Application of Intelligent Computational Models on Computed Tomography Lung Images

H. S. Pheng, S. M. Shamsuddin\textsuperscript{1}, and S. Kenji\textsuperscript{2}

\textit{Soft Computing Research Group}\textsuperscript{1}
Universiti Teknologi Malaysia, 81310 Skudai, Johor.
e-mail: seepheng@gmail.com; mariyam@utm.my

\textit{Department of Radiology}\textsuperscript{2}
The University of Chicago, 5841 South Maryland Avenue, Chicago, Illinois 60637.
e-mail: suzuki@uchicago.edu

\textbf{Abstract}

With computed tomography (CT) scanners, hundreds of slices are generated to visualize the condition of lung per patient. The analysis on slices-by-slices dataset is time-consuming for radiologists. Therefore, automated identification of abnormalities on CT lung images is vital to assist the radiologists to make an interpretation and decision. In this paper, we review the performance of various conventional and computational intelligence algorithms in the segmentation, detection and quantification of lung nodules on CT lung images. The accuracy of lung region segmentation is found important as a preprocessing step to identify the lung nodules. By mean of these computerized systems, the detection and measurement of lung nodules can assist the radiologists to determine whether the lung nodules are benign or malignant.

\textbf{Keywords}: CT scan, lung nodules, computational intelligence, CAD.
1 Introduction

Computed tomography (CT) is one the major noninvasive diagnosis techniques to assess different types of lung diseases namely lung cancer, emphysema, pulmonary embolisms, airway diseases[1]. With CT scan, a series of cross-sections is obtained and analyzed by the radiologist to detect the abnormality of lung. For instances, thin-section CT scan can generate 1mm thickness sections with about 250 to 350 lung images. Therefore, the CT images interpretation by radiologists may require significant time and effort as hundreds of multi-sections per patient can be obtained from a CT scan. The development of computer aided diagnosis (CAD) is increased steadily to provide automated detection and classification of anomaly structure in human body and organ. The automated CAD scheme is normally to use to give second opinion to the radiologists in the identification of abnormalities and nodules’ growth rate on patients’ medical images [2]. Performance of radiologist in lung nodules detection have been reported improve specially in the interpretation of large-scale dataset [3].

In recent years, several researches had been conducted with the aim to improve the performance of CAD in the clinical trial. To provide reliable and accurate detection of nodules in CT images, lung segmentation is known as a necessary pre-processing step in different types of CAD [4]. In the consequent analysis, the process of segmentation is extremely important to confirm that the abnormalities can be detected especially for the lung nodules which exist at edge or border of the lung. Besides, small nodules are normally missed in the detection by naked eyes. Thus, the precise algorithm is required to identify those small nodules as they can be curable lung cancer. The early detection of small lung nodules can improve the prognosis significantly [5]. In this paper, we are focusing on the review of existing intelligence computational algorithms for segmentation of CT lung images and quantification lung nodules. This paper will also introduce the publicly datasets released by Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI). In some previous researches, many proposed methods in CAD showed the good performance. However, the dataset used in these researchers are difference and some do not contain high density of abnormalities[6]. The generalization of algorithms is important to be tested by using the publicly dataset. The paper is organized as follows. In Section 2, overview of CT scans for lung and roles of CAD are discussed. Section 3 is a recent review of lung CT images by CAD. Section 4 focuses on the intelligence techniques used in CAD. Section 5 discusses the current challenges for the research on CAD and Section 6 is a conclusion of this paper.
2 Computed Tomography Scans and Role of CAD

Among several of modalities, computed tomography (CT) has been known as the most common imaging for early detection for pulmonary nodules due to its sensitivity and ability to reconstruct a three-dimensional structure of thorax [5, 7]. Basically, there are two types of CT scans: thin and thick CT scan. Thin slide images show better performance at high resolution where the image is normally captured at 1 to 2 mm interval [8]. For a conventional CT scan, the thickness of thick slides scan is normally 10mm, with the scans obtained at the interval of 10mm [9]. The CT sensitivity can be increased with a reduction of the slice thickness due to a decrease of partial-volume effect, and overlapping images reconstruction [5]. Chest CT scans are able to detect early lung cancer where the change of nodules sizes through the sequential of follow up diagnostic[10].

![Diagram](image)

Fig 1: The roles of CAD in the different analysis of lung CT scans.
The efforts for CAD had been begun since middle of 1980 by physicists and radiologist to obtain the automated computer interpretation on medical imaging [11]. Fig. 1 shows the overview of CAD roles in a common sequence of analysis on lung CT scans. CAD is crucial to perform several diagnoses such as segmentation, registration detection, classification and quantification. The segmentation of organ from CT images plays a role to reduce the redundant images and hence simplified the analysis of region of interest (ROI). In the classification of images, it is important to detect the abnormalities such as cancer or certain diseases on the organ. The development of pattern recognition algorithms in CAD is categorized into two: a) methods based on decision theory, b) computational intelligence methods. The details of computational intelligence developed models will be included in the next sections. The CAD for lung CT images usually performs several steps as shown in Fig. 2. The automated interpretation is used as second opinion by medical radiologists in decision making.

Fig. 2: The process of CAD system for lung nodules detection.
3 Recent Review on Lung CT Images Analysis by CAD

There are recent reviews had been done on the development of CAD for the analysis of lung medical images [3, 12-14]. Lee et al.[13] reviewed the recent existing methods on automated lung nodule detection, and compare the performance on each method. Based on their survey, the authors had introduced a generic structure that consists of data acquisition, pre-processing, lung segmentation, nodules detection and false positives reduction. They claimed the high sensitivity of approaches in CAD was occasionally due to the small datasets used in the evaluation. Besides, efforts on false positives reduction may significantly improve the performance of existing methods. However, their survey emphasized on the performance of method rather than depiction for each automated methods. Bagci et al. [14] reviewed CAD system used on different types of infectious lung diseases. They focused mainly on the features used in various types of methods in CAD. The detection of lung abnormalities on both chest radiography (X-rays) and CT images was discussed. Their study was done not only on detection of lung nodules but also on many other types of diseases related to lung. Girvin and Ko [3] introduced the detection and characterization of lung nodules on chest radiography (X-rays), CT and Magnetic Resonance Imaging (MRI). The management and types of pulmonary nodules had been presented in this paper. They had also reviewed different type of approaches on advanced CAD that use to interpret the lung nodules. These authors concluded that lung nodules grouping are important and the weakness of current CAD is necessary to be assessed. Sluimer et al [12] presented a comprehensive survey on the computer analysis of CT scans of the lung. The authors had discussed the role of CAD in the analysis of different types of lung diseases which included segmentation, characterization, registration, detection and quantification of lung diseases and nodules. The performance of various techniques had been compared by using sensitivity and false positive (FP) rate in this survey paper from year 1999 to 2004. It was found that the datasets that used by those techniques were thoroughly different.

Armato III et al. [15] have described the now-completed publicly available database of lung nodules on CT scans from Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI). The creation of this database was to encourage the CAD development and validation. Furthermore, Early Lung Cancer Action Program (ELCAP)[16] public lung image database and Medical Image Database[17] are also considered as popular public lung nodules database that normally used by researchers in CAD evaluation. In the current and future CAD validation, it is expected that those database will be widely tested to confirm the generalization and performance of algorithms. Table 1 shows the reviews of several techniques used in CAD in the assessment and analysis the lung CT scans, especially for lung nodules.
### Table 1: Summary of reviews on CAD and lung CT scans analysis

<table>
<thead>
<tr>
<th>Authors</th>
<th>Title</th>
<th>Description</th>
<th>Comment</th>
</tr>
</thead>
</table>
| Messay et al. [7] | A new computational efficient CAD system for pulmonary nodule detection in CT imagery | (1) A novel CAD system was developed and assessment its performance by using publicly available data.  
(2) Focus on 3D lung segmentation, local contrast enhancement, down-sampling, orienting and detection.  
(3) 245 features were used in the classifier of CAD system. | This is a relatively latest research which focusing on 3D imaging analysis.                  |
| Wang et al. [18] | Multilevel binomial logistic prediction model for malignant pulmonary nodules based on texture features of CT image | (1) Multilevel binomial logistic prediction model-based computer-aided diagnostic (CAD) was introduced.  
(2) Five texture features was tested statistically: Inertia, Entropy, correlation, difference-mean, sum-entropy.  
(3) Test on how various observed characteristics of patient influence the probabilities of benign and malignant pulmonary nodules. | Authors use the statistical method (multiple regression models) to build a prediction model estimate benign and malignancy of nodules. They focus on the patients and texture of CT images. |
| Lei et al. [19] | Automated lung segmentation algorithm for CAD system of thoracic CT | (1) Updated the traditional optimal thresholding algorithm and use mathematical morphologic to acquire the rough image of lung segmentation.  
(2) 30 patient’s chest CT is used for evaluation.  
(3) The successful rate is 94.8%. | The performance of this approach shows the promising results. However, the datasets used are only based on 30 patients. |
<table>
<thead>
<tr>
<th>Authors</th>
<th>Title</th>
<th>Description</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iwano et al. [20]</td>
<td>Computer-aided differentiation of malignant from benign solitary pulmonary nodules imaged by high-resolution CT</td>
<td>(1) The discriminate analysis was used for two thresholds in differentiating malignant from benign. (2) 107 HRCT images were used, 48 female and 58 males. (3) Two quantitative parameters involved: circularity and second central moment. (4) Sensitivity = 70.9%</td>
<td>The computer analysis was not able to provide a complete diagnosis; final diagnosis must be done by radiologists.</td>
</tr>
<tr>
<td>Delogu et al. [5]</td>
<td>Preprocessing methods for nodule detection in lung CT</td>
<td>(1) Applied combination of techniques such as threshold-based, morphological operators, border detection, border thinning, border reconstruction, region filling, segmentation and 3D filter. (2) Algorithm tested on a dataset of standard and low-dose CT scans where system was applied to the high and low dose with thickness 5mm and 1mm. (3) Algorithm has to be robust to allow the detection of nodules close to pleural.</td>
<td>The results had shown the high FP rate. It can be improved by using wavelet transforms as suggested by authors.</td>
</tr>
<tr>
<td>Suzuki et al. [21]</td>
<td>Computer Aided Diagnosis Scheme for distinction between benign and malignant nodules in thoracic low-dose CT by use of Massive Training Artificial Neural Network</td>
<td>(1) To differentiate between benign and malignant nodules in low dose CT scan. (2) Train and test data on 3 categories of cancer: pure ground-glass opacity (GGO), Mix GGO and solid nodule. (3) Used 2D Gaussian function and MTANN.</td>
<td>The computational model provided satisfied results massive training. There was no segmentation involved.</td>
</tr>
<tr>
<td>Authors</td>
<td>Title</td>
<td>Description</td>
<td>Comment</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Suzuki and Doi [22] | How Can A Massive Training Artificial Neural network (MTANN) be trained with a small number of cases in the distinction between nodules and vessels in thoracic CT? | (1) Trained MTANN with very small numbers (1 to 9) of sub-regions. 
(2) MTANN was able to learn, form a very small number of actual nodule and non-nodule cases, the distinction between nodules and vessels. | The training and learning of the model provided good results through small number of training cases involved. |
| Li et al. [23] | Computer-aided detection of peripheral lung cancers missed at CT: ROC analyses without and with localization | (1) Evaluate whether a difference -image CAD scheme to detect peripheral lung cancers missed at low dose CT. 
(2) Two groups of radiologists, with and without CAD to indicate their confidence in detecting cancer. 
(3) The CAD scheme applied matched filter, multiple-gray-level threshold techniques, and rule-based schemes. 
(4) MTANN used to reduce FPs rates. 
(5) Results were tested by using Receiver Operating Characteristics (ROC) curve. | This paper had proven that CAD play an important role in assisting radiologists especially to detect missed cancers at low – dosed CT scans. |
| Lin et al.[24] | Autonomous detection of pulmonary nodules on CT images with a neural network-based fuzzy system | (1) Automatically identify a set of appropriate fuzzy inference rules, and refine the membership functions through the steepest gradient descent learning algorithm. 
(2) 29 cases were tested and ROC is used for evaluation. | The new fully automated system had provided fast processing and detection with high sensitivity and low FP rate. |
<table>
<thead>
<tr>
<th>Authors</th>
<th>Title</th>
<th>Description</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzuki et al. [25]</td>
<td>Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low dose computer tomography</td>
<td>(1) Used pattern recognition technique based on an artificial neural network. (2) Used plural and multiple MTANN and arranged in parallel. Each MTANN acts as an expert for a specific type of non nodules. (3) MTANN was trained to reduce false positives.</td>
<td>The classification sensitivity is high (98.3%). This model is useful to reduce the false positives rate in lung nodules detection.</td>
</tr>
<tr>
<td>Armato et al. [26]</td>
<td>Lung cancer: performance of automated lung nodules detection applied to cancers missed in a CT screening program</td>
<td>(1) Three main stages involved: 2D processing, 3D analysis and feature extraction. (2) Segmentation, multiple gray-level thresholds, group voxels of contiguous structures, reduce false using rule-based scheme and apply automated classifier were applied.</td>
<td>This paper explained about how radiologists missed the nodules. the sensitivity (84%) can still be improved.</td>
</tr>
<tr>
<td>Retico at al. [27]</td>
<td>Pleural nodule identification in low dose and thin-slice lung computed tomography</td>
<td>(1) Identification of pleural nodules in low dose and thin slice CT. (2) Used directional gradient concentration method, combined with morphological opening-based procedure, analyzed by rule-based filter and neural classifier. (3) 42 CT datasets is used in the evaluation. K-fold cross validation is use to assess the performance of classifier.</td>
<td>The sensitivity is not high and the developed CAD system could not be compared by using publicly datasets as the characterization of nodules types is not determined clearly.</td>
</tr>
</tbody>
</table>
4 Intelligence Computational Models and CAD

From the previous section, it was found that computational intelligence techniques were commonly used by the research in lung nodules detection. Artificial intelligence is considered one of new computer science fields which mimics biological intelligence that involve the computers learn and adapt in new situation [28]. There are several typical computational intelligence approaches such as artificial neural network (ANN) [21-23, 25, 29], genetic algorithm, fuzzy logic [24, 28] and artificial immune system (AIS) [30-32] have been applied in the analysis medical imaging. Although the AIS was seldom to be applied in the detection of lung nodules on CT images, it is expected this new paradigm can also contribute in this domain of pattern recognition. There was several application AIS had been successfully conducted to analyze the medical imaging such MRI and chest tomography (X-rays) for different types of human organ [30-32]

4.1 Artificial Neural Network (ANN) and CAD

Artificial neural network (ANN) is one of the well-known intelligence approaches in identifying and solving problem on medical images. The advantages of ANN algorithms include the parallel process of information and learning from input data or training sets to achieve target results. Basically, the application of ANN in medical imaging are categorized into 3 basic types: Hopfield nets, multilayer feed-forward neural network and self organizing maps (SOM)[1]. In recent years, ANN had been extensively developed to solve different types of problem on CT lung images. For instances, Suzuki et al., [21, 22, 25, 29] had developed an ANN model for pattern recognition in medical imaging, which named as Massive Training Artificial Neural Network (MTANN). This model is built based on multilayer ANN, called neural filters. To differentiate lung nodules and non-nodules (vessels) in low dose CT scans, multiple MTANN was well-trained by a large number of input sub-regions with teaching images [25]. This method was also able to reduce the false positives by eliminating and non-nodules various types of non-nodules with high sensitivity. Subsequently, MTANN was trained to differentiate between benign and malignant nodules in low dose CT scan[21]. Suzuki and Doi [22] had demonstrated that the MTANN could perform well by using a very small number of training cases: 10 nodules and 10 non-nodules. They concluded that the MTANN contributed high generalization due to a huge number of input and teaching pixels with only small amount of cases involved. The uniqueness of this method is no segmentation involved in the process of lung nodules classification. Further, Suzuki [29] developed
a supervised filter in MTANN for the enhancement of actual lesion to improve the sensitivity and specificity of CAD scheme.

5 Current Challenges of CAD

With CT scans, non-invasive visualization of lung condition through medical images has been improved significantly with large numbers of datasets. In the assessment of lung nodules, current CADs are needed to detect the lung nodules and also categorize the types of nodules, for example malignant and benign. However, classification and quantification of small lung nodules (less than 5mm) are still needed to be improved based on the accuracy of segmentation. There are some other challenges in the segmentation of the lung nodules such as nodules close to pleural, low contrast lung nodules that connect to blood vessels and airways, Hounsfield value of nodules similar to blood vessels and airways and both densities as well as shapes of lung are altered by abnormalities [5, 6, 12]. Consequently, the precision of quantification for these small and hidden lung nodules based on CT images will be affected by the partly solid types and ground glass opacities (GGO) [12]. Therefore, the enhancement of segmentation and quantification of lung nodules is crucial to improve the overall performance of CAD scheme.

6 Conclusions

The development of CAD would have a wide and bright future as it is effectively decrease the daily workload of radiologists to interpret of huge datasets of screening [12]. For research purpose, enhancement CAD will be undeniably simplifying and contributing the process of lung cancer detection and diagnosis. The enhancement of automated segmentation and quantification algorithms is expected to contribute higher accuracy and robustness results in the lung nodules analysis. As a preferable approach, computational intelligence method is used increasingly in medical imaging analysis. Besides, current researches are now focusing on detecting small lung nodules and nodules near to pleural. From the reviews, researchers are still using their own datasets for CAD evaluation. Therefore, the performance of this computer system should be tested by using publicly datasets which provided by (LIDC), Image Database Resource Initiative (IDRI), Early Lung Cancer Action Program (ELCAP) and Medical Image Database ELCAP to ensure the generalization of algorithms.
ACKNOWLEDGEMENTS.

Financial support provided by UTM Zamalah, Universiti Teknologi Malaysia is gratefully acknowledged.

References


