# Automatic detection of large pulmonary solid nodules in thoracic CT images

Arnaud A. A. Setio,[a)] Colin Jacobs, and Jaap Gelderblom
*Diagnostic Image Analysis Group, Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen 6525 GA, The Netherlands*

Bram van Ginneken
*Diagnostic Image Analysis Group, Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen 6525 GA, The Netherlands and Fraunhofer MEVIS, Bremen 28359, Germany*

**Purpose:** Current computer-aided detection (CAD) systems for pulmonary nodules in computed tomography (CT) scans have a good performance for relatively small nodules, but often fail to detect the much rarer larger nodules, which are more likely to be cancerous. We present a novel CAD system specifically designed to detect solid nodules larger than 10 mm.

**Methods:** The proposed detection pipeline is initiated by a three-dimensional lung segmentation algorithm optimized to include large nodules attached to the pleural wall via morphological processing. An additional preprocessing is used to mask out structures outside the pleural space to ensure that pleural and parenchymal nodules have a similar appearance. Next, nodule candidates are obtained via a multistage process of thresholding and morphological operations, to detect both larger and smaller candidates. After segmenting each candidate, a set of 24 features based on intensity, shape, blobness, and spatial context are computed. A radial basis support vector machine (SVM) classifier was used to classify nodule candidates, and performance was evaluated using ten-fold cross-validation on the full publicly available lung image database consortium database.

**Results:** The proposed CAD system reaches a sensitivity of 98.3% (234/238) and 94.1% (224/238) large nodules at an average of 4.0 and 1.0 false positives/scan, respectively.

**Conclusions:** The authors conclude that the proposed dedicated CAD system for large pulmonary nodules can identify the vast majority of highly suspicious lesions in thoracic CT scans with a small number of false positives. © *2015 American Association of Physicists in Medicine.* [http://dx.doi.org/10.1118/1.4929562]

Key words: lung nodules, computed tomography, CAD system, detection

## 1. INTRODUCTION

Lung cancer is the leading cause of death from cancers worldwide.[1] The majority of patients are diagnosed with late stage cancer and as a result, 5-yr survival rate is only 16.8%. If cancer is diagnosed at an earlier stage, effective treatment options are available and survival rate increases substantially. Early detection is therefore considered as key to reduce the burden of lung cancer. The national lung screening trial (NLST) showed that screening with low-dose computed tomography (CT) significantly reduces lung cancer and total mortality among high-risk individuals.[2] Compared to x-ray, low-dose CT visualizes more lung cancers and has been proposed to be used as primary modality for early lung cancer screening. However, for human readers, the interpretation of lung CT screening scans, as well as regularly obtained clinical thoracic CT scans, is tedious, error-prone, and can take up to 10 min/scan.[3] Therefore, optimization of thoracic CT screening workflow is becoming an active field of research.[4–11]

Implementation of computer-aided detection (CAD) systems[4–7] has been proposed as a step toward more efficient screening. CAD systems are designed to automatically detect pulmonary nodules, the early manifestation of lung cancers.

In practice, CAD has been used as a second reader.[5,6,12] In the second reader mode, radiologists are required to diagnose CT scans unaided, and thereafter inspect the result of CAD to improve the accuracy of detection. As compared to diagnosis without CAD, this scenario clearly increases the reading time, which remains an issue to be addressed for the implementation of large-scale lung cancer screening program. Aiming to reduce the reading time for diagnosing CT scans, CAD should be used as a first reader. In the first reading mode, radiologists only inspect locations flagged suspicious by CAD, accept or reject the CAD marks, and possibly perform a quick inspection on the scan to identify any nodule completely missed by CAD.

To fully utilize CAD as a first reader in lung CT screening, it is vital that CAD reaches a high sensitivity for all screening relevant nodules. In the past decade, efforts have been made to design generic CAD systems that detect all types of nodules. Although existing CAD systems have shown promising detection performance,[13–22] the ANODE09 study[23] showed that the published CAD systems still often missed important subgroups of suspicious nodules (e.g., subsolid nodules, large nodules). For nodules larger than 5 mm, the participating CAD systems achieved an average detection sensitivity of only 29% at 1 false positive/scan (FP/scan). The low sensitivity can be

explained by the fact that large nodules are much less common than smaller nodules (see Table I). As CAD systems have usually been trained with a set of randomly selected cases from screening or clinical practice, large nodules are often undersampled in the training set and hence are put to lower priority by CAD. Very large nodules may even be missed by the initial candidate detection stage that most CAD systems employ.[7]

In a second reader scenario, the low sensitivity of CAD for large nodules is not problematic since humans are typically able to detect larger nodules and thus compensate the misses of CAD. However, if one would like to move toward using CAD as a first reader, which may be one of the key factors for efficient implementation of forthcoming lung cancer screening, the human will only look at the locations pointed out by CAD. In consequence, the low sensitivity of CAD for large nodules becomes a major impediment. Compared to other nodules, large nodules have a significant higher probability for being malignant[8] and require additional clinical diagnostic workup.[24–27]

In order to increase the reading efficiency of chest CT scans, we aim to develop a CAD solution that combines multiple CAD systems. Each subsystem excels at its own subtask and consequently, all different types of nodules can be detected. In this paper, we address the problem of low detection sensitivity of CAD systems on large nodules. A dedicated CAD system specifically designed for large nodules is presented, aiming to further complete the performance gap of CAD systems in lung nodule detection. The proposed CAD system could thus be combined with published state-of-the-art CAD systems to improve their detection sensitivity. Improving the overall detection performance is an important prerequisite to be able to efficiently use CAD as a first reader.

## 2. MATERIALS AND EVALUATION PROTOCOL

In this study, CT scans from the full lung image database consortium (LIDC–IDRI)[28] were used. LIDC–IDRI is a publicly available image database initiated by the National Cancer Institute (NCI). The database contains heterogeneous set of 1018 cases from seven LIDC institutions. Each scan was annotated by four radiologists in two reading rounds. In the first blinded reading round, suspicious lesions were annotated and each of them was categorized as non-nodule, nodule < 3 mm, or nodule ≥ 3 mm. Nodules ≥ 3 mm were characterized fully,

for which 3-D segmentation, diameter, and morphological characteristics were provided. In the second reading round, each of the four readers reviewed both their annotations and all annotations from other readers, and decided to either accept or reject each annotation. Finally, the results of unblinded review from all readers were compiled and a final set of annotations was acquired.

We included scans with section thickness below or equal to 2.5 mm. Thicker section data were excluded as we deemed these data not of sufficient quality. Although large nodules may still be detectable on scans with a slice thickness of 3 mm, recent clinical guidelines[29,30] recommend the use of thin-slice CT scans and hence, we only included thin-slice CT scans. In addition, scans with inconsistent slice spacing were discarded. This resulted in a set of 888 CT scans suitable for analysis (see the Appendix).

An overview of our dataset is provided in Table I. From 888 scans, 36 378 annotations were made by the LIDC–IDRI readers. As nodules could be annotated by multiple readers, annotations from different readers that were located in close proximity (defined as less than the sum of their radii) were merged. The merging was only performed for the nodule > 3 mm annotations since we used these to define our reference standard. The diameters, volumes, and coordinates of merged annotations were averaged. We selected nodules categorized as potentially malignant nodules as our reference for determining the size threshold for large nodules. Based on the criterion employed by the Dutch-Belgian NELSON lung cancer screening trial,[31] potentially malignant nodules are defined as those with a volume larger than 500 mm³, which corresponds to a diameter of approximately 10 mm, assuming an ideal sphere. Thus, we applied 10 mm as the nodule size threshold. Furthermore, we selected nodules that were accepted by the majority of readers (at least three readers). This resulted in a set of 269 nodules. As we have recently developed a CAD system designed to specifically detect subsolid nodules,[32] we decided to focus on a CAD system that detects only solid nodules. Various morphological characteristics scored by readers were used to define the nodule type. In this study, a nodule was considered solid if the majority of readers gave the texture characteristic score higher than 3 (1 = ground-glass/nonsolid, 3 = part-solid, 5 = solid). Eventually, we obtained 238 large solid nodules, which formed our final set of nodules used for analysis.

During evaluation, a CAD mark was considered as hit if it was located within a radius of annotation. CAD marks were categorized as true positives if they hit nodules in our reference set. Marks on lesions that were not included in the reference set (i.e., nodules ≥ 10 mm accepted by only one or two radiologists, nodules < 10 mm, subsolid nodules, and nonnodules) were considered irrelevant and were not counted as false positive.

## 3. METHODS

Figure 1 shows the main steps of the proposed algorithm. The algorithm is initiated by performing lung segmentation. Lung segmentation refinement is performed to include nodules

TABLE I. Statistics on the number of nodules in the selected set of scans from the LIDC–IDRI dataset. Each row indicates nodules that were accepted by different agreement levels among radiologists and the column shows the subset of nodules in different size or type.

| Agreement levels | Nodules ≥ 3 mm | Nodules ≥ 10 mm | Solid nodules ≥ 10 mm |
|---|---|---|---|
| At least 1 | 2290 | 393 | 322 |
| At least 2 | 1602 | 325 | 277 |
| At least 3 | 1186 | 269 | 238 |
| At least 4 | 777 | 199 | 172 |

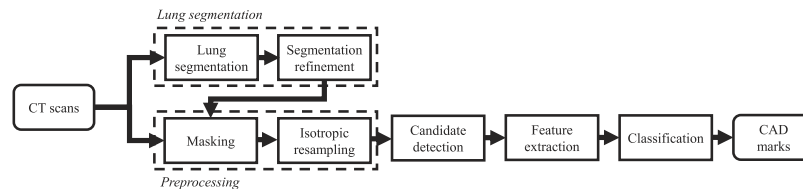Note: Total number of CT scans: 888; total number of annotations: 36 378.

FIG. 1. Overview of proposed CAD system.

attached to lung border. Thereafter, a preprocessing step is applied to remove the background outside the lungs and to resample the image to isotropic resolution. A candidate detection stage identifies the location of large nodule candidates and constructs a segmentation for each candidate. Discriminant characteristics of candidates are obtained using feature extraction, which are used to classify candidates into nodules and non-nodules in the classification stage. All the main parts of the proposed approach are explained in detail in Secs. 3.A–3.E.

### 3.A. Lung segmentation

Lung segmentation is the first step of the CAD system applied to localize the lung as the region of interest. The majority of published algorithms rely on a conventional threshold-based method. Due to similar intensity characteristic, nodules attached to pleural wall often fail to be included inside lung segmentation. In consequence, these nodules are excluded from further processing and remain undetected by the CAD system. In order to include these nodules, the lung segmentation should be corrected.

We applied the algorithm proposed by van Rikxoort *et al.*[33] as initial lung segmentation. The method consists of large airway extraction, lung segmentation, left and right lungs separation, and segmentation smoothing. When the lung segmentation algorithm failed, we corrected the lung segmentation by manually measuring input parameters, such as trachea's seed point.

To include large pleural nodules inside lung segmentation, an additional lung segmentation refinement is employed. The areas where nodules are excluded often appear as holes at the surface of the lung segmentation. We experimented with two lung segmentation refinement methods: (1) applying morphological rolling-ball operation and (2) adding dilation after rolling-ball operation. All morphological operations were performed using spherical structuring element. The diameter of the structuring element $d_{struct}$ is set to a percentage of the $x$-dimension size of the input CT scan. We evaluated $d_{struct}$ = {2%,3%,4%,6%,8%} for rolling-ball and $d_{struct}$ = {0%,1%} for dilation.

### 3.B. Preprocessing

The lung segmentation is used as mask where the area outside the segmentation is set to the average intensity of lung tissue [−900 Hounsfield units (HU)]. This avoids that pleural lesions look very different from internal nodules, and thus would require different or additional features to accu-

rately detect these nodules. The lung image is resampled to an isotropic resolution of 1.0 mm using Gaussian filter.

### 3.C. Candidate detection

The candidate detection step aims to localize all nodules within the lung. This is a difficult task, since nodules have a large variation in morphology, size, and density. Moreover, the computational complexity of the algorithm should also be low as the entire lung volume is processed in a voxel-wise method. In this study, the candidate detection consists of three modules: initial candidate detection, connected component analysis, and nodule segmentation refinement.

#### 3.C.1. Initial candidate detection

As large nodules typically have a very different morphological characteristic compared to other similarly dense structures (mainly vessels), a simple combination of thresholding and morphological opening should be sufficient to detect most of the large nodules. However, due to their size, large nodules are prone to be interconnected to other pulmonary dense structures, mostly vasculature and the pleural wall. This complicates the parameter selection of the detector, specifically the size of the structuring element of morphological opening. Connected structures may have a wide size variability. As an example, a large structuring element is required to remove attachment of large structure (e.g., large airways) on nodules, but leads to failing detection of smaller nodules.

To detect candidates with different sizes, a multistage process of thresholding and morphological opening is applied. An intensity threshold of −300 HU is used to identify solid nodules. In each stage, dilation with different spherical structuring diameters of 9, 7, 5, or 3 mm is applied in order, and an intermediate candidate mask is generated. The candidate detection starts from large structuring diameter and progressively continues with smaller diameter. To preserve prior candidates from being merged with new candidates, a margin border of 3 mm around prior candidates is applied before processing the next stage. After computing the candidate detection in stage $n$, the new intermediate mask is merged with the prior output mask from stage $n-1$ using logical OR operation. Examples of progressing candidate detection outputs are shown in Fig. 2.

#### 3.C.2. Connected component analysis

After applying initial candidate detection, all connected voxels are clustered as candidates using connected component analysis. To remove clusters with size outside the target range,
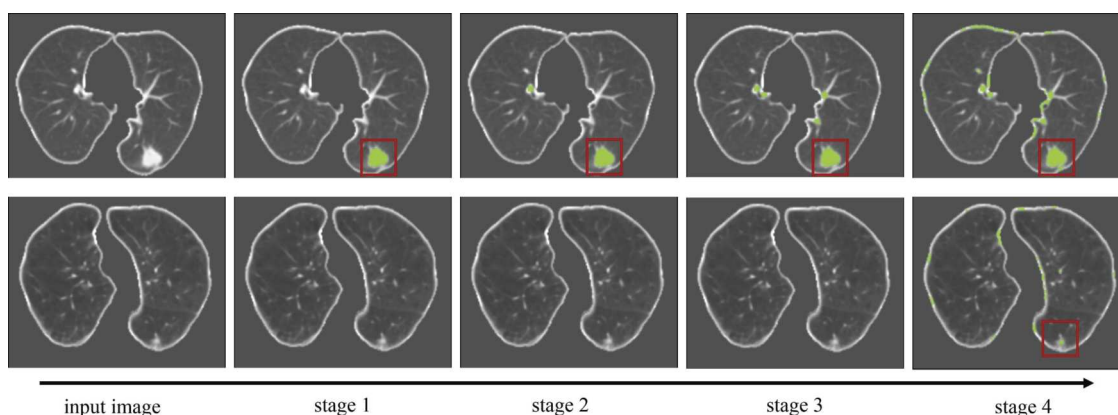
FIG. 2. Examples of multistage candidate detection outputs. In each stage, a specific thresholding and morphological opening is applied and the resulting mask is added to the prior segmentation. Detected candidates are indicated by a green contour while detected nodules are indicated by a red box. The first row shows an example of a large nodule detected in stage 1 and the second row shows an example of smaller nodule detected in stage 4. Applying multistage candidate detection prevents segmentation of the large nodule to be incorrectly interconnected to pleural while ensuring the smaller nodule to be detected.

we discarded clusters with a volume smaller than $268 \text{ mm}^3$ and larger than $33,524 \text{ mm}^3$, which correspond to the volume of an ideal sphere with a diameter of 8 and 40 mm, respectively.

### 3.C.3. Nodule segmentation

Initial candidate detection indicates the coordinates of candidate clusters. As an accurate volumetric and morphological assessment of candidates is important for quantifying nodule characteristics, we applied a robust nodule segmentation proposed by Kuhnigk et al.[34] Given a cubical input volume and seed points on the cluster, the algorithm performs region growing and applies an efficient morphological operation to separate nodules from vasculature and chest wall. For each candidate cluster, seed points are obtained from the major axis of the cluster. The volume of interest is defined as a cube around the cluster with an initial edge length of 60 mm. The size of the cube is automatically adapted, if necessary, to accommodate larger nodules.

To further remove candidates smaller or larger than the defined target, we applied the same size threshold used in Sec. 3.C.2. To prevent duplicates in the set of nodules, a new segmentation located less than 10 mm from any accepted segmentation is discarded. The results of the nodule segmentation were further used for feature extraction and classification.

### 3.D. Feature extraction

Feature extraction was applied to obtain characteristic of candidates that can be used to differentiate nodules from false-positive candidates. We defined four different sets of features: intensity, cluster, blobness, and context features. Table II shows an overview of all features.

### 3.D.1. Intensity features

Intensity features are used to quantify intensity characteristics of regions inside and outside the candidate segmentation. The feature set is directly computed from the isotropically resampled CT scan. Three different regions are used: (1) region inside the candidate segmentation, (2) region inside the bounding box of the candidate segmentation, and (3) region surrounding the candidate segmentation within a distance of 8 mm. For each region, mean, standard deviation, and entropy of voxel intensity are measured.

### 3.D.2. Cluster features

Cluster features are measured from candidate segmentation. These feature sets consist of diameter, volume, sphericity, compactness$_1$, and compactness$_2$.

Diameter is measured by calculating the longest axis of the segmentation in axial plane. Volume is determined by calculating the cluster size in $\text{mm}^3$. To measure sphericity, a sphere $S$ that has the same volume as the candidates is constructed at the center of mass of the candidate. Sphericity is defined as the ratio of candidate volume inside sphere $S$ to the volume of sphere $S$. Compactness is the ratio of candidate volume to the volume of bounding box around the candidate segmentation. Two different bounding boxes are used. Compactness$_1$ uses the bounding box defined as the smallest box which encloses the candidate segmentation in all $x$, $y$, and $z$ dimensions. Compactness$_2$ uses the bounding box defined as a cube with a size of the largest dimension of candidate segmentation.

### 3.D.3. Blobness features

Blobness features are widely used to enhance nodular structures and to improve sensitivity for nodule detection. We applied the nodule enhancement filter developed by Li, Sone, and Doi[35] and extracted blobness features from the filtered image.

The algorithm computes the Hessian matrix by convoluting the input image with second order derivatives of the Gaussian kernel. Given the Hessian matrix, three eigenvalues, which are denoted by $\lambda_1 = f_{xx}$, $\lambda_2 = f_{yy}$ and $\lambda_3 = f_{xy}$, are measured, where $\lambda_1 > \lambda_2 > \lambda_3$. The final output of the enhancement filter is obtained by calculating $z(\lambda_1, \lambda_2, \lambda_3) = |\lambda_3|^2 / |\lambda_1|$, if $\lambda_1 < 0$, $\lambda_2 < 0$, and $\lambda_3 < 0$; otherwise 0. To accommodate a wide range

TABLE II. Overview of features calculated on candidates for characterization and classification.

| ID | Description | Notes |
|---|---|---|
| | *Intensity features* | |
| 1–3 | Density inside candidate segmentation (mean, stdev, entropy) | |
| 4–6 | Density inside bounding box (mean, stdev, entropy) | |
| 7–9 | Density around candidate segmentation (mean, stdev, entropy) | within 8 mm outside segmentation |
| | *Cluster features* | |
| 10 | Volume of candidate segmentation $V_{cand}$ | |
| 11 | Diameter of candidate segmentation $D_{cand}$ | The longest diameter on axial plane |
| 12 | Sphericity: ratio of candidate's volume inside sphere $S$ to the volume of sphere $S$ | Sphere $S$ is centered at the candidate location with diameter $D_{cand}$ |
| 13 | Compactness$_1$: $V_{cand}/(\dim_x \cdot \dim_y \cdot \dim_z)$ | $\dim_i$ is the width of bounding box in dimension $i$ |
| 14 | Compactness$_2$: $V_{cand}/((\max(\dim_x, \dim_y, \dim_z))^3)$ | |
| | *Blobness features* | |
| 15 | Maximum filter response | Features are computed using the filter |
| 16–17 | Filter response inside candidate segmentation (mean, stdev) | response of Li blobness filter (Ref. 35) (scale: 2, 5, and 8 mm) |
| 18–19 | Filter response inside bounding box (mean, stdev) | |
| | *Context features* | |
| 20 | Distance to pleural wall | |
| 21–23 | Distance to carina in *X*, *Y*, and *Z* | |
| 24 | Distance to top of the lung in *Z* | |

of nodule size, multiscale enhancement filters are applied. We defined three scales (2, 5, and 8 mm) for the Gaussian kernel, which are empirically determined based on the size of targeted nodules to be enhanced. This results in three different output images, which are combined by selecting the maximum values from all images.

The blobness features are extracted by measuring: (1) maximum output of the filter; (2) mean and standard deviation of output inside the nodule segmentation; and (3) mean and standard deviation of output inside the bounding box.

### 3.D.4. Context features

Context features describe the position of the candidates relative to other pulmonary structures, i.e., pleural wall, carina, and top of lung. Depending on their location, both lung nodules and false-positive candidates may have different morphological characteristics and malignancy probability. Nodules are more likely to be elongated when attached to more rigid structures, e.g., airways around hilum or pleural wall. Horeweg *et al.*[36] showed that the majority of malignant nodules were found at outer one-third of the costal-hilar diameter and at upper lobes, indicating that these regions should gain more attention. Distance to carina and top of lung is expected to give these suspicious lung regions more weight.

To calculate the distance relative to the pleural wall, we performed a distance transform within the lung segmentation and extracted the value located at the center of gravity of the candidate. The carina point is detected by identifying the axial section where the trachea has split in two parts, selecting the

trachea region around the bifurcation, and extracting its center of gravity. The relative *X*, *Y*, and *Z* world distance of candidates from the carina are used as features. The relative distance of candidate to the top of the lung is measured by calculating the distance from the *Z* world coordinate of the candidate to the *Z* world coordinate of the top slice of the lung segmentation.

### 3.E. Classifiers

For the classification of the feature vectors, support vector machine with radial basis function kernel (SVM-RBF)[37] is used. The *C* and $\gamma$ parameters are optimized in ten-fold cross-validation using a grid search on the training dataset. *C* defines the regularization parameter and $\gamma$ defines the width of the RBF kernel. The search range of $C = \{2^1, 2^2, \ldots, 2^{12}\}$ and $\gamma = \{2^{-12}, 2^{-11}, \ldots, 2^{-5}\}$ is applied. In this paper, SVM is implemented using LIBSVM.[38] Prior to classification, we perform normalization to zero mean and unit standard deviation on all features. In pilot experiments with various classifiers, we found that radial basis SVMs achieved the best result for this application.

## 4. RESULTS

In this section, the performance of the CAD system for detection of large nodules is reported. The evaluation is performed in a patient-level ten-fold cross-validation while the classifier is optimized in a nested cross-validation. We pick nodules accepted by majority of radiologists (three out of four) as primary evaluation reference standard. For completeness,

the performance of CAD system on nodules accepted by other agreement levels is also included.

## 4.A. Lung segmentation

Automatic lung segmentation was performed. In 12 of the 888 scans, lung segmentation failed because the trachea seed point was not correctly detected. These cases were fixed by providing a manual seed point. Table III shows the impact of applying lung segmentation refinement on the CAD system. A lung nodule is categorized as being included in the lung segmentation if the corresponding position is contained within the segmentation. The algorithm is evaluated by measuring the number of detected nodules and the number of candidates. When lung segmentation is computed without the additional refinement algorithm, only 84.9% (202/238) large nodules were included. As shown in Table III, applying the additional refinement algorithm decreases the number of excluded nodules, until eventually all nodules have been included. This suggests that a more aggressive refinement algorithm is desirable for nodule detection. However, as the refinement algorithm globally enlarges the lung segmentation, it may also include irrelevant regions (e.g., lung borders, hilar), which deteriorates the performance of candidate detection. Therefore, a pilot study was performed, for which the sensitivity of candidate detection using different lung segmentation refinement was evaluated. 287 scans containing all nodules ≥ 10 mm annotated by at least three radiologists were used as dataset. The experiment showed that applying rolling-ball with $d_{struct} = 6\%$ and dilation with $d_{struct} = 0\%$ after initial lung segmentation allows candidate detection to achieve the highest sensitivity with a reasonable number of candidates/scan. This configuration was used in all remaining experiments.

## 4.B. Candidate detection

On the full dataset with 888 CT scans, the candidate detection generated an average of 48.3 candidates/scan, includ-

TABLE III. The impact of refining lung segmentation on including nodules inside lung segmentation. The evaluation was performed on 287 scans containing all 238 large nodules annotated by at least three out of four radiologists. Best performing configuration of each performance metric is shown in bold. It should be noted that candidate detection can still detect nodules that are partially included inside the lung segmentation.

| Kernel size (% of image size) | | Inside lung segmentation (%) | After candidate detection (%) | Candidates/ scan |
|---|---|---|---|---|
| Rolling ball | Dilation | | | |
| Without additional lung segmentation refinement | | | | |
| — | — | 84.9 | 87.8 | 39.5 |
| With additional lung segmentation refinement | | | | |
| 2 | — | 84.9 | 87.8 | **39.5** |
| 4 | — | 97.1 | 98.3 | 47.4 |
| 6 | — | 98.3 | **99.2** | 56.6 |
| 8 | — | 99.6 | 98.7 | 63.2 |
| 2 | 1 | 88.7 | 95.4 | 120.7 |
| 4 | 1 | 99.6 | 97.9 | 133.1 |
| 6 | 1 | **100.0** | 96.6 | 145.5 |

TABLE IV. Sensitivity of candidate detection in 888 scans containing nodules annotated by at least 1–4 radiologists. Performance on datasets with different agreement levels is reported.

| Agreement levels | Solid nodules >10 mm | Detected nodules (%) | Candidates/ scan |
|---|---|---|---|
| At least 1 | 322 | 97.2 | 48.3 |
| At least 2 | 277 | 98.9 | 48.3 |
| At least 3 | 238 | 99.2 | 48.3 |
| At least 4 | 182 | 100.0 | 48.3 |

ing 99.2% (236/238) of all large nodules. This set was used for further classification tasks. The performance of candidate detection on sets of nodules with different agreement levels is shown in Table IV. The sensitivity of candidate detection increases for nodules with higher agreement level among radiologists.

## 4.C. Feature extraction and classification

Free-Response Operator Characteristic (FROC) curves of the CAD systems on datasets containing large nodules with different agreement levels are depicted in Fig. 3. Figures on the left and the right side show the performance of the CAD systems with and without "irrelevant" findings considered as false positives. The number of false positives is shown on a logarithmic scale. For quantitative comparison, average sensitivities at different false positive rates is provided in Table V. As shown in Table V, the proposed CAD system identifies 94.1% (224/238) and 98.3% (234/238) large nodules at 1 FP/scan and 4 FPs/scan, respectively. It is worth to note that the maximum sensitivity of the classification stage is bounded to the sensitivity of the candidate detection, which is 99.2% (236/238) (see Table IV). This indicates that relative to the number of detected candidates, the classification stage correctly classifies 94.9% and 99.1% nodules at 1 and 4 FPs/scan, respectively.

To evaluate the performance of feature sets, we repeated the experiments using only individual sets of features. Figure 4 shows the FROC curves of the CAD systems when different feature sets are individually applied. Intensity features, for which density characteristic inside and outside the candidate segmentation is quantified, were found to be the most discriminative set of features. Although the other sets of features have much lower sensitivity at corresponding FPs/scan, the combination of all features improves the classification performance, which suggests that these sets of features are complementary to each other.

## 4.D. Comparison with existing CAD

To put the proposed system in context, we compared it with a previously published CAD system. ISICAD,[13] which achieves the highest rank on ANODE09,[23] is used for comparison. The FROC curves of both CAD systems are shown in Fig. 5. The candidate detection stage detects 99.2% (236/238) and 84.9% (202/238) large nodules using the proposed CAD
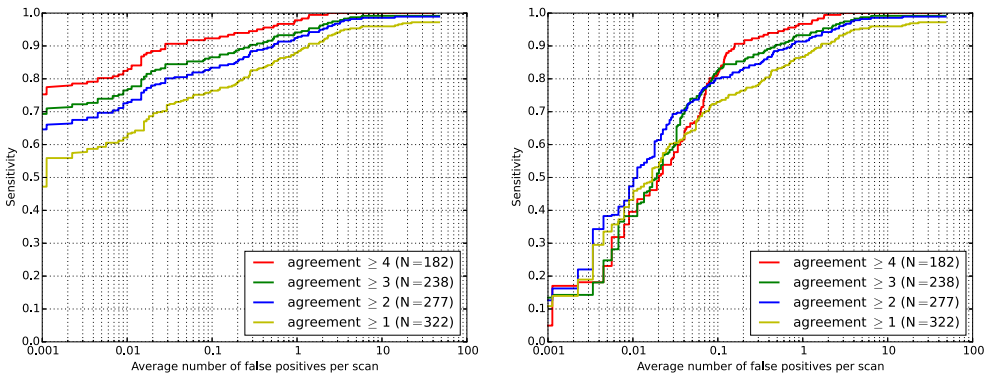
Fig. 3. FROC curves of the CAD system on datasets containing large nodules with different agreement levels. Number of false positives is shown on a logarithmic scale. The left figure and the right figure show the FROC curves of the CAD system on dataset without and with irrelevant findings considered as false positives, respectively. The irrelevant findings are defined as annotated non-nodule lesions, subsolid nodules, solid nodules < 10 mm, and solid nodules ≥ 10 mm that were only detected by < N radiologists.

and ISICAD, respectively. At 1 FP/scan, 63.0% (150/238) more large nodules are correctly classified by the proposed CAD. It is worth to note that the two CADs are designed for different types of nodules.

## 5. DISCUSSION

The proposed CAD system is designed to specifically detect nodules with a diameter larger than 10 mm. These large nodules are more likely to be cancerous and therefore their detection is critical. A major limitation of current CAD systems in the detection of large nodules is the fact that these systems have been usually trained with a dataset containing a very small fraction of large nodules. Supervised classification may put large nodules to lower priority, leading to an average sensitivity of only 29% at 1 FP/scan, for CAD systems participating in ANODE09.[23] Lung segmentation and candidate detection may even exclude large nodules, mostly due to the attachment to other pulmonary structures and a large deviation on morphological characteristics.

Using the largest publicly available CT nodule data set, the LIDC–IDRI database, a CAD system aiming specifically at large nodules was designed. Our dataset contained CT scans collected from different scanners and reconstruction kernels. For evaluation, we used 888 CT scans, for which 238 large solid nodules annotated by the majority of radiologists were selected as reference standard.

Table III demonstrates that optimization of lung segmentation is an important step for nodule detection application.

Relatively low inclusion rate of nodules inside the lung segmentation is likely to degrade the performance of remaining stages of CAD system, as nodules completely excluded from lung segmentation could not be processed. With only morphological based refinement algorithm, the inclusion rate of nodules increases from 84.9% to 98.3%. This improves the sensitivity of our candidate detection from 87.8% to 99.2%. Examples of CT slices where lung nodules were previously excluded from conventional lung segmentation are shown in Fig. 6. It is worth to note that while the inclusion criterion depends on whether nodules in reference standard are located in the lung segmentation, excluded nodules may still be detected by CAD as long as a part of the nodules is inside the segmentation. In our study, we found that three out of four excluded nodules [see Fig. 6(b)] were eventually detected by the candidate detection stage.

We have shown that the proposed CAD system successfully detects 94.1% (224/238) large nodules at a relatively low 1 false positive/scan. Compared to the best CAD system[13] submitted on ANODE09,[23] the proposed CAD system de-

TABLE V. Quantitative results of CAD performance on datasets containing nodules with different agreement levels. Sensitivities at seven average false positive rates of 1/8, 1/4, 1/2, 1, 2, 4, and 8 are reported and averaged to obtain an overall score, as was done in the ANODE09 study (Ref. 23).

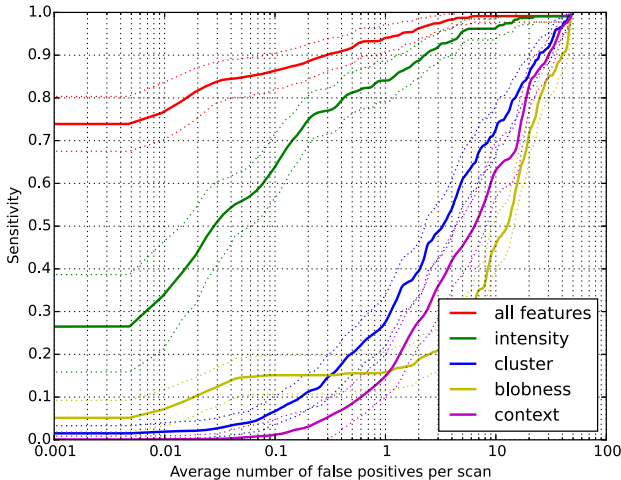| Agreement levels | 1/8 | 1/4 | 1/2 | 1 | 2 | 4 | 8 | Average |
|---|---|---|---|---|---|---|---|---|
| At least 1 | 0.773 | 0.804 | 0.842 | 0.879 | 0.913 | 0.950 | 0.960 | 0.874 |
| At least 2 | 0.841 | 0.866 | 0.895 | 0.924 | 0.949 | 0.978 | 0.982 | 0.920 |
| At least 3 | 0.874 | 0.895 | 0.916 | 0.941 | 0.962 | 0.983 | 0.992 | 0.938 |
| At least 4 | 0.929 | 0.940 | 0.956 | 0.978 | 0.995 | 1.000 | 1.000 | 0.971 |



Fig. 4. FROC curves of different feature sets on datasets containing 238 nodules accepted by at least three radiologists. Number of false positives is shown on a logarithmic scale. Dashed curves show the 95% confidence interval estimated using bootstrapping.
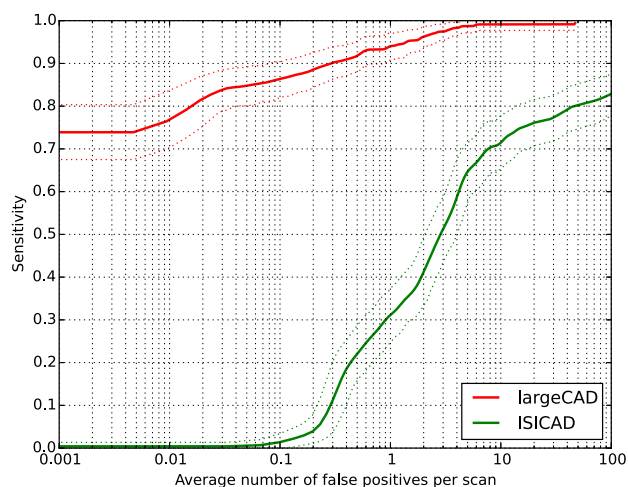
FIG. 5. FROC comparison of proposed CAD with previously published ISICAD (Ref. 13). ISICAD achieved the highest rank on ANODE09 study (Ref. 23). Dashed curves show the 95% confidence interval estimated using bootstrapping.

tected significantly more large nodules (*p*-value < 0.0001), see Fig. 5. Examples of detected nodules are shown in Fig. 7(a). To examine the results further, we performed a visual inspection of false positives, as well as of nodules missed by CAD. Examples of false positives detected at 0.125 false positives/scan are shown in Fig. 7(b). Most of the false positives detected with high score are large vessels, mediastinal structures, scarring, bone structures, or areas of interstitial lung disease. These errors are mainly caused by the refinement of the lung segmentation required to include pleural nodules. Improvement of the lung segmentation refinement step is an interesting topic for further research.

We analyzed nodules that were missed by CAD. After the candidate detection stage, two out of 238 nodules were not detected. These two nodules are shown in the most right examples of Fig. 7(c) and were missed because (1) one nodule is located outside the lung segmentation due to its flat morphology [see also Fig. 6(c)] and (2) other nodule has a cavi-

tation and was not picked up by the candidate detector as its inner region is dark. After classification, most of missed nodules at 0.125 FPs/scan rate present with wider variation in morphology or with leaking segmentation.

One reason for the good performance, specifically at a very low FPs/scan rate, is the fact that other lesions (i.e., non-nodule lesions, subsolid nodules, solid nodules < 10 mm, and solid nodules ≥ 10 mm that were only detected by one or two radiologists) are not counted as false positives in the evaluation. These lesions are marked as irrelevant, and are neither considered as true positive nor false positive when any of these lesions are detected by CAD. The rationale for marking lesions as irrelevant is that they remain, in practical use, interesting findings for screening. Some discarded lesions (e.g., subsolid nodules) may represent lung cancers. Although this is less clear for the non-nodule category, these findings were also considered irregular lesions. We compared the performance on our dataset with and without considering the irrelevant lesions as false positives. This is shown in Fig. 3. At a very low false positive rate (≤0.1 FPs/scan), we found that most of the false positives are large nodules with a lower agreement level among radiologists. At an operating point of 0.5 FPs/scan, the detection sensitivity is higher than 90%, even when irrelevant lesions are included as false positives.

The performance of CAD on the irrelevant findings is shown in Fig. 8. The proposed CAD detects 6.7% of nodules < 10 mm at 1 FP/scan. The detected small nodules are nodules with a size close to the threshold of 10 mm. Although the proposed CAD is not designed to detect subsolid nodules, 38.7% of subsolid nodules are detected at 1 FP/scan, as the solid core is large enough to be captured.

Table VI summarizes statistics of the proposed CAD system and previously published CAD systems that use LIDC–IDRI as dataset. For each CAD system, we listed number of scans used in validation dataset, nodule inclusion criteria (e.g., nodule type, nodule size, agreement levels), number of nodules, and reported CAD performance. Although it is difficult to straightforwardly compare CAD performance due to variability in dataset and evaluation method (e.g., true positive
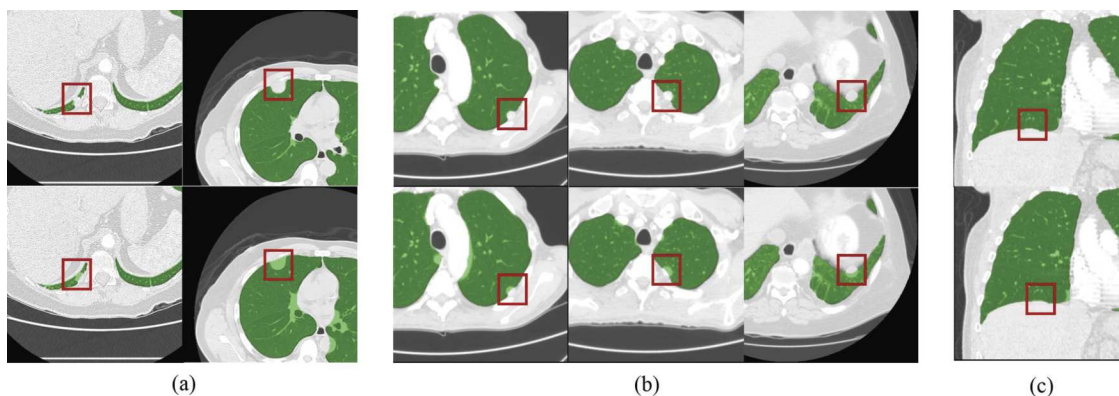


FIG. 6. Examples of lung segmentation refinement outputs and its impact on nodule inclusion. First row shows CT slices with initial lung segmentation, for which lung nodules are excluded from the segmentation. Second row shows the corresponding CT slices with lung segmentation after refinement algorithm. All nodules are detected by the majority of radiologists and are indicated by red box in figures. (a) Example cases where nodules are included after refinement algorithm. (b) Example cases where nodules remain outside the lung segmentation, but are detected by the candidate detection. (c) Remaining case where nodule is neither detected by the candidate detection nor is inside the lung segmentation, as it has a flat morphology.
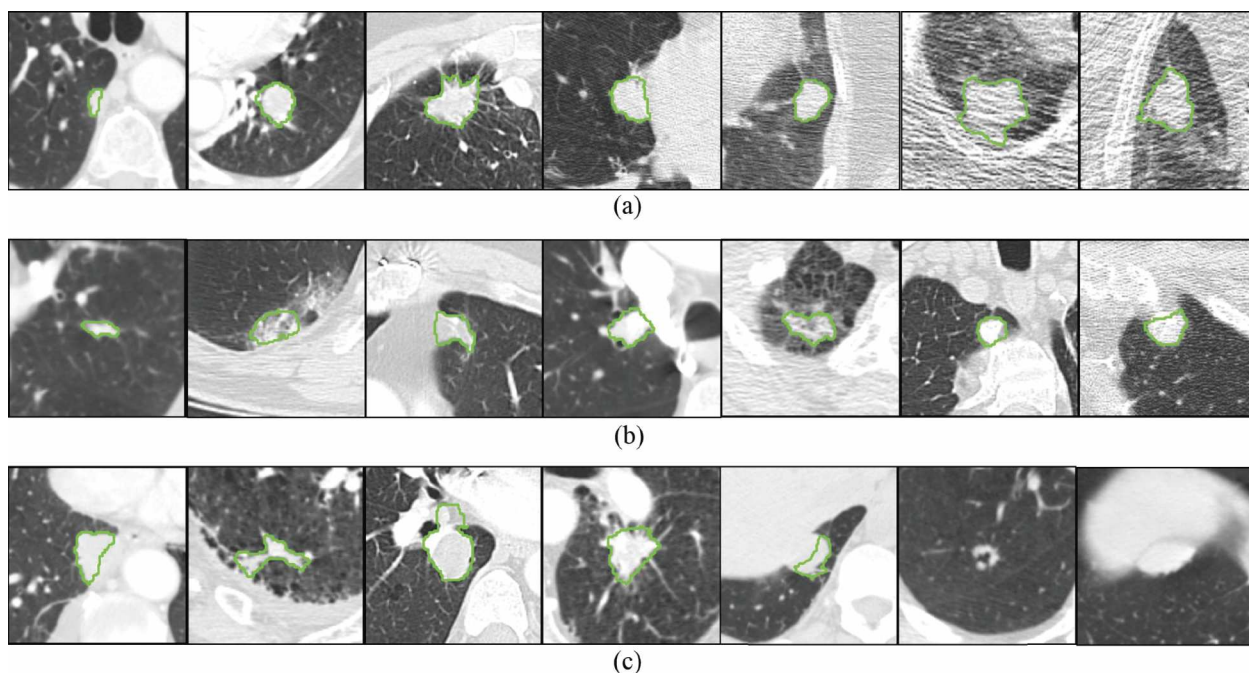
(a)

(b)

(c)

FIG. 7. Examples of lesions detected or missed by CAD system in transverse view. Each lesion is located at the center of the 128×128 patch and the result of the segmentation is indicated by the green contour. Nodules without contour imply that the nodule segmentation is not (successfully) computed. (a) Nodules detected at an average 1 false positive/scan. (b) False positives detected at an average 0.125 false positives/scan. (c) Nodules missed at an average 1 false positive/scan, mostly due to variation in morphology or leaking segmentation. The two most right images show nodules that are completely missed by CAD while the last nodule is the same nodule shown in Fig. 6(c).

and false-positive criteria), the table provides an impression of the existing CAD systems' performance in LIDC–IDRI. We have been one of the few groups, along with Torres *et al.*,[22] that use the complete LIDC–IDRI database and report results on a much larger number of scans. To accommodate the comparison of different CAD systems on the same set of scans, we
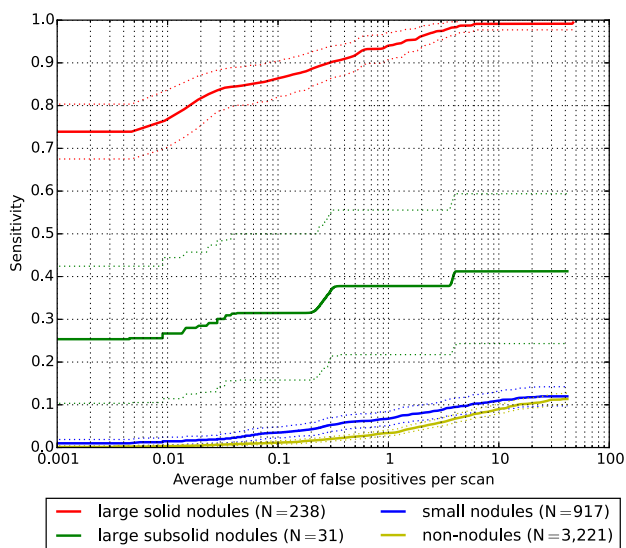


FIG. 8. FROC curves of the proposed CAD system on different lesion categories. The FROC curves show the additionally detected subsolid nodules, small nodules, and non-nodules corresponding to the operating points of the FROC curve of large solid nodules. Small nodules are defined as nodules ≥ 3 and < 10 mm and large nodules are defined as nodules ≥ 10 mm. The proposed CAD does not detect any nodule < 3 mm. Dashed curves show the 95% confidence interval estimated using bootstrapping.

provided the list of LIDC–IDRI scans that were used in this study in the Appendix.

This study has a limitation. Although we used a cross-validation experiment to validate the classifier, we did use LIDC–IDRI cases to optimize the preprocessing steps of the proposed CAD system, specifically lung segmentation refinement, candidate detection, and feature extraction. However, the data were only minimally used to optimize the preprocessing steps and the classifier is by far the most important component of the proposed CAD system. The candidate detection was designed to be as simple as possible; taking a coarse step in both thresholding and morphology kernel size. The features were qualitatively preselected based on the characteristics of the data and were not extensively optimized using the data of our study. The only step where we used the LIDC–IDRI data was the selection of the best lung segmentation refinement parameters. As Table III shows, the lung refinement stage improves the detection sensitivity from 87.8% to a reasonable range of 95.4%–99.2%. We intend to perform validation studies on other datasets in the future.

Although the detection of large nodules may not seem to be a difficult task, the fact that only 177 out of 322 large nodules were accepted with complete agreement among radiologists indicates that there is substantial subjectivity in the decision whether a large lesion is in fact considered to be a nodule. Although an efficient diagnosis of CT scans is the primary goal of developing the dedicated CAD system, the proposed CAD can also be alternatively used as a second opinion in screening. It would be an interesting option for further research to carry out an observer study to assess the impact of our CAD scheme on the detection sensitivity of radiologists.

TABLE VI. Summary of published CAD systems using LIDC–IDRI as dataset. Number of scans, reference standard criteria, and number of nodules used for validation are listed. The reported performance is provided.

| Authors | Years | No. of scans | CT slice thickness (mm) | Nodule size (mm) | Nodule type | Agreement levels | No. of nodules | Sensitivity (%) | | FPs/scan | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Ground-truth criteria | | | | Performance | | | |
| Proposed CAD | — | 888 | ≤2.5 | ≥10 | Solid | At least 4 | 182 | 100.0 | 97.8 | 4.0 | 1.0 |
| | | 888 | ≤2.5 | ≥10 | Solid | At least 3 | 238 | 98.3 | 94.1 | 4.0 | 1.0 |
| | | 888 | ≤2.5 | ≥10 | Solid | At least 2 | 277 | 97.8 | 92.4 | 4.0 | 1.0 |
| | | 888 | ≤2.5 | ≥10 | Solid | At least 1 | 322 | 95.0 | 87.9 | 4.0 | 1.0 |
| Torres *et al.* (Ref. 22) | 2015 | 949 | NA | ≥3 | All | At least 2 | 1749 | 80.0 | | 8.0 | |
| Brown *et al.* (Ref. 39) | 2014 | 108 | 0.5–3 | ≥8 | All | At least 3 | 68 | 79.3 | | 1.0 | |
| | | 108 | 0.5–3 | ≥4 | All | At least 3 | 68 | 75.0 | | 2.1 | |
| Choi and Choi (Ref. 21) | 2013 | 58 | 0.5–3 | 3–30 | All | At least 1 | 151 | 95.3 | | 2.3 | |
| Choi and Choi (Ref. 40) | 2013 | 84 | 1.25–3 | ≥3 | All | At least 1 | 148 | 97.5 | | 6.8 | |
| Tan *et al.* (Ref. 20) | 2013 | 360 | NA | ≥3 | All | At least 4 | NA | 83.0 | | 4.0 | |
| Teramoto and Fujita (Ref. 19) | 2013 | 84 | 0.5–3 | 5–20 | All | At least 1 | 103 | 80.0 | | 4.2 | |
| Cascio *et al.* (Ref. 18) | 2012 | 84 | 1.25–3 | ≥3 | All | At least 1 | 148 | 97.0 | 88.0 | 6.1 | 2.5 |
| Guo and Li *et al.* (Ref. 17) | 2012 | 85 | 1.25–3 | ≥3 | All | At least 3 | 111 | 80.0 | 75.0 | 7.4 | 2.8 |
| Camarlinghi *et al.* (Ref. 41) | 2011 | 69 | 0.5–2 | >3 | All | At least 2 | 114 | 80.0 | | 3.0 | |
| Riccardi *et al.* (Ref. 16) | 2011 | 154 | 0.5–3 | ≥3 | All | At least 4 | 117 | 71.0 | 60.0 | 6.5 | 2.5 |
| Tan *et al.* (Ref. 15) | 2011 | 125 | 0.75–3 | ≥3 | All | At least 4 | 80 | 87.5 | | 4.0 | |
| Messay, Hardie, and Rogers (Ref. 14) | 2010 | 84 | 1.3–3 | ≥3 | All | At least 1 | 143 | 82.7 | | 3.0 | |

## 6. CONCLUSION

In this study, a new automatic CAD system designed specifically for large pulmonary nodules was presented. We showed that the proposed dedicated CAD system can identify the vast majority of highly suspicious lesions in thoracic CT scans at the expense of only a small number of false positives.

## ACKNOWLEDGMENT

## APPENDIX: LIDC–IDRI CASES

LIDC–IDRI cases were downloaded from https://public.cancerimagingarchive.net/ncia/login.jsf. Due to thick slice thickness and inconsistent slice spacing, the following cases were excluded: 0078, 0079, 0085, 0086, 0098, 0101, 0103, 0104, 0107, 0110, 0115, 0117, 0127, 0132, 0134, 0136, 0146, 0150, 0151, 0154, 0156, 0267, 0296, 0303, 0310, 0315, 0323, 0334, 0337, 0346, 0347, 0351, 0354, 0355, 0359, 0360, 0365, 0366, 0368, 0378, 0379, 0381, 0382, 0387, 0389, 0390, 0395, 0396, 0414, 0418, 0421, 0426, 0435, 0436, 0439, 0442, 0449, 0475, 0484, 0499, 0504, 0509, 0514, 0518, 0526, 0527, 0533, 0534, 0541, 0550, 0572, 0577, 0595, 0597, 0611, 0635, 0640, 0642, 0672, 0685, 0691, 0706, 0707, 0709, 0713, 0719, 0724, 0737, 0741, 0749, 0750, 0751, 0773, 0798, 0812, 0855, 0861, 0872, 0886, 0892, 0901, 0904, 0910, 0926, 0935, 0937, 0944, 0945, 0946, 0949, 0957, 0958, 0961, 0962, 0972, 0974, 0979, 0980, 0983, 0993, 0998, 1000, 1008.

[a]Electronic mail: arnaud.arindraadiyoso@radboudumc.nl

[1]American Cancer Society, Cancer Facts and Figures 2014, 2014.
[2]D. R. Aberle, A. M. Adams, C. D. Berg, W. C. Black, J. D. Clapp, R. M. Fagerstrom, I. F. Gareen, C. Gatsonis, P. M. Marcus, and J. D. Sicks, "Reduced lung-cancer mortality with low-dose computed tomographic screening," N. Engl. J. Med. **365**, 395–409 (2011).
[3]G. D. Rubin, J. K. Lyo, D. S. Paik, A. J. Sherbondy, L. C. Chow, A. N. Leung, R. Mindelzun, P. K. Schraedley-Desmond, S. E. Zinck, D. P. Naidich, and S. Napel, "Pulmonary nodules on multi-detector row CT scans: Performance comparison of radiologists and computer-aided detection," Radiology **234**, 274–283 (2005).
[4]K. N. Jeon, J. M. Goo, C. H. Lee, Y. Lee, J. Y. Choo, N. K. Lee, M.-S. Shim, I. S. Lee, K. G. Kim, D. S. Gierada, and K. T. Bae, "Computer-aided nodule

detection and volumetry to reduce variability between radiologists in the interpretation of lung nodules at low-dose screening computed tomography," Invest. Radiol. **47**, 457–461 (2012).

[5]Y. Zhao, G. H. de Bock, R. Vliegenthart, R. J. van Klaveren, Y. Wang, L. Bogoni, P. A. de Jong, W. P. Mali, P. M. A. van Ooijen, and M. Oudkerk, "Performance of computer-aided detection of pulmonary nodules in low-dose CT: Comparison with double reading by nodule volume," Eur. Radiol. **22**, 2076–2084 (2012).

[6]S. Matsumoto, Y. Ohno, H. Aoki, H. Yamagata, M. Nogami, K. Matsumoto, Y. Yamashita, and K. Sugimura, "Computer-aided detection of lung nodules on multidetector CT in concurrent-reader and second-reader modes: A comparative study," Eur. J. Radiol. **82**, 1332–1337 (2013).

[7]B. van Ginneken, C. M. Schaefer-Prokop, and M. Prokop, "Computer-aided diagnosis: How to move from the laboratory to the clinic," Radiology **261**, 719–732 (2011).

[8]A. McWilliams, M. C. Tammemagi, J. R. Mayo, H. Roberts, G. Liu, K. Soghrati, K. Yasufuku, S. Martel, F. Laberge, M. Gingras, S. Atkar-Khattra, C. D. Berg, K. Evans, R. Finley, J. Yee, J. English, P. Nasute, J. Goffin, S. Puksa, L. Stewart, S. Tsai, M. R. Johnston, D. Manos, G. Nicholas, G. D. Goss, J. M. Seely, K. Amjadi, A. Tremblay, P. Burrowes, P. MacEachern, R. Bhatia, M.-S. Tsao, and S. Lam, "Probability of cancer in pulmonary nodules detected on first screening CT," N. Engl. J. Med. **369**, 910–919 (2013).

[9]N. Horeweg, J. van Rosmalen, M. A. Heuvelmans, C. M. van der Aalst, R. Vliegenthart, E. T. Scholten, K. Ten Haaf, K. Nackaerts, J.-W. J. Lammers, C. Weenink, H. J. Groen, P. van Ooijen, P. A. de Jong, G. H. de Bock, W. Mali, H. J. de Koning, and M. Oudkerk, "Lung cancer probability in patients with CT-detected pulmonary nodules: A prespecified analysis of data from the Nelson trial of low-dose CT screening," Lancet Oncol. **15**, 1332–1341 (2014).

[10]R. Yip, C. I. Henschke, D. F. Yankelevitz, and J. P. Smith, "CT screening for lung cancer: Alternative definitions of positive test result based on the national lung screening trial and international early lung cancer action program databases," Radiology **273**, 591–596 (2014).

[11]D. S. Gierada, P. Pinsky, H. Nath, C. Chiles, F. Duan, and D. R. Aberle, "Projected outcomes using different nodule sizes to define a positive CT lung cancer screening examination," J. Natl. Cancer Inst. **106**, 284–291 (2014).

[12]F. Beyer, L. Zierott, E. M. Fallenberg, K. U. Juergens, J. Stoeckel, W. Heindel, and D. Wormanns, "Comparison of sensitivity and reading time for the use of computer-aided detection (CAD) of pulmonary nodules at MDCT as concurrent or second reader," Eur. Radiol. **17**, 2941–2947 (2007).

[13]K. Murphy, B. van Ginneken, A. M. R. Schilham, B. J. de Hoop, H. A. Gietema, and M. Prokop, "A large scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification," Med. Image Anal. **13**, 757–770 (2009).

[14]T. Messay, R. C. Hardie, and S. K. Rogers, "A new computationally efficient CAD system for pulmonary nodule detection in CT imagery," Med. Image Anal. **14**, 390–406 (2010).

[15]M. Tan, R. Deklerck, B. Jansen, M. Bister, and J. Cornelis, "A novel computer-aided lung nodule detection system for CT images," Med. Phys. **38**, 5630–5645 (2011).

[16]A. Riccardi, T. S. Petkov, G. Ferri, M. Masotti, and R. Campanini, "Computer-aided detection of lung nodules via 3d fast radial transform, scale space representation, and zernike mip classification," Med. Phys. **38**, 1962–1971 (2011).

[17]W. Guo and Q. Li, "High performance lung nodule detection schemes in CT using local and global information," Med. Phys. **39**, 5157–5168 (2012).

[18]D. Cascio, R. Magro, F. Fauci, M. Iacomi, and G. Raso, "Automatic detection of lung nodules in CT datasets based on stable 3D mass–spring models," Comput. Biol. Med. **42**, 1098–1109 (2012).

[19]A. Teramoto and H. Fujita, "Fast lung nodule detection in chest CT images using cylindrical nodule-enhancement filter," Int. J. Comput. Assisted Radiol. Surg. **8**, 193–205 (2013).

[20]M. Tan, R. Deklerck, J. Cornelis, and B. Jansen, "Phased searching with neat in a time-scaled framework: Experiments on a computer-aided detection system for lung nodules," Artif. Intell. Med. **59**, 157–167 (2013).

[21]W.-J. Choi and T.-S. Choi, "Automated pulmonary nodule detection system in computed tomography images: A hierarchical block classification approach," Entropy **15**, 507–523 (2013).

[22]E. L. Torres, E. Fiorina, F. Pennazio, C. Peroni, M. Saletta, N. Camarlinghi, M. E. Fantacci, and P. Cerello, "Large scale validation of the M5L lung CAD on heterogeneous CT datasets," Med. Phys. **42**, 1477–1489 (2015).

[23]B. van Ginneken, S. G. Armato, B. de Hoop, S. van de Vorst, T. Duindam, M. Niemeijer, K. Murphy, A. M. R. Schilham, A. Retico, M. E. Fantacci, N. Camarlinghi, F. Bagagli, I. Gori, T. Hara, H. Fujita, G. Gargano, R. Belloti, F. D. Carlo, R. Megna, S. Tangaro, L. Bolanos, P. Cerello, S. C. Cheran, E. L. Torres, and M. Prokop, "Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The ANODE09 study," Med. Image Anal. **14**, 707–722 (2010).

[24]H. MacMahon, J. H. M. Austin, G. Gamsu, C. J. Herold, J. R. Jett, D. P. Naidich, E. F. Patz, S. J. Swensen, and the Fleischner Society, "Guidelines for management of small pulmonary nodules detected on CT scans: A statement from the Fleischner Society," Radiology **237**, 395–400 (2005).

[25]R. J. van Klaveren, M. Oudkerk, M. Prokop, E. T. Scholten, K. Nackaerts, R. Vernhout, C. A. van Iersel, K. A. M. van den Bergh, S. van 't Westeinde, C. van der Aalst, E. Thunnissen, D. M. Xu, Y. Wang, Y. Zhao, H. A. Gietema, B. de Hoop, H. J. M. Groen, G. H. de Bock, P. van Ooijen, C. Weenink, J. Verschakelen, J. J. Lammers, W. Timens, D. Willebrand, A. Vink, W. Mali, and H. J. de Koning, "Management of lung nodules detected by volume CT scanning," N. Engl. J. Med. **361**, 2221–2229 (2009).

[26]J. H. Pedersen, H. Ashraf, A. Dirksen, K. Bach, H. Hansen, P. Toennesen, H. Thorsen, J. Brodersen, B. G. Skov, M. Dossing, J. Mortensen, K. Richter, P. Clementsen, and N. Seersholm, "The Danish randomized lung cancer CT screening trial–overall design and results of the prevalence round," J. Thorac. Oncol. **4**, 608–614 (2009).

[27]N. L. S. T. R. Team, D. R. Aberle, C. D. Berg, W. C. Black, T. R. Church, R. M. Fagerstrom, B. Galen, I. F. Gareen, C. Gatsonis, J. Goldin, J. K. Gohagan, B. Hillman, C. Jaffe, B. S. Kramer, D. Lynch, P. M. Marcus, M. Schnall, D. C. Sullivan, D. Sullivan, and C. J. Zylak, "The national lung screening trial: Overview and study design," Radiology **258**, 243–253 (2011).

[28]S. G. Armato, G. McLennan, L. Bidaut, M. F. McNitt-Gray, C. R. Meyer, A. P. Reeves, B. Zhao, D. R. Aberle, C. I. Henschke, E. A. Hoffman, E. A. Kazerooni, H. MacMahon, E. J. R. V. Beek, D. Yankelevitz, A. M. Biancardi, P. H. Bland, M. S. Brown, R. M. Engelmann, G. E. Laderach, D. Max, R. C. Pais, D. P. Y. Qing, R. Y. Roberts, A. R. Smith, A. Starkey, P. Batrah, P. Caligiuri, A. Farooqi, G. W. Gladish, C. M. Jude, R. F. Munden, I. Petkovska, L. E. Quint, L. H. Schwartz, B. Sundaram, L. E. Dodd, C. Fenimore, D. Gur, N. Petrick, J. Freymann, J. Kirby, B. Hughes, A. V. Casteele, S. Gupte, M. Sallamm, M. D. Heath, M. H. Kuhn, E. Dharaiya, R. Burns, D. S. Fryd, M. Salganicoff, V. Anand, U. Shreter, S. Vastagh, and B. Y. Croft, "The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans," Med. Phys. **38**, 915–931 (2011).

[29]D. P. Naidich, A. A. Bankier, H. MacMahon, C. M. Schaefer-Prokop, M. Pistolesi, J. M. Goo, P. Macchiarini, J. D. Crapo, C. J. Herold, J. H. Austin, and W. D. Travis, "Recommendations for the management of subsolid pulmonary nodules detected at CT: A statement from the Fleischner Society," Radiology **266**, 304–317 (2013).

[30]D. Manos, J. M. Seely, J. Taylor, J. Borgaonkar, H. C. Roberts, and J. R. Mayo, "The lung reporting and data system (lu-rads): A proposal for computed tomography screening," Can. Assoc. Radiol., J. **65**, 121–134 (2014).

[31]D. M. Xu, H. Gietema, H. de Koning, R. Vernhout, K. Nackaerts, M. Prokop, C. Weenink, J. Lammers, H. Groen, M. Oudkerk, and R. van Klaveren, "Nodule management protocol of the NELSON randomised lung cancer screening trial," Lung Cancer **54**, 177–184 (2006).

[32]C. Jacobs, E. M. van Rikxoort, T. Twellmann, E. T. Scholten, P. A. de Jong, J. M. Kuhnigk, M. Oudkerk, H. J. de Koning, M. Prokop, C. Schaefer-Prokop, and B. van Ginneken, "Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images," Med. Image Anal. **18**, 374–384 (2014).

[33]E. M. van Rikxoort, B. de Hoop, M. A. Viergever, M. Prokop, and B. van Ginneken, "Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection," Med. Phys. **36**, 2934–2947 (2009).

[34]J. M. Kuhnigk, V. Dicken, L. Bornemann, A. Bakai, D. Wormanns, S. Krass, and H. O. Peitgen, "Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans," IEEE Trans. Med. Imaging **25**, 417–434 (2006).

[35]Q. Li, S. Sone, and K. Doi, "Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans," Med. Phys. **30**, 2040–2051 (2003).

[36]N. Horeweg, C. M. van der Aalst, E. Thunnissen, K. Nackaerts, C. Weenink, H. J. M. Groen, J.-W. J. Lammers, J. G. Aerts, E. T. Scholten, J. van

Rosmalen, W. Mali, M. Oudkerk, and H. J. de Koning, "Characteristics of lung cancers detected by computer tomography screening in the randomized Nelson trial," Am. J. Respir. Crit. Care Med. **187**, 848–854 (2013).

[37]V. Vapnik, *The Nature of Statistical Learning Theory* (Springer-Verlag, New York, NY, 1995).

[38]C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," ACM Trans. Intell. Syst. Technol. **2**, 1–27 (2011), available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[39]M. S. Brown, P. Lo, J. G. Goldin, E. Barnoy, G. H. J. Kim, M. F. McNitt-Gray, and D. R. Aberle, "Toward clinically usable CAD for lung can-cer screening with computed tomography," Eur. Radiol. **24**, 2719–2728 (2014).

[40]W.-J. Choi and T.-S. Choi, "Automated pulmonary nodule detection based on three-dimensional shape-based feature descriptor," Comput. Methods Programs Biomed. **113**, 37–54 (2013).

[41]N. Camarlinghi, I. Gori, A. Retico, R. Bellotti, P. Bosco, P. Cerello, G. Gargano, E. L. Torres, R. Megna, M. Peccarisi, and M. E. Fantacci, "Combination of computer-aided detection algorithms for automatic lung nodule identification," Int. J. Comput. Assisted Radiol. Surg. **7**, 455–464 (2011).