Classifier Ensemble Based on Computed Tomography Attenuation Patterns for Computer-Aided Detection System

FERNANDO ROBERTO PEREIRA,1, JOÃO MARIO CLEMENTIN DE ANDRADE2, DANTE LUIZ ESCUSSIATO2, AND LUCAS FERRARI DE OLIVEIRA3

1Federal Institute of Education, Science and Technology of Santa Catarina, Canoinhas 89466-312, Brazil
2Department of Internal Medicine, Federal University of Paraná, Curitiba 80060-900, Brazil
3Department of Informatics, Federal University of Paraná, Curitiba 82590-300, Brazil

Corresponding author: Fernando Roberto Pereira (fernando.pereira@ifsc.edu.br)

This work was supported in part by the Coordination for the Improvement of Higher Education Personnel (CAPES).

ABSTRACT Cancer is one of the leading causes of mortality worldwide, specifically lung cancer. Computer-Aided Detection (CADe) systems are being proposed to assist radiologists in the task of pulmonary nodule detection. In this paper, we propose a CADe system that uses Deep Convolutional Neural Network (DCNN). In the Nodule Candidate Detection (NCD) step, we used Mask Region-Convolutional Neural Network (Mask R-CNN) to detect bounding boxes in 2D slices of low-dose Computed Tomography (CT) scans. In the False Positive Reduction (FPR) step, we used a classifier ensemble based on CT attenuation patterns to boost 3D pulmonary nodule classification performance. The final confidence index generated by the CADe system to the pulmonary nodule candidates is the average of the prediction obtained with the NCD and FPR steps. The CADe system was validated on the publicly available LUng Nodule Analysis 2016 (LUNA16) challenge and obtained a sensitivity of 94.90% and an average of 1.0 False Positives per scan (FP/Scan), against 96.90% of the proposal that combines different existing CADe systems. To the best of our knowledge, our proposal has one of the best results of CADe systems, outperforming other state-of-the-art individual methods.

INDEX TERMS Classifier ensemble, computer-aided detection system, convolutional neural network, pulmonary nodule, thoracic computed tomography scan.

I. INTRODUCTION Cancer is one of the biggest public health problems worldwide. In 2018, cancer was responsible for 9.6 million deaths globally, about one-sixth of the fatalities. Lung cancer is one of the most common cancers, with 2.09 million cases, and has caused approximately 1.76 million deaths [1].

According to the Brazilian National Cancer Institute (INCA) [2], lung cancer is the second most common type of this disease in Brazil (not counting non-melanoma skin cancer). It accounted over more than 26 thousand deaths in 2015. The five-year survival rate for patients diagnosed with lung cancer is 18.0% (15.0% for men and 21.0% for women). Only 16% of cancer cases are diagnosed at an early stage (localized lesion). In this case, the five-year survival rate is 56.0%. At the end of the 20th century, lung cancer was one of the leading causes of preventable deaths.

The American Cancer Society estimates about 235,760 new cases of lung cancer and approximately 131,880 deaths in the United States for the year 2021. The number of new cases and deaths from lung cancer would decrease due to people stopping smoking and due to advances in early detection and treatment for this disease. If a pulmonary lesion (nodule) is detected at an earlier stage when it is small and before it has spread throughout the body, it is more likely to be successfully treated [3].

There are several image exams modalities that are able to evaluate pulmonary nodules, such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and radiograph. CT has a relative low cost, wide availability and good sensitivity. It is the preferred exam modality for pulmonary nodule evaluation in many clinical situations. Nowadays
low-dose CT is the lung cancer screening exam recommended for patients with considerable smoking history [4].

CT scan image analysis is challenging and can be exhausting to radiologists. Factors such as distraction, fatigue and limitations of professional experience may result in diagnostic errors [5], [6]. Computer-Aided Detection (CADe) systems are being developed to automate pulmonary nodule detection and measurement to assist the radiologists. These solutions are used as a second opinion in diagnostics [7], [8].

With the development and deployment of images database, such as the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), the validation of CADe systems is possible based on expert diagnostics [9]. The LIDC-IDRI database consists of low-dose thoracic CT scan images and the diagnosis of the lesions found [10]. To evaluate the performance of CADe systems, it was created the Lung Nodule Analysis 2016 (LUNA16) challenge. The metric used is the Competition Performance Metric (CPM) score obtained by the Free Response Receiver Operating Characteristic (FROC) curve (i.e., the average sensitivity at seven predefined False Positives per Scan (FPs/Scan) rates: 1/8, 1/4, 1/2, 1, 2, 4, 8) [11].

In several previous studies, investigators have developed automated CADe systems. Typically, CADe systems consist of five main steps: i) data acquisition; ii) pre-processing; iii) lung parenchyma segmentation; iv) Nodule Candidate Detection (NCD); and v) False Positive Reduction (FPR) [12]. However, some studies summarize only two major steps, NCD and FPR respectively [13]–[15].

CADe systems have been proposed with different approaches. Traditional approaches use hand-crafted features that require domain expertise [16]. These features have limited representation capability, and it is insufficient to deal with the large variations of pulmonary nodules [17]. The pulmonary nodules have variations in sizes, shapes, and locations. Furthermore, the contextual environments around the nodules presented differently for each category, such as solitary nodules, ground-glass opacity nodules, cavity nodules, and pleural nodules [18]. Solutions to pulmonary nodule candidate detection using Deep Convolutional Neural Network (DCNN) are being proposed. These solutions use learning-based features and show better results than traditional approaches [19].

This paper proposes an automated CADe system for pulmonary nodule detection in low-dose thoracic CT scan images with two major steps, NCD and FPR respectively. In the first step, we use an end-to-end object detection method based on two-dimensional (2D) DCNN to identify regions as pulmonary nodules on axial slices of thoracic CT scan. In the second step, a three-dimensional (3D) DCNN classifier ensemble based on CT attenuation patterns removes false positive pulmonary nodule candidates. The main contributions of this paper are the following:

1) We employ an image with three channels and 8 bits per channel composed of slices of thoracic CT scans. In this way, we used an end-to-end object detection method called Mask Region-Convolutional Neural Network (Mask R-CNN) [20] pre-trained with natural images to 2D nodule candidates detection.

2) We proposed an algorithm to discard pulmonary nodule candidate detections that have centroid close to each other in an exam, removing redundant candidates.

3) We designed a 3D DCNN network based on Visual Geometry Group (VGG) [21] that supports patches with different sizes to classify nodule candidates.

4) We proposed a classifier ensemble based on CT attenuation patterns to boost the performance for nodule classification.

The rest of this paper is organized as it follows: Section II introduces the related works about automated CADe systems for pulmonary nodule candidate detection; Section III describes the proposed methodology; Section IV presents the experiments and results obtained and presents an ablation study about the classifier ensemble. Finally, we conclude this paper in Section V.

II. RELATED WORK

The available literature offers acknowledged studies with DCNN and it has achieved excellent results in the field of automated CADe systems for pulmonary nodule detection.

Some of these researches work combining individual systems. Setio et al. [13] proposed a CADe system for detecting pulmonary nodules using ablation study DCNN with two steps. They used three existing CADe systems to detect nodule candidates specifically designed for solid, subsolid and large nodules in the first step. For each candidate, it was extracted a set of 2D patches from differently oriented planes. In the second step, a classifier ensemble based on 2D DCNN was proposed, for which the outputs are combined using a dedicated fusion method to get the final classification. Their CADe system achieved a sensitivity that reaches 85.40% and an average of 1.0 FP/Scan on the LUNA16 dataset. In another work, Setio et al. [11] combined seven pulmonary nodule candidate detection systems and five false positive reduction systems developed per participants of the LUNA16 challenge. Combining these solutions achieved a sensitivity of 96.90% and an average of 1.0 FP/Scan.

Some works collectively use traditional approaches and deep learning models for nodule detection. For example, Zhang et al. [22] used multi-scale Laplacian of Gaussian (LoG) filters and prior shape and size constraints to nodule candidates detection. Further, a densely dilated 3D DCNN was applied to reduce the number of FPs and estimate nodule diameters, so it has obtained a sensitivity of 94.90% and an average of 1.0 FP/Scan on the LUNA16 dataset.

In the last few years, solutions have been proposed exclusively with deep learning models. Dou et al. [23] proposed a solution with two stages. In the first stage, they designed a 3D Fully Convolutional Network (FCN) to screen the nodule candidates. In the second stage, they created a hybrid loss residual network that harnesses the location and size information to the pulmonary nodule detection, which
achieved on the LUNA16 dataset a sensitivity of 86.50% and an average of 1.0 FP/Scan. Zheng et al. [14] proposed a 2D CNN based on the U-net [24] that takes Maximum intensity projection (MIP) images of different slab thicknesses axial section slice as an input. Further, they applied multiple 3D DCNN to reduce false positive candidates and fused the final predictions. Their CADe system achieved a sensitivity that reaches 92.67% and an average of 1.0 FP/Scan on the LUNA16 dataset. In another approach, Ozdemir et al. [25] proposed a system with two main components: i) a CADe module that detects and segments suspicious lung nodules, and ii) a Computer-Aided Diagnosis (CADx) module that performs analyzing suspicious lesions from CADe. The CADe module comprises a 3D FCN based on the V-Net architecture [26] and a 3D scoring network, which computes refined nodule probability estimates for false positive reduction. The sensitivity of their solution is 94.20% and an average of 1.0 FP/Scan on the LUNA16 dataset. Tang et al. [27] proposed a scheme that is composed of many stages. Initially, the pulmonary parenchyma area is segmented by various methods. After, a scheme based on the 3D U-Net [24] and transfer learning method is used to detect pulmonary nodules, so it has obtained on the LUNA16 dataset a sensitivity of 92.40% and an average of 4.0 FP/Scan. Zhu et al. [28] proposed a Generative Adversarial Network (GAN) architecture called Functional-Realistic GAN (FRGAN) to create realistic high-resolution images. Thus, they used data augmentation to balance the ratio of the positive and negative nodule samples. In the experiments, used the CADe system CUMedVis [17] and, the LUNA16 and Non-Small-Cell Lung Cancer (NSCLC) datasets. The CPM score of their scheme is 91.50% with 1304 nodules.

Deep learning techniques have been achieved on the state-of-the-art results for object detection and instance segmentation in computer vision challenges. Most notably are region-based DCNN [20], [29]–[31] techniques capable of achieving great results on a range of object detection tasks, amongst them, nodule detection. For example, Ding et al. [32] proposed a novel pulmonary nodule detection approach based on DCNNs. Initially, they used a deconvolutional structure Faster Region-based CNN (R-CNN) [31] for candidate detection on axial slices. Then, a 3D DCNN is presented for the false positive reduction, which achieved on the LUNA16 dataset a sensitivity of 92.20% and an average of 1.0 FP/Scan. Zhu et al. [19] proposed a fully automated cancer diagnosis system called DeepLung. Two 3D Dual Path Networks (DPN) DCNN were designed for nodule detection and classification respectively. Specifically, a 3D Faster R-CNN is designed for nodule detection. The solution achieved a CPM score of 84.20% without any FP nodule reduction step on the LUNA16 dataset. Cai et al. [33] proposed a detection and segmentation method for pulmonary nodule 3D visualization diagnosis based on Mask R-CNN [20] and ray-casting volume rendering algorithm, respectively. It was applied the Feature Pyramid Network (FPN) with CNN to explore multi-scale feature maps fully. Their CADe system achieved on the LUNA16 dataset a sensitivity that reaches 88.10% and an average of 1.0 FP/Scan. Peng et al. [34] proposed two types of 3D multi-scale DCNN for the NCD and FPR steps, respectively. In the first step, a CNN was used to extract multi-scale feature of pulmonary nodules and a Region Proposal Network (RPN) structure was used to determine region candidates. In the second step, another 3D multi-scale DCNN was used to classify the nodule candidates. The CPM score of their scheme is 92.30% on the LUNA16 dataset. In another approach, Mei et al. [35] proposed a CADe system based on DCNN named SANet. Initially, they used a 3D RPN to generate pulmonary nodule candidates. Then, a false positive reduction module uses multi-scale feature maps. The experiments were carried out on the LUNA16 and PN9 (new dataset proposed) datasets. The sensitivity of their solution is 90.09% and an average of 1.0 FP/Scan on the LUNA16 dataset.

III. METHOD

We proposed a CADe system based on CT attenuation patterns for pulmonary nodule detection. In each step of the CADe system, attenuation ranges were used for highlight different visual patterns. We have used patterns with higher CT numbers to represent consolidation and nodules. Fig. 1 illustrates the CADe system overview proposed. Attenuation ranges were used on Interstitial Lung Diseases (ILD) to improve visibility or visual separation among disease categories [36]. Ideally, in CADe systems with two steps, the NCD step should be designed for screening nodule candidates with low false positives rates and high sensitivity, but that is not always possible. Hence, the false positive nodule candidates may be corrected by the FPR step, generally by classification techniques. However, false negative cases have no opportunity to be corrected after the NCD step. Moreover, systems with high false positives rates in the DCN step need robust techniques in the FPR step. Therefore, we designed the NCD step of our CADe system to get reasonable rates between false positives and sensitivity.

In the NCD step, the 2D Mask R-CNN method was used for object detection and instance segmentation in low-dose thoracic CT scan images (Fig. 2a). This method generates bounding boxes for each interest object in the image, i.e., 2D nodule candidate detection (Fig. 2b) (Section III-C2). Instances generated outside the lung parenchyma were discarded (Fig. 2c). We employ a scheme to segment the lung parenchyma (Section III-C3). Then, we used the algorithm.
proposed in Pereira et al. [37] to group 2D bounding boxes on nodules (Fig. 2d). We used Jaccard overlap bigger than or equal to the threshold of 0.2. Finally, we used the algorithm proposed to discard pulmonary nodule candidate detections that have the centroid close to each other in an exam (Fig. 2e) (Section III-C4). The NCD step output comprises a set of coordinates and confidence indexes of nodule candidates (Fig. 2f).

In the FPR step, we used a classifier ensemble to determine nodule candidate detections generated in the previous step (Fig. 2g). The ensemble combined 3D DCNN models of the same architecture (Fig. 2i). Thus, the final prediction of the classifier ensemble comprises a set of confidence indexes (Fig. 2j). However, the models were trained with different samples regarding attenuation ranges, size 3D patches, and affine transformations (rotation and shifting) in each epoch (Fig. 2h).

The final prediction of the nodule candidates was calculated (Fig. 2k) with the confidence index obtained by NCD (Fig. 2f) and FPR (Fig. 2j) steps.

A. DATASET

Our experiments were performed using the thoracic CT scan images from the publicly available LUNA16 dataset. It contains only CT scans with slice thickness less than or equal to 2.5 millimeters (mm). Inconsistent slice spacing or missing slices were excluded, leaving 888 CT scans out of a total of 1,018 CT scans from the LIDC-IDRI database. It was defined 1,186 positive nodules that at least three out of four radiologists believe that the nodule diameter is bigger than 3 mm [11]. The LUNA16 dataset is split into 10 subsets. Thus, we perform 10-fold cross-validation to evaluate the performance of the CADe system.

B. PRE-PROCESSING

The original thoracic CT scan images were pre-processing with different attenuation ranges in each step. Thus, we applied a linear transformation to the CT number values of the original image in Hounsfield Units (HU) inside a defined attenuation range. CT number values below and over of the range is mapped to a constant. Equation (1) gives the mapping of the CT number values. Fig. 3 illustrates four different attenuation ranges.

\[
\alpha = \begin{cases} 
\xi, & \beta < \gamma \\
(\beta - \gamma) \times \frac{(\delta - \xi)}{(\lambda - \gamma)} + \xi, & \gamma \leq \beta \leq \lambda \\
\delta, & \beta > \lambda 
\end{cases}
\] (1)

FIGURE 2. Flow chart of the NCD and FPR steps. a) CT scan. b) 2D mask R-CNN method (based on [20]). c) Discard bounding boxes (in yellow). d) Group 2D bounding boxes (based on [37]). e) Discard nodule candidate. f) NCD step output (set of coordinates and confidence indexes). g) Nodule candidate. h) Affine transformations. i) Classifier ensemble. j) Classifier ensemble output (set of confidence indexes). k) FPR step output (final confidence index).

FIGURE 3. Pre-processing thoracic CT scan images. a) HU \(\gamma = -1000\) and \(\lambda = -200\). b) HU \(\gamma = -1000\) and \(\lambda = 0\). c) HU \(\gamma = -1000\) and \(\lambda = 200\). d) HU \(\gamma = -1000\) and \(\lambda = 400\).
Among them, $\alpha$ is the output value of the linear transformation, $\beta$ is the input value of the original CT image. The $\zeta$ and $\delta$ are the minimum and maximum of the output CT number values, respectively. The $\gamma$ and $\lambda$ are the minimum and maximum of the input CT number values, respectively.

### C. NODULE CANDIDATE DETECTION

We designed the NCD step to identify and localize pulmonary nodules with the highest possible sensitivity. Thus, we try to minimize false negative, i.e., cases with no opportunity to be corrected. The FPR step tries to correct the false positive cases.

1) DATA ACQUISITION

The LUNA16 dataset ignores nodules annotations larger than 30 mm. Thus, machine learning-based CADe systems trained with this dataset could learn to ignore large nodules [25]. Therefore, due to the limited annotations provided by the LUNA16 dataset, we used additional training data in the NCD step from the one publicly available in Reeves and Biancardi [38], i.e., 2D nodules annotations performed by experts. We also used an additional validation set with 123 CT scans from the LIDC-IDRI database.

Due to the heterogeneity of the low-dose thoracic CT scan images and the diversity of tissues present in the exams, we defined an attenuation range with the CT numbers $\gamma = -1000$ HU and $\lambda = 400$ HU, and rescaled them to $\zeta = 0$ and $\delta = 255$ (Section III-B). Thus, we got slices of CT scan with 8-bit and kept the isotropic resolution of the slices.

Then, we generated images with three channels composed of slices of CT scans. We defined $E = \{e_1, e_2, \ldots\}$ how a thoracic CT scan and $e$ a slice of the exam in the axial plane. We also defined $I = \{e \mid e \in E\}$ as an image, i.e., $I \subseteq E$. Therefore, an image $I$ contemplates three adjacent slices $I = \{e_{i-1}, e_i, e_{i+1}\}$. Besides, the slices that make up the ends of an exam generate exceptions, in this case, two image channels correspond to the same slice.

Thus, the employed image contemplates context information between the slices, an essential aspect for nodule detection since a nodule can contemplate more than one slice. Another reason for using the three-channel images was to accommodate the DCNN architecture pre-trained with natural images to 2D inputs.

2) OBJECT DETECTION METHOD

We used the end-to-end object detection Mask R-CNN method [20] to predict bounding boxes for each 2D nodule candidate in the input 2D image. We employ the architecture ResNeXt as the backbone network, configured with 101 convolutional layers, the cardinality of 64, and the bottleneck width of 4d [39]. Motivated by the success of deep transfer learning in CADe systems [40], [41]. We used the transfer learning technique with pre-trained weights of the ImageNet database (i.e., 2D natural images) [42]. Next, the weights were fine-tuning with our dataset (Section III-C1). The problem was modeled with one class nodule.

At the training stage, the network was optimized by Stochastic Gradient Descent (SGD) algorithm with backpropagation, where the initial learning rate was set to $10^{-3}$ and gradually reduced up to $10^{-5}$, the momentum parameter was set to 0.9, the gamma was set to 0.1, the weight decay was set to $10^{-4}$, and the mini-batch size was set to 4. We regularized our models by the early stopping strategy to avoid overfitting. In the testing stage, we accepted the detection confidence index (i.e., $P_{NCD}$) greater than or equal to the threshold of 0.1.

3) LUNG PARENCHYMA SEGMENTATION

Thoracic CT scan images may contain imaging sensor noises, tissues, cartilages, calcifications, bones, organs, etc. Therefore, we employ a scheme to segment the lung parenchyma, it contemplates many stages, as shown in Fig. 4.

Initially, we defined an attenuation range with the CT numbers $\gamma = -1000$ HU and $\lambda = 400$ HU, and rescaled them to $\zeta = 0$ and $\delta = 1.0$ (Fig. 4b). Next, we used a binary threshold with the mean value of the whole CT scan to divide the chest into the outside and inside regions (Fig. 4c). We also used a binary min/max curvature flow filter to denoise due to processing previous. Then, we segment the thoracic with the neighborhood connected filter using eight seeds, one to each corner from the CT scan (Fig. 4d). Irrelevant objects were removed through the binary dilation filter with the ball structure element of radius seven (Fig. 4e).

Furthermore, we applied the logical operation OR between the resulting image and out of the curvature flow filter (Fig. 4f), followed by the logical operation NOT (Fig. 4g). After, the two largest connected components were selected as the internal lung area. However, there were holes in the mask generated during the segmentation process, such as noise, graininess, vessels, tissues, nodules attached to the pleura, etc. In sequence, we applied the morphological operation DILATION with the ball structure element of radius five (Fig. 4h), and we applied the convex hull algorithm per slice (Fig. 4i).

![FIGURE 4. Steps to generate the lung parenchyma mask. a) Original slice. b) Attenuation range. c) Binary threshold. d) Neighborhood connected filter. e) Binary dilation filter. f) Or logical operator. g) Not logical operator. h) Binary dilation. i) Convex hull.](image-url)
Therefore, the generated parenchyma mask served to discard the 2D nodule candidate detections with the center outside the lung parenchyma generated by the Mask R-CNN method. Fig. 4i illustrates the detections on each slice. In yellow, the detections disposable, and in red, the detections non-disposable, which the algorithm [37] uses to group and generate 3D nodule candidate detections.

4) DISCARD NODULE CANDIDATE DETECTION

In the list of nodule candidates generated from algorithm [37], we identified some candidates whose centroid was too close to each other. Problem repeatedly identified in CADe systems [11], [13], [14]. Therefore, we designed an algorithm to remove redundant nodule candidates. Details of the algorithm are shown in Algorithm 1.

The proposed algorithm aims to select a set of nodule candidates to identify a region of an exam. The relationships between nodule candidates using a directed graph were decomposed into directed subgraphs to distinguish different relationships between nodule candidates. The hypothesis is that each directed subgraph represents a region that possibly identifies a nodule. In summary, given a set of nodule candidates \( C_{in} \) per exam, the algorithm generates a set of directed graphs. Each graph can generate one or more directed subgraphs distinct based on the criterion out-degree (outDegree greater than zero) and path length (pathLength equal one). Thus, each resulting subgraph can generate one or two nodule candidates.

However, there are patients with multiple nodules with the centroids close to each other. For example, the algorithm can generate one or two nodule candidates for each directed subgraph. The algorithm checks if there is a difference between the vertex with a greater out-degree and the vertex with a greater confidence index. Thus, the two vertices have characteristics to identify distinct nodules with close centroids.

Fig. 5 illustrates one hypothetical example of nodule candidate detections of an exam, represented in 2D, to help understand the algorithm iterations. Fig. 5a shows three nodule candidates, \( c_0 \) (black), \( c_1 \) (cyan), and \( c_2 \) (red), with the confidence index 0.9, 0.8, and 0.7, respectively. Fig. 5d shows three nodule candidates, \( c_3 \) (black), \( c_4 \) (cyan), and \( c_5 \) (red), with the confidence index 0.9, 0.8, and 0.7, respectively. Therefore, there are six candidate detections in the exam \( C_{in} = \{ c_0, c_1, c_2, c_3, c_4, c_5 \} \). Each candidate has five data as illustrated in Algorithm 1, e.g., \( c_0 = \{ x, y, z, d, p = 0.9 \} \).

Initially, the proposed algorithm performs all combinations between the candidates of an exam. When comparing two candidates, the algorithm checks overlapped between candidates and used the radius as a criterion. Thus, it can determine the relationships between the candidates and creates a set of directed graphs (line 1 up to 11 of the Algorithm 1).

To exemplify these instructions, Fig. 5b shows oriented graph obtained \( G_0(V_0, A_0) \), composed of vertices \( V_0 = \{ v_0, v_1, v_2 \} \) (mapped from candidates \( c_0, c_1 \), and \( c_2 \), respectively) and edges \( A_0 = \{ \{ v_0 - v_1 \}_0, \{ v_0 - v_2 \}_1, \{ v_1 - v_0 \}_2 \} \). Fig. 5e shows oriented graph obtained \( G_1(V_1, A_1) \), composed of vertices \( V_1 = \{ v_3, v_4, v_5 \} \) (mapped from candidates \( c_3, c_4 \), and \( c_5 \), respectively) and edges \( A_1 = \{ \{ v_3 - v_4 \}_0, \{ v_4 - v_3 \}_1 \} \).

Further, the algorithm creates all possible distinct directed subgraphs that obey the criterions out-degree and path length for each directed graph generated. Therefore, each directed graph can generate one or more directed subgraphs distinct. Among the subgraphs that contain the same vertices, only remains the subgraph with greater out-degree and greater confidence index. When multiple directed subgraphs present the same values for out-degree and confidence index, arbitrary selection occurs (line 12 of the Algorithm 1).

We present through examples the result of the function. Fig. 5c shows the directed subgraphs distinct \( H_0 \) and \( H_1 \), they have the same vertices, and the vertices \( v_0 \) and \( v_1 \) have the same out-degree (outDegree = 2). However, the vertex \( v_0 \) has a greater confidence index. Therefore, the subgraph \( H_1 \) is discarded. Fig. 5f shows the directed subgraphs distinct \( H_2 \) and \( H_3 \), and they do not have the same vertices. Therefore, the two subgraphs remain.

Finally, the algorithm determines if each directed subgraph generates one or two nodule candidates. For decision-making, it evaluates if the vertex with the greater out-degree also has the greater confidence index, in this case, it generates only one candidate (e.g., vertex \( v_0 \) of the subgraph \( H_0 \) − Fig. 5c, and vertex \( v_3 \) of the subgraph \( H_3 \) − Fig. 5f). However, if the vertex with the greater out-degree is not the vertex with the greater confidence index, in this case, it generates a nodule candidate for each vertex (e.g., vertices \( v_3 \) and \( v_4 \) of the subgraph \( H_3 \) − Fig. 5f) (line 13 up to 20 of the Algorithm 1). Besides, if a vertex is listed more than once (e.g., vertex \( v_3 \) is listed in the subgraphs \( H_2 \) and \( H_3 \) − Fig. 5f), it only results in a nodule candidate (e.g., \( v_3 \) − Fig. 5f).

In the example in Fig. 5a, there are three nodule candidates, and after the execution of the algorithm, only remained \( c_0 \).
\begin{algorithm}
\caption{Discard Nodule Candidate Detection}
\textbf{Require:} \( C_{in} \), outDegree, pathLength \{Set of 3D nodule candidate \( C_{in} = \{c_0, c_1, \ldots\} \) to one scan \( E \), where \( c_i \) is a candidate mapped on \([x, y, z, d, p]\). We defined \([x, y, z]\) how location information, diameter \([d]\) and confidence index \([p]\).\}
\textbf{Ensure:} \( C_{out} \) \{Set of 3D nodule candidate resulting.\}
1: \textbf{for all} \( c_a \in C_{in} \) \textbf{do}
2: \hspace{1em} \textbf{for all} \( c_b \in C_{in} \) \textbf{do}
3: \hspace{2em} \textbf{if} \( c_a \neq c_b \) \textbf{then}
4: \hspace{3em} \( db \leftarrow \text{getDistanceBetween}(c_a, c_b) \) \{Returns the distance between \( c_a \) and \( c_b \).\}
5: \hspace{3em} \textbf{if} \( db < c_a \) \textbf{then}
6: \hspace{4em} \( V \leftarrow \text{setVertexIfNotExists}(V, c_a, c_b) \) \{Inserts the vertices \( c_a \) and \( c_b \) in \( V \), if not exists.\}
7: \hspace{4em} \( A \leftarrow \text{setEdge}(A, \{c_a - c_b\}) \) \{Inserts the edge \([c_a - c_b]\) in \( A \).\}
8: \hspace{3em} \textbf{end if}
9: \hspace{3em} \textbf{end if}
10: \hspace{1em} \textbf{end for}
11: \textbf{end for}
12: \( H \leftarrow \text{getAllDirectedSubgraphDistinct}(V, A, \text{outDegree}, \text{pathLength}) \) \{Creates all the directed subgraphs distinct with out-degree greater than \( \text{outDegree} \) and path length equal \( \text{pathLength} \). Among the subgraphs that contain the same vertices, only remains the subgraph with greater out-degree and greater confidence index.\}
13: \textbf{for all} \( h \in H \) \textbf{do}
14: \hspace{1em} \( c_{od} \leftarrow \text{getVertexGreaterOutDegree}(h) \) \{Returns the candidate with greater out-degree on \( h \)\}
15: \hspace{1em} \( c_{ci} \leftarrow \text{getVertexGreaterConfidenceIndex}(h) \) \{Returns the candidate with greater confidence index on \( h \)\}
16: \hspace{1em} \( C_{out} \leftarrow \text{setNodeCandidate}(C_{out}, c_{od}) \) \{Inserts \( c_{od} \) in \( C_{out} \)\}
17: \hspace{1em} \textbf{if} \( c_{od} \neq c_{ci} \) \textbf{then}
18: \hspace{2em} \( C_{out} \leftarrow \text{setNodeCandidate}(C_{out}, c_{ci}) \) \{Inserts \( c_{ci} \) in \( C_{out} \)\}
19: \hspace{1em} \textbf{end if}
20: \hspace{1em} \textbf{end for}
21: \( C_{out} \leftarrow \text{getComplement}(C_{out}, C_{in}, V) \) \{Inserts candidates who have no relationship with other candidates, i.e., the relative complement of sets \( C_{in} \) and \( V \), denoted by \( C_{in} - V \).\}
22: \textbf{return} \( C_{out} \)
\end{algorithm}

In the example in Fig. 5d, there are three nodule candidates and after the execution of the algorithm, just remained \( c_3 \) and \( c_4 \).

\textbf{D. FALSE POSITIVE REDUCTION}

In order to face the nodule candidates classification problem, we used multiple 3D DCNN classifiers, specifically, the ensemble method with the Bootstrap Aggregation (Bagging) [43] technique. Therefore, we trained classifiers with samples of different sizes and attenuation range.

\textbf{1) DATA ACQUISITION}

In the FPR step, the subset for training was split, 80\% were used for training and 20\% for validation. We highlight that the validation set was only used to determine the best models with the deep learning techniques.

Due to significant variations of nodule sizes, types, texture, and geometry characteristics, we defined four attenuation range with the CT numbers \( G = \{\{\gamma = -1000 \text{ and } \lambda = -200\}, \{\gamma = -1000 \text{ and } \lambda = 0\}, \{\gamma = -1000 \text{ and } \lambda = 200\}, \{\gamma = -1000 \text{ and } \lambda = 400\}\} \) expressed in HU and rescaled them to \( \zeta = 0 \) and \( \delta = 1.0 \) (Section III-B). We performed for each scan to anisotropic resampling. Therefore, we rescaled the Z-direction to 1.0 mm and kept the original spacing in the X and Y directions. Thus, we normalized the size of scans.

We extracted 3D patches with two different sizes for each candidate, the first one with \( 32 \times 32 \times 32 \) pixels and another with \( 48 \times 48 \times 48 \) pixels, i.e., \( L = \{32^3, 48^3\} \), and with one channel centered around the candidate location. We used them as inputs to the networks.

\textbf{2) 3D DCNN ARCHITECTURE}

We proposed a 3D DCNN architecture based on the VGG [21], which provided better image classification performance and is commonly used for feature extraction [14]. The architecture was defined empirically with the results obtained in the validation set. The problem was modeled with two classes \( W = \{\text{nodule, irrelevant finding}\} \). Each model generated by 3D DCNN architecture is a classifier, denoted by \( K(a) = \{P(w \mid a) \mid \forall w \in W\} \), where the classifier \( K \) determines confidence index \( P(w \mid a) \) for input sample \( a \) to each class \( w \) of \( W \). Fig. 6 shows the 3D DCNN architecture proposed to classify nodule candidates.

The architecture supports 3D patches with different sizes as inputs due global max-pooling layer, which was used to maintain only the most essential features. The network consisting of six convolution blocks, each block has the 3D
Specifically, we used the best checkpoint based on the checkpoint ensemble method [45], which employs ranges \(G\) and size \(L\) of 3D patches \(m\). We employed a classifier ensemble proposed to combine \(3\) \(D\) DCNN by employing a Softmax activation function. Besides, the maximum number of channels is 128. In total, the network consists of 12 layers.

3) TRAINING PROCESS
The number of false positive candidates obtained of the NCD step was many times bigger than the number of true nodules of the LUNA16 dataset, which means the training sets were highly imbalanced. Therefore, we used data augmentation methods to balance the set. Thus, we address the data skewness problem and prevent overfitting by adding invariances to the set.

We construct an online scheme to apply randomly affine transformations for all training samples on-the-fly. In each training epoch, the center of each 3D patch was slightly shifted to a random position up to \(\pm 3\) pixels along each direction (i.e., X, Y, and Z). Moreover, the 3D patch was rotated at angles 0°, 90°, 180°, and 270° randomly using three possible directions. Therefore, in each epoch, the composition of the training set was different. This procedure was applied to all 10 subsets.

At the training stage, each network was optimized by the SGD algorithm with back-propagation and cross-entropy as the loss function. The initial learning rate was set to \(10^{-3}\), the momentum parameter was set to 0.9, the weight decay was set to \(5 \times 10^{-4}\), and the mini-batch size was set to 16. We regularized our models by the early stopping criterion the operative point 1 FP/Scan and CPM metric. Experiments were realized using two NVIDIA Titan Xp graphics card with 12 GB.

In order to get the final classification result of the testing stage, we combined the confidence index obtained by the NCD and FPR steps through average, hence enhance discrimination capability. Equation (3) presents the calculation formula.

\[
P = \frac{1}{2}(P_{NCD} + P_{FPR})
\]

4) TESTING PROCESS
The classifier ensemble proposed to combine \(3\) \(D\) DCNN models trained with different samples regarding attenuation ranges \(G\) and size \(L\) of 3D patches \(m\). Besides, we employ the checkpoint ensemble method [45], which employs checkpoints stored in the middle of the training stage. Specifically, we used the best checkpoint based on the validation set \(m_{cp}\) and two more checkpoints as predecessors \(M = \{m_{cp}, m_{cp-1}, m_{cp-2}\}\). Also, each classifier performed several classifications about the same candidate with different angle rotations, i.e., for each candidate, four augmented copies were generated and rotated at angles \(O = \{0°, 90°x, 180°y, 270°z\}\). Therefore, the final prediction of the classifier ensemble comprises the fused of the output of several \(3\) \(D\) DCNN to nodule class through average. The calculation formula is presented in Equation (2).

\[
P_{FPR} = \frac{1}{|L||G||M||O|} \sum_{l=1}^{L} \sum_{g=1}^{G} \sum_{m=1}^{M} \sum_{o=1}^{O} \text{class} \ K_{\text{ign}}(a_o)
\]
TABLE 1. Performance comparison of substep of NCD step.

| Metric                               | Substep 1 | Substep 2 | Substep 3 |
|--------------------------------------|-----------|-----------|-----------|
| True positives                       | 1.169     | 1.169     | 1.164     |
| False positives                      | 44.613    | 40.005    | 37.010    |
| False negatives                      | 17        | 17        | 22        |
| Total number of candidates           | 52.805    | 48.230    | 44.111    |
| Double detections on a nodule       | 205       | 205       | 22        |
| Sensitivity (%)                      | 98.57     | 98.57     | 98.15     |
| Avg. no. of candidates per scan     | 59.47     | 54.31     | 49.67     |

TABLE 2. Sensitivities at seven operative points of FROC curve.

| FFs/Scan | 1/8 | 1/4 | 1/2 | 1   | 2   | 4   | 8   | CPM |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|
| Sensitivity (%) | 81.19 | 87.49 | 91.84 | 94.90 | 95.74 | 96.98 | 97.57 | 92.24 |

nodule candidate detections that have centroid close to each other in an exam, the system achieved a sensitivity of 98.15% and an average of 49.67 FFs/Scan, and thus it reduced the number of candidates by 8.54% and losing only 5 TPs (Table 1 - Sub-step 3). As mentioned above, this step reached the goal to identify and localize pulmonary nodule candidates with reasonable rates between false positives and sensitivity.

We also evaluated the performance of the CADE system proposed at seven operative points of the FROC curve and calculated the CPM score. The results are shown in Table 2.

Table 3 shows the performance comparison CADE system proposed and that of other published CADE systems. The works of Setio et al. [11] and Zhang et al. [22] obtained better sensitivities on nodule candidate detections, 98.30%, and 100.00%, respectively. However, in order to localize as many candidates as possible, these CADE systems obtained more FPs. The CADE system proposed by Zhang et al. [22] achieved the same result as we did, a sensitivity of 94.90% and an average of 1.0 FP/Scan. However, the CADE system proposed by Setio et al. [11] achieved a sensitivity of 96.90% and an average of 1.0 FP/Scan. Although, they combined the results from seven pulmonary nodule candidate detection systems and five false positive reduction systems.

Newer studies have been proposed with relevant results. For example, Zhu et al. [28] treated the quality of the images, and Peng et al. [34] and Mei et al. [35] proposed new DCNN architectures. However, they do not surpass all our results in the different steps (NCD and FPR) with the evaluation metrics average sensitivity at FFs/Scan or CPM.

Our method achieved a CMP score of 92.24, a result comparable with current researches. We highlight that the CADE system proposed to have achieved a sensitivity of 94.90% and an average of 1.0 FP/Scan, outperforming other state-of-the-art methods. Fig. 7 shows the detection results, where we only visualize a slice where the detection center is located. We crop the square area with the detection center with the size of 48 pixels. Fig. 7a shows the detected true positive nodules (cyan rectangles), and Fig. 7b shows the detected false positive nodules (yellow rectangles), and Fig. 7c shows the false negative nodules (magenta circles).

Besides, our proposal has a low computational cost. In the NCD step, we used 2D DCNN, and in the FPR step, we used 3D DCNN to classify small patches. Despite using a classifier ensemble, the training process is performed offline.

Most CADE systems have their operative point set somewhere between 1 to 4 FFs/Scan [23], [32]. Therefore, our proposal is in line with current usage practices and it has the potential to be used in clinical decision support.

A. ABLATION STUDY

To assess the gain in performance caused by the classifier ensemble, we compared the individual performance of each classifier. Fig. 8 presents the FROC curve of the CADE system proposed, also exposes the individual performance of each classifier. For each classifier, we used only a sample $O = \{0^\circ\}$, and the best checkpoint $M = \{m_{cp}\}$. Thus, each classifier exclusively uses a sample size $L$ and an attenuation pattern $G$.

We evaluate each classifier ensemble based on an attenuation pattern to demonstrate the performance individually and...
TABLE 3. Performance comparison of published CADe systems.

| CADe systems     | Year | Datasets | Number of scans | Number of folds | Number of nodules | Sensitivity (%) | NCD step | FPs/scan | Sensitivity (%) | FPR step | CPM     |
|------------------|------|----------|-----------------|-----------------|-------------------|----------------|----------|----------|----------------|----------|---------|
| Zhu et al. [19]  | 2018 | LUNA16   | 888             | 10              | -                 | 95.80          | -        | -        | -              | -        | 94.20   |
| Zhu et al. [28]  | 2021 | NSCLC (146) | LUNA16 (888) | -              | 1304             | -              | -        | -        | -              | -        | 91.50   |
| Peng et al. [34] | 2021 | LUNA16   | 888             | 10              | 1186              | -              | -        | -        | -              | -        | 92.30   |
| Tang et al. [27] | 2020 | TIANCHI17 (800) | LUNA16 (888) | 1688           | -                 | -              | -        | 92.40    | 4.0             | -        | -       |
| Seto et al. [13] | 2016 | LUNA16   | 888             | 5               | 1186              | 93.25          | 239041   | 269.20   | 85.40          | 1.0      | 82.80   |
| Don et al. [22]  | 2017 | LUNA16   | 888             | 10              | 1186              | -              | -        | 86.50    | 1.0             | 83.90    |         |
| Cai et al. [33]  | 2020 | TIANCHI17 (800) | LUNA16 (888) | 888           | 10               | 1186              | -        | 88.10    | 1.0             | 79.60    |         |
| Mei et al. [35]  | 2021 | LUNA16   | 888             | 10              | -                 | 94.60          | -        | 90.09    | 1.0             | 79.39    |         |
| Ding et al. [32] | 2017 | LUNA16   | 888             | 10              | -                 | 94.60          | -        | 92.20    | 1.0             | 89.10    |         |
| Zheng et al. [14]| 2020 | LUNA16   | 888             | 10              | 1186              | 95.36          | 18116    | 20.40    | 92.67          | 1.0      |         |
| Ozdemir et al. [25]| 2020 | LUNA16 | 888             | 10              | 1186              | 96.50          | -        | 94.20    | 1.0             | 92.10    |         |
| Zhang et al. [22]| 2018 | LUNA16   | 888             | 10              | 1186              | 100.00        | 45939    | 51.73    | 94.90          | 1.0      | 94.70   |
| Seto et al. [11] | 2017 | LUNA16   | 888             | 10              | 1186              | 98.30          | 754975   | 850.20   | 96.90          | 1.0      | 95.20   |
| Our proposed     | -    | LUNA16   | 888             | 10              | 1186              | 98.15          | 44111    | 49.67    | 94.90          | 1.0      | 92.24   |

TABLE 4. Performance comparison of the classifier ensemble based on attenuation patterns.

| Sensitivity (%) | FPs/Scan | 1/8 | 1/4 | 1/2 | 1   | 2   | 4   | 8   | CPM |
|-----------------|----------|-----|-----|-----|-----|-----|-----|-----|-----|
| λ = −200       | 77.69    | 84.12 | 89.67 | 93.76 | 95.31 | 96.46 | 97.38 | 90.63 |     |
| λ = 0          | 80.64    | 85.40 | 90.95 | 93.64 | 94.70 | 95.97 | 97.23 | 91.22 |     |
| λ = 200        | 79.71    | 85.56 | 90.24 | 93.51 | 95.38 | 96.87 | 97.40 | 91.24 |     |
| λ = 400        | 80.94    | 86.37 | 91.01 | 93.51 | 95.44 | 96.60 | 97.38 | 91.61 |     |
| All            | 81.36    | 87.00 | 91.50 | 94.40 | 95.57 | 96.60 | 97.47 | 91.99 |     |

TABLE 5. Impact of the checkpoint ensemble method and the strategy to classify candidates with different angle rotations in the proposed methodology.

| O   | M       | mεp | mεp, mεp−1 | mεp, mεp−1, mεp−2 |
|-----|---------|-----|------------|-------------------|
| 0°  | 94.40   | 94.67 | 94.66      |
| 0°, 90°x | 94.36 | 94.75 | 94.70      |
| 0°, 90°x, 180°y | 94.30 | 94.71 | 94.68      |
| 0°, 90°x, 180°y, 270°z | 94.51 | 94.85 | 94.90      |

the boost performance when combining different patterns. In the experiments, we used \( L = \{32^3, 48^3\} \) and combine the different parameters of \( G = \{\gamma = -1000, \lambda = 0\} \). Table 4 shows quantitative results at seven operative points of the FROC curve and calculated the CPM score.

We also measure the impact of the checkpoint ensemble method against the strategy to classify the same candidate with different angle rotations. Table 5 shows quantitative results given by sensitivity and average of 1.0 FP/Scan. For pulmonary nodule candidate detection, we used an end-to-end object detection method. Next, we used an algorithm to group 2D bounding boxes on 3D nodules. Then, we designed the algorithm to discard nodule candidate detections that have centroid close to each other in an exam. Also, we designed a classifier ensemble for false positive reduction, which fuses the individual classification through average to improve nodules classification sensitivity. Our CADe system achieved a sensitivity of 94.90% and an average of 1.0 FP/Scan on the publicly available LUNA16 dataset. We believe that the CADe system proposed is in line with the current practices and it can be a powerful computational tool for clinical decision support on pulmonary nodule detections.

V. CONCLUSION
This paper proposes a CADe system that uses a classifier ensemble based on CT attenuation patterns to detect pulmonary nodules in low-dose 3D CT scans automatically. For pulmonary nodule candidate detection, we used an end-to-end object detection method. Next, we used an algorithm to group 2D bounding boxes on 3D nodules. Then, we designed the algorithm to discard nodule candidate detections that have centroid close to each other in an exam. Also, we designed a classifier ensemble for false positive reduction, which fuses the individual classification through average to improve nodules classification sensitivity. Our CADe system achieved a sensitivity of 94.90% and an average of 1.0 FP/Scan on the publicly available LUNA16 dataset. We believe that the CADe system proposed is in line with the current practices and it can be a powerful computational tool for clinical decision support on pulmonary nodule detections.

ACKNOWLEDGMENT
The authors would like to thank NVIDIA Corporation for the donation of the two Titan XP graphics cards used in their experiments.

REFERENCES
[1] World Health Organization. (2018). Cancer. Accessed: Nov. 20, 2020. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer
[2] Instituto Nacional de Câncer. (2020). Câncer de Pulmão. Accessed: Dec. 28, 2020. [Online]. Available: https://www.inca.gov.br/tipos-de-cancer/cancer-de-pulmao
[3] American Cancer Society. (2021). Lung Cancer. Accessed: Mar. 1, 2021. [Online]. Available: https://www.cancer.org/cancer/lung-cancer/about/key-statistics.html
[4] A. H. Krist, K. W. Davidson, C. M. Mangione, M. J. Barry, M. Cabana, A. B. Caughey, E. M. Davis, K. E. Donahue, C. A. Doubeni, M. Kubik, and C. S. Landefeld, “Screening for lung cancer: US preventive services task force recommendation statement,” J. Amer. Med. Assoc., vol. 325, no. 10, pp. 962–970, Mar. 2021, doi: 10.1001/jama.2021.1117.
K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” in W. Zhu, C. Liu, W. Fan, and X. Xie, “DeepLung: Deep 3D dual path nets for pulmonary nodule detection,” New England J. Med., vol. 348, no. 25, pp. 2535–2542, Jun. 2003, doi: 10.1056/NEJMoa031329-0.

S. G. Armato III, G. McLennan, L. Bidaut, M. F. McNitt-Gray, and D. R. Aberle, “Toward clinically usable CAD for lung cancer screening with computed tomography,” Eur. Radiol., vol. 24, no. 2, pp. 2719–2728, Jul. 2014, doi: 10.1007/s00330-014-3329-0.

A. A. A. Setio, A. Traverso, T. De Bel, M. S. N. Berens, C. Van Den Bogaard, P. Cerello, H. Chen, Q. Dou, M. E. Fantacci, B. Geurts, and R. van der Guten, “Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge,” Med. Image Anal., vol. 42, pp. 1–13, Dec. 2017, doi: 10.1016/j.media.2017.06.015.

Z. Wu, R. Ge, G. Shi, L. Zhang, Y. Chen, L. Luo, Y. Cao, and H. Yu, “MD-NDNet: A multi-dimensional convolutional neural network for false-positive reduction in pulmonary nodule detection,” Phys. Med. Biol., vol. 65, no. 23, Dec. 2020, Art. no. 235053, doi: 10.1088/1361-6560/abaa7c.

B. L. Torres, E. Fieschi, F. Pennazio, C. Peroni, M. A. L. M. W. Wille, M. Naqibullah, I. C. Sánchez, and B. Van Ginneken, “Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks,” IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 1160–1169, May 2016, doi: 10.1109/TMI.2015.2536809.

S. Zheng, J. Guo, X. Cui, and R. N. J. Veldhuis, “Automatic pulmonary nodule detection in CT scans using convolutional neural networks based on maximum intensity projection,” IEEE Trans. Med. Imag., vol. 39, no. 3, pp. 797–805, Mar. 2020, doi: 10.1109/TMI.2019.2935553.

Q. Dou, H. Chen, Y. Jin, H. Lin, J. Qin, and P.-A. Heng, “Automated pulmonary nodule detection via 3D convnets with online sample filtering and hybrid-loss residual learning,” in Medical Image Computing and Computer Assisted Intervention—MICCAI 2017, M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. L. Collins, and S. Duchesne, Eds. Cham, Switzerland: Springer, 2017, vol. 10379, pp. 787–795, doi: 10.1007/978-3-319-71167-9_72.

O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in Medical Image Computing and Computer Assisted Intervention—MICCAI 2015, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham, Switzerland: Springer, 2015, vol. 9351, pp. 234–241, doi: 10.1007/978-3-319-24574-4_27.

F. Milletari, N. Navab, and S.-A. Ahmadi, “V-Net: Fully convolutional neural networks for volumetric medical image segmentation,” in Proc. 4th Int. Conf. 3D Vis. (3DV), Oct. 2016, pp. 565–571, doi: 10.1109/3DV.2016.79.

T. J. Ding, A. Li, Z. Hu, and L. Wang, “Accurate pulmonary nodule detection in CT scans using convolutional neural networks based on maximum intensity projection,” IEEE Trans. Med. Imag., vol. 39, no. 3, pp. 797–805, Mar. 2020, doi: 10.1109/TMI.2019.2947595.

S. Wang, H. Sun, and Y. Guo, “3D multi-scale deep convolutional neural networks for realistic CT image super-resolution for early-stage pulmonary nodule detection,” Future Gener. Comput. Syst., vol. 115, pp. 475–485, Feb. 2021, doi: 10.1016/j.future.2020.09.020.

R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Columbus, OH, USA, Jun. 2014, pp. 580–587, doi: 10.1109/CVPR.2014.81.

S. Chen, H. Ke, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” in Proc. 28th Int. Conf. Neural Inf. Process. Syst., Montreal, QC, Canada, vol. 1, 2015, pp. 91–99.

J. Ding, A. Li, Z. Hu, and L. Wang, “Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks,” in Medical Image Computing and Computer Assisted Intervention—MICCAI 2017, M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. L. Collins, and S. Duchesne, Eds. Cham, Switzerland: Springer, 2017, pp. 559–567, doi: 10.1007/978-3-319-66179-7_64.

R. Girshick, C. Farabet, P. Dollár, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Columbus, OH, USA, Jun. 2014, pp. 580–587, doi: 10.1109/CVPR.2014.81.

R. Girshick, “Fast R-CNN,” in Proc. IEEE Int. Conf. Comput. Vis., Santiago, Chile, Dec. 2015, pp. 1440–1448, doi: 10.1109/ICCV.2015.169.

S. Auyeung, K. H. Ke, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” in Proc. 28th Int. Conf. Neural Inf. Process. Syst., Montreal, QC, Canada, vol. 1, 2015, pp. 91–99.

L. Cai, T. Long, Y. Dai, and Y. Huang, “Mask R-CNN-based detection and segmentation for pulmonary nodule 3D visualization diagnosis,” IEEE Access, vol. 8, pp. 44400–44409, Feb. 2020, doi: 10.1109/ACCESS.2020.2976432.

H. Peng, H. Sun, and Y. Guo, “3D multi-scale deep convolutional neural networks for pulmonary nodule detection,” PLoS ONE, vol. 16, no. 1, pp. 1–14, Jan. 2021, doi: 10.1371/journal.pone.0244406.

J. Mei, M. M. Cheng, G. Xu, L. R. Wan, and H. Zhang, “SANet: A slice-aware network for pulmonary nodule detection,” IEEE Trans. Pattern Anal. Mach. Intell., early access, Mar. 9, 2021, doi: 10.1109/TPAMI.2021.3065086.

G. M. Uo, B. Bagci, L. Lu, A. Wu, M. Buty, H.-C. Shin, H. Roth, G. Z. Papadakis, A. Depeursinge, R. M. Summers, and Z. Xu, “Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks,” Comput. Methods Biomech. Biomed. Eng. Imag. Visualizat., vol. 6, no. 1, pp. 1–6, Jun. 2016, doi: 10.1111/cmb.12181.
F. R. Pereira et al.: Classifier Ensemble Based on CT Attenuation Patterns

[41] Z. Shi, H. Hao, M. Zhao, Y. Feng, L. He, Y. Wang, and K. Suzuki, “A deep CNN based transfer learning method for false positive reduction,” Multimedia Tools Appl., vol. 78, no. 1, pp. 1017–1033, Jan. 2019, doi: 10.1007/s11042-018-6082-6.

[42] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, and A. C. Berg, “ImageNet large scale visual recognition challenge,” Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Apr. 2015, doi: 10.1007/s11263-015-0816-y.

[43] L. Breiman, “Bagging predictors,” Kluwer Academic Publishers, vol. 24, no. 2, pp. 123–140, Aug. 1996, doi: 10.1023/A:1018054314350.

[44] V. Nair and G. E. Hinton, “Rectified linear units improve restricted Boltzmann machines,” in Proc. 27th Int. Conf. Int. Conf. Mach. Learn., Haifa, Israel, 2010, pp. 807–814.

[45] H. Jung, B. Kim, I. Lee, J. Lee, and J. Kang, “Classification of lung nodules in CT scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method,” BMC Med. Imag., vol. 18, no. 1, pp. 1471–2342, Dec. 2018, doi:10.1186/s12880-018-0286-0.

FERNANDO ROBERTO PEREIRA received the B.Sc. degree in information systems from the Con- testado University (UnC), in 2007, and the M.Sc. degree in science from the Federal University of Technology–Paraná (UTFPR), in 2011. He is currently pursuing the Ph.D. degree in computer science with the Federal University of Paraná (UFPR). He is also a Computer Professor with the Federal Institute of Education, Science and Technology of Santa Catarina (IFSC), Brazil. His research interests include computer vision, pattern recognition, and machine learning. He is also focusing on convolutional neural network applied to medical image.

JOÃO MARIO CLEMENTIN DE ANDRADE graduated in medicine from the Federal University of Paraná (UFPR), in 2017. He received the master’s degree in internal medicine and health sciences from UFPR. From 2018 to 2020, he was a Resident in radiology with the Hospital De Clínicas, UFPR. He is currently a Radiologist with UFPR. He is also a fellow in radiology with the Clínica Diagnóstico Avançado por Imagem (DAPI). His research interest includes computer-aided diagnosis in medical imaging.

DANTE LUIZ ESCUISSATO received the B.Sc. degree in medicine and the M.Sc. degree in internal medicine from the Federal University of Paraná, Brazil, in 1983 and 1993, respectively, and the Ph.D. degree in medicine (concentration in radiology) from the Federal University of Rio de Janeiro, in 2004. He performed the Medical Residency in radiology with the Hospital de Clínicas, Federal University of Paraná, from 1984 to 1986. He held Fellowship in CT and MRI at the Hospital da Beneficência Portuguesa de São Paulo, in 1987. He is currently an Associate Professor with the Department of Internal Medicine, Federal University of Paraná. His research interests include image evaluation in thoracic diseases and in applications of quantitative image and artificial intelligence in lung diseases.

LUCAS FERRARI DE OLIVEIRA received the B.Sc. degree in computer science from the Marília University (UNIMAR), Marília, in 1997, the M.Sc. degree in electrical engineering from São Paulo University (USP), São Carlos, in 2000, and the Ph.D. degree in clinical science from São Paulo University (USP), Ribeirão Preto, in 2005. He is currently a Professor with the Department of Informatics, Federal University of Paraná (UFPR), Brazil. His research interests include medical imaging processing, computer-aided diagnosis/detection, pattern recognition, and machine learning.