Computer Analysis of Computer Tomography Images for Lung Nodule Detection

Publication details
Maria Evelina Fantacci, Alessandra Retico
Published online on: 14 Dec 2017

How to cite: Maria Evelina Fantacci, Alessandra Retico. 14 Dec 2017, Computer Analysis of Computer Tomography Images for Lung Nodule Detection from: Handbook of X-ray Imaging, Physics and Technology CRC Press
Accessed on: 05 Dec 2023
62

Computer Analysis of Computer Tomography Images for Lung Nodule Detection

Maria Evelina Fantacci and Alessandra Retico

CONTENTS

62.1 Introduction .............................................................................................................................................. 1259
62.2 Designing a CAD System for Lung Nodule Detection: The Main Building Blocks ......................................................... 1260
   62.2.1 Defining the Target: Nodule Types ........................................................................................................ 1260
   62.2.2 Image Preprocessing .......................................................................................................................... 1262
   62.2.3 Nodule Candidate Identification ....................................................................................................... 1262
   62.2.4 Classification or False-Positive Reduction ....................................................................................... 1263
62.3 Analysis of the CAD System Performance ............................................................................................... 1264
   62.3.1 How to Evaluate the Performance .................................................................................................. 1264
   62.3.2 Public Resources for Algorithm Training and Validation: Data Collection and Challenges ............. 1264
   62.3.3 Survey of CAD Performance ........................................................................................................... 1264
   62.3.4 Winning Strategy in Nodule Detection: Combining Different Algorithms Leads to Improved Performance .......................................................................................................................... 1266
62.4 The Use of a CAD System in Clinical Practice: First or Second Reader? ......................................................... 1267
62.5 Open Issues in CAD Research .................................................................................................................... 1268
62.6 Nodule Tracking ....................................................................................................................................... 1269
62.7 Conclusions and Perspectives .................................................................................................................... 1269
References ...................................................................................................................................................... 1270

62.1 Introduction

Computed tomography (CT) is the imaging modality of choice for screening and staging of lung cancers (see Section III, Chapter 32). Low-dose chest CT is the preferred test for screening members of the asymptomatic population who are at risk of developing lung cancer. Often, early lung cancers originate as small sub-centimeter size lung nodules. Based on their location and size, as well as the experience of the radiologists, identification of small lung nodules on low dose chest CT can be a challenging task, and radiologists may benefit from the use of dedicated computer-aided detection (CAD) systems (see Section IV, Chapter 59). Here, we provide a description of the basic algorithmic components of a CAD system, of the methodological issues to consider in CAD performance evaluation, and of the potential impact of a CAD on the clinical radiology workflow. An overview of the CAD systems recently reported in the literature and a comparison of their performance are also provided. Open issues in this field of research are presented with the view of shortening the path to the release of accessible, fast, and high-performing lung CAD systems.

CAD systems integrated in routine clinical practice of a radiology department could assist physicians in the diagnosis of a large variety of pathologies. A CAD can process the complex information encoded in biomedical images and highlight abnormalities to a human reader who is in charge of annotating the exams.

With the fast advance of imaging technology, the number of diagnostic images to be reviewed by radiologists is increasing, together with the demand for automated tools for image analysis. This happens in many fields of diagnostic radiology and, in particular, in the framework of screening protocols for early detection of lung cancer with low-dose CT (LDCT). CAD systems devoted to the automated identification of lung nodules in chest CTs have been greatly investigated in recent decades (Doi 2005, 2007, Sluimer et al. 2006, Li 2007, Giger 2008, Sprindzuk 2010, Goo 2011, Lee et al. 2012, Suzuki 2012b, Retico 2013, Valente et al. 2016). It is still a matter of debate whether they are more effective in triaging screening CT scan (first reader usage) or they are more valuable to provide radiologists with a second opinion (second reader usage). The National Lung Screening Trial (NLST) showed that screening with LDCT, as compared with chest radiography, reduced lung cancer mortality. In fact, a significant reduction in lung cancer mortality in study participants screened with LDCT with respect to participants screened with X-rays was reported in 2011 (NLST 2010, Kramer et al. 2011). The cost-effectiveness of screening with LDCT in the NLST has also been more recently evaluated (Black et al. 2014). This result encouraged the organization of large-scale screening for lung cancer and constituted a boosted motivation for lung CAD development. For example, the U.S. Preventive Services Task Force (USPSTF) recently updated the recommendation on screening for lung cancer after reviewing
the evidence on the efficacy of LDCT, chest radiography, and sputum cytologic evaluation for lung cancer screening in asymptomatic persons who are at average or high risk for lung cancer (current or former smokers) and the benefits and harm of these screening tests and of surgical resection of early-stage non-small cell lung cancer (Moyer and U.S. Preventive Services Task Force 2014). The USPSTF also commissioned modeling studies (de Koning et al. 2014) to provide information about the optimum age at which to begin and end screening, the optimum screening interval, and the relative benefits and harm of different screening strategies. The result is that the USPSTF recommends (Moyer and U.S. Preventive Services Task Force 2014) annual screening for lung cancer with LDCT in adults aged 55–80 years who have a 30 pack-year smoking history and currently smoke or have quit within the past 15 years.

In screening programs for lung cancer with LDCT, radiologists are required to review a large number of images (section thickness ≤1 mm) for each patient and to identify subtle abnormalities on a very noisy background. Although most solid nodules less than 5 mm do not change recommendation for annual LDCT for lung cancer screening, larger nodules are indeterminate and warrant follow up CT or other imaging based on their size. Key features of lung nodules on LDCT include their size, location, attenuation (such as solid, ground-glass, mixed attenuation, calcified), and change in size over serial exams. Several CAD schemes for automated lung nodule identification have been developed in recent years, and many of them show robust enough performance to be successfully integrated in a lung screening workflow (Bellotti et al. 2007, Dehmeshki et al. 2007, Enquobahrie et al. 2007, Wang et al. 2007, Li et al. 2008, Pu et al. 2008a,b, Retico et al. 2008, 2009, Golosio et al. 2009, Murphy et al. 2009, Suárez-Cuenca et al. 2009, Ye et al. 2009, Messay et al. 2010, Cerello 2010b, Camarlinghi et al. 2011, Tan et al. 2011, Riccardi 2011, Cascio et al. 2012, Lopez Torres et al. 2015).

The research on CAD systems for lung nodule detection and diagnosis has been covered by detailed reviews since early developments (Doi 2005, 2007, Sluimer et al. 2006, Li 2007), including information about the historical context where the CAD concept originated (Doi 2007, Sprindzuk 2010). Different CAD strategies developed in recent years have been presented in detailed review papers (Giger 2008, Sprindzuk 2010, Goo 2011, Lee et al. 2012, Suzuki 2012b, Retico 2013, Valente et al. 2016).

The current availability of extensive computing power resources has allowed computationally-demanding approaches to become executable: for example, decisional algorithms can now be trained on large databases more accurately resembling the underlying target population. In this chapter, we aim to summarize and complement the existing review information in order to provide an overview of the analysis methods implemented to process CT data to automatically extract useful information. Moreover, we will also stress the importance of conducting appropriate tests to assess whether or not CAD systems are effective aids for radiologists, giving an overall idea of the long process, from the design and development of a CAD system for lung nodule detection to its use in clinical practice. We thus start with the description of the basic components required to build a CAD system, then we illustrate the appropriate methods to carry out a CAD performance evaluation and to estimate the potential impact of a CAD system on the workflow of a radiology department. We provide an overview of recent papers describing full systems devoted to lung nodule automated identification and report on recent studies of the impact of CAD systems as first or second readers. Open issues in this field of research are finally highlighted to make the reader confident of the long road traveled to come to this point and the missing final steps that could make a lung CAD a tool the clinicians can no longer do without.

### 62.2 Designing a CAD System for Lung Nodule Detection: The Main Building Blocks

The algorithmic strategies that can be adopted in the automatic analysis of lung CTs to detect lung nodules are many and various, also depending on the kind of nodules being identified. First of all, it is necessary to define the target of a lung CAD. Lung nodules can be categorized according to their shape, location, and possible connections with anatomical structures. Dedicated algorithm combinations have to be developed to build a CAD system with optimized sensitivity to a specific category of nodules in a defined size range. To account for the variability of the nodule shapes and for extended ranges of nodule sizes, combinations of dedicated modules of analysis and multiscale solutions can be designed.

#### 62.2.1 Defining the Target: Nodule Types

Lung nodules belong to two broad categories (Hansell et al. 2008): internal nodules, that is, those embedded in the lung parenchyma, and juxtapleural nodules, that is, those adjacent to the pleural surface of the lung. Internal nodules can be further categorized as: isolated nodules, peri-fissural nodules, and perivascular nodules. Isolated nodules are rather spherical abnormalities completely embedded within the lung parenchyma, without any contact with vessels, fissures, pleura, or other nodules. Peri-fissural nodules show a connection with a pleural fissure. Perivascular nodules appear to be connected to the vascular tree and can be further categorized as nodules characterized by a weak apparent connection to vessels, that is, nodules whose connection to vessels is not obvious from a visual inspection, or nodules attached to vessels whose intensity on CT is similar to or lower than nodular intensity (weakly perivascular nodules), and nodules characterized by a strong apparent connection to vessels, that is, nodules connected to vessels whose intensity visually is greater than nodular intensity (strongly perivascular nodules). Juxtapleural nodules include both hemispherical pleural nodules originating from the pleura and growing towards the lung parenchyma, and pulmonary nodules originating from the lungs with almost spherical shape, connected to the pleura through a tail or with abutment of pleural surface. Some examples of different nodule types extracted from CT exams of the Lung Image Database Consortium (LIDC) database are shown in [Figure 62.1](#).

Lung nodules can also be distinguished according to their CT contrast into solid and part-solid or ground-glass nodules (ground-glass opacities, GGOs). As lung nodules may differ either in shape or intensity, a CAD developer should characterize them by a large variety of features that the algorithms can interpret. Regardless of the type of nodule a CAD algorithm is optimized for, it can generally be schematized in three steps, as described in [Figure 62.2](#): first the lung CT volume undergoes a preprocessing step, including
filtering and resampling algorithms if necessary and, in many cases, the lung parenchyma segmentation (i.e., the identification of the lung tissue with respect to the surrounding different anatomical structures); then, the initial selection of nodule candidates is performed; finally, as many false-positive findings as possible are eliminated from the list of nodule candidates. Each step of this basic scheme is discussed in detail in the following sections.

In many cases, CAD systems are developed with the aim of detecting all types of nodules at the same time. However, specific approaches can be targeted to each nodule type, for example, CAD$_I$ for internal nodule detection (Retico et al. 2008) and CAD$_{JP}$ for juxtapleural nodule detection (Retico et al. 2009), with a final integration of the synergic algorithms (Camarlinghi et al. 2011), as depicted in Figure 62.3.
62.2.2 Image Preprocessing

Many CAD developers implement image filters as the first step of the analysis. Filters span from Gaussian filters used to reduce noise in the images, to resampling strategies adopted to have isotropic voxel sizes. Procedures to enhance interesting structures may also be applied at this stage, such as spherical-shaped object enhancement filters (Ochs et al. 2007) or multiscale enhancement filters, to suppress blood vessels and highlight nodule-like structures (Li et al. 2003, 2008; Retico et al. 2008). Sophisticated filtering strategies have also been adopted to account for the large variability of real nodules with respect to the nodule models adopted in conventional filtering procedures. In particular, a supervised filter based on the massive-training artificial neural network approach (Suzuki 2003) was trained with actual nodules to enhance actual patterns of nodules.

A widely-used procedure in the preprocessing step is lung segmentation: this consists of identifying the target volume for the automated search of nodules. The aim of this segmentation task is both to reduce the computational costs of the CAD execution and to avoid CAD marks pointing at anatomical structures external to the lungs. There are studies fully dedicated to the development of specific algorithms for accurate lung parenchyma identification (Pu et al. 2008a,b, Golosio et al. 2009, van Rikxoort 2009, De Nunzio et al. 2011) and an extensive review on automated segmentation of different lung structures (van Rikxoort and van Ginneken 2013). The identification of lung parenchyma is a non-trivial task, especially when an elderly population is investigated, due to the complexity of the anatomical structures, especially in the presence of underlying lung diseases (Pu et al. 2008a,b, van Rikxoort 2009, van Rikxoort and van Ginneken 2013). The accuracy of the lung segmentation step is crucial, especially when juxtapleural nodules have to be detected. In this case, the lung segmentation step is required either to include or to exclude the pleural abnormalities (including juxtapleural nodules) in the segmented volume, depending on the specific design of the detection algorithm. The performance of the lung segmentation step limits the maximum sensitivity of the CAD systems. Keeping in mind that the sensitivity is defined as the percentage of nodules correctly identified by the CAD system, it is evident that lung nodules located outside of the segmented volume at this stage will be definitely missed by the detection systems. The validation of lung segmentation algorithms is far from being trivial, requiring an extremely time-consuming manual segmentation task by an experienced radiologist to define a ground truth for comparison. A large public database with a well-defined ground truth would be needed to evaluate lung segmentation algorithms and compare each other’s performance.

The segmentation of structures in medical images still represents a challenging task many research groups try to deal with, and the implementation of innovative methods is more than welcome (Pu et al. 2011, Masala et al. 2013, van Rikxoort and van Ginneken 2013).

62.2.3 Nodule Candidate Identification

As shown in Figure 62.1, most CAD programs perform an initial selection of nodule candidates, which is a crucial task whose performance, in addition to that stated for the segmentation step, puts an upper limit on the nodule detection sensitivity. Image processing techniques, devoted to highlighting suspicious regions of the images, are implemented. The approaches can be broadly divided into the following main categories: intensity-based algorithms, shape-based methods, and template-matching procedures.

Intensity-based algorithms exploit the CT intensity information, which is quantitative information as it is expressed in Hounsfield Units (HU). Such algorithms are based on the information that lung nodules present with a higher relative intensity with respect to lung parenchyma. Among the intensity-based methods applied to identify nodule candidates are the multiple gray-level thresholding techniques (Armato, Giger and MacMahon 2001, Messay, Hardie and Rogers 2010). In these techniques, the segmented lung volume is applied to a series of gray-level thresholds, leading to a series of thresholded lung subvolumes, which are selected as initial lung nodule candidates if they satisfy a criterion on volume size. The intermediate candidate masks obtained for each threshold in the multiple-threshold approach can then be processed by a specific morphological opening operation and finally combined by the logical disjunction (OR) operation to obtain the final nodule candidate masks (Messa, Hardie and Rogers 2010). The initial selection of nodule candidates can also be achieved by means of a multi-threshold surface-triangulation approach (Golosio et al. 2009). In this case, the surface triangulation is performed at different threshold values over a wide range. At each threshold value, a nodule candidate is defined as the volume inside a connected component of the triangulated isosurface (Suárez-Cuenca et al. 2009). Multi-threshold nodule candidates are defined as a path of the tree-like structure that represents the evolution of a nodule candidate as a function of the threshold values. Instead of applying multiple thresholds, it is also possible to execute a local-adaptive thresholding approach (Lee et al. 2001, Messay et al. 2010); in this approach, to detect nodule candidates the segmented lung volume is processed by an adaptive threshold that automatically converges to the optimum gray-level value that separates the higher density regions from the background (Suárez-Cuenca et al. 2009). Intensity thresholding can be combined with morphological processing (Bellotti, De Carlo and Gargano 2007, Messay, Hardie and Rogers 2010) or rule-based pruning (Choi and Choi 2012) to identify and segment nodule candidates simultaneously. Another intensity-based method for nodule candidate detection is the region-growing algorithm, where seed points are iteratively chosen from the segmented lung volume (Bellotti et al. 2007). In this method, the segmented volume is scanned until a voxel satisfying the region-growing inclusion rule is found; that voxel is used as a seed point and the growth starts; once the region is completely grown, it is removed from the CT and stored for further analysis; then, the search for new seed points is iterated until no more seed points satisfy the inclusion rule. The second category consists of techniques that exploit the shape properties of the objects to be detected, in addition to their intensity. Nodules are considered to be nearly spherical objects and thus suitable features that highlight the local sphericity characteristic of image regions are computed to identify nodule candidate locations. Since the pioneering work of Li, Sone, and Doi (2003), many authors have implemented the selective filter to enhance nodules and suppress normal anatomic structures.
Computer Analysis of Computer Tomography Images for Lung Nodule Detection

62.2.4 Classification or False-Positive Reduction

Once the list of nodule candidates has been populated by the algorithms described in the previous section, a dedicated analysis is necessary to prune from the list candidates not corresponding to real nodules (i.e., what we will call false-positive findings at the end of the CAD execution). Nodule candidates have to be characterized in terms of image properties, often referred to as features, which are generally investigated by machine-learning approaches. This procedure is referred to as nodule candidate classification: each entry of the initial list of candidate nodules is classified either as a CAD finding, according to a certain degree of suspicion, or as a portion of normal tissue. The amount of nodule candidates corresponding to false-positive detections in the list of nodule candidates can be quite large, depending on the quality of the candidate identification algorithm and on the total amount of list entries a CAD developer decides to maintain up to the next stage to avoid limiting the CAD sensitivity a priori. The classification step is required to reach a compromise between the need to have high sensitivity to lung nodules and the need to avoid showing too many confounding CAD marks to the radiologists. The aim of this step of the analysis is to improve the specificity of the CAD system, while keeping the sensitivity as high as possible.

Among the numerous different strategies that are generally implemented, we can recognize two main categories of nodule characterization approaches: the first one aims to describe the nodule candidates by global shape- and/or intensity-based features (Bellotti et al. 2007, Dehmeshki et al. 2007, Enquobahrie et al. 2007, Osman et al. 2007, Li et al. 2008, Murphy et al. 2009, Golosio et al. 2009, Retico et al. 2009, Suárez-Cuenca et al. 2009, Ye et al. 2009, Messay et al. 2010, Tan et al. 2011, Riccardi 2011, Cascio et al. 2012, Cao et al. 2014); the second one is based on the characterization of the nodule candidate at the voxel level (Gori et al. 2007, Retico et al. 2008, 2009, Suzuki 2003, 2009) (see Figure 62.4). Both approaches require the segmentation of nodule candidates, the feature extraction, and finally, the feature classification. The accuracy of the nodule segmentation step is particularly important in the first approach, where the computed features strictly depend on the shape of the segmented nodule candidate. The accurate segmentation of lung nodules is even more important to reliably evaluate the nodule doubling time; thus, research on this particular issue is very active...
The approaches to false-positive reduction based on single voxel analysis are, in principle, less affected by nodule segmentation accuracy. Thanks to the current increasing availability of computational power, the current trend is to classify every single voxel of a diagnostic image, also bypassing the candidate selection and characterization steps (Suzuki 2012a). However, this choice is still not efficient in the case of 3D CT data, such as the CT data acquired in lung cancer screening. The global features usually computed on the nodule candidates are gray-level based features, texture features, and morphological features. They are all encoded in a vector of features and finally classified by machine learning-based methods. In contrast, when characterizing nodule candidates at the voxel level, each voxel is assigned a number of features extracted from its neighborhood, then each voxel is classified by a decisional system and a majority criterion has to be implemented to finally assign the nodule candidate to either the nodule or healthy tissue class.

The nodule candidates are usually represented in terms of the vectors of features extracted from either each nodule candidate or each voxel of each nodule candidate. For their classification, pattern-recognition techniques are generally implemented. Among them, artificial neural networks (ANNs) are very popular (Bishop 1990), and linear discriminant analysis (LDA) (Fukunaga 1990) and support vector machines (Vapnik 1995) are also widely used.

The choice of the minimal number of features to properly characterize the nodule candidates and carry out an efficient reduction of false-positives is a nontrivial task, and it is still a matter of investigation. If many features are considered, possible drawbacks are high computational time and overfitting of the classifiers. Within-class imbalance and high-dimensionality problems may reduce the classification performance, thus suitable learning strategies have to be adopted (Cao et al. 2014). By contrast, if an overly restrictive set of features is chosen to characterize the nodule candidate, it could not encode all necessary information to get an accurate classification performance. The choice of the number of features to characterize the nodule candidates and the design of the appropriate classifier to handle them require practical considerations to be made customized to the available set of training cases. CAD developers have to take into account the finite size of available cases to train, test, and validate the CAD system, and that the sample size affects the classifier performance (Sahiner 2008).

62.3 Analysis of the CAD System Performance

62.3.1 How to Evaluate the Performance

Once a CAD system has been designed and developed, it is crucial to make a reliable estimate of its standalone performance. Although CAD systems have been developed to assist radiologists in their task, it is necessary to fully characterize the CAD performance before its impact on radiologists’ diagnostic ability can be evaluated. Moreover, although research on CAD has been very active in recent years, it is difficult to carry out a fair performance comparison among the systems developed by different research groups. A reliable comparison of the nodule detection ability of different CAD systems could only be done on common databases. The rapid growth of CAD algorithms over recent decades has simultaneously triggered the development of the appropriate methodology to carry out CAD performance assessment and also highlighted the necessity of the availability of a large, clinically relevant and annotated database to CAD developers to test and validate their systems.

62.3.2 Public Resources for Algorithm Training and Validation: Data Collection and Challenges

CAD researchers very often develop their algorithms on private lung CT databases collected and annotated in medical centers accessible within specific research networks and projects. However, it is currently also possible to access public databases of annotated CT images. Useful repositories of annotated lung CT images are those released by the Early Lung Cancer Action Program in 2003 (ELCAP 2016) and by the LIDC (LIDC 2016), funded in the USA by the National Cancer Institute, in 2000. The aim of the LIDC was to develop consensus guidelines for a spiral CT lung image resource and to construct a database of spiral CT lung images (Armato et al. 2004). Another dataset of lung CT scans is accessible within the ANODE09 (ANODE 2009) initiative (see Section 62.3.4) (van Ginneken et al. 2010), the first challenge for lung CADs organized in 2009 in the framework of the SPIE (Society of Photo-Optical Instrumentation Engineers) Conference, which contains data from the NELSON study, the largest CT lung cancer screening trial in Europe. In this case, the CT annotations are not available on the website; however, CAD developers can submit the output of their systems for each case to the organizers and the CAD performance will be evaluated and published on the website. After the ANODE09 experience, other challenges on common datasets were organized in 2015 (LUNGx 2015) in the framework of the SPIE Conference, with the support of the American Association of Physicists in Medicine (AAPM) and the National Cancer Institute (NCI), and in 2016 (LUNA 2016) in the framework of the International Symposium of Biomedical Imaging (ISBI) Conference. All these initiatives promote the early detection of lung cancers by supporting the research of image processing tools and CAD and the diagnosis of lung nodules.

62.3.3 Survey of CAD Performance

A number of relevant CAD systems were developed before public samples of lung CT scans became available. We will report on examples of those systems, as well as on more recent studies where the CAD systems were validated on public data sources. It is evident that, when a CAD system is developed and validated on private datasets, a fair comparison of its performance with preexisting literature is not possible. Although largely encouraged, it is not yet common practice for all CAD developers to validate their CAD systems on a public dataset and to discuss the results achieved in comparative research papers.

Many valuable CAD approaches have been implemented and tested by means of private datasets (Bellotti et al. 2007, Dehmeshki et al. 2007, Li et al. 2008, Pu et al. 2008a,b, Murphy et al. 2009, Suárez-Cuencia et al. 2009, Camarlinghi et al. 2011). Bellotti et al. developed a CAD system based on the region growing method and an original active contour model with a local...
Rule-based classifiers and ANNs are implemented in a leave-one-out cross-validation protocol. The system reached a sensitivity of 88.5% with 6.6 false-positives per scan (FP/scan) on 15 CT scans containing 26 nodules, acquired within the ITALUNG-CT trial, the first Italian randomized controlled trial for the screening of lung cancer (Lopes Pegna et al. 2009). Dehmeshki et al. implemented a shape-based genetic algorithm template-matching method to detect nodules with spherical elements (Dehmeshki et al. 2007). Lung nodule phantom images were used as reference images for template matching. This method has been validated on a private clinical dataset of 70 thoracic CT scans containing 178 nodules, achieving a sensitivity of 90% with 14.6 FP/scan. Pu et al. developed a detection procedure in the signed distance field of the image (Pu et al. 2008a, b). Nodule candidates are identified by searching local maxima of signed distances; detected candidates are then classified according to the similarity distance of their medial axis-like shapes achieved through a progressive clustering strategy combined with a marching cubes algorithm from a sphere-based shape. This system reached a sensitivity of 81.5% with 6.5 FP/scan on a dataset of 52 CT scans with 184 nodules, including 16 nonsolid nodules. Li et al. implemented the selective nodule enhancement filter in the nodule candidate identification procedure and automated rule-based classifier to reduce false-positive findings (Li, Li and Doi 2008). A case-based fourfold cross-validation testing method led to an overall sensitivity of 86% with 6.6 FP/scan on a dataset of 117 CT scans with 153 nodules, including both solid and nonsolid nodules. Suárez-Cuenca et al. developed a CAD system based on the capability of an iris filter to discriminate between nodules and false-positives (Suárez-Cuenca et al. 2009). Suspicious regions were characterized with features based on the iris filter output, gray level, and morphological features, and finally classified by LDA. The system reached a sensitivity of 80% with 7.7 FP/scan on an independent validation dataset of 22 CT scans containing 77 nodules. Murphy et al. proposed an algorithm based on local image features of shape index and curvedness to detect nodule candidates and two successive k-nearest neighbor classifiers to reduce false-positive findings (Murphy et al. 2009). The CAD was trained and tested on three datasets extracted from a large-scale experimental screening study. The system performance was evaluated on a random selection of 813 scans, leading to a sensitivity of 80% with an average 4.2 FP/scan. The extensive training and validation of the system on large datasets of nodules of varying sizes, types, and textures allows for a realistic measure of the CAD system performance in low-dose screening CT studies. Camarlinghi et al. developed two dedicated and integrated procedures to detect separately isolated and juxtapleural lung nodules (Camarlinghi et al. 2011). The selective nodule enhancement filter and the directional gradient concentration approach are used in the internal and pleural nodule candidate selection steps, respectively. A sensitivity of 70% with an average of 3 FP/scan was reached on an independent validation set of 20 CT scans of the ITALUNG-CT screening trial, containing 38 nodules.

Several other CAD systems were recently developed using the LIDC database (Ge 2005, Golosio et al. 2009, Sahiner et al. 2009, Messay et al. 2010, Riccardi 2011, Tan et al. 2011, Cascio et al. 2012, Choi and Choi 2012, 2014, de Carvalho Filho et al. 2014, Lopez Torres et al. 2015, Jacobs et al. 2016). Sahiner et al. used adaptive 3D clustering and a 3D active contour algorithm to detect nodule candidates; clustering initialized by k-means was then used to segment the nodule candidates, further characterized by three groups of features (Sahiner et al. 2009). An automated discrimination between internal and juxtapleural nodules was performed before applying rule-based classification and LDA to both the juxtapleural and internal nodule candidates. Validation of the CAD system on an independent dataset of 52 CT scans containing 241 nodules in the 3–18.6 mm diameter range achieved 54% sensitivity with 5.60 FP/scan. Golosio et al. detected the nodule candidates by means of a multi-threshold surface-triangulation approach; shape- and intensity-based features were then computed at each threshold on nodule candidates and classified by ANNs (Golosio et al. 2009). Following a twofold cross-validation protocol, the system achieved a 79% sensitivity at 4.0 FP/scan in the detection of nodules with a diameter greater than or equal to 3 mm on a dataset of 84 CT scans containing 148 nodules. Messay et al. implemented a sequential forward selection process to select the optimum features for LDA and quadratic discriminant analysis (Messay et al. 2010). They reached a sensitivity of 83% with 3 FP/scan on a dataset of 84 CT scans containing 143 nodules in a sevenfold cross-validation test. Riccardi et al. developed a 3D fast radial filtering system to select nodule candidates and a heuristic approach based on geometric features, followed by a support vector machine for classification (Riccardi 2011). Following a twofold cross-validation protocol, the CAD system achieved a sensitivity of 71% with 6.5 FP/scan on a dataset of 154 CT scans containing 117 nodules. Tan et al. implemented a feature-selective classifier based on a genetic algorithm and ANNs for classification (Tan et al. 2011). They achieved a sensitivity of 87.5% with 4.0 FP/scan on an independent validation set of 125 CT scans containing 80 nodules. Choi and Choi developed a classifier based on genetic programming to process 2D and 3D features of nodule candidates detected by optimal multiple thresholding and rule-based pruning applied to the lung volume segmented by thresholding and 3D-connected component labeling (Choi and Choi 2012). They obtained 94.1% sensitivity with 5.45 FP/scan on a subset of 32 CT scans containing 76 nodules in the 3–30 mm range. Cascio et al. used a stable 3D mass-spring model combined with a spline curve reconstruction process to identify nodule candidates, taking into account both intensity and shape information (Cascio et al. 2012). A double-threshold cut on candidate features and a neural classifier were then implemented to reduce false-positives. The CAD performance was evaluated through a cross-validation procedure on a dataset of 84 CT scans containing 148 nodules with a diameter greater than or equal to 3 mm, obtaining sensitivity values of 97% and 88% with 6.1 and 2.5 FP/scan, respectively, at two different operative points the CAD system can run on. Choi et al. refined the method previously described (Choi and Choi 2014) implementing a novel 3D shape-based feature descriptor. After lung volume segmentation, nodule candidates are detected using multi-scale dot enhancement filtering in the segmented lung volume. Next, feature descriptors are extracted from the detected nodule candidates, and these are refined using an iterative wall elimination method (Choi and Choi 2014). Finally, a support vector machine-based classifier is trained to classify nodules and...
non-nodules. This method has been validated on a dataset of 84 CT scans with 148 nodules with a diameter greater than or equal to 3 mm achieving 97.5% sensitivity with 6.76 FP/scan. de Carvalho Filho et al. (2014) developed a three-stage method: in the first stage, the extraction and reconstruction of the pulmonary parenchyma is carried out and then enhanced to highlight its structures. In the second stage, nodule candidates are segmented. In the third and final stage, shape and texture features are extracted, selected, and then classified using a support vector machine. To validate this system, 140 CTs from the LIDC database have been used (80% of which were used for training and 20% used for testing), obtaining a sensitivity of 85.91% with a false-positive rate of 1.82 per CT. Lopez Torres et al. (2015) implemented a CAD system (MSL lung CAD) by combining two independent subsystems (Camarlinghi et al. 2012) based on the Channeler Ant Model as a segmentation tool and on the Voxel-Based Neural Approach for classification. The lungCAM (Channeler Ant Model) CAD was upgraded with a scan equalization module and a new procedure to recover the nodules connected to other lung structures; its classification module, which makes use of a feed-forward neural network, is based on a small number of features (13), so as to minimize the risk of lacking generalization, which could be possible given the large difference between the size of the training and testing datasets, containing 94 and 1019 CT scans, respectively. The MSL lung CAD performance was extensively tested on 1043 CT scans from three different datasets, including a detailed analysis of the full Lung Image Database Consortium/Image Database Resource Initiative database, obtaining a performance consistent across the databases, with a sensitivity of 80% at eight false-positive findings per scan, despite the variable annotation criteria and acquisition and reconstruction conditions. A reduced sensitivity is found for subtle nodules and GGO structures. Jacobs et al. (2016) report the performance of two commercial and one academic CAD system tested on 888 thoracic CT scans with a section thickness of 2.5 mm or lower from the LIDC database. The CAD systems compared are: the commercial CAD system Visia (MeVis Medical Solutions AG, Bremen, Germany), a commercial prototype CAD system Herakles (MeVis Medical Solutions AG, Bremen, Germany), and an academic nodule CAD system ISICAD (Utrecht Medical Center, Utrecht, The Netherlands). The updated commercial prototype showed the best performance with a sensitivity of 82% at an average of 3.1 FP/scan.

From the previously reported results, we can see that even when the publicly available LIDC database is used for CAD development and validation, the number of considered cases is not consistent among different studies. Not considering the different cardinality of the training and test sets among the various studies, the performance achieved on the LIDC datasets spans the 54\%—97.5\% range of sensitivity and 1.82—6.5 FP/scan. On private datasets of lung CT scans, values of sensitivity are in the 70\%—90\% range and values of FP/scan of 3—14.6 have been reported.

Among the most challenging tasks reported by CAD developers is the difficulty in the identification of nodules with low contrast with respect to the surrounding tissue, for example, GGO nodules, especially when strongly connected to the vasculature or to the pleura surface. These types of nodules often happen to be false-negatives of CAD systems and still represent an open issue in lung CAD research.

### 62.3.4 Winning Strategy in Nodule Detection: Combining Different Algorithms Leads to Improved Performance

The large variety of CAD strategies for lung nodule detection adopted in recent years strongly suggests the absence of a single method that clearly outperforms the others. Each approach has its own strengths and weaknesses, and the complexity and variety of nodule appearance with respect to anatomical location, contrast, and geometry causes each detection algorithm to fail to reach 100\% sensitivity. Despite this consideration, such systems still represent useful aids in the diagnostic image reviewing process. They have to be accurately set up and their performance has to be evaluated according to common criteria shared by different research groups. In general, among the large variety of CAD approaches, those that are less affected by the presence of empirical thresholds set on available data are preferable because they guarantee a better generalization ability. Now that large datasets of CT data are available, it is hoped that researchers will reach an agreement on which data to use in the different steps of the training procedures to carry out highly instructive comparisons of different methods, with the possibility of also mixing and merging the different procedures implemented at each stage of the analysis.

Analysis systems implementing different approaches to detect lung nodules may generally be characterized by high sensitivity to a particular nodule type, while being less accurate in detecting nodules with different characteristics. Mimicking the multiple reader algorithm often used in diagnostic protocols, lung CT data can be processed by different CAD systems based on different analysis strategies, and their outputs can be compared and possibly automatically combined (van Ginneken et al. 2010, Niemeijer et al. 2011, Camarlinghi et al. 2012). The combination of different systems extends the concept of the multiple classification problem, a well-known issue in pattern recognition research (Kittler 1998). The classification step is only the final task of a CAD algorithm, and the maximum sensitivity of the system has often already been reduced at that stage. The potential beneficial impact of the combination of different approaches lies in the principle that even less-performing systems can contribute to the enhancement of the overall detection accuracy.

To avoid the mere summation of false-positive findings, the outputs of different analysis systems have to be merged according to an appropriate criterion.

When two independent systems analyze two non-overlapping or partially overlapping regions of interest, such as the lung parenchyma and the pleura surface, searching for different kinds of nodules (Retico et al. 2008, 2009), the output of the two systems could, in principle, merely be merged. However, in the choice of the operating point at which each system works, the relative weight between the two systems has to be at least empirically determined (Camarlinghi et al. 2011).

General practical methods to combine many system outputs have been proposed by van Ginneken et al. (2010) and Niemeijer et al. (2011). Those authors organized a lung CAD challenge, the ANODE09 competition (ANODE 2009) where CAD developers were asked to blindly validate their CAD systems on a wide lung screening dataset (55 CT scans acquired in the framework of the NELSON study) by uploading their CAD findings on a web-based framework for CAD performance evaluation. As reported
in van Ginneken et al. (2010), the combination of the outputs of algorithms, even characterized by different performances, led to an overall improvement in nodule detection ability.

A similar study conducted on 138 cases of the LIDC database demonstrated that the combination of the output of three different CAD systems (Bellotti et al. 2007, Cerello 2010a,b, Camarlinghi et al. 2011), contributes to the reduction of false-positives by exploiting the CAD complementarity (Camarlinghi et al. 2012). A dedicated plug-in for the OsiriX open-source DICOM viewer (OSIRIX 2016) has also been developed to interactively review each CAD output and their combination at different operating points (Camarlinghi et al. 2011).

More recently, C. Jacobs, A.A. Adiyoso Setio, A. Traverso, and B. van Ginneken organized the LUNA (Lung Nodule Analysis) 2016 Challenge (LUNA 2016), a completely open challenge in which images and reference standards are publicly available. The goal of LUNA16 was to provide an opportunity for participants to test their algorithm on a common database with a standardized evaluation protocol. For this challenge, the publicly available LIDC/IDRI database has been used. In total, 888 CT scans have been included. Scans with a slice thickness greater than 3 mm have been excluded, along with scans with inconsistent slice spacing or missing slices. The complete dataset has been divided into 10 subsets that should be used for the 10-fold cross-validation.

The research community has been invited to participate in one or two of the following challenge tracks:

1. **Nodule detection**: using raw CT scans, the goal is to identify locations of possible nodules, and to assign a probability for being a nodule to each location. The pipeline typically consists of two stages: candidate detection and false-positive reduction.

2. **False-positive reduction**: given a set of candidate locations, the goal is to assign a probability of being a nodule to each candidate location. Hence, one could see this as a classification task: nodule or not a nodule. Candidate locations will be provided in world coordinates.

3. **False-positive reduction (extended)**: after discussions during the dedicated workshop at the ISBI conference, a new set of candidates has been released. This new set achieves a substantially higher detection sensitivity (1162/1186 nodules) to further improve the overall performance of the submissions.

The first results of this challenge were presented in a dedicated workshop (ISBI 2016) of the ISBI 2016 Conference, and the overall results of the combinations of the outputs of algorithms are published in (Adiyoso Setio et al. 2017).

### 62.4 The Use of a CAD System in Clinical Practice: First or Second Reader?

The extremely large population that could be enrolled in screening programs for lung cancer with low-dose CT makes the use of CAD systems to assist the radiologists quite necessary to make the programs cost-effective and thus affordable for the National Health Systems. CAD systems have the potential to improve the radiologist’s performance in detecting lung cancer. Especially in large-scale screening programs of an asymptomatic population, subtle early-stage lung nodules can be overlooked (Li et al. 2002). The efficacy of CAD systems in detecting nodules missed in CT screening programs has been demonstrated (Armato et al. 2002, Li et al. 2005).

CAD systems can, in general, be used as first or second readers of diagnostic images. In both cases, to assess the potential clinical usefulness of a CAD system, observer studies have to be conducted. They consist of estimating the diagnostic accuracy of radiologists with and without the use of CAD systems.

When the CAD is used as first reader, the CAD acts as a pre-selector of the most suspicious cases to be directly pointed to further diagnostic investigation. The recent work by Ritchie et al. (2016) reports on the possibility of a technician carrying out the first reading of lung cancer screening CT scans, assisted by computer vision tools. The technician reviewed 828 scans, achieving a sensitivity for abnormal scans of 97.8% and a specificity of 98%. Moreover, aided by computer vision, the technician correctly detected 92.9% of the 112 prevalent nodules identified as malignant in the follow up, compared with the 84.8% of the study radiologists.

To evaluate the impact of CAD as a second reader, the observer performance has to be evaluated before and after the CAD output is shown. Usually, two to ten radiologists with different levels of experience in image annotation participate in the study. The possible improvement of the radiologist’s performance when using CAD has to be evaluated according to statistically relevant protocols. For example, the jackknife free-response receiver operating characteristic (JAFROC) analysis can be implemented (Chakraborty 2004, 2006). The JAFROC figure of merit is the probability that lesions are rated higher than false-positive marks on normal images. With respect to other statistical comparison procedures, the JAFROC method has shown an improved statistical power that can allow for a reduced sample size requirement for the observer study.

Both radiologists and CAD systems have to be evaluated against a dataset of annotated CT cases. The consensus achieved by two experienced radiologists or by an expert panel is usually implemented as the reference standard.

Numerous studies report that the use of CAD systems improved the radiologist’s diagnostic accuracy (Beyer et al. 2007, Das et al. 2008, Hirose et al. 2008, White et al. 2008, Sahiner et al. 2009). The impact of CAD systems on the diagnostic workflow has been demonstrated to be beneficial, especially in lung screening settings (Beigelman-Aubry et al. 2007, Fraioli et al. 2007, Bogoni et al. 2012). The observer study of Beigelman-Aubry et al. on 54 pairs of CT scans containing nodules with diameters ≥4 mm, involving two radiologists and a commercial CAD system, demonstrated an improvement in the radiologist’s sensitivity, while the overall reading time on the CAD workstation and on the clinical workstation was comparable (Beigelman-Aubry et al. 2007). Beyer et al. investigated the impact of a commercial CAD system on radiologists’ sensitivity and reading time (Beyer et al. 2007). The study involved four radiologists reviewing 50 CT scans containing lung nodules with diameters ≥4 mm. When using CAD as a second reader, the radiologists’ sensitivity improved from 68% to 75% at the cost of longer reading times. Fraioli et al. conducted an observer performance study on 200 CT scans with screening characteristics (Fraioli et al. 2007). The study involved
three radiologists with variable experience, whose sensitivity to lung nodules had been found to be 57, 68, and 46%. With CAD used as a second reader, sensitivity significantly improved up to 94, 96 and 94%, respectively. Hirose et al. conducted a JAFROC study on six expert radiologists reviewing the CT exams of 21 subjects (Hirose et al. 2008). The standalone CAD sensitivity to lung nodules was 71.4% with 0.95 FP/scan. The average radiologist sensitivity improved significantly from 39.5% to 81.0% when using CAD. Das et al., in an observer study involving two radiologists reviewing 77 CT exams, showed a benefit in the sensitivity of the two radiologists of between 8 and 24% when using CAD (Das et al. 2008). White et al. conducted a multicenter observer performance study involving 10 radiologists with variable levels of experience reviewing 109 CT cases acquired with different CT scanners, containing nodules with diameters ≥4 mm (White et al. 2008). The average increase in the performance for the 10 readers with CAD software was 1.9% for a 95% CI (0.8%–8.0%). Sahiner et al. carried out a JAFROC experiment on a dataset of 85 CT exams involving six radiologists (Sahiner et al. 2009). A statistically significant improvement in radiologists’ average sensitivity from 56% at 0.67 FP/scan to 67% at 0.78 FP/scan was achieved on nodules with diameters ≥3 mm, whereas no significant improvement was obtained on nodules with diameters ≥5 mm, thus indicating an enhanced sensitivity to small nodules, which are those more easily overlooked on visual inspection. Bogoni et al. evaluated the impact on radiologists’ performance of a commercial CAD system integrated with the picture archiving and communication system (PACS) environment (Bogoni et al. 2012). The average sensitivity of five radiologists evaluated on 48 CT exams improved from 44 to 57% for nodules with diameters ≥3 mm, whereas the number of false-positive detections significantly increased for only two radiologists. The integration of CAD into the PACS increased reader sensitivity with minimal impact on interpretation time.

Two radiologists and three different CAD systems were studied to find the best pairing in terms of the highest sensitivity of first and second reader in the study by Christe et al. (2013). Highest sensitivities (between 97 and 99.0%) were achieved by combining any radiologist with any CAD (i.e., second reader use). Combining any two CAD systems (i.e., first reader use), sensitivities were significantly lower (between 85 and 88%) than for radiologists combined with CAD.

Recent work by Matsumoto et al. (2013) compared the performance of six radiologists, including three experienced radiologists, and a CAD on 50 CT scans. It is the first study addressing CAD performance as first and second reader using the JAFROC analysis. The study has not revealed significant differences between the performance in the two cases, with an average figure of merit of 0.70 for first reader use and of 0.72 for second reader. The average time taken for the diagnosis was significantly shorter in the first case (132 seconds) than in the second (210 seconds).

### 62.5 Open Issues in CAD Research

While tremendous work has been accomplished in the domain of CAD for lung nodules in chest CT, currently, the cost of applying commercial CAD systems is not reimbursable in the United States. This implies that the radiology departments have to absorb the CAD costs. This should also be an important issue for all CAD related research.

Nodules with a GGO appearance (part-solid and part-nonsolid nodules), as described before, are particularly difficult to detect. They are characterized by low contrast with respect to the background and ill-defined boundaries, thus they appear subtler with respect to solid nodules and are often missed by analysis algorithms. The likelihood of malignancy for GGO nodules is much higher than that for solid nodules (Henschke 2002).

Despite some CAD systems including nonsolid nodules among the detected nodules (Dehmeshki et al. 2007, Li et al. 2008), dedicated approaches have been also attempted (Kim et al. 2005, Ye et al. 2007, Jacobs 2011). Kim et al. proposed a method to detect pure nonsolid nodules based on the analysis of overlapping regions of a CT scan by computing 2D texture and Gaussian fitting features, afterwards classified by ANNs (Kim et al. 2005). Ye et al. implemented geometric shape features (e.g., shape index and dot enhancement) calculated for each voxel and then thresholded to detect nodule candidates; false-positives were removed by rule-based filtering (Ye et al. 2007). On a dataset of 50 CT scans containing 52 GGO nodules, the system achieved a sensitivity of 92.3% at 12.7 FP/scan. This dedicated approach has been further refined by Ye et al., where a CAD system of a dataset of clinical lung CT was developed and validated to detect both solid and GGO nodules (Ye et al. 2009). Jacobs et al. based their dedicated method on the computation of a rich set of intensity, shape, and context features to accurately describe the appearance of this type of nodule (Jacobs 2011). A two-stage classification method based on LDA and GentleBoost classifier led to a detection sensitivity of 73% at only 1 FP/scan on a dataset of 140 CT scans acquired in a screening trial. Although dedicated approaches to GGO detection are needed due to the different appearance of GGOs with respect to solid nodules, it is desirable that GGO-dedicated analysis modules are integrated into existing CAD approaches for other nodule types in order to get the most from the complementarity of these systems.

Another issue related to particular kinds of nodules is that of peri- fissural nodules (PFN). Ciompi et al. (2015) tackled the problem of automatic classification of PFNs. The problem has a high relevance in the context of lung cancer screening, since PFNs have been proven to be benign nodules, though they represent 30% of nodules detected in the screening population. The problem has been formulated as a machine learning approach, where nodule candidates are classified as [PFN, non-PFN]. Supervised learning was used, where a classification framework is trained to automatically label the detected nodule. Authors collected examples of PFNs and non-PFNs from baseline scans of 1729 participants from one center of the NELSON Study, the Dutch–Belgian randomized lung cancer screening trial (NP 2016), discarding all the nodules with a size <6 mm and obtaining 568 nodules consisting of 65 typical PFNs, 19 atypical PFNs, and 484 non-PFNs. They used both Random Forests and Support Vector Machines and they evaluated the performance via ROC (Receiver Operating Characteristics) analysis, using the AUC (Area Under Curve) to measure the performance. The best result in terms of AUC was obtained with Support Vector Machine classifiers, achieving a value of AUC = 0.847.

Moreover, automated analysis of lung CTs can also constitute a valuable help in the case of images acquired at very low doses.
and, in particular, of those reconstructed by means of the recently available IR (iterative reconstruction) algorithms instead of FBP (Fourier back projection) based methods. Christe et al. (2013) used an anthropomorphic lung phantom and artificial lung nodules to simulate screening CT-examination at standard dose (100 mAs, 120 kVp) and eight different low dose levels and tested the sensitivity of two radiologists and three different CAD software, concluding that the combination of a human observer with any of the tested CAD systems provides optimal sensitivity for lung nodule detection, even at a reduced dose of 25 mAs/80 kVp. More recently, Den Harder et al. (2016) evaluated the effect of radiation dose reduction and IR on the performance of CAD for pulmonary nodules, concluding that the sensitivity of CAD does not decrease with both FBP and IR at sub-milliSievert dose levels, but the use of IR did result in an increased number of false-positive findings at reduced dose levels.

In addition to the automated research of particular kinds of nodules and to the analysis of very low dose and noisy images, another open issue in the research field of lung nodule automated detection is the appropriate setting of the analysis algorithms and their use in the clinical workflow. Many complete CAD systems show high standalone performance and have been demonstrated to enhance the radiologists’ ability in detecting lung cancers, but their use is still not widespread in clinical practice. There are many obstacles to be overcome before such analysis systems can be used daily in the clinical reading environment (Goo 2005, van Ginneken et al. 2011). Some are related to the reliability and generalizability of CAD performance. Moreover, medical experts and technicians involved in this field would often prefer not to have to handle unexpected failures or update requests.

To be considered valuable aid instruments in clinical practice, CAD systems should be characterized by very high sensitivity and specificity levels for all nodule types and sizes, and they should be fast and easy to use. Depending on the task they are demanded to accomplish, they should allow the possibility to browse through CAD marks according to type, size, and degree of suspicion, and CAD marks should be accessible from any workstation used by the radiologists and connected to the PACS.

Valente et al. (2016) recently presented a systematic review of techniques for the 3D automatic detection of pulmonary nodules in CT images. They believe that further research is needed to develop CAD systems, or to optimize existing systems, and that a closer relationship between researchers and the medical community is necessary, since the lack of recognition of some specific needs has hindered the wide use of CAD systems. This will only be possible through a joint effort between the involved parties in the process, including government, physicians, patients, engineers, scientists, and technicians.

A possible solution to making CAD algorithms easily accessible by clinicians could be the use of web-based on-demand CAD services (Berzano et al. 2012, Lopez Torres et al. 2015). This service allows for CT data uploading through secure web protocols. Then the service provider executes the CAD algorithm by means of cloud computing techniques and notifies the end-user via e-mail that the CAD output is ready for downloading. CAD output may be available in standard formats agreed with the end-user. In this scenario, a short algorithm execution time is not mandatory because the CAD output in any case can be available before the radiologist starts to review the case. In addition, if the web-based on-demand CAD service paradigm is adopted, the quality of the CAD output can be continuously optimized. In fact, the CAD service provider can, in principle, own more than one CAD system; thus different algorithms can be simultaneously executed, and CAD findings can be compared and combined to enhance their nodule detection ability. Moreover, hospitals and diagnostic centers do not have to acquire powerful workstations to carry out huge CAD computations or own CAD software licenses, and the maintenance and the upgrade of CAD systems and the computing power are fully managed by the CAD service provider.

### 62.6 Nodule Tracking

In addition to the detection aspect of lung nodules with CAD systems, nodule size determination and change over serial examinations is another crucial aspect. Increase in nodule size over time raises concerns of malignancy and warrants additional testing whereas consistent stability and regression over time triggers conservative management. Manual measurements are prone to both inter- and intra-observer variations. Most commercially available CAD systems enable users to automatically derive minimum, maximum, and average diameters for the nodules. Lesion volume, which is a better indicator of change in lesion size than diameters, can also be estimated with CAD systems, although this often requires users to threshold or edit the lesion contours, particularly for perivascular nodules.

### 62.7 Conclusions and Perspectives

Analysis systems for lung nodule identification on CT images have a great potential to detect early-stage lung cancers and may represent a valuable aid when used either as first or as second readers in the workflow of a radiology department. The research on the CAD algorithms targeting any nodule type has reached satisfactory results after two decades of very active work by many research groups. However, a standardization of many procedures in CAD analysis, in classification system training, and in performance evaluation and comparison is still to be reached. In particular, CAD systems’ performance should be validated against large annotated databases, including the complexity and variability of a screening population or, even worse, of a clinical population. The initiative carried out by the LIDC is extremely useful for CAD research, as it made a large repository of annotated CT scans available to CAD developers for designing, validating, and comparing systems (Armato et al. 2004). The combination of different detection algorithms in a more complex system was demonstrated to be beneficial. The optimal combination lies in the complementarity among the detection strategies of different CAD systems. Studies devoted to improving CAD validation procedures must go on simultaneously with the evaluation of the impact of CAD on the radiologist’s performance. When used as a first reader of CT scans, a CAD system is required to have high sensitivity to nodule detection. When a CAD is used as a second reader, it is not required to have comparable or better performance in nodule detection with respect to human readers, but to be complementary to the radiologist’s ability. Many studies demonstrated that the use of CAD systems could improve...
the daily radiologic workflow, however these systems are not yet widespread instruments in clinical routine practice.

CAD systems still need a further improvement, especially with new algorithms for detecting nodules attached to other structures and nodules with subtler appearance, such as GGOs. The main issue in this research field is to maintain high levels of sensitivity to all types of nodules while keeping a limited number of false-positive marks to minimize the number of unnecessary subject follow ups in a screening program where the CAD is used as first reader or to preserve the radiologists’ reading time in reviewing or discarding the CAD findings when the CAD is used as second reader.

To facilitate the integration of CAD into the workflow of radiology departments, CAD should be connected to the PACS environment. An original solution to facilitate the use of CAD systems by medical experts could be access to a web-based CAD service. In this case, the CAD end-users would be relieved from facing CAD system installation, upgrade, maintenance, and possible unexpected failures during the algorithm execution. A provider of a lung CAD service would guarantee huge computational power, usually not available in medical centers, and also the possibility to simultaneously execute several CAD algorithms for comparisons and combinations.

Research in the field of CAD systems and their potential role in supporting the radiologist’s daily work is still a very active area. Once the main technical obstacles to the widespread implementation of these systems are removed, lung CAD systems are expected to become useful tools to assist radiologists daily in lung CT scan interpretation.

REFERENCES


