set . . . go. *Magnetom Flash* **46**, 6–15 (2011)
6. Biederer, J., Beer, M., Hirsch, W., Wild, J., Fabel, M., Puderbach, M., Van Beek, E.: MRI of the lung (2/3). why . . . when . . . how? *Insights Imaging* **3**(4), 355–371 (2012)
7. Boettger, T., Kunert, T., Meinzer, H.P., Wolf, I.: Interactive constraints for 3d-simplex meshes. In: *Proceedings of SPIE Medical Imaging 2005. Image Processing*, vol. 5747, pp. 1692–1702. International Society for Optics and Photonics (2005)
8. Böttger, T., Grunewald, K., Schöbinger, M., Fink, C., Risse, F., Kauczor, H., Meinzer, H., Wolf, I.: Implementation and evaluation of a new workflow for registration and segmentation of pulmonary MRI data for regional lung perfusion assessment. *Phys. Med. Biol.* **52**(5), 1261 (2007)
9. Böttger, T., Kunert, T., Meinzer, H.P., Wolf, I.: Application of a new segmentation tool based on interactive simplex meshes to cardiac images and pulmonary MRI data. *Acad. Radiol.* **14**(3), 319–329 (2007)
10. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(11), 1222–1239 (2001)
11. Cohen, J., Cohen, P.: *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Lawrence Erlbaum, Hillsdale (1975)
12. Delingette, H.: General object reconstruction based on simplex meshes. *Int. J. Comput. Vis.* **32**(2), 111–146 (1999)
13. Devaki, K., MuraliBhaskaran, V.: Study of computed tomography images of the lungs: A survey. In: *2011 International Conference on Recent Trends in Information Technology (ICRTIT)*, pp. 837–842. IEEE (2011)

14. Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology* **26**(3), 297–302 (1945)
15. Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*. Wiley, New York (2000)
16. Fenchel, M., Requardt, M., Tomaschko, K., Kramer, U., Stauder, N.I., Naegele, T., Schlemmer, H.P., Claussen, C.D., Miller, S.: Whole-body MR angiography using a novel 32-receiving-channel MR system with surface coil technology: First clinical experience. *J. Magn. Reson. Imaging* **21**(5), 596–603 (2005)
17. Gonzalez, R., Woods, R.E.: *Digital Image Processing*. Prentice Hall, New York (2002)
18. Hegenscheid, K., Kühn, J., Völzke, H., Biffar, R., Hosten, N., Puls, R.: Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based ship study. *Rofo* **181**(8), 748–759 (2009)
19. Heydarian, M., Kirby, M., Wheatley, A., Fenster, A., Parraga, G.: Two and three-dimensional segmentation of hyperpolarized ^3He magnetic resonance imaging of pulmonary gas distribution. In: *Proceedings of SPIE Medical Imaging 2012*, vol. 8317. International Society for Optics and Photonics (2012)
20. Horn, B.K.: Closed-form solution of absolute orientation using unit quaternions. *J. Opt. Soc. Am. A* **4**(4), 629–642 (1987)
21. Ibanez, L., Schroeder, W., Ng, L., Cates, J.: *The ITK Software Guide*. Kitware Inc., Clifton Park (2003)
22. Ivanovska, T., Hegenscheid, K., Laqua, R., Kühn, J.P., Gläser, S., Ewert, R., Hosten, N., Puls, R., Völzke, H.: A fast and accurate automatic lung segmentation and volumetry method for mr data used in epidemiological studies. *Comput. Med. Imaging Graph.* **36**(4), 281–293 (2012)
23. Jaccard, P.: The distribution of the flora in the alpine zone. 1. *New Phytol.* **11**(2), 37–50 (1912)
24. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. *Int. J. Comput. Vis.* **1**(4), 321–331 (1988)
25. Kauczor, H.U.: *MRI of the Lung*. Springer, Berlin (2009)
26. Kirby, M., Heydarian, M., Svenningsen, S., Wheatley, A., McCormack, D.G., Etemad-Rezai, R., Parraga, G.: Hyperpolarized ^3He magnetic resonance functional imaging semiautomated segmentation. *Acad. Radiol.* **19**(2), 141–152 (2012)
27. Kullberg, J., Johansson, L., Ahlström, H., Courivaud, F., Koken, P., Eggers, H., Börner, P.: Automated assessment of whole-body adipose tissue depots from continuously moving bed MRI: a feasibility study. *J. Magn. Reson. Imaging* **30**(1), 185–193 (2009)
28. Lelieveldt, B.P.F., van der Geest, R.J., Ramze Rezaee, M., Bosch, J.G., Reiber, J.H.C.: Anatomical model matching with fuzzy implicit surfaces for segmentation of thoracic volume scans. *IEEE Trans. Med. Imaging* **18**(3), 218–230 (1999)
29. Lelieveldt, B.P.F., Sonka, M., Bolinger, L., Scholz, T.D., Kayser, H.W.M., van der Geest, R.J., Reiber, J.H.C.: Anatomical modeling with fuzzy implicit surface templates: application to automated localization of the heart and lungs in thoracic MR volumes. *Comput. Vis. Image Underst.* **80**(1), 1–20 (2000)
30. Lichy, M.P., Mugler, B.M.W.J., Horger, W., Menzel, M.I., Anastasiadis, A., Siegmann, K., Niemeyer, T., Knigsrainer, A., Kiefer, B., Schick, F., Claussen, C.D., Schlemmer, H.P.: Magnetic resonance imaging of the body trunk using a single-slab, 3-dimensional, T2-weighted turbo-spin-echo sequence with high sampling efficiency (SPACE) for high spatial resolution imaging: initial clinical experiences. *Investig. Radiol.* **40**(12), 754–760 (2005)
31. Lorensen, W.E., Cline, H.E.: Marching cubes: a high resolution 3d surface construction algorithm. In: *ACM Siggraph Computer Graphics*, vol. 21, pp. 163–169. ACM, New York (1987)
32. Lowe, D.G., et al.: Fitting parameterized three-dimensional models to images. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(5), 441–450 (1991)
33. Lui, J.K., LaPrad, A.S., Parameswaran, H., Sun, Y., Albert, M.S., Lutchen, K.R.: Semi-automatic segmentation of ventilated airspaces in healthy and asthmatic subjects using hyperpolarized ^3He MRI. *Comput. Math. Methods Med.* **2013**, Article ID 624683 (2013)
34. MacQueen, J.B.: Some methods for classification and analysis of multivariate observations. In: *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297 (1967)

35. Marieb, E.N., Hoehn, K.: *Human Anatomy and Physiology*. Pearson Education, London (2007)
36. Martin Bland, J., Altman, D.: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **327**(8476), 307–310 (1986)
37. McGraw, K.O., Wong, S.: Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* **1**(1), 30 (1996)
38. Memon, N.A., Mirza, A.M., Gilani, S.: *Limitations of Lung Segmentation Techniques*, vol. 27, chap. 76, pp. 753–766. Springer, Berlin (2009)
39. Middleton, I., Damper, R.I.: Segmentation of magnetic resonance images using a combination of neural networks and active contour models. *Med. Eng. Phys.* **26**(1), 71–86 (2004)
40. Mills, G., Wild, J., Eberle, B., Van Beek, E.: Functional magnetic resonance imaging of the lung. *Br. J. Anaesth.* **91**(1), 16–30 (2003)
41. Modersitzki, J.: *Numerical Methods for Image Registration (Numerical Mathematics and Scientific Computation)*. Oxford university press, Oxford (2004)
42. Möller, H.E., Chen, X.J., Saam, B., Hagspiel, K.D., Johnson, G.A., Altes, T.A., de Lange, E.E., Kauczor, H.U.: Mri of the lungs using hyperpolarized noble gases. *Magn. Reson. Med.* **47**(6), 1029–1051 (2002)
43. Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. *J. Comput. Phys.* **79**(1), 12–49 (1988)
44. Plathow, C., Ley, S., Fink, C., Puderbach, M., Heilmann, M., Zuna, I., Kauczor, H.U.: Evaluation of chest motion and volumetry during the breathing cycle by dynamic MRI in healthy subjects: comparison with pulmonary function tests. *Investig. Radiol.* **39**(4), 202–209 (2004)
45. Plathow, C., Schoebinger, M., Fink, C., Ley, S., Puderbach, M., Eichinger, M., Bock, M., Meinzer, H.P., Kauczor, H.U.: Evaluation of lung volumetry using dynamic three-dimensional magnetic resonance imaging. *Investig. Radiol.* **40**(3), 173–179 (2005)
46. Pratt, J.H.: Long-continued observations on the vital capacity in health and heart disease. *Am. J. Med. Sci.* **164**(6), 819–831 (1922)
47. Ray, N., Acton, S.T., Altes, T., De Lange, E.E.: Mri ventilation analysis by merging parametric active contours. In: *Proceedings. 2001 International Conference on Image Processing*, vol. 2, pp. 861–864. IEEE (2001)
48. Ray, N., Acton, S.T., Altes, T., De Lange, E.E., Brookeman, J.R.: Merging parametric active contours within homogeneous image regions for MRI-based lung segmentation. *IEEE Trans. Med. Imaging* **22**(2), 189–199 (2003)
49. van Rijsbergen, C.J.: *Information Retrieval*. Butterworth, London (1979)
50. Sensakovic, W.F., Armato III, S.G.: Magnetic resonance imaging of the lung: automated segmentation methods. In: *General Methods and Overviews, Lung Carcinoma and Prostate Carcinoma*, pp. 219–234. Springer, Berlin (2008)
51. Sensakovic, W.F., Armato III, S.G., Starkey, A.: Automated lung segmentation in magnetic resonance images. In: *Proc. SPIE* **5747**, 1776–1781 (2005)
52. Sensakovic, W.F., Armato III, S.G., Starkey, A., Caligiuri, P.: Automated lung segmentation of diseased and artifact-corrupted magnetic resonance sections. *Med. Phys.* **33**, 3085 (2006)
53. Sluimer, I., Schilham, A., Prokop, M., van Ginneken, B.: Computer analysis of computed tomography scans of the lung: a survey. *IEEE Trans. Med. Imaging* **25**(4), 385–405 (2006)
54. Sodickson, D.K., McKenzie, C.A., Ohliger, M.A., Yeh, E.N., Price, M.D.: Recent advances in image reconstruction, coil sensitivity calibration, and coil array design for smash and generalized parallel MRI. *Magn. Reson. Mater. Phys., Biol. Med.* **13**(3), 158–163 (2002)
55. Soille, P.: *Morphological Image Processing: Principles and Applications*. Cambridge University Press, Cambridge (1999)
56. Sonka, M., Hlavac, V., Boyle, R.: *Image Processing, Analysis, and Machine Vision*. Thomson, Toronto (2008)
57. Tetzlaff, R., Schwarz, T., Kauczor, H.U., Meinzer, H.P., Puderbach, M., Eichinger, M.: Lung function measurement of single lungs by lung area segmentation on 2d dynamic MRI. *Acad. Radiol.* **17**(4), 496–503 (2010)

58. Tokuda, J., Schmitt, M., Sun, Y., Patz, S., Tang, Y., Mountford, C.E., Hata, N., Wald, L.L., Hatabu, H.: Lung motion and volume measurement by dynamic 3d MRI using a 128-channel receiver coil. *Acad. Radiol.* **16**(1), 22–27 (2009)
59. Tustison, N.J., Gee, J.C.: N4itk: Nick's n3 itk implementation for MRI bias field correction. *Insight J.* (2009)
60. Tustison, N.J., Avants, B.B., Flors, L., Altes, T.A., de Lange, E.E., Mugler, J.P., Gee, J.C.: Ventilation-based segmentation of the lungs using hyperpolarized ^3He MRI. *J. Magn. Reson. Imaging* **34**(4), 831–841 (2011)
61. Udupa, J.K., LaBlanc, V.R., Schmidt, H., Imielinska, C., Saha, P.K., Grevera, G.J., Zhuge, Y., Currie, L.M., Molholt, P., Jin, Y.: Methodology for evaluating image-segmentation algorithms. *Med. Imaging 2002 Image Process.* **4684**(1), 266–277 (2002)
62. van Beek, E.J., Wild, J.M., Kauczor, H.U., Schreiber, W., Mugler, J.P., de Lange, E.E.: Functional MRI of the lung using hyperpolarized 3-helium gas. *J. Magn. Reson. Imaging* **20**(4), 540–554 (2004)
63. Vincent, L., Soille, P.: Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(6), 583–598 (1991)
64. Virgincar, R.S., Cleveland, Z.I., Sivaram Kaushik, S., Freeman, M.S., Nouis, J., Cofer, G.P., Martinez-Jimenez, S., He, M., Kraft, M., Wolber, J., et al.: Quantitative analysis of hyperpolarized ^{129}Xe ventilation imaging in healthy volunteers and subjects with chronic obstructive pulmonary disease. *NMR Biomed.* **26**(4), 424–435 (2012)
65. Warfield, S.K., Zou, K.H., Wells, W.M.: Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation. *IEEE Trans. Med. Imaging* **23**(7), 903–921 (2004)
66. Woodhouse, N., Wild, J.M., Paley, M.N., Fichele, S., Said, Z., Swift, A.J., van Beek, E.J.: Combined helium-3/proton magnetic resonance imaging measurement of ventilated lung volumes in smokers compared to never-smokers. *J. Magn. Reson. Imaging* **21**(4), 365–369 (2005)
67. Wu, N.Y., Cheng, H.C., Ko, J.S., Cheng, Y.C., Lin, P.W., Lin, W.C., Chang, C.Y., Liou, D.M.: Magnetic resonance imaging for lung cancer detection: experience in a population of more than 10,000 healthy individuals. *BMC Cancer* **11**(1), 242 (2011)